I-MEDIA ’07
I-SEMANTICS ’07

International Conferences on New Media Technology and Semantic Systems

I-MEDIA '07

I-SEMANTICS ’07

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

I-MEDIA '07 and I-SEMANTICS '07
International Conferences on New Media Technology and Semantic Systems

Preface

This volume contains the proceedings of I-MEDIA '07 and I-SEMANTICS '07, which are part of the TRIPLE-I conference series. TRIPLE-I reflects the increasing importance and convergence of knowledge management, new media technologies and semantic systems.

I-MEDIA '07 addresses the latest scientific achievements in new media technologies. The contributions published in these proceedings range from success factors for user generated content to interactive TV services to semantic multi-media tagging.

I-SEMANTICS '07 addresses latest scientific results in semantic systems and complements these topics with brand new research challenges in the area of social software, semantic content engineering, logic programming and Semantic Web technologies.

Both conferences welcome leading researchers and practitioners who present their ideas to more than 500 TRIPLE-I attendees. An international program committee selected 50 full papers and six short papers to be included in the conference proceedings.

In addition to a high-quality program, networking is considered an important element of the conference. Community building is fostered in special interactive events designed around particular themes. Deliberately long breaks throughout the conference and social events in the evenings provide excellent opportunities for meeting people from all over the world.

The program of I-MEDIA '07 and I-SEMANTICS '07 is structured as follows. In thematically focused events, community building is supported. Experts and an interested audience are brought together for an in-depth discourse in focused thematic areas. Two special tracks on *Semantics in Life Science* and *Reasoning and Deduction Systems* as well as a workshop on *Multimedia Metadata Applications* are examples for such events.

These special events are complemented by presentations covering current trends and latest developments in new media technologies and semantic systems. The presentations include but are not limited to the following areas:

- Semantic Desktop
- Semantic Social Software
- Semantic Modeling
- Media Convergence
- Media Semantics
- Social Networks

Events such as the I-MEDIA '07 and I-SEMANTICS '07 require active support at different levels: We are grateful to our three invited keynote speakers, *Martin*

Eppler (University Lugano, Switzerland), *Peter Reiser* (SUN Microsystems, Switzerland) and *Marc Smith* (Microsoft Research, USA) for sharing with our attendees their ideas about the future development of knowledge management, new media technologies and semantic technologies. Many thanks go to all authors who submitted their papers and of course to the program committee for their careful reviews. The contributions selected by our program committee are published as a special issue of the *Journal of Universal Computer Science* (J.UCS) which supports the open access initiative for scientific literature and thus ensures the knowledge transfer towards the community. Revised versions of the best contributions will also appear in the *Computing Intelligence Series* of Springer New York.

We would like to thank the sponsors: insiders Wissensbasierte Systeme, punkt. netServices, Sun Microsystems, UMA Information Technology, edicos, Pars Group and Go International – an initiative of the Austrian Federal Economic Chamber und Federal Ministry of Economics and Labour. Special thanks also go to Dana Kaiser for preparing and ensuring the high quality of the conference proceedings. We also would like to thank our staff at the Know-Center, the Semantic Web School and Salzburg NewMediaLab for their continuous efforts and motivation in organizing these two conferences. The local organization of I-MEDIA and I-SEMANTICS would not have been possible without the support of Anke Beckmann, Gisela Dösinger, Patrick Höfler, Alexander Stocker and Anita Wutte – many thanks to all of them.

We are convinced that I-MEDIA '07 and I-SEMANTICS '07 will provide you with new ideas for your research and with new opportunities for partnerships with other research groups.

Sincerely yours,

Klaus Tochtermann, Werner Haas,
Frank Kappe, and Arno Scharl
Conference Chairs of I-MEDIA ´07

Tassilo Pellegrini and
Sebastian Schaffert
Conference Chairs of I-SEMANTICS ´07

Graz, Salzburg, Vienna August 2007

Program Committee I-MEDIA '07

- Yannis Avrithis, NTUA/ITI, Germany
- Thomas Becker, Buhl Data Service, Germany
- Susanne Boll, University Oldenburg, Germany
- Elizabeth Chang, Curtin University of Technology, Australia
- Gisela Dösinger, Know-Center Graz, Austria
- Werner Haas, Joanneum Research, Austria
- Jörg Heuer, Siemens Corporate Technologies, Germany
- Paola Hobson, Motorola, United Kingdom
- Frank Kappe, Graz University of Technology, Austria
- Ralf Klamma, RWTH Aachen University, Germany
- Stefanos Kollias, NTUA, Germany
- Yannis Kompatsiaris, NTUA, Germany
- Harald Kosch, University of Passau, Germany
- Wei Liu, University of Western Australia, Australia
- Robin Mansell, London School of Economics, United Kingdom
- Jamie Murphy, University of Western Australia, Australia
- Markus Rohde, University of Siegen, Germany
- Stefan Rüger, Open University, United Kingdom
- Arno Scharl, MODUL University Vienna, Austria
- Peter Scott, Open University, United Kingdom
- Peter Stanchev, Kettering University, USA
- Marc Smith, Microsoft Research, USA
- Alexander Stocker, Know-Center Graz, Austria
- Martin Wolpers, KU Leuven, Belgium
- Volker Wulf, University of Siegen, Germany
- Raphael Troncy, CWI, Netherlands
- Jörg Westbomke, HDM Stuttgart, Germany

Additional Reviewers for I-MEDIA '07

- Astrid Dickinger, MODUL University Vienna, Austria
- Martin Ebner, Graz University of Technology, Austria
- Daniel Peintner, Siemens AG, Austria
- Hermann Stern, Graz University of Technology, Austria
- Steffen Zorn, University of Western Australia, Australia

Program Committee I-SEMANTICS ‘07

- Sören Auer, University Leipzig, Germany
- Sacha Berger, Ludwig-Maximilians-Universität München, Germany
- Chris Bizer, Freie Universität Berlin, Germany
- Andreas Blumauer, Semantic Web School, Austria
- Tobias Buerger, Salzburg Research, Austria
- Jorge Cardoso, University of Madeira, Germany
- Isabel Cruz, University of Illinois at Chicago, USA
- Jörg Diederich, L3S Research Center, Leibniz Universität Hannover, Germany
- Peter Dolog, Department of Computer Science, Aalborg University, Denmark
- Erich Gams, Salzburg Research, Austria
- Norberto Fernandez, Garcia Telematic Engineering Department. Carlos III University of Madrid, Espania
- Georg Güntner, Salzburg Research, Austria
- Siegfried Handschuh, DERI Galway, Ireland
- Tom Heath, KMI, The Open University, United Kingdom
- Knud Hinnerk Möller, DERI Galway, Ireland
- Josef Holy, Sun Microsystems, Czech Republik
- Malte Kiesel, Deutsches Forschungszentrum Künstliche Intelligenz, Germany
- Jiri Kopsa, Sun Microsystems, USA
- Christoph Lange, Jacobs University Bremen, Germany
- Enrico Minack, L3S Research Center, Leibniz Universität Hannover, Germany
- Daniel Mitterdorfer, Salzburg Research, Austria
- Eyal Oren, DERI Galway, Ireland
- Paula Patrajan, University of Munich, Germany
- Tassilo Pellegrini, Semantic Web School, Austria
- Axel Rauschmayer, Ludwig-Maximilians-Universität München, Germany
- Harald Sack, Friedrich-Schiller-Universität Jena, Germany
- Roman Schindlauer, Vienna University of Technology, Austria
- Daniel Schwabe, Dept. of Informatics, PUC-Rio, Brasilia
- Elena Simperl, DERI Innsbruck, Austria
- Max Völkel, FZI Forschungszentrum Informatik Karlsruhe, Germany
- Denny Vrandecic, AIFB, University of Karlsruhe, Germany
- Christoph Wieser, Salzburg Research, Austria

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Social Media Maps of Online Communities

Keynote Speaker
Marc Smith
(Microsoft Research, USA
marc.smith@microsoft.com)

Social media, the products of the growing read/write web, leaves traces behind that can be gathered, processed and visualized. These pictures of individual and collective authorship patterns illuminate the ecologies of online communities, highlighting the roles people play as leaders, answer people, discussion participants, questioners, spammers, and flame warriors. An ecological model of these roles focuses attention on the need for integration of multiple roles for successful communities. Mobile devices are changing the community scene further, creating new forms of input generated by passive sensors, leaving behind detailed trails that can themselves be shared, searched, and aggregated.

About Marc Smith

Marc Smith is a Senior Research Sociologist leading the Community Technologies Group at Microsoft Research in Redmond, WA. His group focuses on computer-mediated collective action. Marc Smith studies and designs enhancements for social cyberspaces, in particular he is interested in the emergence of social organizations like communities in online conversation and annotation environments. With Peter Kollok he co-edited the book Communities in Cyberspace (Routledge). Co-edited with Peter Kollock, the book explores identity, social order and control, community structures, dynamics, and collective action in cyberspace.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Community Equity – How to Implement and Measure a Social Capital System in Enterprise Communities

Keynote Speaker
Peter Reiser
(Sun Microsystems, Switzerland
peter.reiser@sun.com)

Web 2.0 is a user centric phenomenon. People can express themselves through blogs, videos, podcasts etc. The community decides what is hot through means like rating, voting and commenting, creating a dynamic social value system were good content bubbles up and less interesting content bubbles down. But what are the opportunities of Web 2.0 for an enterprise? How can a company take advantage of the social dynamics of Web 2.0 and turn it into a business value and business advantage? Peter Reiser's talk describes the building blocks of a Community Equity Measurement system that SUN Microsystems is implementing as part of its Customer Engineering CE2.0 project.

The architecture is based on reusable web services, standard protocols (Restful, ATOM, WebDAV) and extensible set of web widgets. The community life cycle methodology describes the tools, roles and psychological dynamics of a community and provides a cook book how to build, sustain and archive measurable communities. The objective is to build a dynamic Social Capital system by measuring the contribution, participation, skills and roles equity a person can gain by actively engage in communities.

About Peter Reiser

Peter H. Reiser is a Principal Engineer at Sun Microsystems. He is currently leading the Web2.0 and social network implementation for the global technical community at Sun. Prior to this he was responsible for the Knowledge & Intellectual Capital Management for Sun's Global Sales Organization. In Sun Europe he established the first community based solution practices and was head of the e-Finance Competency Center, where he was responsible for the architecture and implementation of some of largest Internet Banking and e-Commerce solutions in Europe.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Collaborative Knowledge Visualization: Researching and Augmenting Visual Practices in Knowledge Work

Keynote Speaker
Martin Eppler
(University of Lugano, Switzerland
martin.eppler@lu.unisi.ch)

How can knowledge workers integrate their diverse knowledge more productively? How can experts from different fields share what they know efficiently and create breakthrough innovations? To answer this question, this keynote address focuses on one feasible solution: facilitating knowledge integration through collaborative real-time visualization. Recent research results as well as field tests with new tools show that visual practices can improve collaborative knowledge work significantly. But not all interactive visualization methods are equally conducive to knowledge creation or sharing. The framework and concepts presented in this keynote help to identify the key levers (and barriers) for effective visual practices in knowledge work. An outlook on future developments concludes the talk.

About Martin Eppler

Martin Eppler is Professor at the University of Lugano, prior vice director of the institute of media and communications management at the University of St. Gallen where he also led the Competence Center Enterprise Knowledge Medium, and 'Privatdozent' at the University of St. Gallen. His research focus is on information quality and knowledge management, his work experiences regard consulting and the media industry. His latest book publication is "Managing Information Quality" (Springer). A lot of his publications concern knowledge visualisation.

I-MEDIA '07

International Conference on New Media Technology

Editors

Klaus Tochtermann
Werner Haas
Frank Kappe
Arno Scharl

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Perceived Simultaneous Consumption of Media Content Services among Media Aware University Students

Ester Appelgren
(Media Technology and Graphic Arts, KTH CSC, Sweden
ester@kth.se)

Abstract: This exploratory study aims to discuss issues on media measurement in relation to concurrent media consumption of digital media news services. A survey was conducted using a sample consisting of students from the Media Technology program at the Royal Institute of Technology (KTH), Stockholm, Sweden. The data was mainly analyzed qualitatively as the majority of the survey questions were open-ended. The respondents stated that a combination of the computer with the Internet where the media most common to consume simultaneously with other media such as TV or the mobile phone. The phone was the media most mentioned as prioritized when consuming media simultaneously. The findings thus indicate that simultaneous consumption is common among the media aware technology students in the sample. Simultaneous media consumption might imply for the media companies that competition for the audience attention takes on new forms. This consumption might then affect media companies as their advertisers are becoming more and more aware of the phenomenon. Measures for media consumption have to be adapted to this behavior and previous research suggests that observation methods might be used to detect simultaneous media consumption.

Keywords: Simultaneous media consumption, multitasking, concurrent media consumption, digital media, news services, on-line services
Categories: A.1, H.0, H.m, J.4, J.m.

1 Introduction

There have been reports from researchers in Sweden claiming that time spent with the media, as well as people's daily routines and behaviour in Sweden, has not changed notably since the 80's [Cepaite 06, Bergström 05]. With the introduction of the Internet however, people have found time to use the various services available with, and on the Internet. Nilsson describes how the usage of the Internet in Sweden has increased over the past years [Nilsson 04]. And while people are on the Internet, they try to make the most of it by multitasking. In the nature of multitasking activities lies, that attention may shift from one task to another. In the context of consuming several media at a time, the shifting of attention may furthermore have an effect on less attention to the advertisements presented in the different media channels.

This paper aims to contribute to the discussion on simultaneous media consumption. As media measurements are important when predicting future use of media and also concerning exposure to different content and advertisements, and as current measurements often avoid to describe concurrent media consumption, the primary focus of this study has been to investigate this particular type of simultaneous media usage.

Based on a survey with Swedish Media Technology university students, this study therefore explores how people perceive that they are consuming media simultaneously, and how users' believe they are prioritizing among the different media channels, when consuming several media at the same time.

2 Method

As a first step to detect the occurrence of simultaneous media consumption, a survey with questions on the perceived simultaneous media consumption was conducted using a sample from the population consisting of the students of the Media Technology Masters degree program ranging from first year students to the fourth year students, at the Royal Institute of Technology, Stockholm, Sweden.

The web based survey was carried out during May 2006 and was created in the software tool Bilda[1], mainly used for education purposes. The survey contained questions both open-ended and with multiple options and was mainly analyzed from a qualitative perspective. It was not possible to answer the survey several times, and the participants had to log into the system, thus preventing people from outside of the sample to participate in the survey.
The course list of spring 2006 consisted of 273 students from all four levels of the program. Each student received an instruction e-mail cover letter from the course leader encouraging him or her to participate in the survey. The response data was stored anonymously for each question. A total of 103 students logged into the system. Of these, 86 students responded to the entire survey. The response rate of the survey was therefore 31% (86 of 273). The respondents ranged in age from 19 to 34 years old, with an average age of 23. 36% of the respondents were female and 64 % male.

3 Studies on simultaneous media use

The majority of available statistics in Sweden is based on traditional measures concerning minutes spent on each media on the average day without presenting any overlapping use. The companies MMS[2], Sifo Research International[3], do however measure simultaneous media consumption, focusing primarily on TV consumption in combination with other media. According to Weibull during the 80ies, momentary simultaneous consumption was reported on in the Swedish National SOM[4] investigations [Weibull 07]. Currently however, not many Swedish studies report on this topic. The Swedish institute MMS [Tavakoli 07], measuring people's TV viewing habits, did however recently publish a report on viewer habits concerning all types of moving images, describing multitasking while watching TV or other moving images. In the report a "multitasking index" is presented, describing the number of simultaneous activities taking place while someone is watching moving images at a given time. TV was found to be very flexible in terms of multitasking, and the

[1] www.bilda.kth.se, 2006-06-22.
[2] Mediemätning i Skandinavien, www.mms.se
[3] Sifo Research International, www.research-int.se/
[4] The research institute Society, Opinion and Media at Gothenburg University, Sweden

viewers did on average three different things at the same time, while watching TV [Tavakoli 07].

In North America, several studies have for a few years, repeatedly reported on the growth of acknowledging simultaneous media consumption [Papper et al. 04, Pilotta et al. 04, Alperstein 05]. In this context, these studies emphasize existing problems with the current methods of measuring media use as most media consumption research evaluates use of one medium at a time, these studies also recommend observation methods as a suitable research method for detecting simultaneous media use.

In a review on literature on simultaneous media usage focusing on activities taking place while watching TV, Alperstein [05] reported on multitasking activities such as eating, drinking, dressing, playing or fighting. Many individuals also keep the TV on, while sometimes not even being in the same room as the TV. Alperstein further reports on studies having found that around 40 % of the time people are in front of the TV, they are not looking at the screen. In order to find out how simultaneous media consumption might have an effect on attention, Alperstein therefore conducted a survey where 200 American college students expressed how they engage in concurrent activities while watching television. He found that 88 % of his sample used media simultaneously with other activities and that the most frequent concurrent behaviours while watching TV were: eating (35,5 %) and socializing (22 %). While watching TV as the primary activity, 36 % of the students indicated that they also went online and 5,5 % read a newspaper while watching TV. [Alperstein 05]

Pilotta and Schultz remind us that even though the media environment today is fragmented with a number of alternatives to choose from, people still only have 24 hours a day to spend. According to Pilotta and Schultz, this development thus calls for a need to consume various media simultaneously. Such simultaneous consumption which also can be seen as a form of multitasking, is according to Pilotta and Schultz not new to the society nor to advertisers in media, however it appears to be relatively new to media researchers and planners. Research conducted in the US under the name of the SIMM Studies by the company BIGResearch concerning simultaneous media usage, has found that such usage is widespread in the range of 40 to 65 percent, depending on the specific media combinations [Pilotta and Schultz 2005]. Pilotta and Schultz [2005:20] suggest that as simultaneous media consumption occurs, one media becomes background, both pass one through the other or there is a dissonance. In their study, 48.9 % stated that they pay attention to one medium more than to other(s), 32.1 percent indicated that they attend to each media form equally at the same time, while 19 percent said that they do not engage in simultaneous media usage. An example of media consumed in the background is TV, which according to Pilotta and Schultz is a medium that people do not have to look at to “watch”. Drawing parallels to McLuhan, Pilotta and Schultz emphasize that a generation raised on newspapers would develop minds that work linearly, like print, engaging in one activity at a time such as reading a book, then talk on the phone before having dinner. However, a younger generation raised with the TV, where flipping channels or interrupting activities with advertisements are commonplace, would experience the world nonlinearly, and thus prefer to consume media simultaneously. They describe consuming simultaneous media as a nonnarrative form with no beginning, middle, or end, with one media flowing through the other, denoted as synesthesia, where a person can be a “viewer”,

”reader” and “listener” at the same time.

In the American Midwest area, media multitasking has been thoroughly explored in the “Middletown Media Studies” [Papper et al. 2004]. These studies concern media use, and involve three data collecting methods: telephone surveys, diary studies and observation studies. Comparing the data from the diary and telephone studies with the observational data, the researchers draw the conclusion that people spend more than double the time with the media than they think they do. Furthermore, the occurrence of multitasking two media or more at the same time was observed at one quarter of the media day in 23.7 % of the observed population. Papper et al. found that people tend to be better at identifying their use of printed media than their use of broadcast media and computers. The researchers found that people use several media while reading to a greater extent than they had expected, however participants that focused primarily on the TV tended not to read at the same time. Contradictory to what the researchers first believed, while primarily watching the TV, people are not multitasking other media at all. The study also investigated media use outside the home such as at work, in the car and at “a friend’s house” and found that there is significant media use taking place in these locations.

4 Results

This section reviews the results from the survey. First, the respondents’ preferences concerning what media is perceived to be consumed in combination with other media is presented. Second, the results concerning what media is perceived as prioritized while consuming media simultaneously is presented.

4.1 Simultaneous media consumption

The respondents were asked if they ever consume and use several media simultaneously, such as using the computer while the TV is on, and talking on the mobile phone. The 91 % that responded that they do, were furthermore asked to state the media they use, and order these depending on the media they put most of their attention to (Figure 1).

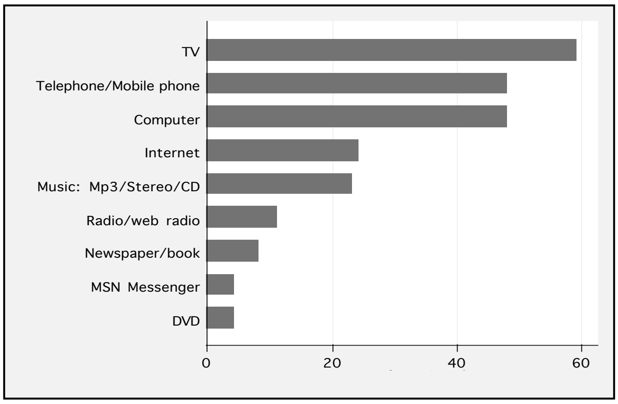


Figure 1: What media do you consume in combination with other media? (N=70)

These results are based on the media types listed by the respondents in free text. Due to the free nature of how the respondents could state the media types, some chose to part activities involving the computer as several activities, such as using the Internet, msn[5], web radio or watching DVDs on the computer. Therefore, even though figure 1 shows that TV was the most frequently mentioned media, the computer with the Internet where the media most common to consume simultaneously with other media, such as TV or the mobile phone. Notable is also that none of the respondents mentioned computer games. The most common combination of several media consumed simultaneously was a combination of the mobile phone, the computer/Internet and the TV, as 69 % of the respondents listed this combination. Of these 69 % respondents, several added a fourth or even a fifth media type to the combination such as mp3, stereo, radio or newspaper.

4.2 Prioritized media when consuming media simultaneously

In first place of the media that the respondents prioritize when consuming media simultaneously came telephone or mobile phone (Figure 2). Other first positions were the computer and the Internet, the TV, the mp3-player and the DVD player. The mobile phone was prioritized over other media in most of the cases, however some respondents stated that this was depending on the person calling. The same argument was made for TV, then concerning how interesting the program was. Several respondents wrote that the prioritized media depends on where you are located in the room, if you are in front of the computer or sitting by the TV.

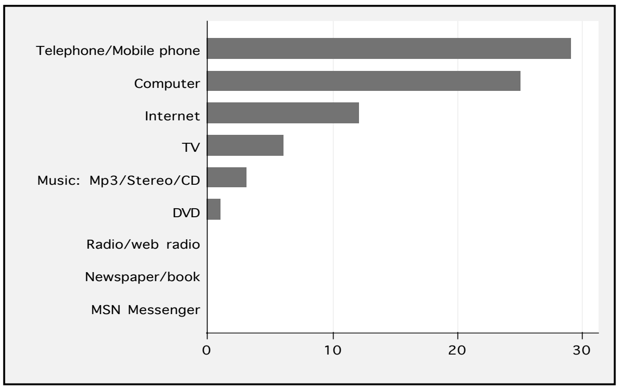


Figure 2: What media do you prioritize when consuming several media simultaneously? (N=70)

Furthermore, some of the respondents described how they only listen to the TV in the background while consuming other media. Music was also described as something to listen to in the background.

[5] Microsoft Network, a messenger service provided by Microsoft.

5 Discussion

Multitasking or simultaneous media consumption is not a new phenomenon, however the notion among researchers of simultaneous media consumption and its implications for the media industry seem to be. The findings of this study show that the sample are aware about their simultaneous consumption of media and that it occurs among the majority of the respondents in the sample. Furthermore, the respondents listed what media they consume simultaneously as could be seen in figure 1, which were: the TV, the telephone/mobile phone, the computer/Internet, music players, radio/web radio, newspaper/book, messenger services and DVD.

Previous studies such as the study performed by Alperstein [05], the Middletown Media studies [04] and the studies made by BIGResearch [05] have focused on a combination of two media consumed simultaneously. Pilotta and Schultz [05] also discussed the use of two media simultaneously, distinguishing between which medium was consumed in the foreground respectively the background. The findings of this study shows that not only do the respondents consume two media simultaneously but often combinations of four media at the same time, such as mp3, stereo, radio or newspaper in combination with mobile phone, computer and the TV where the most frequently mentioned combination of media consumed simultaneously listed by 69 % of the respondents was the combination of the three types of media: the mobile phone, the computer/Internet and the TV, often with a fourth and fifth media added.

Few of the respondents stated that they read newspapers or books while consuming other media. Strangely, of the 7 respondents that stated the media types newspapers and books, none stated that these where prioritized while consuming several media at the same time. Instead, these two media types when listed were almost always listed at the end of the prioritized order. One reason for that newspapers and books seldom were listed in the combinations could be that they the respondents did not consider these two media types as media to consume simultaneously, even if it in reality is highly possible to do so.

Comparing the answers from the sample of the survey to the corresponding statistics of the Swedish population [Nordicom - Sverige 06, Tidningsutgivarna 06] shows that there are differences in terms of media usage on the average day. The average time spent reading a morning newspaper was approximately the same as for the statistics for the Swedish population. The time spent on watching TV on the average day, was lower than the time spent by the Swedish population in the same age group. All the students had access to Internet at the university and most of them at home, and they therefore spent more time on the Internet than the Swedish population in the same age group. Concerning radio listening, the time spent was considerably lower than the average time for the Swedish population in the same age group. Reading on-line newspapers varied across the sample, some do not read on-line newspapers at all while others have been regular readers since the late nineties. The differences with the average population, can be due to that the sample used for the survey are persons interested particularly in media technology as they are studying for a master’s degree in the area. Furthermore, they sample consisted of university students, which also is a group that may differ from the rest of the population in for example interests, time, habits and economy.

Previous research have stressed the importance of using observation methods while studying simultaneous media consumption as participant observation methods might solve the problems with detecting simultaneous media use. Other relevant research methods when detecting the perceived simultaneous media consumption instead of the consumption really taking place are diary methods, letting the respondents describe their media consumption in their own words. This might however not reflect the actual media consumption taking accurately place as previous research has shown that people tend to underestimate their simultaneous media consumption.

6 Conclusion

Studying new information services poses a special problem for researchers since they have no track record [Carey and Elton 1996]. Castells [2007:2-3] describe how futurologists and visionaries have made prophecies about the diffusion and use of wireless technology based on dreams and fears, having caused people, institutions and business to suffer. They recommend researchers to observe the present, and use standard tools of scholarly research to analyze and understand social implications of communication technology. The quote: "Those who have knowledge do not predict. Those who predict do not have knowledge" by Lao-Tzu further emphasizes what Castells [2007:3] describe as visionaries projecting the future based on whatever comes into their minds on the basis of anecdotal observation and ill-understood developments.

Research on media usage and news consumption has previously described usage from a perspective where the media is used one at a time. The findings of this study indicate however that simultaneous consumption is common among the media aware technology students of the sample. The computer in combination with the Internet were at first place as media mentioned as media consumed simultaneously, closely followed by the TV and the mobile phone. In first place of the media that the respondents prioritize when consuming media simultaneously was the telephone or mobile phone. Simultaneous media consumption might imply for the media companies that competition for the audience attention takes on new forms. This consumption might then affect media companies as their advertisers are becoming more and more aware of the phenomenon. Furthermore, aggressive competition in order to keep the audience interested will affect not only how the advertisements are presented to the audience but also how the editorial content is presented. In a longer perspective, simultaneous media consumption further affects the audience concerning non-linear time use.

The results obtained from this study can at this point not be used to draw accurate conclusions from. The sample is too small, it is a snap shot and not based on longitudinal data, and it is not reflecting the population in general. However, it can still be used to discuss opinions and aspects considering simultaneous media usage. Future research should investigate simultaneous media consumption in a more representative sample, in comparison with other countries, and by a deeper analysis of how the consumption takes place, for example using observation methods in combination with interviews. Furthermore, analysis of what types of media publishing channels are best suited for simultaneous media consumption would be of interest. This study is the one of the seven studies concerning future convergence in news production and consumption. The findings will therefore be analyzed further, in comparison with the author's previous case studies at Swedish newspaper companies, concerning the companies' strategies for usage of new digital publishing channels.

References

[Alperstein, 05] N., Alperstein, Living in an age of distraction: Multitasking and simultaneous media use and the implications for advertisers. Communication Dep. Loyola College , Maryland, US, 2005, Available online: http://scholar.google.com/url?sa=U&q=http://mrdadamo.tripod.com/sitebuildercontent/sitebuilderfiles/age_of_distraction.pdf

[Carey and Elton 96] J., Carey and M., Elton, Forecasting demand for new consumer services: challenges and alternatives. In: Dholakia RR, Nundorf N, Dholakia N (eds) New infotainment technologies in the home. Erlbaum, Mahwah, NJ, 1996, pp 35–57

[Castells, 07] M., Castells et al, Mobile communication and society, a global perspective. The MIT Press, Massachusetts Institute of Technology, Cambridge, Massachusetts, 2007.

[Cepaite, 06] A., Cepaite, Medieproducenter och mediekonsumenter i otakt? (Media producers and media consumers out of step?). Thesis for the degree of Doctor at The Royal Institute of Technology, Stockholm, Sweden, 2006.

[Nilsson, 04] Å., Nilsson, Vanliga och mer ovanliga dagars medieanvändning. (Media consumption on usual and more unusual days) In: Ju mer vi är tillsammans, eds Weibull and Holmberg, Nordicom – Sverige, Göteborgs Universitet, 2004.

[Nordicom - Sverige, 06] Nordicom – Sverige, Nordicom-Sveriges Internet barometer 2005, MedieNotiser, Nordicom-Sverige, Göteborgs Universitet, Nr.2, 2006.

[Online Publishers Association, 06] Online Publishers Association, A Day in the Life: An Ethnographic Study of Media Consumption, The OPA White Papers, July 2006.

[Papper et al., 04] R., Papper, et al., Middletown Media Studies. The International Digital Media & Arts Association Journal, Vol. 1, NO. 1, Spring 2004, pp. 3-61.

[Pilotta and Schultz, 05] J., Pilotta and D., Schultz, Simultaneous Media Experience and Synesthesia. Journal of Advertising Research, March 2005, Vol 45: 01, Cambridge University Press. pp 19-26

[Pilotta et al., 04] J., Pilotta et al., Simultaneous Media Usage: A Critical Consumer Orientation to Media Planning. Journal of Consumer Behaviour, Pre-Publication Release, Big Resarch, 2003, 2004.

[Tavakoli, 07] S., Tavakoli, Rörliga bilder 2007, MMS studie om konsumtion på olika plattformar (Moving images 2007, a MMS study in consumption on different platforms). MMS Mediamätning I Skandinavien, March 1, 2007.

[Tidningsutgivarna, 06] Tidningsutgivarna, Statistics from Nordicom-Sveriges Mediebarometer 2005 (adapted), Svensk Dagspress 2006, Fakta om Marknad och Medier. Tidningsutgivarna.

[Weibull, 07] L., Weibull, Professor of Massmedia, Personal Communication, March 28, 2007.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

A Semantics-aware Platform for Interactive TV Services

Antonis Papadimitriou, Christos Anagnostopoulos
Vassileios Tsetsos, Sarantis Paskalis, Stathes Hadjiefthymiades
(Pervasive Computing Research Group, Department of Informatics & Telecommunications
Athens, Greece
{a.papadimitrioy, bleu, b.tsetsos, paskalis, shadj}@di.uoa.gr)

Abstract: Interactive digital TV is becoming a reality throughout the globe. The most important part of the picture is new services for the user, in terms of audio-video quality, but mostly in terms of entirely new content and interacting experience. To this end, we explore the potential of introduction of semantics in the distribution, processing and usage of the media content. We propose a smart iTV receiver framework capable of collecting, extending and processing (reasoning) semantic metadata related to the broadcast multimedia content. System architecture is presented along with an example service to illustrate the combination of semantic metadata, user preferences and external data sources.

Keywords: multimedia, metadata, semantic processing, interactive TV, MPEG-7, MHP, OSGi
Categories: H.5.1 Multimedia Information Systems

1 Introduction

Over the past years, there has been a coordinated effort from multiple organizations towards the establishment of standard transmission technologies and application execution environments for digital interactive TV (iTV). This has led to the development of several standards such as Digital Video Broadcasting (DVB) [1] transmission specifications (for satellite, cable and terrestrial TV) and the Multimedia Home Platform (MHP) [2] as middleware for interoperable interactive applications. There is already an extensive network of digital TV infrastructure based on these specifications, numbering a multitude of homes subscribed to iTV services from different broadcasters. Hence, an open market is evolving with great prospects of benefit both in terms of provider profit and user satisfaction.

The next step to the future of digital TV is the enhancement of interactive applications provided to the iTV subscribers. The platform described in this paper is named POLYSEMA, which is a prototype system which focuses on the development of a "smart iTV receiver" (smart residential gateway) capable of collecting, extending and processing semantic metadata related to the broadcast multimedia content, in order to offer pioneer services to users. In this context, content providers have to supply metadata descriptions along with multimedia content, in a certain metadata standard format (MPEG-7). Broadcasters multiplex metadata information into their transport streams so that smart receivers can take advantage of it. Receivers download metadata from the transport stream and/or the web sites of content providers, and process the information, taking into account user-defined rules and preferences.

POLYSEMA is based on several standards widely adopted by international research and industry communities. Particularly, the DVB-T standard [3] is chosen as the multimedia transmission specification, the MHP standard is chosen as an application execution environment, the OSGi platform [4] as a service-oriented integration platform, the MPEG-7 standard [5] as a metadata description standard and Semantic Web [6] as the base for semantics representation and logic-based inference.

The rest of the paper is organized as follows: Section 2 describes the overall architecture of the POLYSEMA system. In Section 3, we present how the proposed system can be used to support interactive services based on content provider information and user preferences. The following sections consider prior work to that research area and conclude the article by giving some insight for future work.

2 The POLYSEMA Architecture

The core of the POLYSEMA system is the receiver. Yet, in order to exhibit application and video synchronization at the subscriber's receiver, there is a need to add synchronization information at the broadcaster side. In this section, we describe both the novel architecture of the receiver and the additional operations running at the server. It is worth mentioning that POLYSEMA focuses on compatibility with the standards introduced in Section 1. This enables seamless interaction of the system with existing products through the addition of a single component.

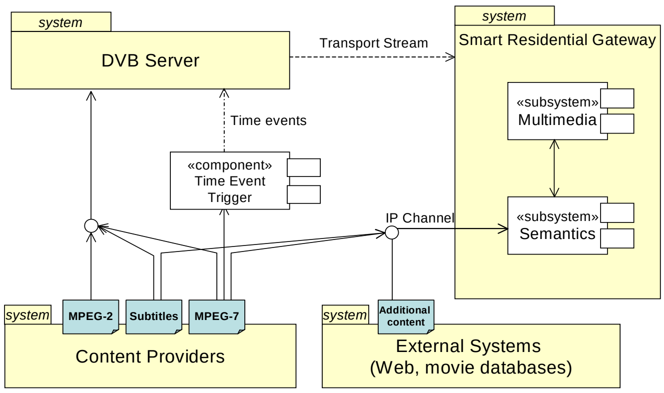


Figure 1: Overall POLYSEMA architecture

Moreover, the media player displaying the multimedia content at the receiver is a common MHP-DVB player. The system interacts with it by using its "default" interfaces (e.g. return channel) and by loading specially tailored MHP applications. Fig.1 depicts the generic components of the POLYSEMA system and their relationship with the broadcasting server, content provider and other external systems.

2.1 Extending the Broadcasting Server

MPEG-7 files can describe multimedia content on a scene-by-scene basis, by providing distinct descriptions for different sections of the video separated by clearly defined time markers. This model allows for a wide range of interactive applications, which adjust their behavior as the content presentation advances. A remarkable problem inhibiting such applications is that, in broadcast environments, the receiver cannot maintain a value depicting the concept of “default media time”. That happens because iTV subscribers may tune to a specific program at any time during the broadcast. In such cases, applications are unaware of the “absolute media time” of a program event. The only way to synchronize video with an application is to transmit *stream events* within the stream, as described in the DSM-CC specification [7].

To overcome this problem, POLYSEMA utilizes a software module to convert MPEG-7 timestamps to stream events sent by the broadcasting server. The module parses the MPEG-7 document, which describes the content to be broadcast, and performs a mapping of MPEG-7 time elements to stream events, at the granularity of video segments. A time event signals the transition to a new scene of the video, which is associated to an MPEG-7 segment. A stream event consists of an *identifier*, *name* and *data*. Data can contain the MPEG-7 segment identifier. Any time the receiver tunes to a channel, the application can determine which part of the video is being displayed by listening for stream events and consulting the MPEG-7 document.

Summing up, the transport stream produced at the server side and sent to the receiver contains: AV content, MHP applications, and the broadcast file system mounted on the transport stream (the broadcast stream may include MPEG-7 files), stream event objects that convey synchronization information, and files in the DSM-CC object carousel, containing a list of URLs, so that the receiver can locate and obtain the corresponding metadata files.

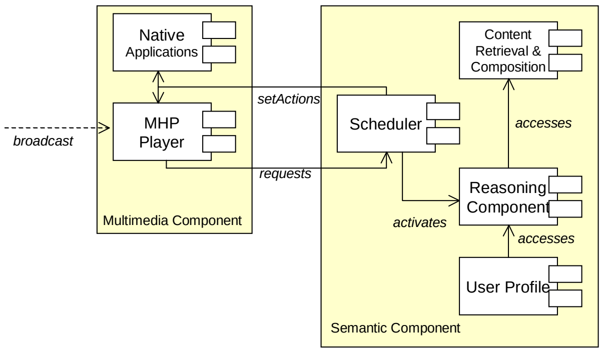


Figure 2: POLYSEMA Software Components

2.2 The POLYSEMA Smart Receiver Architecture

Each interactive service consists of a presentation part and a logic part. The former part refers to actions that the media player ought to perform when it comes to display the results of the latter part. Such actions may be implemented either in a platform-independent way over MHP and Java Virtual Machine, or using native applications. The software module responsible for the media presentation is the Multimedia Component, illustrated in Figure 2. The Semantic Component of POLYSEMA determines the actions which comprise each service. The Semantic Component is also in charge of retrieving and integrating any external resources in the POLYSEMA system. All components are built in an OSGi-based service-oriented fashion, in order to provide maximum flexibility in composing interactive TV services.

2.2.1 The Multimedia Component

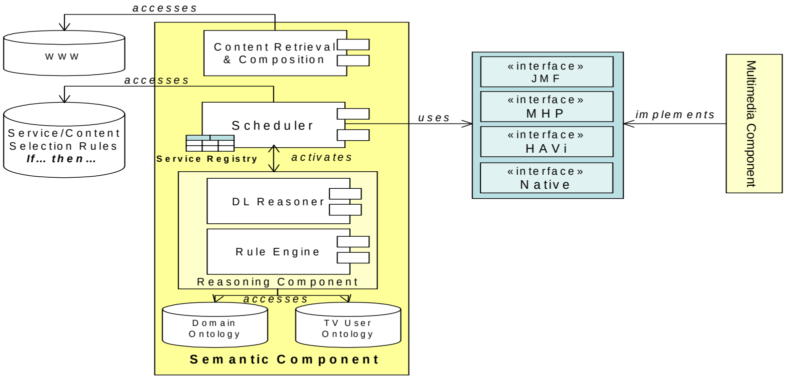
The Multimedia component is the basic constituent of the proposed system that produces multimedia content, interacts with the users and adapts to their preferences accordingly. Specifically, this component refers to the MHP application environment, in which, the MHP application instances run (e.g., on a Set-Top-Box). A user interface is supported through which the user is able to define and place several rules or actions related to their preferences (e.g., on demand multimedia content presentation and retrieval). Moreover, the Multimedia component communicates with the Scheduler of the Semantic component (Figure 2) when further contextual information is required to be fetched and processed. The Semantic component is responsible for a set of actions denoting the specific type of activity that the MHP player has to perform, e.g., change content presentation style and format, tune sound volume, display retrieved info or record part of a program. The Multimedia component can be envisaged as the ingredient that realizes certain actions derived from the reasoning tasks in the Semantic component, as discussed below.

2.2.2 The Semantic Component

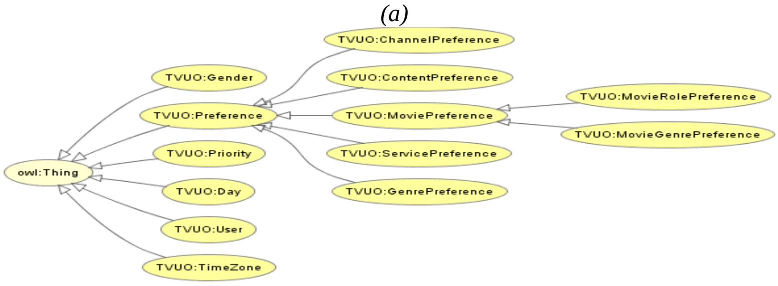
The Semantic component refers to the semantic processing of relevant metadata. A more detailed view of its internal architecture is presented in Figure 3. A basic assumption for our system is that every AV content item is described by an MPEG-7 document. In order to reason over this document we transform it to a corresponding ontology, which was based on that proposed in [13]. In fact, we have developed a stripped down version of the ontology in order to eliminate any unused elements. Moreover, the user "defines" their service preferences by combining templates of possible actions and declaring rules about when such services should launch and how they should be presented. The user input is based on the TV User Ontology (Figure 3b). The Reasoning component of the system uses the MPEG-7 document of the TV program and the user ontology (along with other domain ontologies such as the TV-Anytime classification schemes of MPEG-7) to infer which services should be activated during the broadcast. The Scheduler tells the Content Retrieval & Composition and the Multimedia components what actions they should perform.

The Reasoning component wraps the functionality of a reasoning and a rule engine. The first engine is required mainly for classifying the multimedia content and the user preferences to predefined categories, while the second one is used for deciding which services should be executed given the TV program metadata and the user profile. Bossam [14] is used as both a reasoning and a rule engine. Once the appropriate services have been selected for execution by the Reasoning component, it is the responsibility of the Scheduler component to coordinate the execution of the respective application logic. Such logic is registered in the Service Registry module through procedures specified by OSGi (Figure 3).

The Content Retrieval & Composition component is a framework for registering and managing interfaces with external information sources. For each new source that is registered (e.g., Web site, multimedia database, RSS feed), the available content is described along with its type (e.g., text, video) and the invocation details (e.g., URL, parameters).



(a)



(b)

Figure 3: (a) The Semantics Component, and (b) the TV User Ontology

2.2.3 System Integration with OSGi

POLYSEMA uses both OSGi services and MHP applications. The two platforms exhibit different properties because they follow different design principles, see [8] for a discussion. In order to retain full compatibility with industry standards such as MHP and DVB, and still harvest the service management flexibility of OSGi, we decided to base the component interaction on the IP return channel of the receiver. We incorporated a delegate bundle in the OSGi platform that conceals the nature of MHP applications from the rest of the system. If any other OSGi-based component of POLYSEMA wishes to interact with an MHP application, the MHP-delegate OSGi service should be accessed.

3 Provision of Interactive Services

This section gives an example of an interactive application that could be provided by POLYSEMA, and outlines its implementation. A content provider is assumed to supply the metadata describing the content, while the user enters their profile. POLYSEMA undertakes the responsibility of semantically matching descriptions with profiles, and activate appropriate services. Consider a user who includes in their profile their interest in cars. It could be requested that, in case of appearance of a car in a TV program the system should collect information about it from the Web (e.g.

Wikipedia). Figure 4 depicts an indicative screenshot of such an application. This is just one of the various possible alternative applications, but it illustrates how complex services can be supported by the proposed system.

Figure 4: Screenshot of the interactive service which collects web information.

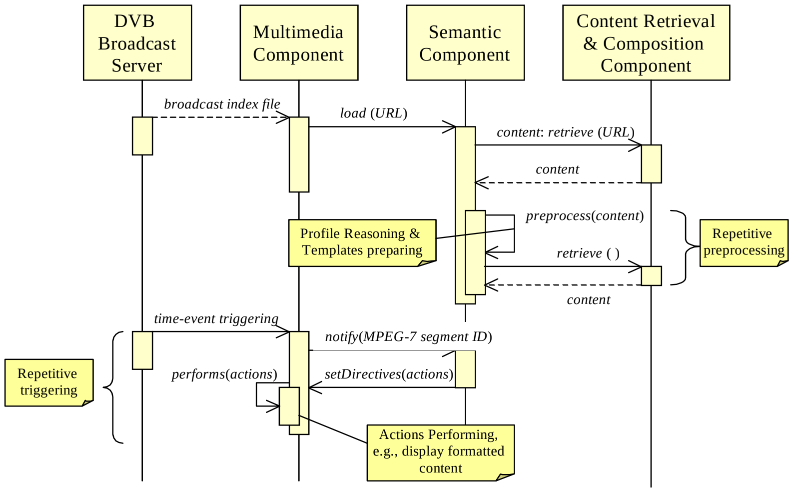


Figure 5: Sequence Diagram of POLYSEMA components functionality

Figure 5 depicts a sequence diagram illustrating the interaction of the Semantic and Multimedia components in order to outline the implementation of the aforementioned application. Specifically, the MPEG-7 document is assumed to contain detailed metadata that annotate the scene in which a car chases a small aircraft (Figure 4). A URL link to the MPEG-7 document is transmitted through the transport stream. At the beginning of the film, the MHP application running at the receiver requests that the Semantic component downloads the MPEG-7 from the Internet. Afterwards, the Semantic component creates an MPEG-7 ontology from the document, processes the descriptions of the film's scenes to match them with the user rules and produces the respective actions. As the broadcast advances, each film scene annotated by the MPEG-7 metadata (e.g., a description of a car make), is eventually displayed on screen. The corresponding time event is triggered at the receiver and the MHP application requests from the Scheduler component to return the desired actions, i.e. the outcome of the preprocessing of the Reasoning component for that media scene. Subsequently, the Semantic component, supplies the desired content through certain layout templates gathered and formatted by the Content Retrieval & Composition component.

4 Prior Work

The AVATAR project discussed in [9] utilizes most of the techniques and technologies that the POLYSEMA project uses, such as video annotation, ontology-based modeling, multimedia metadata, and user profiling/personalization through semantics reasoning. The main objective of this project is to create a personalized digital TV program "recommender", based on the use of TV-Anytime formats and on techniques that have been widely used in the Semantic Web. The project has partial overlap with POLYSEMA, which is focused on a more broad range of applications.

MediaNET [10] is one of the major research efforts concerning multimedia networking. This project is divided in several sub-projects, each of which covers a significant area of the multimedia content creation, service providers and network operators. MediaNET emphasises on the broadcasting issues of the multimedia content and, as far as interactive services are concerned, it delivers the AmigoTV [12] service and a Personal Video Recorder (PVR). The project does not investigate the benefits from using metadata during the various phases of the multimedia content lifecycle. Although POLYSEMA and MediaNET share some infrastructure design issues their focus is on different aspects of the provision of iTV services.

The SAMBITS project [11] is implemented by some of the major telecasting organizations across Europe. The two main objectives of the project is the development of a set of tools for the creation of new interactive multimedia services that use MPEG-4 and MPEG-7 technologies, as well as the development of the technologies that would enable user terminals to access those services. On the other hand, POLYSEMA develops tools for creating multimedia semantics and focuses on the manipulation of such metadata *inside* the residential gateway.

There have been previous attempts to design systems combining both MHP and OSGi platforms. The work in [8] introduces a low-level implementation of a system, which is both OSGi and MHP compliant. The problem with this approach, however, is that, the reference implementation of the MHP platform had to be modified, and, thus, the system does not retain full compatibility with industry standards.

Recent work (see [15] for a comprehensive presentation of the respective ISO ammendment) defines different methods to carry metadata over MPEG-2. Metadata can be sent either by using private sections of MPEG-2, PES packets or the Broadcast File System. The latter approach was preferred, as it allows for in advance loading of the complete metadata, so that it can be timely preprocessed by the semantic component, before respective video scenes arrive. Moreover, in our design, the server can only send links to the metadata and not the entire resource. We believe that this model of metadata transmission is more appropriate, because it saves bandwidth for AV information in the TS, while the metadata can be fetched concurrently from an Internet connection available at the receiver.

5 Conclusions and Future Work

The research work carried out in the context of the POLYSEMA project is driven by the great importance of metadata in providing future iTV services and the need to manage them efficiently. Moreover, the POLYSEMA platform supports applications which adapt their behavior as the content presentation advances, allowing thus for innovative iTV services. Additionally, we believe that more effective personalization can only occur if the preferences of each user are known. This can only be achieved if semantic reasoning process takes place in the end-user premises. Future research may include an even more generic framework for designing services and integrating a variety of external web resources into the TV watching experience.

Acknowledgements

This work was partially funded by the Greek General Secretariat for Research and Technology (GSRT) under the Operational Program “Information Society”, co funded by the EU.

References

[1] The Digital Video Broadcasting Project (DVB), http://www.dvb.org/.

[2] ETSI TS 102 812 V1.2.1, “DVB - Multimedia Home Platform (MHP) Specification 1.1.1”.

[3] ETSI TR 101 190, “Implementation guidelines for DVB terrestrial services; Transmission aspects”, November, 2004.

[4] “About the OSGi Service Platform - Technical Whitepaper Revision 4.0”,www.osgi.org/

[5] Martinez, J, “Standards: Overview of MPEG-7 description tools”, IEEE Multimedia, 2002.

[6] Antoniou, G., and van Harmelen, F. “A Semantic Web Primer”, MIT Press, 2004.

[7] ETSI TR 101 202, v1.2.1, “DVB - Implementation guidelines for Data Broadcasting”.

[8] Vilas, A. F. et al.,“MHP-OSGi convergence: a new model for open residential gateways”, Software Practice and Experience, Vol. 36(13), November, 2006.

[9] Y. Blanco Fernández et al., “AVATAR: An Improved Solution for Personalized TV based on Semantic Inference”, IEEE Trans. on Consumer Electronics. Vol. 52(1), February, 2006.

[10] Serge Travert, Michel Lemonier, “THE MEDIANET PROJECT”, Proc. Image Analysis for Multimedia Interactive Services, Portugal, April, 2004.

[11] SAMBITS Project, “Deliverable 1: Project Description and Plan”, IST1999-12605/Brunel/WP0/PU/P/001, 2000.

[12] AmigoTV: a social TV experience through triple-play convergence, Alcatel/Lucent, 2005.

[13] C. Tsinaraki, P. Polydoros and S. Christodoulakis, “Interoperability support for Ontology-based Video Retrieval Applications” C. In Proc. of Image and Video Retrieval, 2004.

[14] M. Jang, J Sohn, Bossam: an extended rule engine for the web, Proc. of RuleML 2004.

[15] A. Lopez et al. "Synchronized MPEG-7 Metadata Broadcasting over DVB networks in an MHP Application Framework". Proc. of Inrernational Broadcasting Convention, 2003..

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Evaluation of available MPEG-7 Annotation Tools

Mario Döller and Nikolaus Lefin
University of Passau, Germany
Mario.Doeller@uni-passau.de and nikolaus@lefin.net

Abstract: Today the availability of digital media content is well established and widespread. Not only commercial content distribution is a big market, but also user driven digital multimedia content is produced and shared in big communities. There is an increasing need to automatically categorize media such as music or pictures and more general to provide intelligent mechanisms for easier navigation, search and retrieval within huge distributed content repositories. One of the metadata standards that has been established to describe multimedia content via metadata is MPEG-7. This international standard facilitates many application domains and is probably the richest multimedia metadata set available today. However, one of the key issues is to provide an almost automatic annotation of multimedia data in respect to low-level (e.g., color, shape, audio signature, etc.) as well as high-level (e.g., object recognition, event recognition, etc.) features. This work investigates and analyzes currently available MPEG-7 annotation tools and summarizes their applicability and limitations.
Key Words: MPEG-7, MPEG-7 Annotation Tools, Multimedia Metadata
Category: H.3.1, H.5.1

1 Introduction

The ever-growing availability of digital audiovisual material to the user via new media and content distribution methods resulted in an increasing need to automatically categorize digital media. Descriptive information about digital media which is delivered together with the actual content represents one way to facilitate retrieval immensely. The aims of so-called meta data (data about data) are to e.g. detect the genre of a video, specify photo similarity or perform a segmentation on a song, or simply recognize a song by scanning a database for similar meta data. One of the meta data standards that has been established to describe multimedia content via meta data is MPEG-7 [MKP02]. This international standard facilitates many application domains and is probably the richest multimedia meta data set available today.

But, a rich multimedia meta data description on its own is only a first small step to an ultimate multimedia search and retrieval system. Currently, some research concentrates on how multimedia meta data can be stored and accessed effectively (e.g., PTDOM [WK06], MPEG-7 MMDB [KD07]). In addition, efforts are raised for defining standardized access to those databases (see [DWGK06]). A missing part is still an automatic extraction of high-level information (see [HSL+06]) within multimedia data (e.g., object or event recognition, etc.) whereas the extraction of low-level features is well understood. Therefore, we are able to describe multimedia data and to store and access those descriptions. Nevertheless, we are still facing the problem that there is a lack on effective tools assisting a user in annotating multimedia material. For this purpose, this paper evaluates available MPEG-7 annotation tools and summarizes their applicability and limitations.

2 Evaluation Criteria

The criterions presented below serve as basis for the evaluation of all described MPEG-7 annotation tools. The main intention of the chosen criteria is to provide basic decision capabilities to users facing the need for annotating multimedia data with MPEG-7.

1.) **Media:** This evaluation criteria deals with the type of media the respective annotation tool is able to process. Here, on the one side any restriction in the use of different *media formats* (e.g., video, audio, image) is analyzed. On the other side a closer look to supported *compression formats* (e.g., JPEG, MPEG, etc.) is given.

2.) **Domain:** In general, MPEG-7 can be used to support any application domain (e.g., sports, cultural events, etc.). Nevertheless, it has been investigated whether there is any *application dependent* restriction among the annotation tools. Furthermore, the *extensibility* in respect to different application domains has been tested.

3.) **Annotation:** One of the central points of MPEG-7 annotation tools is their ability to annotate any kind of media. Here, the evaluation concentrated on the type of annotation such as *manual, semi-, and/or automatic*. Furthermore, what kind of *feature descriptions* can be extracted. Besides, the *correctness* of the resulting MPEG-7 descriptions has been evaluated.

4.) **Development status:** An important factor for the use of annotation tools is their *system requirement* and how stable (*version*) and advanced the software is.

5.) **User interface:** Available user interfaces are analyzed according its *usability* and the elaborateness of the provided *documentation.*

6.) **Ability to integrate:** As annotation is only one part in an enhanced multimedia process chain, one has to check how the annotated *data* can be *transferred* and what kind of *interfaces* (e.g., for extensibility or database integration, etc.) are provided.

Table 1 provides an overall overview of all annotation tools according to their provided media formats and their supported type of annotation.

Category	Tool	Images	Video	Audio	auto	semi	manual
Visual	Caliph & Emir	★				★	
	M-Ontomat	★				★	
	VideoAnnEx		★			★	
	VIZARD		★		★		
	MARVel	★	★		★		
Audio	Audio Encoder			★	★		
	Audio Analyzer			★	★		
	Audio Low Level			★	★		
Frameworks	MPEG-7 Library	★	★	★	★	★	★
	VizIR	★	★	★	★	★	★

Table 1: Short Overall Analysis

3 Presentation and Evaluation of MPEG-7 Annotation Tools

This section provides an overview of analyzed MPEG-7 annotation tools and states a short description of every application and framework. Due to length restrictions, this paper concentrates only on a subset of available tools. An extended list can be found in [Lef06].

3.1 Visual Tools

- **Caliph & Emir**

  Caliph & Emir [LKG04] are two applications for the annotation and retrieval of images with MPEG-7 descriptors. The project has been initiated at the University of Graz [1]. The tool allows automatic extraction of low-level features of images and supports the user by manual annotation of semantic descriptions. Furthermore, an automatic conversion of EXIF [2] and IPTC IIM [3] is supported as well as semantic annotation by the integrated *Semantic Annotation Panel*.

  Evaluation: As illustrated in table 1, this tool is specialized for annotating images whereas only jpeg and png formats are allowed. The annotation is semi-automatic where all low level features (e.g., color) and available EXIF and IPTC IIM information are extracted automatically. Manual annotation for semantic information is needed whereas a good support is provided by the semantic annotation panel. Unfortunately, the resulting MPEG-7 description

[1] `http://www.tugraz.at/`
[2] `http://www.exif.org/`
[3] `http://www.iptc.org/IIM/`

is not valid (but well-formed) due limitations in the header and *Semantic-Time* objects. The tool runs in Java 5.0 and therefore is available for all operating systems. Furthermore, adequate documentation can be found.

– **M-Ontomat-Annotizer**

The M-Ontomat-Annotizer is part of the *aceMedia* project [HSL+06] for analyzing and annotating multimedia data. It bases on the CREAM framework [HSC02] and is an extension of the OntoMat-Annotizer [HS03]. The tool uses ontologies for supporting the annotation process. These ontologies can be domain specific such as Visual Description Ontology (VDO) for videos or Multimedia Structure Ontology (MSO) which bases on MPEG-7 MDS.

Evaluation: The M-Ontomat-Annotizer provides mechanism for annotating images as well as videos. It is strong in supporting various compression formats (e.g., jpeg, mpeg, mov, etc.) and is highly extensible by the integration of various ontologies. Nevertheless, there are several drawbacks. The extraction of low level features is automatic but every descriptor must be chosen separately. All resulting MPEG-7 descriptors are well-formed but only two are valid. Furthermore, the result of every descriptor (e.g. DominantColor) is stored in a separate file, which reduces usability. The user interface is complex and the software is limited to Windows operating system due the use of Java Media Framework (JMF).

– **IBM Annotation Tool VideoAnnEx**

The IBM MPEG-7 Annotation Tool (short VideoAnnEX) [TLS02] allows the annotation of videos with MPEG-7 descriptions. The granularity of the annotation focuses on shot and frame elements where shots are detected automatically during the load process. Objects within frames can be tagged by bounding boxes which allows a separate region based annotation. The tool supports descriptions about static scenes, events and key objects within shot sequences.

Evaluation: As its name suggests, VideoAnnEx only supports annotation of videos and is limited to MPEG-1/2/4 formats. The annotation process is semi-automatic where only a shot detection is performed automatically and all semantic description has to be realized manually. There is no domain restriction at all and it is arbitrary extensible by events, key objects, etc. All resulting MPEG-7 descriptions are well-formed and valid. The user interface is intuitive and simple but no sufficient documentation is available. The application only runs on Windows operating systems with an adequate hardware setting. Furthermore, the tool does not provide any integration facilities at all.

– **VIZARD - Video Wizard**

The VIZARD tool [RK01] is a video publishing tool which targets on novice users for processing their home videos. It facilitates the annotation of videos by introducing a novel video-book model that provides a structuring in chapters, sections, index, conclusion and so forth.

Evaluation: VIZARD supports annotation of videos of MPEG and AVI compression formats. Currently, the tool is in beta status and run on Windows operating system. The user interface is easy to handle and intuitive but no documentation is available. The annotation mode is semi-automatic (automatic shot detection) whereas no low level features can be extracted. The tool provides support for any domain and is arbitrary extensible by its Lense and Tag concept. The resulting MPEG-7 descriptions are well-formed but not valid due to a *MediaUri* description which is not standard conform. The tool provides no interfaces but enables data transfer by MPEG-7 instance documents and Annotator XML files.

– **IBM MARVel**

MARVel [4] is an image and video retrieval tool that provides automatic indexing and categorizing of image and/or video collections based on integrated content analyze techniques. For instance, the tool assigns the concept *outdoor* when an airplane is detected, and so on. Furthermore, MARVel provides retrieval strategies based on feature descriptors or semantic concepts.

Evaluation: MARVel allows annotation of image and video files where a rich set of compression formats is supported. The tool performs automatic annotation and feature extraction but provides no means for any data transfer as all data is stored in an internal database. Furthermore, the annotation domain is limited to available concepts (e.g., indoor, outdoor) and no extensibility facilities have been documented. The tool runs on Windows operating systems and requires some memory space at the local disk. MARVel does not provide any documentation. The user interface is comfortable and well balanced.

3.2 Audio Tools

– **MPEG-7 Audio Encoder**

The MPEG-7 Audio Encoder [5] is a Java based library for low level feature description of audio data. It has been developed at the *RWTH-Aachen* university in cooperation with *Universita Politecnica delle Marche*, Italy. The tool is available as a console based stand alone version as well as a web based

[4] `http://mp7.watson.ibm.com/marvel/`
[5] `http://mpeg7audioenc.sourceforge.net/`

application.

Evaluation: The Audio Encoder is one of the most sophisticated audio annotation tools among the evaluated ones. It supports a large number of compression formats and provides an automatic annotation and low level feature extraction (all MPEG-7 audio descriptors are available). The resulting MPEG-7 descriptions are well formed and valid. The project is still active and uses Java. It's jar file can be integrated to other projects and the data transfer is realized by MPEG-7 instance documents.

– **MPEG-7 Audio Analyzer**

The MPEG-7 Audio Analyzer [6] has been developed at the *Technische Universität Berlin* and allows the extraction and annotation of all available MPEG-7 low level audio descriptors. Unfortunately, only a web based online version is available.

Evaluation: The Audio Analyzer is restricted to WAVE and MP3 compression formats and supports an automatic extraction of all low level audio features. Unfortunately, the input audio file is restricted in size (1 MB for WAVE and 300kb for MP3). The resulting MPEG-7 description is well-formed but not valid. The project is still active but does not provide any documentation. In addition, there are no interfaces available and the data transfer is realized by MPEG-7 instance documents. There are no restrictions in regard to domain or extensibility.

– **MPEG-7 Audio Low Level Descriptors**

The MPEG-7 Audio Low Level Descriptors calculator [7] has been developed at the University of Wollongong, Australia. Unfortunately, no further information (papers, documentation) about its mode of operation could be found.

Evaluation: This project is very similar to the Audio Analyzer but is more restrictive. It does not support the annotation and extraction of all audio low level feature descriptions. The resulting MPEG-7 descriptions are not valid and one can found the same limitations for the input audio files. Further, the project is not active anymore and no documentation can be found.

3.3 Frameworks

– **Joanneum MPEG-7 Library**

The MPEG-7 Library [8] is a C++ implementation of the MPEG-7 standard ISO/IEC 15938:2001 freely provided by the *Joanneum Research Forschungs-*

[6] `http://mpeg7lld.nue.tu-berlin.de/`
[7] `http://www.whisper.elec.uow.edu.au/mpeg7/`
[8] `http://iiss039.joanneum.at/cms/index.php?id=84`

gesellschaft mbH at Graz, Autria. The current implementation supports Windows as well as Unix operating systems. The library focuses on manipulating, validating, creating etc. of MPEG-7 descriptors. It does not contain feature extractors. Therefore, the library supports developers in XML-DOM programming of MPEG-7 documents.

Evaluation: The library supports annotation features for all available media (image, video, audio) and almost all MPEG-7 description schemes and descriptors (around 1200 classes). The implementation is sophisticated and arbitrary extensible. Besides, the use of different compression formats must be implemented individually. The project bases on C++ and the .NET 2003 platform and is currently available for Windows (the source code also compiles at UNIX). Furthermore, it includes a rich documentation and provides a large example file repository for almost every implemented descriptor.

- **VizIR**

  The VizIR [EB02, EB03] framework has been developed at the *Technische Universität Wien*, Austria. The project provides an open extensible framework containing approaches for feature extraction, distance metrics and components of user interfaces. The main targets of the framework are the creation of a basis for content based retrieval applications and for research on new approaches for automatic content extraction within images and videos.

  Evaluation: The VizIR framework is at an early alpha stage and bases on Java and JMF. It supports all media (image, video, audio) and provides a rich set of compression formats (due the use of JMF). Currently, the project offers an API for around 800 MPEG-7 types (all audio, visual and MDS descriptors and description schemes). The documentation is sufficient (API based) and data transfer is guaranteed through MPEG-7 instance documents and the Mind Reading Markup Language (MRML).

4 Conclusions

This paper evaluated all found MPEG-7 annotation tools and summarized their applicability and limitations. The existing tools were divided into three groups: tools only supporting visual data, tools supporting audio data and frameworks. In general, the support for automatic or at least semi-automatic annotation of low- and high level features is rather poor. This is especially the case for all audio annotation tools where no semantic annotation is supported. In addition, only few tools are able to produce valid MPEG-7 descriptions which limit their usability. Some highlights present Caliph in annotating images and VideoAnnEx for videos, whereas the annotation of video is in most cases limited to shot detection. In the audio domain, the Audio Encoder is most sophisticated among the tested tools. Recently, the development of MPEG-7 frameworks promise to improve the current situation as they can serve as basis for more advanced tools as they do not specialize on small subareas of the standard. Furthermore, several EU supported projects exists (e.g., K-space [9]) whose main goal is research and development of tools for semi-automatic annotation and retrieval of multimedia content.

References

[DWGK06] Mario Döller, Ingo Wolf, Matthias Gruhne, and Harald Kosch. Towards an MPEG-7 Query Language. In *Proceedings of the International Conference on Signal-Image Technology and InternetBased Systems (IEEE/ACM SITIS'2006)*, pages 36–45, Hammamet, Tunesia, 2006.

[EB02] Horst Eidenberger and Christian Breiteneder. A Framework for Visual Information Retrieval. In *Proceedings of Visual Information Systems Conference*, pages 105–116, HSinChu, Taiwan, 2002.

[EB03] Horst Eidenberger and Christian Breiteneder. VizIR - a framework for visual information retrieval. *Journal of Visual Languages & Computing*, 14(5):443 – 469, 2003.

[HS03] Siegfried Handschuh and Steffen Staab. CREAM - CREAting Metadata for the Semantic Web. *Computer Networks*, 42:579–598, 2003.

[HSC02] Siegfried Handschuh, Steffen Staab, and Fabio Ciravegna. S-CREAM Semi-automatic CREAtion of Metadata. In *Proceedings of SAAKM 2002 - Semantic Authoring, Annotation & Knowledge Markup, ECAI 2002 Workshop*, pages 27–34, Lyon, France, 2002.

[HSL+06] Jonathon S. Hare, Patrick A. S. Sinclair, Paul H. Lewis, Kirk Martinez, Peter G. B. Enser, and Christine J. Sandom. Bridging the Semantic Gap in Multimedia Information Retrieval: Top-down and Bottom-up approaches. In *Proceedings of the First International Workshop on Semantic Web Annotations for Multimedia (SWAMM)*, Edinburgh, Scotland, 2006.

[KD07] Harald Kosch and Mario Döller. The MPEG-7 Multimedia Database System (MPEG-7 MMDB). *Journal of Systems and Software, Accepted for publication, In Press by Elsevier. To appear in spring 2007*, 2007.

[Lef06] Nikolaus Lefin. Vergleich bestehender MPEG-7 Annotierungstools. Technical report, University of Passau, Passau, Germany, 2006.

[LKG04] Matthias Lux, Werner Klieber, and Michael Granitzer. Caliph & Emir: Semantics in Multimedia Retrieval and Annotation. In *Proceedings of the 19th International CODATA Conference 2004: The Information Society: New Horizons for Science* , pages 64–75, Berlin, Germany, 2004.

[MKP02] J. M. Martinez, R. Koenen, and F. Pereira. MPEG-7. *IEEE Multimedia*, 9(2):78–87, April-June 2002.

[RK01] H. Rehatschek and G. Kienast. Vizard - an innovative tool for video navigation, retrieval, annotation and editing. In *Proceedings of the 23rd Workshop of PVA Multimedia and Middleware*, Vienna, Austria, 2001.

[TLS02] Belle L. Tseng, Ching-Yung Lin, , and John R. Smith. Video Personalization and Summarization System. In *Proceedings of the SPIE Photonics East 2002 - Internet Multimedia Management Systems*, Boston, USA, 2002.

[WK06] Utz Westermann and Wolfgang Klas. PTDOM: a schema-aware XML database system for MPEG-7 media descriptions. *Software: Practice and Experience,*, 36(8):785–834, 2006.

[9] `http://www.k-space.eu/`

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Applying Media Semantics Mapping in a Non-linear, Interactive Movie Production Environment

Michael Hausenblas
(Institute of Information Systems & Information Management,
JOANNEUM RESEARCH Forschungsgesellschaft mbH
Steyrergasse 17, 8010 Graz, Austria
michael.hausenblas@joanneum.at)

Abstract: In this work we propose how to deal with the Semantic Gap in closed domains. That is, we propose to bridge the Semantic Gap by means of mapping well-known low-level feature patterns found in MPEG-7 descriptions to formal concepts. The key contributions of the proposed approach are (i) the utilisation of ontologies, and rules to enhance the retrieval capabilities (effectiveness), and (ii) the realisation of the feature matching process being carried out on the structural level through indexed MPEG-7 descriptions (efficiency). We discuss advantages and shortcomings of our approach, and illustrate its application in the realm of non-linear, interactive movie productions.
Key Words: Non-linear interactive media, Media Semantics, Semantic Gap
Category: H.5.1, H3.1

1 Introduction

One key problem of multimedia content understanding—bridging the **Semantic Gap**—still is not satisfactory solved. Following [Smeulders et al. 2000], the Semantic Gap is "the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation". In the realm of non-linear, interactive movie productions, one major challenge is the dynamic matching of appropriate clips based on a formal expression describing the desired content. In our setup, a movie is—based on the users interaction—assembled on-the-fly, requiring the retrieval of the audio-visual content to be performed in near-real-time. For example, in one point of the narration, there could be a query for some material that *is about soccer*, *has an interview* in it and *starts with a PAN LEFT* camera motion.

We believe that our approach—driven by requirements that stem from a non-linear, interactive production environment—can offer a sound solution to the problem stated above.

In the next section we start with a discussion on related work and give a short overview on the production environment. We then discuss the theoretical underpinnings of our proposed solution, the *Media Semantics Mapping* in Section 3. In the following Section 4 we present the current (prototypical) application of the approach. Finally, we conclude on the work done so far and report on the next planned steps in Section 5.

2 Related Work and Environment

Thoughts on how to conceptually bridge the Semantic Gap are probably as old as the multimedia content itself [Grosky 1994]. Most of the work in the realm of multimedia content representation focuses on the integration of multimedia metadata—as MPEG-7 [MPEG-7 2001]—with logic-based ontology languages, typically OWL [OWL 2004].

The constitutive work of Hunter et.al. [Hunter 2001] has led to numerous related efforts [Troncy 2003, Garcia and Celma 2005] that all share the translational approach of mapping MPEG-7 to OWL. For the field of ontology-based video retrieval, for example [Tsinaraki et al. 2004] reports a methodology to support interoperability of OWL with MPEG-7[1].

Media Streams—developed by Davis [Davis 1995] in his PhD thesis— is a system for annotating, retrieving, repurposing, and automatically assembling digital video. It uses a stream-based, semantic representation of video content with an iconic visual language interface of hierarchically structured, composable, and searchable primitives. Nack and Putz presented the Authoring System for Syntactic, Semantic and Semiotic Modelling (A4SM) framework [Nack and Putz 2001] that includes the creation, and retrieval of media material. The project goal was to have a framework at hand that would allow for semi-automated annotation of audiovisual objects, and to demonstrate the applicability in a news production environment. Both the Media Streams system and the A4SM can be understood as precursor to our proposed architecture.

Motivated by the promising work reported in [Little and Hunter 2004] and [Hollink et al. 2005] the proposal presented inhere is based on our experiences with MPEG-7 annotation and retrieval [Bailer et al. 2005].

The environment. The New Media for a New Millennium (NM2) project [Rehatschek et al. 2006] targets at the creation of technologies for non-linear, interactive narrative-based movie production. NM2 is an Integrated Project of the EU's 6th Framework Programme running till summer 2007 with 13 partners from eight European countries.

The tools for personalised, reconfigurable media productions are elaborated in six audio-visual productions that range from news reporting and documentaries through a quality drama serial to an experimental television production. Targeted end-user devices are Windows Media Centre-PCs, game consoles, and mobile phones. For a detailed overview on the project objectives, system capabilities and the productions, the reader is referred to [Williams et al. 2006].

[1] For an overview on related MPEG-7 formalisations and multimedia ontologies, the reader is referred to [Hausenblas et al. 2007].

3 Media Semantics Mapping

In this section, we discuss the *Media Semantics Mapping* (MSM) foundations, the terminology used, and the possibilities gained from using this approach.

Modality in our understanding is a path of communication between the human and the computer; major modalities are vision and audition (others are tactition, olfaction, etc.). In this work, audio-visual data is referred to as *essence*, i.e., the actual piece of data that resides e.g. in the file system. A *media item* (MI) is a proxy for some essence and acts as a pivot for attaching low-level features as well as annotations stemming from the domains semantics. In NM2, MPEG-7 [MPEG-7 2001] is utilised for representing low-level features of the essence, as colour descriptors, etc. We head after extracting as much as possible automatically from the essence to produce comprehensive MPEG-7 descriptions based on technologies of our Multimedia Mining Toolbox [Bailer et al. 2005, Section 6]. In the visual domain we use the Dominant Color Descriptor and the Color Layout Descriptor to capture colour features. To describe textures, we make use of the Edge Histogram Descriptor. Shapes can be recognized via the Contour-Based Shape Descriptor. The Camera Motion Descriptor is utilised to describe camera movements (pan, tilt, zoom, etc.). Although a representation of low-level features on the ontological level would be possible, we do not lift MPEG-7 descriptions and description schemes onto the logical level, rather MPEG-7 fragments are referenced from within the ontology.

OWL-DL [OWL 2004] is used to formalise the domain semantics and functions as the interface to the Narrative Structure Language [Ursu and Cook 2005]. A *logical entity* (LE) is anything contained in a MI that can either directly be recognised w.r.t. a modality, or that is not directly observable. A more formal account of the terms is given below.

3.1 Media Semantics

What are media semantics? According to [Harel and Rumpe 2004], any language definition comprises syntax, semantic domain, and a semantic mapping from the syntactic elements to the semantic domain. When talking about media semantics inhere, we subscribe to this point of view. In our understanding the essence itself does not “have” semantics. A piece of essence may be consumed or manipulated, nevertheless, essence “carries” the semantics and it is up to the consumer of the essence to interpret what she understands from it. Hence we do not try to define what in the general case an object “looks like” or “sounds like”. We therefore understand that the ontological constructs in combination with the rules are our syntactical framework, further the semantic domain is conceived as being the domain of the LE that can occur in the essence, and finally define the semantic mapping as described below.

3.2 Spaces of Abstraction

We allow for two orthogonal conceptual paradigms to model media semantics: spaces and the well-known class/instance pattern. A *space* represents a certain level of abstraction, ranging from low-level, as colour or shape to abstract entities such as human feelings. Classes and instances are used to define the actual LE. Therefore "the soccer ball" instance in the context of a soccer game is defined to be black, white, and round but this does not mean that "a ball" in general—referring to the class level—has these properties. Fig. 1[2] depicts the

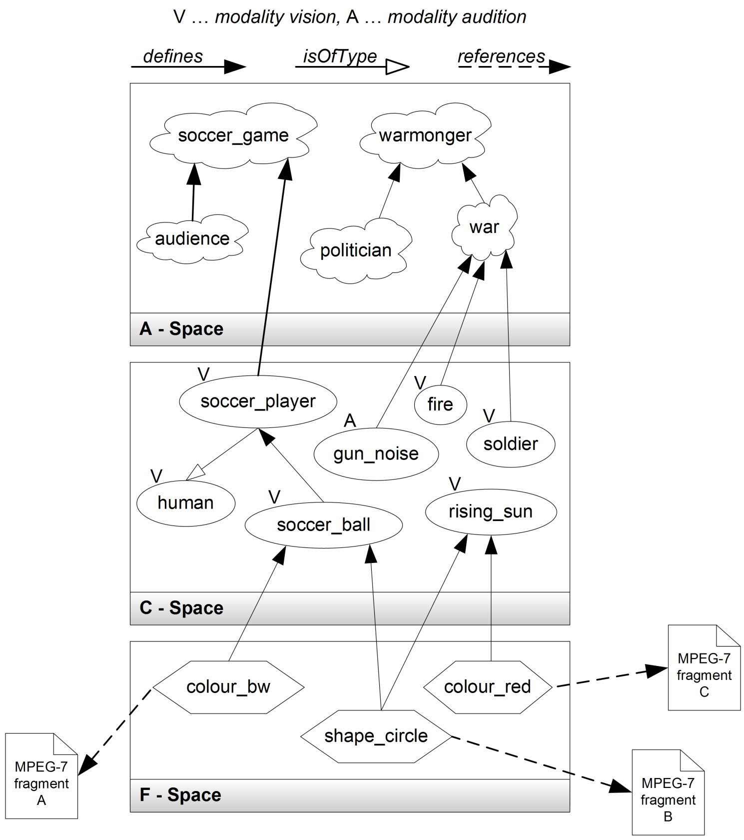


Figure 1: The Media Semantics Modelling Spaces.

spaces available in our approach, with V denoting the modality vision, and A the modality audition. Following [Grosky 1994] we introduce relationships that can

[2] The reader is invited to note that this figure has already been presented at a tutorial—What you Mean is What you Watch: Multimedia and the Semantic Web, cf. `http://gate.ac.uk/conferences/eswc2006/multimedia-tutorial/`—at the ESWC 2006.

hold between an essence and some logical entities. The spaces are defined—listed by increasing level of abstraction—as follows.

1. The Feature Space—**F-Space**. Contains LE that represent low-level features. A low-level feature (LLF) is a single aspect of a certain (spatio-temporal) part of a media item. For example the dominant colour of a spatial region (black and white) is represented as a LLF.

2. The Concrete Logical Entity Space—**C-Space**. Contains LE that can directly be recognized in the essence. A concrete logical entity (CLE) is a distinct object being defined by a combination of low-level features and their respective values (simple CLE) or using other concrete logical entities (composite CLE). For example, in the soccer domain, the CLE soccer ball may be defined by the LLF dominant colour black and white, and circular shape, i.e., a simple CLE. A table could be defined consisting of a CLE tabletop and four CLE table legs, resulting in a composite CLE.

3. The Abstract Logical Entity Space—**A-Space**. Contains LE that are not directly observable. An abstract logical entity (ALE) can be defined by a combination of CLE (simple ALE) or other ALEs (composite ALE). For example the ALE soccer game may be defined by the simultaneous presence of an ALE audience and a number of CLE soccer player.

Within each space, the class-instance modelling can be used to add further semantics, as taxonomies, object relations, etc. In each of the six NM2 productions, a domain-specific ontology is defined covering concepts and instances. This might be 'church', or 'painting' in the case of the documentary production about England's Golden Age in the 16^{th} century, or certain actors, moods, and keywords, as found in the drama Accidental Lovers [Hausenblas and Nack 2007].

Built-in rules. Due to the well-known limitations of DL-ontology languages [Horrocks et al. 2005] we utilise rules in addition to DL-ontologies (see also [Staab et al. 2003]) to define the semantics of a logical entity in the context of a production. However, using rules can lead to serious problems w.r.t. organisational issues. We therefore only provide a minimalistic set of so called built-in rules, and automatically generate the actual rules as described below.

Two properties, defined in the NM2 core ontology, enable the incorporation of rules, hence assisting to define the semantics of a logical entity. The `defines` property allows a combination of `ConcreteLogicalEntity` instances to define either another `ConcreteLogicalEntity` instance (composite pattern) or an `AbstractLogicalEntity` instance, hence an inter-space mapping. For each (partial) `defines`-property in the ABox of the ontology appropriate atoms are added to the corresponding rule.

A media item `contains` a number of `LogicalEntity` instances along with `LLFeature` instances representing an occurrence of a logical entity in a media item. Equally as above, for each occurrence atoms are added accordingly.

An exemplary built-in rule defining the mapping from the F-Space to the C-Space is shown below. Given that a set of low-level features $\{llf_1 \ldots llf_i\}$ `defines` a certain logical entity *cle* (line 1), and it is known that a certain media item *mi* `contains` this set of low-level features (line 2), it can be inferred that *mi* `contains` *cle*.

1 $contains(mi, cle) \leftarrow defines(llf_1, cle) \wedge \ldots \wedge defines(llf_i, cle) \wedge$
2 $\quad contains(mi, llf_1) \wedge \ldots \wedge contains(mi, llf_i)$

However, to ensure the correctness of the definition, some constraints must be put on the variables: $\forall llf_i \in LLFeature$, further $cle \in ConcreteLogicalEntity$, and $mi \in MediaItem$, which highlights the connection to the NM2 core ontology that defines each of the concepts. To enhance the domain-specific ontologies further, so called *user-defined rules* can be manually defined by the user.

The ontology and the rules together form the knowledge base $\mathcal{KB}_{MSM}$, which further is used to annotate the essence automatically. $\mathcal{KB}_{MSM}$ is defined as being a tuple $\langle \mathcal{O}_D, \mathcal{R} \rangle$, with $\mathcal{O}_D$ being an ontology that consists of an ABox and a TBox, and $\mathcal{R}$ a rule-base comprising *built-in rules* and *user-defined rules*.

4 Applying the Media Semantics Mapping

The **M**edia **S**emantics **M**apping Utility (MSM-Utility) is used to define instances based on the built-in rules, described above to generate $\mathcal{KB}_{MSM}$. For managing MPEG-7 documents we use our MPEG-7 Document Server [Bailer et al. 2005, Section 5.2], which provides access to MPEG-7 documents for a number of clients and allows the exchange of whole documents or fragments thereof utilising XPath. Access to parts of documents is crucial for the efficiency of the system, as MPEG-7 documents of larger media items tend to have considerable size. The MPEG-7 documents used in the system are compliant with the Detailed Audiovisual Profile (DAVP) [Bailer et al. 2005].

For processing the ontological information, we use a performant RDF-library, the Redland RDF library[3], wrapped up in an Object-Oriented-API (C++) that enables manipulation and query on the ontological level. Applying $\mathcal{R}$ onto $\mathcal{O}_D$ is done utilising Prolog. $\mathcal{O}_D$ represented in OWL-DL is converted into a number of SWI-Prolog[4] facts.

[3] `http://librdf.org/`
[4] `http://www.swi-prolog.org/`

Typically, users of the NM2 toolkit lay out their production-specific ontologies by means of creating concepts and instances. Through $\mathcal{KB}_{MSM}$ the system is then able to automatically tag the essence in two subsequent steps. Firstly, the low-level features are extracted automatically on the MPEG-7 level. Secondly, $\mathcal{KB}_{MSM}$ is used to match against the generated description of the essence, triggering an update of the ABox of $\mathcal{O}_D$.

5 Conclusion and Outlook

We have shown in this paper how to map low-level features extracted from multimedia essence to logical entities. This enables an effective and efficient retrieval of the essence. Another source for the entity definition process are scripts, shot-logs, etc., which are incorporated through the ingestion process. We also plan to include the support for *guided definitions*. This means to extract MPEG-7 features from a reference image or audio-clip, display the extracted values and let the user select a combination of the extracted values for definition purposes, quite similar to [Little and Hunter 2004].

To allow for queries as “find me all MI with an interview as establishing shot, **followed** by a ZOOM_IN onto a painting”, we currently work on the integration of so called *temporal annotations* to be used within a media item, based on [Allen and Ferguson 1994].

Acknowledgements

The work reported herein was undertaken in the realm of the “New Media for a New Millennium” project (FP6-004124), partially funded under the 6th FP of the European Commission. Special thanks to Werner Bailer and Peter Schallauer for their decent support regarding MPEG-7, and to Rudi Schlatte for useful discussions, as well as to the NM2 production teams. The author would also like to gratefully acknowledge the comments from the reviewers.

References

[Allen and Ferguson 1994] Allen, J. F. and Ferguson, G. (1994). Actions and Events in Interval Temporal Logic. Technical Report TR521, University of Rochester.

[Bailer et al. 2005] Bailer, W., Schallauer, P., Hausenblas, M., and Thallinger, G. (2005). MPEG-7 Based Description Infrastructure for an Audiovisual Content Analysis and Retrieval System. In *Proceedings of SPIE - Storage and Retrieval Methods and Applications for Multimedia*, volume 5682, pages 284–295.

[Davis 1995] Davis, M. (1995). *Media streams: representing video for retrieval and repurposing*. PhD thesis, Massachusetts Institute of Technology.

[Garcia and Celma 2005] Garcia, R. and Celma, O. (2005). Semantic Integration and Retrieval of Multimedia Metadata. In *5^{th} International Workshop on Knowledge Markup and Semantic Annotation (SemAnnot'05)*, Galway, Ireland.

[Grosky 1994] Grosky, W. I. (1994). Multimedia Information Systems. *IEEE MultiMedia*, 1(1):12–24.

[Harel and Rumpe 2004] Harel, D. and Rumpe, B. (2004). Meaningful Modeling: What's the Semantics of "Semantics"? *Computer*, 37(10):64–72.

[Hausenblas and Nack 2007] Hausenblas, M. and Nack, F. (2007). Interactivity = Reflective Expressiveness. *IEEE MultiMedia*, 14(2):1–7.

[Hausenblas et al. 2007] Hausenblas, M., Boll, S., Bürger, T., Celma, O., Halaschek-Wiener, C., Mannens, E., and R. Troncy (2007). Multimedia Semantics on the Web: Vocabularies *W3C Incubator Group Report*.

[Hollink et al. 2005] Hollink, L., Little, S., and Hunter, J. (2005). Evaluating the application of semantic inferencing rules to image annotation. In *3^rd International Conference on Knowledge Capture (K-CAP 2005)*, pages 91–98, Banff, Alberta, Canada. ACM.

[Horrocks et al. 2005] Horrocks, I., Patel-Schneider, P. F., Bechhofer, S., and Tsarkov, D. (2005). OWL Rules: A Proposal and Prototype Implementation. *Journal of Web Semantics*, 3(1):23–40.

[Hunter 2001] Hunter, J. (2001). Adding Multimedia to the Semantic Web - Building an MPEG-7 Ontology. In *First International Semantic Web Working Symposium (SWWS'01)*, Stanford, California, USA.

[Little and Hunter 2004] Little, S. and Hunter, J. (2004). Rules-By-Example - A Novel Approach to Semantic Indexing and Querying of Images. In *3^rd International Semantic Web Conference (ISWC'04)*, volume 3298 of *Lecture Notes in Computer Science*, pages 534–548, Hiroshima, Japan.

[MPEG-7 2001] MPEG-7 (2001). Multimedia Content Description Interface. Standard No. ISO/IEC n15938.

[Nack and Putz 2001] Nack, F. and Putz, W. (2001). Designing annotation before it's needed. In *ACM Multimedia*, pages 251–260.

[OWL 2004] OWL (2004). Web Ontology Language Reference. W3C Recommendation.

[Rehatschek et al. 2006] Rehatschek, H., Hausenblas, M., Thallinger, G., and Haas, W. (2006). Cross media aspects in the areas of media monitoring and content production. In *5th International Conference on Language Resources and Evaluation (LREC) 2006, cross-media indexing workshop*, pages 25–31, Genoa, Italy.

[Smeulders et al. 2000] Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., and Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380.

[Staab et al. 2003] Staab, S., Angele, J., Decker, S., Grosof, B., Horrocks, I., Kifer, M., and Wagner, G. (2003). Where are the rules? *IEEE Intelligent Systems, Trends & Controversies*, 18(5):76–83.

[Troncy 2003] Troncy, R. (2003). Integrating Structure and Semantics into Audiovisual Documents. In *2^nd International Semantic Web Conference (ISWC'03)*, volume 2870 of *Lecture Notes in Computer Science*, pages 566–581, Sanibel Island, Florida, USA.

[Tsinaraki et al. 2004] Tsinaraki, C., Polydoros, P., and Christodoulakis, S. (2004). Interoperability support for Ontology-based Video Retrieval Applications. In *3^rd International Conference on Image and Video Retrieval (CIVR'04)*, Dublin, Ireland.

[Ursu and Cook 2005] Ursu, M. F. and Cook, J. (2005). D5.3: Languages for the representation of visual narratives. Deliverable to EC (permission required), NM2 consortium.

[Williams et al. 2006] Williams, D., Ursu, M., Cook, J., Zsombori, V., Engler, M., and Kegel, I. (2006). ShapeShifted TV – Enabling Multi-Sequential Narrative Productions for Delivery over Broadband. In *The 2^nd IET Multimedia Conference, 29-30 November 2006*. ACM Press.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Imagesemantics: User-Generated Metadata, Content Based Retrieval & Beyond

Marc Spaniol, Ralf Klamma
(Lehrstuhl Informatik 5, RWTH Aachen University, Germany)
{spaniol|klamma}@i5.informatik.rwth-aachen.de

Mathias Lux
(ITEC, Klagenfurt University, Austria)
mlux@itec.uni-klu.ac.at

Abstract: With the advent of Web 2.0 technologies a new attitude towards processing contents in the Internet has emerged. Nowadays it is a lot easier to create, share and retrieve multimedia contents on the Web. However, with the increasing amount in contents the retrieval process becomes more complex and often leads to inadequate search results. One main reason is summarized easily: Approaches to image clustering and retrieval usually either stick solely to the images' low-level features or their semantic tags. However, this is frequently inappropriate since the "real" semantics of an image can only be derived from the combination of low-level features and semantic tags. Consequently, we investigated a more holistic view on semantics based on a system called *Imagesemantics* that tries to close the gap between both approaches by combining them.

Key Words: Web 2.0, Social Media Platform, User-Generated Content, MPEG-7
Category: H.3.3, H.3.4, H.3.5, H.5.1

1 Introduction

"What the heck do these images have to do with what I'm looking for?" That is a question many of us frequently ask when querying for images on the Web. Del Bimbo calls this the "semantic gap", the difference between technical extraction of data and the semantically correct interpretation of content [DelBimbo 1999]. Regardless of searching for pictures via Google Image Search [Google 2007], Yahoo! Photos [Yahoo 2007] or Flickr [Flickr 2007], the retrieval results are usually widespread and thus unsatisfying. This is usually a result of employing a website, where an image is associated with, as single source of metadata for images. Even more, common search interfaces are mostly restricted to textual queries. Contrariwise, the open source projects LIRe and Caliph & Emir allow content based image retrieval given an image (dataset) [Lux et al. 2003, Lux et al. 2006]. The combination of both strategies (text and image analysis) is rather rare but has for example been applied in IBM's Proventia Web Filter [IBM 2007]. For that purpose, the Proventia Web Filter is only suitable for "defensive" content blocking instead of "active" ad-hoc searches. Another approach for the combination of low-level and high-level metadata has been made in *Magick*, an application for cross media visual analysis. Textual data and images are clustered and visualized based on feature spaces they have in common. It relies on weighted averaging of similarity metrics that cannot be applied to an ad-hoc retrieval scenario. An application for common retrieval tasks, which supports ad-hoc search not based on filtering is to the best of our knowledge currently not existing. For that purpose, we have developed *Imagesemantics*: A speedy and concise system that allows searching for images in order to narrow (or even close) the "semantic gap" between low-level content based retrieval and high-level metadata annotations.

In this paper, we first give an overview on state-of-the-art image retrieval techniques. Then, we introduce related image retrieval systems. After that, we present of *Imagesemantics* system, which incorporates OWL-based rules for the combination of high-level features (vocabulary independent keywords, called tags) and low-level image features. The paper closes with conclusions and gives an outlook on further research.

2 Image Retrieval Techniques Compared

In general two different types of image retrieval can be distinguished: (i) Retrieval based on content-dependent metadata and (ii) retrieval based on content-descriptive metadata [DelBimbo 1999]. Content-dependent metadata includes low-level features generated automatically from the image content such as the image's color or shape feature vectors. Content-descriptive metadata are semantically high-level annotations. In the following, we will introduce both concepts, with a focus on standards compliant information processing. In this aspect, we will stick to MPEG-7 (also called the *Multimedia Content Description Interface*) [ISO 2002, ISO 2003] because it offers the semantically richest metadata model for describing the content of audio-visual media.

MPEG-7 [Chang et al. 2001] was standardized by the Moving Pictures Expert Group (short MPEG) and proposes a rich set of description schemata as descriptors organized in descriptor schemes. For the storage and processing of low-level, content-dependent metadata MPEG 7 basically provides three different types of structural features for content-based analysis of images: **Color**, **texture** and **shape** descriptors. MPEG-7 **color** descriptors provide means to analyze images based on their color distribution. In addition, MPEG-7 offers three descriptors that aim the **texture** and three **shape** descriptors. The applicability for photo retrieval and redundancy issues are discussed in [Eidenberger 2004].

In order to overcome the problems of interpreting the semantics of audio-visual contents correctly high-level metadata annotations are used. While the extraction of low-level features can be done automatically, the annotation of images with high-level metadata annotations is mainly a manual procedure. High-level metadata annotations are a means of classifying, organizing and (finally)

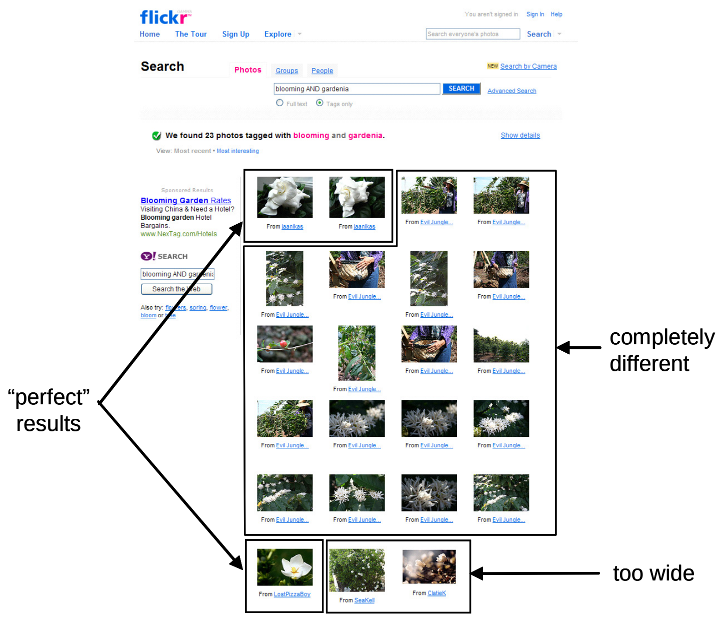


Figure 1: Flickr image search results for “blooming AND gardenia”

retrieving audio-visual contents. Similarly as with the content-based analysis of images by low-level features the MPEG-7 standard provides dedicated descriptors for high-level metadata annotations. These annotations reach from textual content descriptions up to ontologies. Thus, MPEG-7 offers a wide range for high-level interpretable and interoperable metadata annotations.

3 Image Retrieval Systems

There are several image retrieval systems. Most of them are based on either textual (metadata) descriptions or the image’s low-level features. Here, we confine our comparison onto the (probably) most prominent systems in the field of metadata driven image retrieval (Flickr.com) and content based image retrieval (Caliph & Emir). As IBM’s Proventia Web Filter can not be actively queried for ad-hoc image retrieval the search results of the previous systems will be directly compared with our *Imagesemantics* system. For the sake of comparability, we perform our query in any system on the same kind of picture: A blooming flower called *Gardenia*.

Flickr is a typical Web 2.0 representative. It provides its users with functionalities to describing, tagging and arranging images in web-based collections. Even more, the whole community might contribute to the stored images. For that reason, the tags are frequently misleading as different users have different perspectives onto a certain picture or focus on different semantic levels (the image in general or a certain detail). Similarly, images can only be retrieved via the images' metadata descriptions or tags. Thus, users can specify search terms. In the case of our comparative search for a "Blooming Gardenia" our initial query "Gardenia" returned an unmanageable 3000 pictures, of which "only" 1.200 were explicitily tagged as "Gardenia". Therefore, we refined our query to "Blooming AND Gardenia" which returned a reasonable amount of 23 pictures only. However, the result set is quite disappointing (cf. figure 1). Our query returned only three pictures we were searching for, while the remaining 20 were quite different. Two of the images dealt with "Gardenia" but were "too wide" while the remaining 18 showed completely different pictures of coffee plants.

Caliph & Emir are MPEG-7 based Java applications for image annotation and retrieval applying content based image retrieval techniques using MPEG-7 descriptors [Lux et al. 2003]. Besides extraction of existing information in digital photographs and transformation of these contents to MPEG-7, Caliph supports the creation of semantic dependencies. For that purpose, the MPEG-7 descriptions in Caliph comprise metadata description, creation information, media information, textual annotation, semantics and visual descriptors [Lux et al. 2006].

On top of it, Emir supports content based image retrieval in local image repositories created with Caliph. The most sophisticated retrieval techniques applied in Emir are backed by MPEG-7 descriptors ColorLayout, EdgeHistogram and ScalableColor. For the sake of comparability, the our reference query was evaluated on a local copy of all those 3.000 images of flickr, which contained "Gardenia'" in their description. Given the fact that a pre-selection of images based on their flickr-tags has manually been performed beforehand, the images retrieved are of much better quality than in the tag only search presented before. Nevertheless, about the half of the result size is quite different from what we have been querying for. Thus, there are two main drawbacks: First, Emir requires a manual preprocessing of Flickr images by Caliph in order to create a local image collection. Second, the comparison of the reference image's feature vector with all the other image vectors is very time consuming.

The main challenge in image retrieval is a fast and concise interpretation of the query's semantic. As we demonstrated before, a solely text or content based analysis mostly leads to unsatisfying results. While the results obtained from a time consuming content based analysis performed on manually pre-selected images based on their Flickr tags proved to be much more precise, the ultimate goal seems to be the combination of an accelerated text and image analysis. Therefore, our *Imagesemantics* system links both approaches by a k-means clustering algorithm. By comparing the reference image with the cluster vectors, *Imagesemantics* allows its users a fast and concise opportunity to formulate search queries based on search terms and by specifying a reference image.

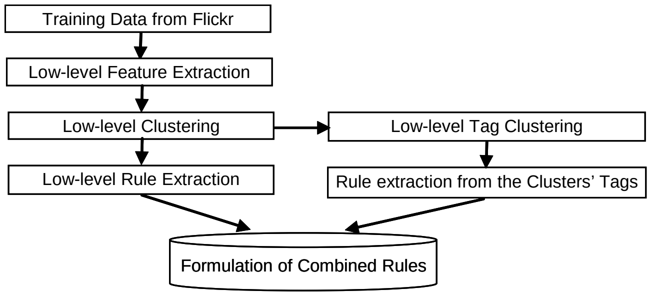


Figure 2: Rule-based image clustering process

4 *Imagesemantics*: Rule-based Image Clustering & Retrieval

In order to close the so-called "semantic gap" our approach is a rule-based image clustering and retrieval system. Starting from our multimedia information systems MECCA [Klamma et al. 2005] and MEDINA [Spaniol and Klamma 2005] developed to foster the annotation, classification, discussion and thus, collaboration about multimedia contents in communities of professionals, we explain how these high-level annotations can also be applied for the structured retrieval of contents.

Imagesemantics links text and content based image retrieval for a concise query processing. For speed-up reasons, *Imagesemantics* is based on a k-means clustering algorithm. By comparing the reference image with the cluster vectors, this procedure image has not to be performed with all n images, but only k times with the reference vectors instead (keeping in mind that usually $n >> k$). Next, we step-by-step describe our rule-based clustering process (cf. figure 2).

In an initialization step *Imagesemantics* extracts the low-level feature vectors of a test collection of images. Here, we apply Flickr's Java API Jickr to obtain a relevant set of test images [Jickr 2007]. Subsequently, the feature vectors of the images are extracted. In order to ensure the interoperability of our data, we make use of the MPEG-7 metadata standard, particularly those having proven to perform concise and fast on image sets: *Edge Histogram*, *Scalable Color* and *Color Layout* (cf. section 2). From these feature vectors we create k image clusters, where k is arbitrary and can be defined upon initialization. The underlying algorithm is a modified k-means clustering process. In order to form preferably homogeneous clusters we apply Ward's minimum variance linkage [Fasulo 1999], in order to obtain the k cluster centroid feature vectors. In the next step, two operations are being performed. On the one side, low-level rules are extracted in order to express the maximum distance from a centroid's feature vector allowed for images belonging to it. On the other side, the members' tags are extracted as a tag-cloud of terms. The tag cloud vectors rules are derived for each cluster so that a sub-clustering based on the high-level semantic annotations takes place.

```
<owl:Class rdf:about="LowlevelCluster_k">
    <Centroid> Values </Centroid>
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:onProperty>
                <owl:ObjectProperty rdf:about="Distance"/>
            </owl:onProperty>
            <owl:allValuesFrom>
                <owl:Class rdf:about="Interval_kmin_kmax"/>
            </owl:allValuesFrom>
        </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>

<owl:DatatypeProperty rdf:about="Centroid">
   <rdfs:domain rdf:resource="LowlevelCluster_k"/>
   <rdfs:range rdf:resource=
   "http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
```

```
<owl:Class rdf:about="LowlevelCluster_k_ Tagx">
    <Hastag>Tagname</Hastag>
    <rdfs:subClassOf rdf:resource="LowlevelCluster_k"/>
</owl:Class>

<owl:DatatypeProperty rdf:about="Hastag">
    <rdfs:range rdf:resource=
    "http://www.w3.org/2001/XMLSchema#string"/>
    <rdfs:domain rdf:resource=" LowlevelCluster_k_ Tagx"/>
</owl:DatatypeProperty>

<LowlevelCluster_k rdf:about="260407965_5c177d3703.mp7.xml">
    <rdf:type rdf:resource="LowlevelCluster_k_ Tagx "/>
    <rdf:type rdf:resource="LowlevelCluster_k_ Tagy"/>
</LowlevelCluster_k>
```

Figure 3: Low-level feature rules (left) and high-level tag rules (right)

In a final step, both low-level feature rules and high-level tag rules are combined. Thus, the gap between purely low-level content analysis and high-level metadata annotations can be bridged.

In order to make the previously extracted rules understandable and interpretable by a reasoner we will now describe how the rules are represented in an OWL ontology stored in an eXist XML database [Meier, 2003]. The class $LowlevelCluster\_k$ is the representative of the k^{th} cluster. This class contains the information about the centroid's feature vector as well as the cluster's range $Interval\_k_{min}\_k_{max}$ (cf. left hand side of figure 2). Based on these information it can now be specified whether an image belongs to a certain cluster or not. Similarly, the extracted rules from the clusters' tags can be expressed in OWL-classes. For instance, the class $LowlevelCluster\_k\_Tag_x$ contains the Tagname as a value of the x^{th} tag in cluster k (keeping in mind that each image and, thus, every cluster may be assigned with more than a single tag). As a result, for each cluster the associated high-level tag are formulated as a rule (cf. figure 3). In order to apply the inference mechanisms of an OWL reasoner, for each image an instance is being created (cf. figure 3).

In retrieval, the instances are in a first step being queried for a certain Tagname x. All those clusters are being identified, which contain this value. In our previous example the cluster $LowlevelCluster\_k$ would be one of the candidates. Then, the reference image's feature vector is being compared with the cluster's centroid vector. In case the difference is below a pre-defined threshold the dedicated cluster is prepared for the result. In a final step a selection takes place so that only those images of the chosen clusters will be shown, which are tagged by Tagname x (cf. figure 4).

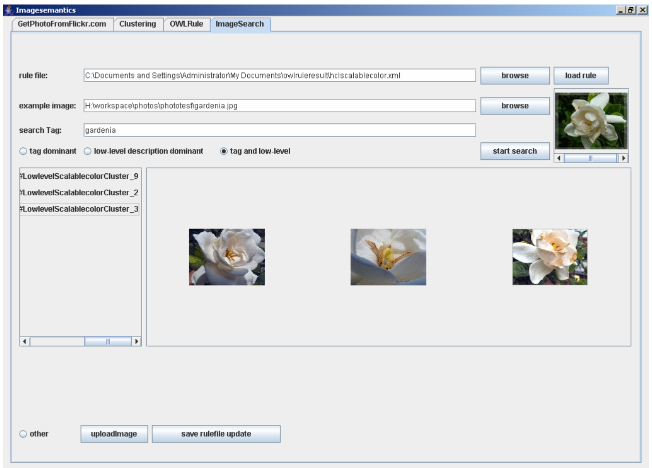


Figure 4: *Imagesemantics* search result based on low-level features <u>and</u> tags

5 Conclusions & Outlook

In this paper we have presented *Imagesemantics*, a system, which combines low-level and high level features for retrieval. Compared to a simple combination of low level and high level descriptors through weighted averaging of similarity functions, *Imagesemantics* relies on a cluster based index, which combines descriptors of both sides of the semantic gap. *Imagesemantics* is currently work in progress and a large scale evaluation is still missing, but first heuristic evaluations have shown the that the results of our system is subjectively better than approaches solely relying on single level descriptors. Therefore our system is a promising approach to narrow (or even close) the "semantic gap" between low-level content based retrieval and high-level metadata annotations in image retrieval. By comparing the search results of Flickr and Caliph & Emir, *Imagesemantics* proves that a combined approach leads to more accurate results.

In future, we will provide the functionalities of *Imagesemantics* via Web Services so that the system needs not to be used as a stand-alone Java application. In addition, we intend to enhance our Imagesemantics from solely image retrieval support to any other multimedia contents, particularly videos. In this aspect we are focusing our research on efficient processing mechanism for content-based key frame analysis and high-level textual descriptions on a large scale. Finally, we plan to evaluate the accuracy and performance obtained by Imagesemantics on a larger scale. Hence, we want to evaluate our retrieval results with other systems based on standardized multimedia test sets we are currently developing in a community of professionals (`www.multimedia-metadata.info`) in the field of multimedia.

Acknowledgements

This work was supported by the German National Science Foundation (DFG) within the collaborative research centers SFB/FK 427 "Media and Cultural Communication", within the research cluster established under the excellence initiative of the German government "Ultra High-Speed Mobile Information and Communication (UMIC)" and by the 6^{th} Framework IST programme of the EC through the NoE on Professional Learning (PROLEARN) IST-2003-507310. We thank Ming Cheng and our colleagues of the Multimedia Metadata Community (`www.multimedia-metadata.info`).

References

[Chang et al. 2001] Chang, S.F., Sikora, T., Puri, A. (2001). Overview of the MPEG-7 standard. Special Issue on MPEG-7 IEEE Transactions on Circuits and Systems for Video Technology, IEEE, pp. 688-695.

[Chen 2006] Chen, M. (2006). Regelextraktion und -darstellung von Multimedia Semantiken. Diploma Thesis, RWTH Aachen University.

[DelBimbo 1999] Del Bimbo, A. (1999). Visual Information Retrieval. Morgan Kaufmann.

[Eidenberger 2004] Eidenberger, H. (2004). Statistical analysis of MPEG-7 image descriptions. ACM Multimedia Systems journal, 10, (2), pp. 84-97.

[Fasulo 1999] Fasulo, D. (1999). An Analysis of Recent Work on Clustering Algorithms. Technical Report 01-03-02, Department of Computer Science and Engineering, University of Washington, Seattle, WA 98195.

[Flickr 2007] flickrTM (2007). http://www.flickr.com/ [10.7.2007].

[Google 2007] Google Images (2007). http://images.google.com/ [10.7.2007].

[IBM 2007] IBM Proventia Web Filter: Overview (2007). http://www.iss.net/products/Proventia_Web_Filter/product_main_page.html, [10.7.2007].

[ISO 2002] ISO ISO/IEC (2002). Information Technology - Multimedia Content Description Interface - Part 3: Visual. ISO/IEC 15938-3:2002, ISO.

[ISO 2003] ISO/IEC (2003). Information technology - Multimedia content description interface - Part 5: Multimedia description schemes. ISO/IEC 15938-5:2003, ISO.

[Jickr 2007] Jickr - Flickr Java API (2007). https://jickr.dev.java.net/ [10.7.2007].

[Klamma et al. 2005] Klamma, R., Spaniol, M., Jarke, M. (2005). MECCA: Hypermedia Capturing of Collaborative Scientific Discourses about Movies. informing science: The International Journal of an Emerging Discipline, 8, pp. 3 - 38.

[Lux et al. 2003] Lux, M., Becker, J., Krottmaier, H. (2003). Semantic Annotation and Retrieval of Digital Photos. In Proc. Of CAiSE 03 Forum Short Paper Proceedings Information Systems for a Connected Society, http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-74/files/FORUM_22.pdf [10.7.2007].

[Lux et al. 2004] Lux, M., Granitzer, M., Kienreich, W., Sabol, V., Klieber, W. , Sarka, W. (2004). Cross Media Retrieval in Knowledge Discovery. In Proc. of the 5th International Conference on Practical Aspects of Knowledge Management PAKM 2004, Vienna, Austria, December http://mathias.lux.googlepages.com/mlux-pakm04-preprint.pdf [18.5.2007].

[Lux et al. 2006] Lux, M, Klieber, W., Granitzer, M. (2006). On the Complexity of Annotation with the High Level Metadata. J.UKM, 1, (1), pp. 54-58.

[Meier, 2003] Meier, W.(2002). eXist : An Open Source Native XML Database. In: Web, Web-Services, and Database Systems, NODe 2002 Web and Database-Related Workshops, Erfurt, Germany, October 7-10, Revised Papers, volume 2593 of LNCS, Springer-Verlag, Berlin Heidelberg, pp. 169 - 183.

[Spaniol and Klamma 2005] Spaniol, M., Klamma, R. (2005). MEDINA: A Semi-Automatic Dublin Core to MPEG-7 Converter for Collaboration and Knowledge Management. In: Multimedia Repositories. In Proc. of I-KNOW '05, Graz Austria, J.UCS Proceedings, LNCS 1590, Springer-Verlag, pp. 136 - 144.

[Yahoo 2007] Yahoo! Photos (2007). http://photos.yahoo.com/ [10.7.2007].

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

The Need for Formalizing Media Semantics in the Games and Entertainment Industry

Tobias Bürger
(Digital Enterprise Research Institute (DERI), Innsbruck, Austria
tobias.buerger@deri.at)

Herwig Zeiner
(Joanneum Research, Graz, Austria
herwig.zeiner@joanneum.at)

Abstract: The digital media and games industry is one of the biggest IT based industries worldwide. Recent observations therein showed that current production workflows may be potentially improved as multimedia objects are mostly created from scratch due to insufficient reusability capacities of existing tools. In this paper we provide reasons for that, provide a potential solution based on semantic technologies, show the potential of ontologies, and provide scenarios for the application of semantic technologies in the digital media and games industry.
Key Words: multimedia semantics, digital entertainment, ontologies, semantic search
Category: H.5.1

1 Introduction

As initial investigations in the SALERO[1] project showed, most multimedia objects are created from scratch due to insufficient reusability capabilities of existing tools. One reason for that is their poor self-description-ability and the lack of formal representations of the properties of multimedia objects, their context and intended meaning. In order to reach a sufficient solution for that problem, several steps need to be solved: Besides the need for algorithms and frameworks to automatically extract high-level semantics from low-level features[2] which is well known as the "Semantic Gap" [Smeulders et al. 2000], annotation support for users, a solution for metadata interoperability across the content lifecycle, and cross-media adaptation is needed.

In this paper, we first introduce the intention of (multimedia) ontologies, its purpose and potential benefits for media production, sketch scenarios in which they can be applied, and finally conclude with an outlook.

[1] The overall goal of the integrated project SALERO (http://www.salero.info) is to define and develop intelligent content with context-aware behaviours for self-adaptive use and delivery across different platforms, building on and extending research in media technologies, web semantics to reverse the trend toward ever-increasing cost of creating media

[2] eg. "this picture depicts a scene in a football game" is inferred from the low level features "white circle AND green background color"

2 Using Ontologies for Semantic Representation of Media Items

As defined by Gruber [Gruber 1993], an ontology is an explicit specification of a (shared) conceptualization. The term ontology has been in use for many centuries and ontologies are widely used in applications related to information integration, information retrieval, knowledge management or in the Semantic Web. Ontologies are usually used to establish a common understanding of a domain and to capture the domain knowledge. This is usually done by modeling basic terms and relations which hold between terms, and by providing rules stating restrictions on the usage of both terms and relations.

In SALERO, we try to establish a **multimedia ontology framework** that combines declarative descriptions of

1. *Low-level physical and semantic features* through the use of multimedia description standards like MPEG-7 [Martinez et al. 2002] or essence internal formats
2. Domain specific *high-level semantic features* through the use of ontology languages like WSML [de Bruijn et al. 2005] or OWL [Dean et al. 2006]
3. *context information and rules* using WSML or RIF [Boley and Kifer 2007]

By using multimedia ontologies, recent research initiatives in the multimedia domain try to overcome the commonly known drawbacks of existing multimedia metadata standards for the descriptions of the semantics of multimedia content (see e.g. [Bloehdorn et al. 2005, Troncy et al. 2006, Benitez et al.al. 2002]).

Multimedia ontologies are mostly designed to serve one or more of the following purposes [Eleftherohorinou et al. 2006]:

- *Annotation*, which is in most cases motivated by the need to have high-level summarizations of the content of multimedia items
- *Automated semantic analysis*, i.e. to support the analysis of the semantics and syntax of the structure and content of multimedia items
- *Retrieval*, ie. to use rich formal descriptions to enable context-based retrieval and recommendations to users. The use of semantics enables automatic matching of content properties with user properties
- *Reasoning*, ie. the application of reasoning techniques to discover previously unknown facts of multimedia content or to enable question answering about properties of the content.

– *Personalized filtering*, ie. the delivery of multimedia content according to user-, network- or device-preferences.
– *Meta-Modeling*, ie. to use ontologies or rules to model multimedia items and associated processes.

3 The Purpose of Multimedia Ontologies in the digital games and entertainment industries

The potential benefits of formalizing media semantics were summarized and highlighted already before [van Ossenbruggen et al. 2004, Nack et al.. 2005]. In order to highlight the benefits for the digital games and entertainment industries, we especially try to summarize the purpose of the ontology framework that will be built in SALERO, sketch expected benefits for traditional media production and point out to important problems that we see in each point.

3.1 Semantic Search

To enable semantic search is the prime advantage of using ontologies in media production. Semantic Search aims to improve recall and precision of search results for multimedia objects. A prime pre-requisite for this is a way to attach hidden or contextual features to media items which are not visually embedded in them. By encoding such information using ontologies, the ontologies can then be used for the retrieval process and to present the results. This helps to clearer present contextual information, and helps to find more accurate results. A semantic search facility also provides a high-level means to explore collections with a high precision and recall.

The *most important problems* that we intend to solve with this facility are

1. The *Semantic Gap*, i.e. how to assign meta-data (semi-)automatically to multimedia data?
2. Low Precision/Recall for search in large multimedia collections or how to increase the amount of true positives for multimedia information retrieval?
3. How to match context of data with context of users?

3.2 Annotation Support

Mmetadata[3] has an important role in the multimedia lifecycle which was already highlighted by many others [Smith and Schirling 2006]. A recurring problem in

[3] metadata is data about data

industrial settings is however the task of creating metadata and keeping it up to date. Therefore one of the biggest issues we see is how to support creative people in creating annotations and how in turn existing workflows are only minimally disturbed by the implementation of new annotation facilities. We aim to develop ontology-based ways of supporting the user with these important tasks by developing a suite of ontology tools which can be used in daily work. An important point – as already emphasized above – is the integration of the ontology-enhancements into the work process with a minimum amount of disturbing the workflow. We intend to solve this by providing APIs that allow to include the functionality into existing multimedia authoring tools.

The *most important problems* that we intend to solve with this facility are

1. How can creative – and sometimes non-technical – people create ontological annotations?
2. How to maximize support and simultaneously minimize disturbance of current production workflows?

3.3 Unifying Disparate Metadata Formats

At present, different metadata standards are used to annotate in- and output of different steps in the multimedia production lifecycle. This fact is highlighted for the broadcasting domain in [Smith and Schirling 2006]. One result of SALERO will be to show the feasibility of using ontologies to unify these partially disparate metadata formats and the vocabularies used in them. Using ontologies is promising because of their modeling power, their formal background and strong their semantics compared to ad-hoc and informally specified data models. We try to use the arising multimedia ontology in order to enhance metadata integration and as a further step to improve the reusability of multimedia items.

The *most important problems* that we intend to solve with this facility are

1. How to process and re-use assets in different production tools?
2. How to integrate different standards that are used to describe different asset types? (eg. to provide a unifying search infrastructure on top of asset repositories)
3. How to create a scalable mapping/mediation layer between each of the standards? (ie. a bidirectional mapping between each of the standards does not scale!)

3.4 Support for Cross Media Adaptation

Another cost driver of today's media productions is the huge manual effort necessary for the adaptation of media to different target platforms and output formats (e.g. cinema, games, print or Internet). Ontologies offer the possibility to model and capture a rich set of metadata including the context of images or other assets. Therefore they can be used to cross purpose multimedia assets as automatically as possible. A wide range of multimedia objects is used by different parties in different media productions. This has to be considered when choosing or developing applications to create, manage or use ontologies for description of multimedia data like image sequences (TV recordings, computer generated content, film sequences), audio objects (recorded sound as well as synthesized speech), 3D Objects (Animated 3D objects in the application domains of interactive games, special effects for film), or any combination of the mentioned types.

The *most important problems* that we intend to solve with this facility are

1. How to ensure a common technological basis for (originally diverse) production tools?
2. How to establish declarative descriptions of workflows in order to match process- with content-descriptions for Cross-Media-Production?

4 Scenarios - Using Multimedia Ontologies in Media Production

In this section we briefly sketch examples how a multimedia ontology framework may be applied in media production environments.

4.1 Ontologies to Support the Re-Use of Assets

One big problem that media production companies are facing is the lack of possibilities to re-use material across productions. This is mainly due to the lack of definitions of methods and rules how assets can be reused and how certain assets can adapt themselves to new environments, eg. it needs to be considered how characters are able to interact with the re-used elements or how the elements can be adjusted to fit in a particular scene. In order to recognize if assets are re-useable their properties have to be stored explicitly (ie. the usage context of the asset, the rights to re-use, or rules about how to extract specific parts of an asset in order to be transferable to other animations. Using ontologies to describe the (usage) context of assets could help to clearly identify which methods are needed to transfer assets between different productions and to automatically identify the parts of the animations that one wants to re-use. This demands for advanced asset management systems that are able to store rich metadata together with the assets, perform fast and reliable searches and to access assets across productions.

4.2 Integrated Production and Cross-Media Delivery of Assets

In the increasingly fragmented media distribution marketplace there is a great need to be able to produce different delivery formats in parallel. On example for that is the sharing of datasets between film asset creation and tie in-game production. There are different problems attached to that wish: First of all there is the problem of reusing media objects in new functions: e.g. media objects in films and games have different roles and functions: in films they are actors, in games they are avatars: in films, they can be seen and heard, in games they are used in various interactive ways. This change of function affects the design of media objects. This demands for explicit descriptions of the story, storytelling, style and medium of a production. There ontologies could help. In turn, rules apply on how to convert between different genres. This rules could be modelled using ontology languages in order to turn them into actionable knowledge that makes an automatic conversion possible.

4.3 Ontologies to Aid Multimedia Information Retrieval

A major research problem in multimedia information retrieval is the "Semantic Gap" [Smeulders et al. 2000], the large gulf between the low level image features which can typically be processed in a multimedia document, and the high level concepts which a user is typically interested in. For example, a user may want to search for a video showing "Bing and Bong on a trip to planet XY"[4]. The high level concepts implicit in this query may be stated as the characters "Bing" and "Bong" , plus the abstract action "traveling". The action "traveling" in this case is almost impossible to recognize and could be derived from other recognizable features like "sitting on a couch"[5] or "galaxy"[6] which has to be explicitly modeled in an ontology capturing the knowledge of this special application domain. Automatic annotation of videos and images is currently an active research topic, and allows data-driven techniques to be used together with large training sets. However, automatic processing is not sufficient to annotate multimedia documents with the features like the ones used in the example above. Therefore annotation tools need to be developed that allow designers to annotate the material during the production to ease a latter retrieval-task.

[4] Bing and Bong virtual characters and part of Tinyplanets (see http://www.tinyplanets.com) which is a UK television show aimed at pre-schoolers
[5] This is Bing and Bong's favorite vehicle
[6] Bing and Bong usually travel in the galaxy

4.4 Ontologies as an Aid to Personalising Search

Search tasks, such as those supported by Multimedia Information Retrieval (MIR) systems, are typically subtasks of some main work task, which may be the creation of a new cartoon character, or some other aspect of the user's work. As such there are many contextual factors which may be captured about users, encoded in an ontology, with the aim of providing better retrieval results for the user in that particular situation. For example, the role of a user within an organisation may imply different search preferences: managers and administration staff may be less technical minded than engineering staff, and therefore an information retrieval engine should attempt to find documents of a less technical difficulty than may be presented to an engineer or scientist. Likewise, the immediate work context of the user engaged on the creation of characters for a particular film, may be used to condition searches, with the aim of interpreting queries relative to this topic.

5 Conclusions

In general, formal semantics can support the annotation, analysis, retrieval or reasoning about multimedia assets. With this paper we aimed to stress the importance of the use of formal semantics in the digital games and entertainment industry by pointing out to expected benefits and by sketching scenarios illustrating their intended application in media production in general and the SALERO project in particular.
Our next steps in SALERO are to engineer the first version of the ontologies and implement a first version of the ontology management framework to support the different media production scenarios in the SALERO project.

Acknowledgements

The research leading to this paper was partially supported by the European Commission under contract IST-FP6-027122 "SALERO".

References

[Benitez et al.al. 2002] Benitez, A. B., Rising, H., Jörgensen, C., Leonardi, R., Bugatti, A., Hasida, K., Mehrotra, R., Tekalp, A. M., Ekin, A., and Walker, T.: "Semantics of Multimedia in Mpeg-7"; Proceedings of the IEEE International Conference on Image Processing (ICIP 2002), pp. 137–140, 2002.

[Bloehdorn et al. 2005] Stephan Blöhdorn, Kosmas Petridis, Carsten Saathoff, Nikos Simou, Vassilis Tzouvaras, Yannis Avrithis, Siegfried Handschuh, Yiannis Kompatsiaris, Steffen Staab, and Michael G. Strintzis: "Semantic Annotation of Images and Videos for Multimedia Analysis"; Proceedings of the Second European Semantic Web Conference, ESWC 2005, Heraklion, Greece, pp. 592–607. Springer, May 2005.

[Boley and Kifer 2007] Harold Boley and Michael Kifer (eds.): "RIF Core Design", W3C Working Draft 30 March 2007, 2007. vailable online: http://www.w3.org/TR/rif-core/ (last accessed: 01.04.2007)

[de Bruijn et al. 2005] de Bruijn (ed.): "The Web Service Modeling Language WSML"; WSML Final Draft, 20 March 2005, Deliverable D16.1v0.2, 2005. Available online: http://www.wsmo.org/TR/d16/d16.1/v0.2/ (last accessed: 01.04.2007)

[Buerger and Westenthaler 2006] Bürger Tobias and Rupert Westenthaler: "Mind the Gap - Requirements for the Combination of Content and Knowledge"; Proceedings of the first international conference on Semantics And digital Media Technology (SAMT), December 6-8, 2006, Athens, Greece.

[Dean et al. 2006] Dean M. and Schreiber G (eds.): "OWL Web Ontology Language Reference"; W3C Recommendation, 2006. Available online: http://www.w3.org/TR/owl-ref/ (last accessed: 01.04.2007)

[Eleftherohorinou et al. 2006] Hariklia Eleftherohorinou, Vasiliki Zervaki, Anastasios Gounaris, Vasileios Papastathis, Yiannis Kompatsiaris, and Paola Hobson: "Towards a Common Multimedia Ontology Framework"; Analysis of the Contributions to Call for a Common multimedia Ontology Framework Requirement; April, 2006. Available online: http://www.acemedia.org/aceMedia/files/multimedia_ontology/cfr/MM-Ontologies-Reqs-v1.3.pdf (last accessed: 01.04.2007)

[Gruber 1993] Gruber, T. R.: "A translation approach to portable ontology specifications"; Knowledge Acquisition, vol. 5, no. 2, pp. 199–220, Juni, 1993.

[Martinez et al. 2002] Martinez-Sanchez, J. M., Koenen, R., and Pereira, F.: "Mpeg-7: The generic mul- timedia content description standard, part 1"; IEEE MultiMedia, vol. 9, no. 2, pp 78–87, 2002.

[Nack et al.. 2005] Nack, F., van Ossenbruggen, J., and Hardman, L.: "That Obscure Object of Desire: Multimedia Metadata on the Web"; Part 2 IEEE MultiMedia, IEEE Computer Society, 2005, 12, 54-63

[van Ossenbruggen et al. 2004] van Ossenbruggen, J., Nack, F., and Hardman, L.: "That Obscure Object of Desire: Multimedia Metadata on the Web"; Part 1 IEEE MultiMedia, IEEE Computer Society, 2004, 11, 38-48

[Smeulders et al. 2000] Arnold W.M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, Ramesh Jain: "Content-Based Image Retrieval at the End of the Early Years"; IEEE Transactions on Pattern Analysis and Machine Intelligence ,vol. 22, no. 12, pp. 1349–1380, December, 2000.

[Smith and Schirling 2006] John R. Smith and Peter Schirling: "Metadata standards roundup"; IEEE Multimedia, vol. 13, no. 1, 2006, pp. 84–88, 2006.

[Troncy et al. 2006] R. Troncy, W. Bailer, M. Hausenblas, P. Hofmair, R. Schlatte.: "Enabling Multimedia Metadata Interoperability by Defining Formal Semantics of MPEG-7 Profiles"; Proceedings of the 1st International Conference on Semantics And Digital Media Technology (SAMT 06), Athens, Greece, December 2006.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Quantitative Analysis of Success Factors for User Generated Content

Christian Safran
(Institute for Information Systems and Computer Media,
Graz University of Technology, Austria
csafran@tugraz.at)

Frank Kappe
(Institute for Information Systems and Computer Media,
Graz University of Technology, Austria
frank.kappe@iicm.edu)

Abstract: User generated content published via weblogs (also known as blogs) has gained importance in the last years, and the number of globally available weblogs increases. However, a large fraction of these show low publishing activity and are rarely read. This paper is a quantitative analysis of success factors in a community of over 15.000 weblogs, hosted by a local Austrian newspaper. We looked at publishing activity by content type, community activity and writing style. Also, the interconnectedness of the community was analyzed.

Keywords: User Generated Content, Weblog Analysis, E-Communities, Blogging
Categories: H 3.5, H.4.3, H.5.1

1 Introduction

In the age of the so-called Web 2.0, user generated content is credited increasing importance, as *participation* is one of the key characteristics of this concept [O'Reilly, 05]. Such content is gradually complementing and, in some cases, even contesting information provided in a classic, published environment. Especially in the news domain, weblogs (also known as blogs) have become an important source of information beside newspaper websites. This paper will focus on such *user generated content* published via weblogs.

The key question we asked ourselves in this context is: What makes a weblog successful? In the context of this research, success is measured by the number of visits to a blog. For the use in this analysis, visits were counted as unique IP addresses, counting a new visit after 20 minutes of inactivity. So, the question can be reformulated as: what are the factors that lead to a high number of visits of the content created by the users?

Another interesting aspect of user generated content is the community of the involved users. Possible questions in this context concern the structure of this community. Is the community divided into smaller communities or is there a central group of active users?

2 Related Work

Quite a lot of research on weblogs has been published in the recent years. Interesting in this context are publications on ranking weblogs like [Kritikopoulos, 06]. The authors present a modification of the PageRank algorithm [Page, 98] designed to take into account the links between the blogs and the similarity of the users, as well as links to non-weblog URLs.

[Du, 2006] tries to answer the question for success factors of weblogs from a technology perspective. The study analyzed the impact of technology used on the success of 126 blogs taken from the top 100 listings of the Technorati[1] website. In the case of Technorati, success is measured by the number of inbound links to a blog.

As far as the analysis of weblogs communities is concerned, [Cohen, 06] counts connections by hyperlinks and relation by type or topic as possible relations forming communities of weblogs. This point of view focuses on the relations provided in the content of the weblog. A complementing approach, as presented in [Li, 07] is to take into account the information available from the guest comments to the entries.

To the best of our knowledge, our work is the largest quantitative analysis of success factors for weblogs to date.

3 The Analysis

The "Meine Kleine" Weblogs[2] of the local Austrian "Kleine Zeitung" newspaper offer a promising possibility to take a closer look at a large number of weblogs in a relatively closed environment. Users of this environment have the ability to publish text and images (photos) to their own weblogs. In addition, comments as guestbook entries can be written to the weblogs of other users.

By November 21st 2006, the blogspace consisted of 15702 active blogs, ranging from topical weblogs by the newspaper's editors to private diaries of individual readers. In our research, we had access to the servers log files as well as the database holding the weblog entries. The log files and entries from Oct 7th to Nov 21st 2006 (about 6 weeks) were analyzed in the course of this project.

3.1 Basic Statistics

The community of Kleine Zeitung readers consist of a large number of more than 118,000 registered users. Of those, 15702 have activated the weblog for their account and had at least one visit in the analyzed 6-week period. 560 of those bloggers were active publishers in this period, meaning they added content to their weblog in these 6 weeks. This, together with the fact that only the top 1730 weblogs had five or more visits in this period, resulted in our decision to take only the 2000 most visited weblogs into account for all further examinations (see Figure 1).

The publishing activity of the examined users was quite incoherent. The number of (text) entries published ranged from 0 to 45 for the individual blogs. The most active poster of images published 469 images in the examined period. As far as the

[1] http://www.technorati.com
[2] http://www.meinekleine.at

activity in the community is concerned, the most active users posted 80 comments and 137 guestbook entries in other users' weblogs. Figure 2 shows the different forms of publishing activity for the user with the 20 most-visited weblogs.

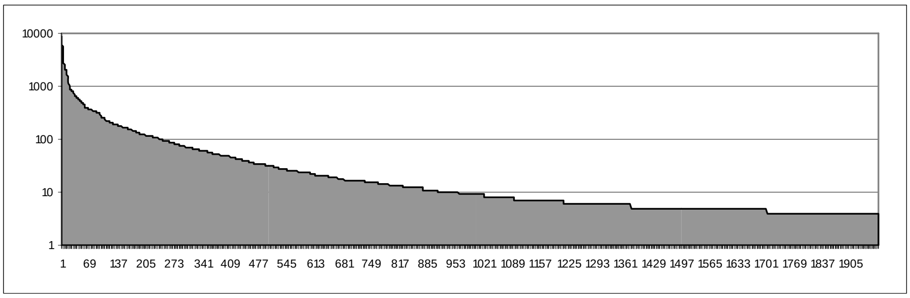


Figure 1: Distribution of visits of the top 2000 weblogs

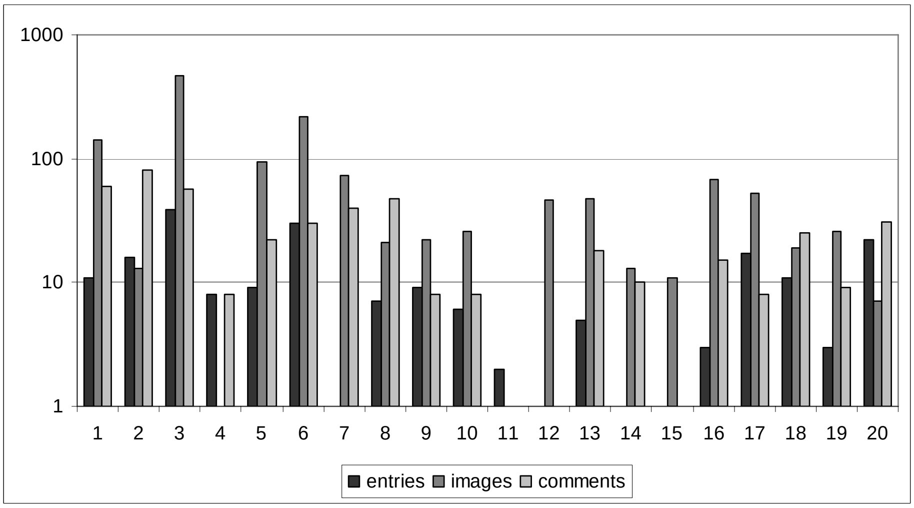


Figure 2: Publishing Activity of the 20 most-visited weblog users

3.2 Activity of Authors vs. Readers by Time of Day

Beside the basic statistics of the weblogs, an investigation of the user activity in the course of a day has been carried out. On the one side the activity of the content authors, those users creating text entries or uploading photos, shows an almost even distribution throughout the day, with one peak at noon and one in the early evening. This is in contrast to our assumption that most users would be contributing content to their blogs in the evening and in a home environment. This assumption is true, however, for the posting of comments, which mainly occurs between 19:00 and 22:00. Figure 3 shows the significantly different graphs for publishing own content versus activity in the community (i.e. writing comments and guestbook entries in other users' weblogs).

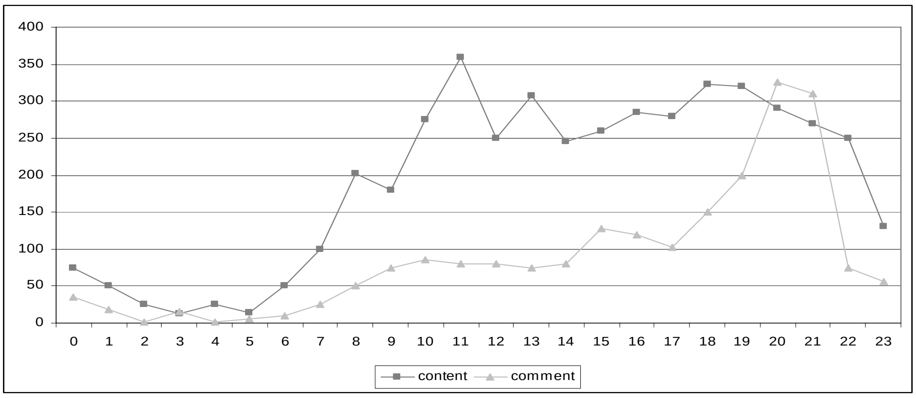


Figure 3: Creation of Content and Comments in the course of a day

On the other hand, the activity of the visitors of the blogs in the course of a day was also investigated and viewed separately for the different types of content. As with the authoring activity, reading is almost evenly distributed from 7:00 to 23:00. Peak usage is from 19:00 to 22:00, which corresponds to the peak in commenting. The visits of guestbook entries vary from this general observation, as they have a peak in the early afternoon and none in the evening (see Figure 4).

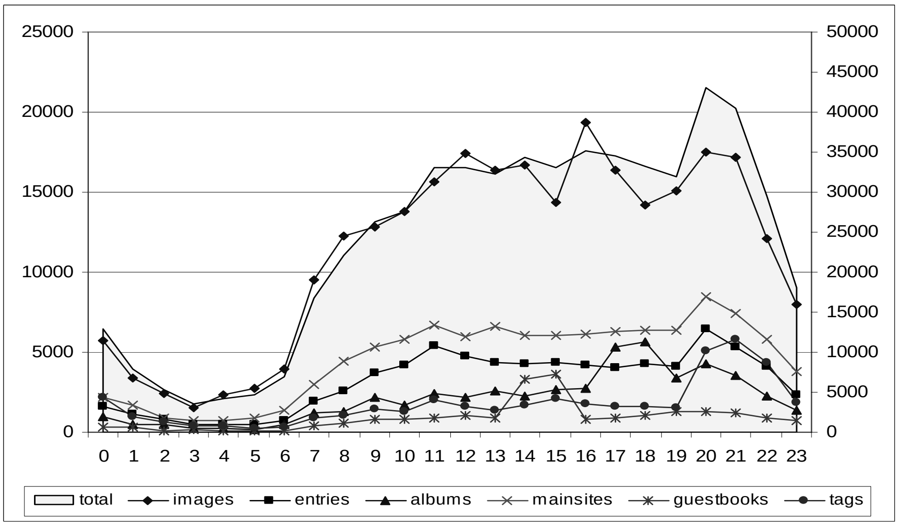


Figure 4: Views of different content types in the course of a day

4 Influences on Popularity

The main focus of the research on the “Kleine Zeitung” blogspace was to find and verify factors for the success of weblogs. Based on the available data, we decided to analyze nine possible criteria, arranged into three groups.

4.1 Influence of Content Types provided

In a first step the different types of content composing the blog were analyzed on their impact on the popularity. The visits of the individual blogs were compared to the number of textual entries, photos and the days the user was active in the period investigated. The highest correlation could be found with the user’s active days, being 0.68. The correlation to the entries and images were lower, at 0.60 and 0.53, respectively.

4.2 Influence of Community Activity

Secondly, the influence of community activity was investigated. We decided to analyze the correlation of comment and guestbook activity, outbound as well as inbound. The highest value was found for the obvious correlation of received comments and visits to the blog with 0.77. The number of guestbook entries received correlates with 0.69. For own comments and guestbook entries in other blogs the correlations are 0.70 and 0.68, respectively (see Figure 5).

It should be noted that these correlations coefficients are higher than those of publishing content. In other words, in order to have a highly visible weblog, it is even more important to be active in the community than to publish own content regularly! This true for the individual correlations as well as for the summary of content provided respectively own community activity. There is a total correlation of 0.61 of content provided to the number of visits, while the correlation of community activity to number of visits is 0.71.

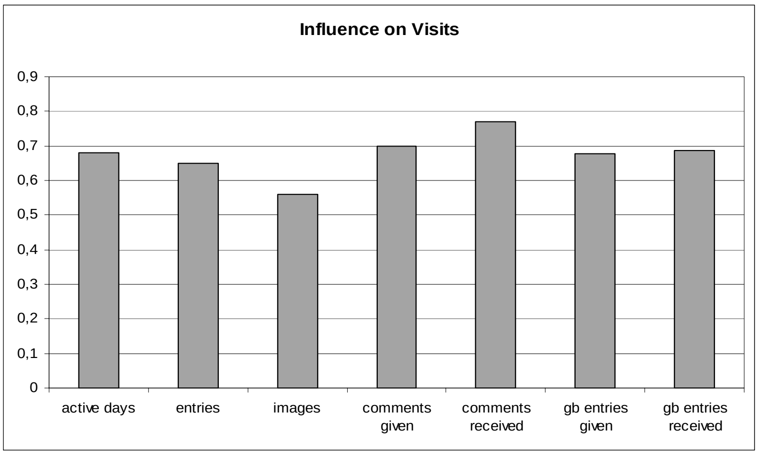


Figure 5: Influence of content provided and community activity on popularity

4.3 Influences of Writing Style

A further aspect of the research investigated upon the contents of the Kleine Zeitung blogspace is the influence of the author's writing style. For this purpose, the similarity computation of the Autonomy Search Engine, which is based on Bayesian Inference [Autonomy, 07], was used.

The blog entries of the top 2000 blogs were compared to the editorial blogs of the Kleine Zeitung (written by professional journalists) and to the top 5 blogs. In both cases the correlation was very low with coefficients of determination of under 0.20. Thus no reliable statement is possible.

5 Communities in "Meine Kleine" Weblogs

As the previous chapter showed, the activity within the community is a crucial factor for the success of a weblog. In a final step of our research, we tried to visualize the communication between the community members by comments and guestbook entries. In the resulting graphs, the members of the community are depicted as nodes and links resulting from comments or guestbook entries as edges. A Fruchterman-Reingold force directed placement algorithm [Fruchterman, 91] was chosen for the graph for a comprehensive visualisation of the interconnectedness of the community.

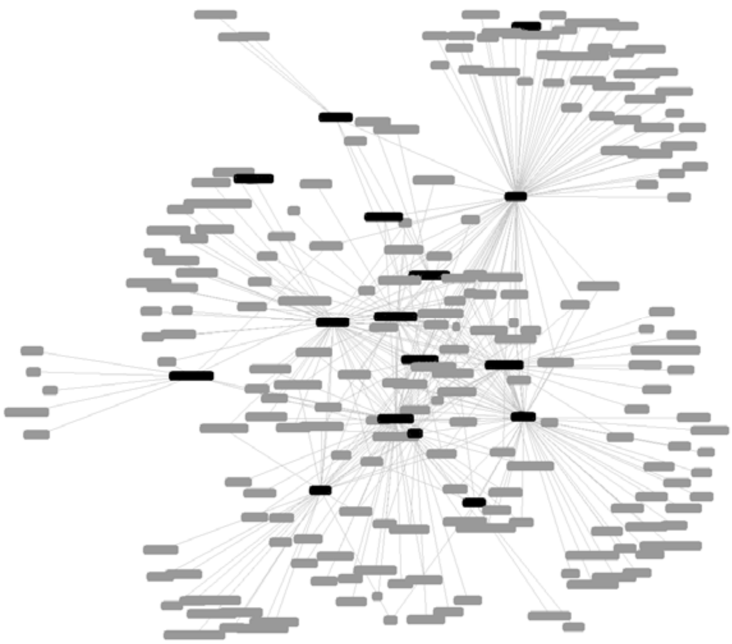


Figure 6: Force Directed Placement of the 20 most-visited Users and their Contacts[3]

[3] For privacy reasons the user names have been blackened in the community graphs.

In order to produce clearly arranged graphs, the number of community members taken into account was reduced. As the twenty users with the highest numbers of visits were responsible for 91% of the community activity, only these twenty and the corresponding conversational partners were taken into account.

In the resulting Figure 6 those nodes that represent the 20 most-visited weblogs are presented in darker shade than the others. Four of these top-twenty users have did not give any comments or create guestbook entries in the analyzed period and were thus removed from the graph. The remaining 16 top-scorers form a tight network with several communities unique to individual users.

As a next step to clarify the social network of the Kleine Zeitung blogspace, only those edges were taken into account that represented three or more communication activities. The resulting graph shows a tight network of eleven of the top score users, while the remaining 9 have no connections to the graph (Figure 7). Four of these eleven users also build their own sub graphs of community activities.

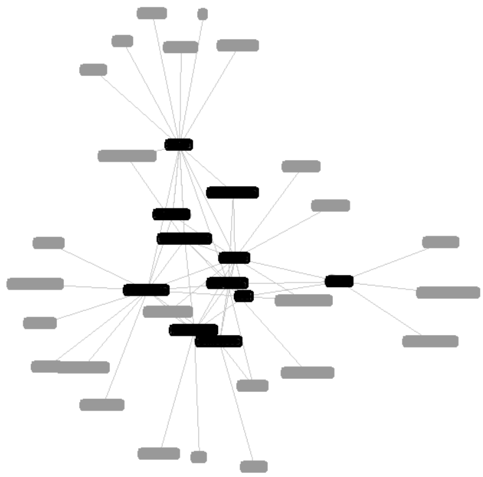


Figure 7: The Community of the most-visited users

6 Conclusions

The analysis on the blogspace of the Kleine Zeitung online community was conducted to find key factors for success of weblogs. The factors activity, number of textual entries, number of images, comments given, comments received, guestbook entries given and guestbook entries received were analyzed in this context. The comparison of the influence of these 7 factors showed that the most important of these factors are the community activities of the authors, i.e. writing comments and guestbook entries in other blogs.

7 Future Work

The research on the Kleine Zeitung blogspace gave a first hint on the success factors for weblogs. Anyway some open questions remain. In order to better understand the relation of community activity and weblog success a follow-up analysis over a longer period of time is planned for this year.

Acknowledgements

Financial support of this research by Styria Media AG, in conjunction with the Endowment Professorship for Innovative Media Technology at Graz University of Technology, is gratefully acknowledged.

References

[Autonomy, 07] Autonomy Inc., Autonomy Technical Overview, 2007, http://www.autonomy.com/content/Technology/index.en.html

[Cohen, 06] Cohen, E. and Krishnamurthy, B., A short walk in the Blogistan. *Comput. Networks* 50, 5, Apr. 2006, 615-630.

[Du, 2006] Du, H. S. and Wagner, C., Weblog success: Exploring the role of technology. Int. J. Hum.-Comput. Stud. 64, 9, Sep. 2006, 789-798.

[Fruchterman, 91] Fruchterman, T. M. and Reingold, E. M. 1991. Graph drawing by force-directed placement. *Softw. Pract. Exper.* 21, 11, Nov. 1991, 1129-1164.

[Kritikopoulos, 06] Kritikopoulos, A., Sideri, M., and Varlamis, I., BlogRank: ranking weblogs based on connectivity and similarity features. In *Proceedings of the 2nd international Workshop on Advanced Architectures and Algorithms For internet Delivery and Applications* (Pisa, Italy, October 10 - 10, 2006).

[Li, 07] Li, B., Xu, S., and Zhang, J. 2007. Enhancing clustering blog documents by utilizing author/reader comments. In *Proceedings of the 45th Annual Southeast Regional Conference* ,Winston-Salem, North Carolina, March 23 - 24, 2007

[O'Reilly, 05] O'Reilly, T., What Is Web 2.0 - Design Patterns and Business Models for the Next Generation of Software, 30.09.2005, http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html

[Page, 98] Page, L., Brin, S., Motwani, R. & Winograd, T., The pagerank citation ranking: Bringing order to the web. Technical report, Stanford, USA. 1998

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Online Crowds – Extraordinary Mass Behavior on the Internet

Christian Russ
(University Klagenfurt, Institute of Applied Informatics, Austria
chr@ifit.uni-klu.ac.at)

Abstract: In this paper a novel form of online users, the "Online Crowds", is described. "Online Crowds" gather virtually, behave and act collectively and produce effects and phenomena which would not be possible without the Internet [Hof 05]. A remarkable example is "The Million Dollar Homepage"[1] where a clever student made more than £ 100.000 only by offering a simple but unique online idea. He used the dynamics between online users and companies to make money with almost no effort. To understand these "social contagion" processes, an interdisciplinary conceptual and psychological model of "Online Crowds" is introduced. The model is based on the principles of "Other directedness", "Critical mass", "Positive feedback loops" and the accelerating impact of network effects on the Internet.
Some recommendations are sketched how such "Online Crowds" can be actively formed for promising online business models. If the behavior and the characteristics of "Online Crowds" are better understood, decision makers and providers will be better capable to predict and promote successful online communities and services.
Moreover a look at the positive and negative effects of these phenomena is taken and their challenges, as well as the implications for the affected society are analyzed. Especially the domain of New Media Technology (NMT) and the particular area of online recommender- and personalization technology are facing a potential for exploiting these Internet phenomena.
Finally, a list of related work in the field and an outlook on further improvements in the discussed approach are given.

Keywords: Internet phenomena, collective and extraordinary behavior, social contagion, societal and business implications
Categories: H.5.1, H.5.4, J.4., K.4

1 Introduction

Today's Internet is rapidly changing. Starting in the early years as a research project and later as web-based information system for company homepages and simple product catalogues, we can find today different variants of e-collaboration and communication on the web. Especially the upcoming consumer-to-consumer (C2C) and peer-to-peer (P2P) business models are a point of interest, because in these cases the online users themselves populate and facilitate their online platforms and form phenomena that have not seen before [Hof 05]. Typical examples are auction

[1] http://www.milliondollarhomepage.com/

overbidding at eBay.com, masses of video clips on YouTube.com, social buddy networks at MySpace.com, free social encyclicals like Wikipedia.org or unlimited social photo sharing with Flickr.com and many more. These “New Media” online business models use one unique and emerging principle of the Internet: The power of “Online Crowds”.

The paper starts with an explanation of the origins of “Online Crowds” in the real world by collective behavior and “social contagion” [Gladwell 01]. Based on the model of crowd psychology [Pelzmann 00] and their major properties a simplified model of the “Online Crowds” is sketched, trying to understand why some Internet ideas are getting blockbusters while others are no-starters.

2 The roots: Social contagion of the “Real Crowds”

Mass phenomena are situations where a lot of people seem to overreact in similar ways and the collective result of these processes are often not known before. Typical examples can be mass movements, trends and fads, as well as hypes and panics [Shiller 00]. The outcome can have a positive or a negative nature and frequently mass media even amplifies this process in many ways. Occasionally sensitive people or organizations make use of such an emergent phenomenon and profit from it by offering information, products or services for and during this process. Furthermore politics and corporations are often cautious about this social behavior, because it can change and influence political situations and markets in a rapid and unpredictable way (see [Le Bon 82], [Kindleberger 78], [Bonabeau 04]).

The actors of such extraordinary mass behavior are termed here as “Real Crowds”. These are participants (agents) of a social contagion process where a self-enforcing movement leads to unexpected results. The agents may not act and behave rational, they behave irrational and “other directed” ([Schelling 78], [Pelzmann 00]). This means, they do not take decisions on the basis of facts and experience, but they rather observe and follow the behavior of others and “run after” them like herds. The best examples are fashion trends which are heavily “other directed” and can change quickly depending on a common “Zeitgeist”. Self-enforcing is defined as the reciprocal observation and acting of agents which can lead to information cascades of misguided behavior [Bikhchandani et al. 98]. This spiral phenomenon is accompanied by “Positive feedback loops”[2] and “Path dependence”[3] mechanisms where small changes at the beginning can lead to unpredictable outcomes at the end.

Additional examples of “Real Crowds” are overheated finance bubbles and their inevitable bursts [Kindleberger 78], the Internet hype in 2000, pop stars and commercial blockbusters like the book Harry Potter or Apple’s Smartphone called iPhone.

Although the emergent behavior of these social interactions varies from case to case, the overall development shows similar patterns and phases. An idealized

[2] Positive feedback loops are circular loops with no setback and which can act self-enforcing with exponential output ([Arthur 94], [Ossimitz and Lapp 06]).
[3] Path dependence is a past dependency, some random events, a customer grove-in or an early advantage of a product which can lead to market dominance [Arthur 94].

development model of the "Real Crowds" can illustrate the typical phases of these phenomena.

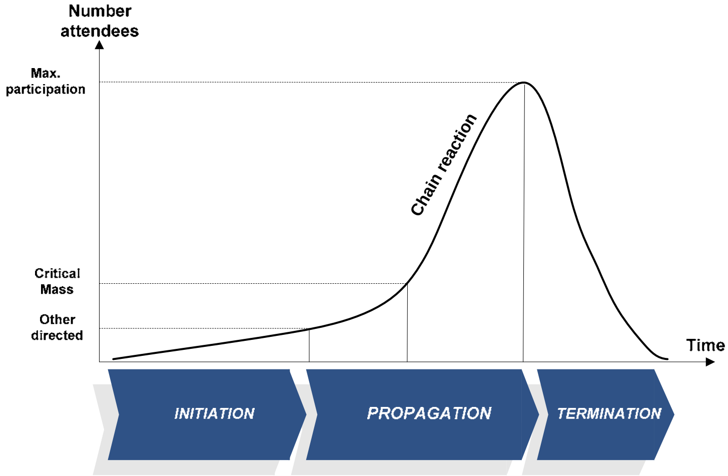


Figure 1: Idealized social contagion model of the "Real Crowds"

Figure 1 shows the most important phases of the social contagion process and their development, based on the number of new affected agents (in-flow). The first phase, called **"Initiation"** is triggered by some psychological attractors who focus the audience on some event or environmental change. It might be some cutting-edge technology, innovation or displacement in the social or economic world for instance. If the attention of the audience becomes keener then the psychological catalysts like media, opinion leaders, the government and others step in and accelerate the activation of these agents (initiated agents). Normally there are some additional restrictions in the general environment like limited knowledge of the targeted situation, uncertainty or some instability in the social relationship and the process starts to grow.

Those agents that are already involved, as well as new agents tend to step into the following phase termed **"Propagation"**. People evaluate the potential risks by observing involved agents and start to ignore facts more and more by imitating the actions they notice by others ("other directedness"[4]). The more agents are acting like this, the more will follow and take decisions based on others: The social contagion is started [Pelzmann 00]. But not all agents behave the same; it is the dominance of the super spreaders like trend setters, opinion leaders and pioneers who serve as an object for imitation at first, however later on more and more ordinary peers are imitated as well. This process could still drain or get stabilized by some external influence, however if the crowd grows and reaches the important point of **"Critical mass"**[5], the phenomenon may not be stopped any longer. When exceeding this threshold of "mind

[4] The behavior of "other directedness" is also known in the fauna, e.g. "Cleaner fish" are selected by clients by observing the choice of other clients [Bshary and Grutter 06].

[5] The term "Critical mass" was coined in nuclear chain reaction processes.

infected” agents, the whole process turns into a self-enforcing chain reaction of imitating agents. The emerging “Positive feedback loops” additionally boost the reactions and “Path dependence” factors start to become effective. At this stage the process can be compared with some nuclear chain reaction of colliding neutrons or some infectious diseases where a virus spreads through exposure with others. In this stage the “Real Crowds” are fully developed and spread rate can reach exponential increase. Because of this overreaction, it can produce irrational behaviors which are not foreseen nor are the consequences known.

Each of those overheated “Real Crowds” end up at some point of time; it is just a question when and how. Typically, if all energy of the drivers is burned out, the process dramatically swaps to the final phase called **“Termination”**. In this phase the euphoria breaks off and panic may step in. Sometimes the agents can save themselves and a controlled exit is feasible, but many times it is a crash which ends up the irrational exuberance. The best known and investigated examples for such social contagion processes are the finance markets with their booms and their crashes signifying the “Termination” phase (see [Kindleberger 78], [Shiller 00])

3 A simplified model of “Online Crowds”

More and more users are populating the World Wide Web and they have accepted this virtual space as a part of their personal living. It is used as a coexisting environment for everyday tasks, like communicating and collaborating with friends, colleagues or business partners, for e-Commerce, e-Learning and e-Government or just for fun and entertainment.

With this new medium, the old limits of time and space do not exist any longer. People can inform and communicate with others, hence ideas and messages can be spread within seconds all over the world. Even more, the nature of digital content assist the “copy cat effect” [Bonabeau 04] and concepts like “Creative Commons”[6] stimulate the imitation nature of humans and the economical model of increasing returns[7]. Moreover, new computer techniques like information retrieval, collaborative filtering and personalization technology support online users in their information gathering and utilization process. So social software like Wiki’s, social tagging and web logs are offering information and opinions directly and unfiltered to a broad audience, giving prior unknown communities a voice or breaking down established institutions (see [Hof 05]).

The leverage lies in the effect of the “Online Crowds” which acts behind the scene as facilitator of these Internet phenomena. The origin of the “Online Crowds” can be found in the laws and mechanisms of the “Real Crowds”, but varies in speed, size and scope. Every day new online services are published and others disappear from the online space. New technologies like Web 2.0 are emerging and offering ordinary online users an easy “click-and-go” way to make their ideas and activities

6 Creative Commons: Share, reuse, and remix — legally, http://www.creative-commons.org/

7 Increasing returns: Instead of classical economic theory of perfect markets with returns (e.g., profits) that decrease and tend toward equilibrium, “increasing returns” offer several unpredictable equilibria but can scale dramatically in growth and range [Arthur 94].

work [Rollett et al. 07]. In addition online users are overloaded with web information and they are searching for orientation and affiliation.

Sensitive business actors and innovators have recognized this emerging trend and took the advantage on their side. They set up online platforms where seeking online users can flock together in order to cultivate their social ties and lifestyle. Spectacular examples are Google.com, eBay.com, MySpace.com, YouTube.com and many more. Some of them evolved by luck and coincidence, but some others knew how to design "Online Crowds" readiness. They noticed the new laws of the "Online Crowds" as an enhancement of the "Real Crowds" because of the supporting network effects of the Internet.

Transforming the building blocks of the Real Crowds into the Internet, there are three major phases of an "Online Crowd" development process (Figure 2). One can recognize that the first two are identical with the "Real Crowds", but the model is extended with a third one called the "Amplification". "Online Crowds" can possibly emerge if an online service or an Internet phenomenon endorses the following phases, but they are not completely sufficient:

- *Initiation (best content):* Similar to "Real Crowds" psychological attractors are needed, so that many get noticed about the service. The most important factor of this phase is the valuable content. Without any useful seed content or attractor nobody will be willing to spend time on the online service, neither will he or she come back. Some years ago there was the slogan "Content is king", but this is here only partially true. Additionally usability and simplicity of the user interface are essential to open the online service to interested visitors.
- *Propagation (best members):* In this phase users start to populate the online platform and the decision on whether somebody else is joining depends on the existing members. The phase transition is triggered by the "other directedness" threshold. So the most important factors in this phase are the best active "core members" who are attracting others. Moreover they deliver new content (user generated content), minimize risk for others (trust) and assist newcomers on daily problems on the platform (confidence). But the core members are not enough to grow big; the virtual word of mouth must spread further.
- *Amplification (best social facilitation):* In the last stage of growth it is important to offer the online agents not only a stable and scalable online platform, but also an environment where they can almost live their social lives. The phase transition is marked by reaching the "critical mass". Hence social facilitation features which support their daily practices and social behaviors are essential to keep the service expanding. The facilitation mechanisms can be divided into technical features like tagging, annotating, blogging, collaborating content and social features like adding collective context and emotion to the content. If available and applied, positive feedback loops and the path dependence become operative and the viral diffusion is fully accomplished.
- *Termination:* Last but not least there is also the final stage in which some online services end or break up. Normally this happens because of saturation

or a substitution with a new service or because the hype evaporates and the agents hop off or switch to another trend.

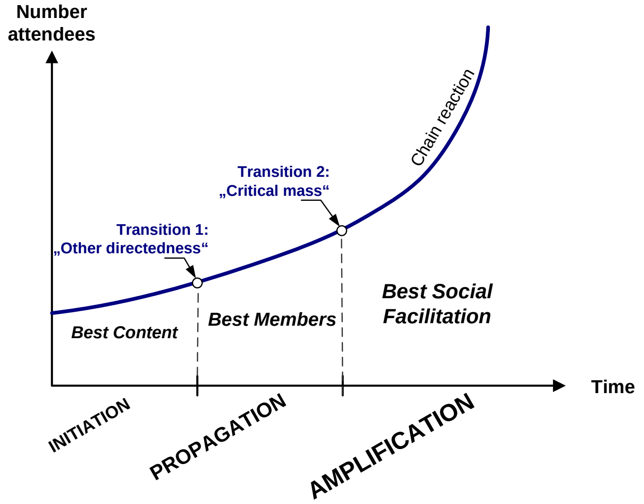


Figure 2: Simplified phase model of "Online Crowds"[8]

So the structure of social networks and the Internet are excellent surroundings to foster the emergence of "Online Crowds". For sure not all presented amplification effects are cumulated 100%, there are also attrition rates and forces that countervail the leveraging effect [Dodds et al. 03]. But this generalized model could help to reinterpret some Internet phenomena seen over the last years and could outline a bit more what is happening behind the scenes.

4 Consequences and changes for online business models

Earlier online business models relied on a classical supplier and consumer relationship where online users consume products and services of an online company. These business models normally can not scale beyond phase two (Propagation) of Figure 2, because it is limited to the number of customers power 2 (N^2). If online ideas are focusing on new types of businesses, which involve the customer more directly into the sales and the value adding processes, then C2C and P2P models are the right choice. In this case phase three (Amplification) is capable to enhance the online business and improve valorization dramatically.

Therefore an "Epidemic Online Model" (EOM) is introduced, which is based on the "User centric innovation" concept of [Hippel 05], where the user himself plays the role of consumer, innovator and producer (co-creator) of the business. Here this concept is extended to the needs of the online world with their digital assets.

---

[8] The "Termination" phase, which is less important at this point, is dimmed out for now.

Like presented in Figure 3 there are two main components in the EOM which feed each other and speed up the growth process (reciprocal acceleration by contagion):

- *New and existing content:* Existing content can be consumed by the present users and can attract new users who contribute additional ideas and content.
- *New and existing users:* On the other side, existing users can create new content, connecting existing user and content, as well as attract new users.

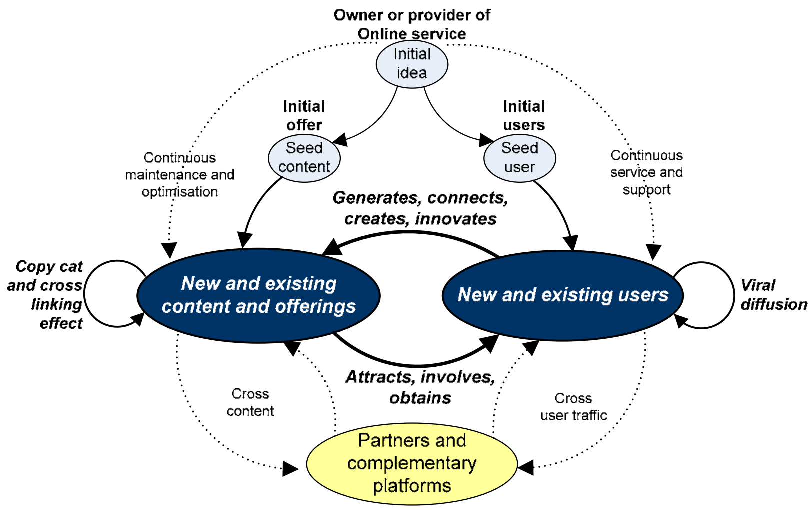


Figure 3: The "Epidemic Online Model" as a driver of online business models[9]

Moreover, at the "Initiation" phase some kind of seed is needed, which helps to start up the spin and can be separated into the two areas:

- *Seed content:* The seed content attracts the first users (early adopters) and stimulates them to produce further content and activity.
- *Seed users:* The initial agents can be seen as pilot or beta users, who verify the functioning and the usability of the system. In addition they signalize safety for successors, because they minimize the risk for late adopters and doubters.

Furthermore for many online ideas it makes sense to think about syndication and integration of content and services from partners and complementary online platforms. Current Web 2.0 technologies help to foster and resolve this attempt in a promising way [Rollett et al. 07]. This amplifies once again the number of users and the evolving value which increases the overall business value of the system. All in all the online models must change from a push concept, like typical online shops to a pull concept, driven by the self enforcing undertow of the "Online Crowds".

A vivid example is the online auctioning system eBay. It started as small online marketplace for used goods and grew dramatically because eBay offered online users

---

[9] Basic concepts of Causal Loop Diagramms based on [Ossimitz and Lapp 06] are used.

a platform to deal and sell in a C2C style. eBay was only the enabler of interaction and collaboration for their communities, who "reproduced" themselves with content and new participants. Additionally psychological attractors like the auctioning concept, fast auction profits and extraordinary auction stories initiated and resumed the hype.

5 Chances and challenges of "Online Crowds"

New and emerging phenomena create opportunities as well as challenges for the society, for the government and for the business world. As discussed in the previous chapter, "Online Crowds" can accelerate a business and generate remarkable online growth rates. But on the other side, uncontrolled and elusive online phenomena can create Internet fads, rumors, misguided collective behavior and forces which also influence everyone's daily life [Shiller 00].

Some opportunities of Online Crowds (OC):

- Applied on appropriate online business models they can increase usage, publicity and size of the service.
- Even if products or services are not clearly block busters at the beginning, OC can initiate and facilitate buzz on them.
- OC can build up emerging online platforms with high traffic without big marketing budgets; minimizing costs for content and customer support because of some kind of self-organization.
- OC are promising applications and research domains for the next generation of "Social Software", recommender technologies and personalization.

As mentioned in the last items, OC open new potential on the technical utilization of their behaviors. As people tend to make decisions based on others, this knowledge can be used to better design interactive selling tools ([Adomavicius and Tuzhilin 05]) which can benefit from the "collective patterns of the crowds". The future of proactive and more intelligent online sales systems for products and services will supply hybrid recommendation technologies, offering online customers advise and help during the whole purchase process. Combinations of content-based, user-based [Zanker et al. 07] and knowledge-based [Felfernig et al. 06] reasoning systems will help to better fulfill the customers' requirements and ease the maintenance of the sales systems themselves [Burke 02]. Furthermore Folksonomies[10] and social ontologies can help to improve product categorization- and description tasks [Rollett et al. 07]. In all these areas "Online Crowds" can provide useful and aggregated sales-, behavioral- and customer related data. If accessible and canalized in an appropriate way this can be valuable knowledge created by the collective "intelligence" of the crowd.

[10] Folksonomies are evolving superimposed structures which are created collectively by online users e.g. with social tagging [Rollett et al. 07].

Some challenges and perils of “Online Crowds” are:

- Misguided online bubbles and fads, initiated by dubious initiators
- Susceptibility of online users to follow trends and phenomena which are disadvantageous for themselves
- Misuse and intrusion of user privacy because of careless “Online Crowds”
- Simulated pseudo-democracy and mob behavior on the Web

The listed shortcomings are becoming more and more evident these days. “Online Crowds” are hard to predict and even harder to control if they are fully developed. So anti-democratic movements, cyber-crime, character assassination in politics and business are already present in the online space [Groebel et al. 01]. In some cases the “collective intelligence” seems to end up in irrational exuberance and misguided results. Moreover intensive user-tracking and profiling is performed by leading online enterprises without any objections of the affected online users. It is absolutely open what will happen with this sensitive information in the future. Indicating examples are global search engines and their privacy philosophies [Russ 07].

6 Related Work and Conclusion

Mass phenomena are nothing new, they have already been investigated in the 19th century by scientists like [Le Bon 82] and in the last century by [Drucker 87] and others. This sociological research concentrated on individual and collective behavior, political movements and protest waves. The origin of this research can be found in the social psychology and the crowd research [Marx and McAdam 94].

The differences today are mainly the quality and the speed how collective behavior emerges. The Internet and the attached properties like speed, real-time, virtuality, interactivity, duplication and low communication cost help to grow such phenomena in a much more aggressive and unpredictable way [Bonabeau 04].

One facet of this topic is discussed by many researchers who work on building strategies of online communities. They mainly try to identify the interrelationship of users and contextual characteristics which facilitate this process [Preece 01]. But normally communities grow more slowly and are more predictable than Online Crowds. Another view on crowds was defined by [Henein and White 06] with the “Swarm Information Model”, which focuses on intrinsic factors of individuals as main drivers of simple crowds effects.

Agent based systems and virtual crowd simulations are additional fields which share some similarities with “Online Crowds”. In detail, the structure of crowd phenomena and their composition, as well as its visualization, are research topics in this area [Goldstone and Janssen 05], [Pettre et al. 06]. Positive as well as negative effects of self-structuring social organization through technology mediated behavior are discussed by [Rheingold 03] and others.

This paper presented an interdisciplinary approach how some “social contagion” and extraordinary mass phenomena can emerge on the Internet. Based on the concepts and theories of the crowd psychology (Real Crowds) a model of “Online Crowds” has been developed. Real people as well as online users take decisions based on the rationale of others (other directedness), if the situation is unpredictable or new. Because communication and collaboration technologies of the Internet are fertile environments to develop extraordinary waves of collective behavior, it can sometimes end up in irrational herd behavior [Huang and Chen 06].

Based on this knowledge it is possible to design and set up new online business models which increase the user traffic as well as the profitability of online ideas. The "Epidemic Online Model" may help C2C-, P2P- and related platforms to build up online inductees in an economical and low cost way. This is also a chance for new and small innovators to take over market shares and to outdate today's online giants of the Web if they facilitate "Online Crowd"-enabled strategies.

For sure the model is only a simplified approach to tackle the complexity of a multi-dimensional social phenomenon and more work has to be done to validate this concept. In particular the question about computability of the "Critical mass" still remains open and a mathematical tool set to predict the probability of prevailing online ideas is far away from being available. Consequently additional research has to be done on the factors to ensure sustainability and robustness of an "Online Crowd" over some time. Finally it has to be clarified, whether the suggested model is also applicable on other NMT like mobile, digital TV and so forth. This remains to be done in future work.

References

[Adomavicius and Tuzhilin 05] Adomavicius G. and Tuzhilin, A.: "Towards the next generation of recommender systems: A survey of the state-of-the-art and possible extensions"; IEEE Transactions on Knowledge and Data Engineering, 17 (6) (2005), pp.734-749

[Arthur 94] Arthur W. B.: "Increasing Returns and Path Dependence in the Economy"; University of Michigan Press (1994)

[Barabasi and Bonabeau 03] Barabási A.L., Bonabeau E.: "Scale-free networks"; Scientific American 288, (2003), pp.60-69

[Bikhchandani et al. 98] Bikhchandani S.,Hirshleifer D., Welch I.: "Learning from the Behavior of Others: Conformity, Fads, and Information Cascades"; American Economic Association in Journal of Economic Perspectives, (1998), pp.151-170

[Bonabeau 04] Bonabeau E.: "The Perils of the Imitation Age"; Harvard Business Review Article, Jun 1, pp.45-47, 49-54 (2004)

[Bshary and Grutter 06] Bshary R., Grutter A.S.: "Image scoring and cooperation in a cleaner fish mutualism"; Nature 441, pp.975 - 978 (22 Jun 2006) Letter

[Burke 02] Burke, R.: "Hybrid recommender systems: Survey and experiments"; User Modeling and User-Adapted Interaction. 12 (4), (2002), pp.331-370

[Dodds et al. 03] Dodds P.S., Muhamad R.,Watts D.J.: "An Experimental Study of Search in Global Social Networks"; Science 8 August 2003: Vol. 301. no. 5634, pp. 827 - 829

[Drucker 87] Drucker P.: "Die Chance des Unternehmers. Signale für das Management von morgen"; Econ, München (1987)

[Felfernig et al. 06] Felfernig, A., Friedrich, G., Jannach, D., Zanker, M.: "An Integrated Environment for the Development of Knowledge-Based Recommender Applications"; Intl. Journal of Electronic Commerce, Special issue on Recommender Systems, 11(2), Winter 2006-7, pp. 11-34

[Gladwell 01] Gladwell M.: "The Tipping Point: How Little Things Can Make a Big Difference"; Little Brown (2001)

[Goldstone and Janssen 05] Goldstone R.L., Janssen M.A.: "Computational Models of Collective Behavior"; Trends Cogn Sci. 2005 Sep. 9 (9) (2005), pp.424-30

[Groebel et al. 01] Groebel J., Metze-Mangold V., van der Peet J., Ward D.: “Twilight Zones in Cyberspace: Crimes, Risk, Surveillance and User-Driven Dynamics”; Stabsabteilung der Friedrich-Ebert-Stiftung, Bonn (2001)

[Henein and White 06] Henein C.M., White T.: "Information in Crowds: The Swarm Information Model"; Lecture Notes in Computer Science, Springer Berlin / Heidelberg (2006)

[Hippel 05] von Hippel E.: "Democratizing Innovation"; The MIT Press (2005)

[Hof 05] Hof R.D.: "The Power Of Us - Mass collaboration on the Internet is shaking up business"; The Future of Tech, Business Week, June 20, 2005 (2005)

[Huang and Chen 06] Huang J.H., Chen Y.F.: "Herding in online product choice"; Psychology and Marketing, Volume 23, Issue 5 , Pages 413 - 428, Special Issue: Marketing and E-commerce . Issue Edited by Ming-Hui Huang (2006)

[Kindleberger 78] Kindleberger C. P.: “Manias, Panics, and Crashes: A History of Financial Crises”; Basic Books, New York (1978)

[Le Bon 82] Le Bon G.: "Psychologie der Massen"; 15. Auflage, Kröner Verlag, Stuttgart (1982)

[Marx and McAdam 94] Marx G.T., McAdam D.: "Collective Behavior and Social Movements: Process and Structure"; Prentice Hall, New York (1994)

[Ossimitz and Lapp 06] Ossimitz G., Lapp C.: “Das Metanoia Prinzip – Eine Einführung in systemgerechtes Denken und Handeln“; Franzbecker Verlag, Hildesheim, Berlin (2006).

[Pelzmann 00] Pelzmann L.: „Wirtschaftspsychologie, Behavior Economics, Behavioral Finance“; Arbeitswelt, 3., erw. Aufl. Wien-New York: Springer (2000)

[Pettre et al. 06] Pettré J., Ciechomski P.H., Maïm J., Yersin B., Laumond J.P., Thalmann D.: "Real-time navigating crowds: scalable simulation and rendering"; Computer Animation and Virtual Worlds, Volume 17, Issue 3-4 , Pages 445 - 455, Special Issue: CASA 06 (2006)

[Preece 01] Preece, J.: "Online communities: Usability, Sociability, Theory and Methods."; In R. Earnshaw, R. Guedj, A. van Dam and T. Vince (Eds) Frontiers of Human-Centred Computing, Online Communities and Virtual Environments. Springer Verlag: Amsterdam, pp263-277 (2001)

[Rheingold 03] Rheingold H.: “Smart Mobs: The Next Social Revolution”; Perseus Books (2003)

[Rollett et al. 07] Rollett H., Lux M., Strohmaier M., Dösinger G., Tochtermann K.: "The Web 2.0 way of learning with technologies"; Int. J. Learning Technology (2007)

[Russ 07] Russ C.: „Heute schon “gegoogelt”? Privatsphäre und Internet-Suchmaschinen“; DACH-Security 2007, Hsg. Patrik Horster, Klagenfurt (2007)

[Schelling 78] Schelling T. C.: “Micromotives and Macrobehavior”; Norton, W. W. & Company, Inc (1978)

[Shiller 00] Shiller R. J.: „Irrationaler Überschwang“; Campus Verlag GmbH, Frankfurt/Main (2000)

[Watts 04] Watts D. J.: “The "New" Science of Networks”; Annual Review of Sociology Vol. 30: 243-270, (2004)

[Zanker et al. 07] Zanker M., Jessenitschnig M., Jannach D., Gordea S.: "Comparative evaluation of different recommendation strategies in a commercial context"; to appear IEEE Intelligent Systems, Special Issue on Recommender Systems (2007)

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

WordFlickr: A Solution to the Vocabulary Problem in Social Tagging Systems [1) 2)]

Josef Kolbitsch
(Graz University of Technology, Austria
josef.kolbitsch@tugraz.at)

Abstract: Allowing users to publish and share photos on the Internet makes Flickr one of the most popular tagging services currently available. The organisation of images in Flickr is based on Folksonomies, where users attach loose metadata—instead of well-defined terms from a controlled vocabulary—to their images. Although this lowers the barrier to participation it has a number of negative effects and can make searching, for instance, more difficult.
This paper offers a solution to a particular issue that can be encountered in Flickr—the Vocabulary Problem. The suggested approach is based on the use of a semantic lexical database for expanding Flickr queries. WordFlickr, a prototype implementation of this concept, is presented together with FlickrClustr, a related tool for clustering Flickr search results. Results of informal tests with these tools are provided, and characteristics of tag usage are derived.

Keywords: Social Tagging, Vocabulary Problem, Flickr, Folksonomies, Web 2.0.
Categories: H.3.3, H.3.5, H.3.7.

1 Introduction

Social bookmarking, file sharing, and similar services are becoming increasingly popular. One of the underlying concepts of these services is tagging, which refers to the users' attaching loose metadata to the information items they store (e.g., [Mathes 2004; Hammond et al. 2005; Tonkin 2006]).

This paper gives a brief overview of the nature of tagging systems. More importantly, a solution to a frequently encountered issue is offered. The following sub-sections describe tagging systems and two popular examples: del.icio.us and Flickr. With WordFlickr, a response to the Vocabulary Problem in tagging systems, and Flickr in particular, is presented in [section 2]. [Section 3] discusses WordFlickr's concept and the results of an informal comparison between Flickr and WordFlickr.

1.1 Tagging Systems

In conventional libraries all items including books and magazines are stored in a catalogue. For every resource, the catalogue usually retains information on the author, a primary category, optional secondary categories, a list of keywords, and other metadata. The categorisation in such catalogues is frequently based on strict, hierarchical classification systems, and the keywords often originate from controlled

---

1) This paper is dedicated to Jennifer A. Lennon. You will be missed.
2) An extended version of this paper is available at http://www.kolbitsch.org/research/.

vocabularies. In many of the taxonomies currently in use every category has exactly one parent, i.e., it can be sub-concept of exactly one super-concept.

Traditional libraries employ experts for generating accurate keywords and for classifying resources. A slightly different approach is to let content authors assign keywords and categories to the information they produce, and have experts approve the authors' input.

A recent trend in metadata generation and information classification is the participation of the community. Users may attach loose, unstructured metadata to resources. These metadata fields, named tags, serve both as keywords and as "classes". When a user attaches the tags "Sherlock Holmes" and "fiction" to a resource other users can find the resource, for instance, by searching for "Sherlock Holmes" or by browsing the all resources tagged "fiction". If a tag is assigned to a resource several times by different users its weight increases, and it might be more significant than other tags also assigned to the same information item.

In systems based on tagging, any string can be used as a tag. Hence, metadata creation does not rely on controlled vocabularies and does usually not yield classic, hierarchical taxonomies.

1.1.1 Structure of Tagging Systems

A tagging system has at least three types of entities: information items, tags attached to these items, and users creating information items and tags. Every information item has (at least) one author and has usually at least one tag assigned. Depending on the service, items can be links to resources on the Web, images, video clips, etc. Although tags may, in theory, be arbitrary strings many tagging system restrict the users' freedom and impose rules on the use of tags. In some systems, every tag may only consist of a single word, while other systems allow multiple terms per tag; in some systems tags are case sensitive, while in others they are not; etc.

1.2 del.icio.us

Numerous tagging systems with varying functionality exist today. Popular examples are del.icio.us, a social bookmark manager, CiteULike, an online service for managing academic literature, and Flickr, a photo sharing service (see [del.icio.us 2006; CiteULike 2006] and [section 1.3]).

In del.icio.us, users can store bookmarks and assign tags to every bookmark (see [del.icio.us 2006]). Users can make their bookmarks private or public. People exploring a bookmark in del.icio.us are presented a list of all users that tagged this reference as well as all tags assigned by all users. When selecting a user's name people can browse all public bookmarks of this user. On choosing a tag, all bookmarks stored with the same tag in del.icio.us are displayed.

1.3 Flickr

Flickr is a community-based web-site for organising and sharing photos (see [Flickr 2006]). Users can upload images to Flickr's server, attach tags, and insert notes into pictures. Whenever an image is displayed all associated tags and the photographer's name are provided. Similar to del.icio.us, selecting a user's name shows all public photos of this user. Selecting a tag results in a list of all images that have the same tag attached.

1.4 Motivation

On close inspection, it can be seen that the concept of del.icio.us and its consequences are rather different from Flickr's approach. Although both services are tagging services, del.icio.us allows multiple users to store the same item in the system and lets users attach potentially different tags. Hence, eventually numerous tags with varying frequencies might be assigned to the same item. While a small number of tags might be dominant, a wide range of disparate terms can be available (see [Golder and Huberman 2005]). This means that users would be capable of finding a given item using many different queries. This is a type of "divergence criterion".

In Flickr, however, the same item (image) is usually only stored once, and the tags initially assigned by the photographer are rarely modified or expanded (e.g., [Marlow et al. 2006]). Therefore the divergence criterion explained above cannot be satisfied.

This leads to a variation of the Vocabulary Problem. [Furnas et al. 1987] describes that users interacting with computing systems frequently employ very different terms for naming the same objects, tasks, commands, actions, etc. When asked, for example, to assign a term to an object in an information system user A might use an expression not obvious to user B and vice versa. This is the Vocabulary Problem.

In Flickr, a variant of this problem can be encountered. When users upload photos to Flickr they perceive the tags they assign to their images as apparent. Other users, however, might not think of these tags when searching for photos and might therefore not be able to find them. The Vocabulary Problem is widely considered a problem in tagging system (e.g., [Golder and Huberman 2005; Furnas et al. 2006; Mathes 2004]). In this paper, an attempt is made to counter this deficiency. [Section 2] presents a potential solution for Flickr and similar systems. Moreover, a prototype implementation of the concept is provided.

2 Enhancing Flickr Queries

2.1 WordFlickr

WordFlickr is a dictionary-based approach to solving the Vocabulary Problem in Flickr (see [section 1.1]). Since this approach requires some type of dictionary and the majority of tags in tagging systems are in English (see [Guy and Tonkin 2006]), the concept and implementation focus on the English language.

2.1.1 Concept

Users querying Flickr for the term "shoe", for instance, are usually presented with photos of all kinds of shoes. However, if the photographers publishing photos use tags such as "slippers" or "boot" (instead of "shoe") users would not be able to find these photos with their query for "shoe".

WordFlickr offers a solution by allowing users to expand their queries with semantically related terms. In WordFlickr, the users' queries are analysed and expanded using a database called WordNet. WordNet is a sort of a lexical ontology developed at Princeton University. It contains words of the English language together with semantic relations and is both a dictionary and a thesaurus (see [WordNet 2006]). However it does not only contain relations such as synonyms and antonyms but also more complex relations including holonyms, meronyms, hypernyms, and hyponyms.

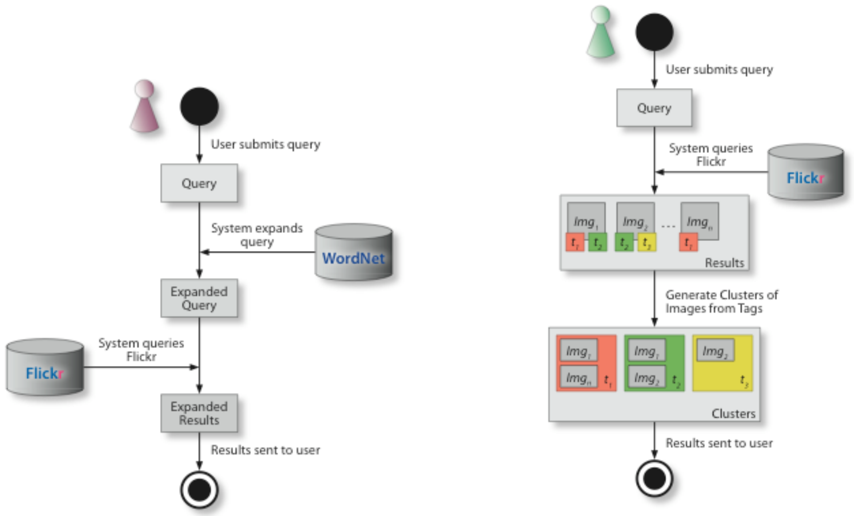


Figure 1: *Structures of WordFlickr (left) and FlickrClustr (right). In WordFlickr, the user's query is expanded using the WordNet database and is sent to Flickr. With FlickrClustr, a user's query is sent directly to Flickr. The results' tags are used for generating clusters of images.*

When users submit queries to WordFlickr they can choose which types of relations are used for expanding their initial queries. As depicted in [Fig. 1], the user's query is analysed and looked up in WordNet. After word stemming is applied, terms such as synonyms and hypernyms are retrieved from the WordNet database according to the user's settings. The modified, expanded query is submitted to Flickr, and Flickr's response is sent to the user. The WordFlickr prototype is available online at [Kolbitsch 2006a].

2.1.2 Discussion of the Implementation

WordFlickr's solution to the Vocabulary Problem depends on the content stored in tags. Therefore a dictionary such as WordNet required, which has a negative effect on the scalability of the service. On the one hand, WordNet offers a lexical database with powerful semantic relations that can be used well for finding semantically equivalent and similarly related terms. On the other hand, WordNet restricts WordFlickr to the

English language. Words from other languages as well as technical terms and personal tags cannot be expanded.

2.2 FlickrClustr

First informal experiments with WordFlickr showed that it was hard to "analyse" the results of the queries (see also [section 3]). It was not possible, for instance, to tell if images found through WordFlickr queries were more closely related to the terms searched for than in traditional Flickr searches. Therefore another tool was implemented that can make it easier to analyse search results from both Flickr and WordFlickr—FlickrClustr.

2.2.1 Concept

FlickrClustr is an approach to make use of simple clustering in Flickr. FlickrClustr analyses the search results of Flickr queries and the tags assigned to the images in the results. The most frequently used tags are the basis for forming image clusters.

The aim of FlickrClustr is to produce clusters from any search results in Flickr. This enables users to find the most frequently used tags within the results, to group large numbers of search results into a small set of clusters, and also gives them easy access to potentially related tags.

At the moment only identical tags are clustered, so even tags spelt uppercase and lowercase are regarded as separate tags. Clusters could be made "more tolerant" by transforming all terms lowercase and by employing a word stemming mechanism. However, this functionality is deliberately not incorporated in the current research prototype in order to be able to collect information on the usage of tags—upper- vs. lowercase, singular vs. plural forms, etc. Details are provided below.

3 Discussion

3.1 Concept and Implementation

With WordFlickr, the impact of the Vocabulary Problem can be reduced. Hence, a solution to a single, yet common issue in tagging systems can be offered. Moreover, the basic level variation described in [Golder and Huberman 2005] can be dealt with by including hypernyms and hyponyms in the expansion of the users' queries.

However, WordFlickr cannot consider further shortcomings such as the use misspelt tags, personal tags such as "my dog", and individually encoded tags like "VacationInHongKong" (see [Guy and Tonkin 2006]). Moreover, the concept of WordFlickr is not capable of providing a solution to tag ambiguity, polysemy, and synonym control (see [Mathes 2004]).

3.2 Informal Comparison

An informal comparison was performed between Flickr and WordFlickr. The sample consists of a pseudo-random selection of 21 tags. These terms were selected by the author of this paper, which can be seen as a methodological concern. However, the intention is not to conduct a formal experiment and to present quantitative results. The aim is to find out in which way results from WordFlickr differ from those of traditional Flickr queries.

For each of the words both Flickr and WordFlickr were queried from November 11th to 13th, 2006. “Query by tag”, “order by relevance”, and “≤ 200 results” were chosen as search options. Synonyms, hypernyms, and hyponyms were included in WordFlickr queries. For all queries, clusters of the ten most frequently used tags in the search results were generated with FlickrClustr. The results for three representative tags are listed in [Tab. 1].

Query	*wine*		*rock*		*shoe*	
Sys	**Flickr**	**WordFlickr**	**Flickr**	**WordFlickr**	**Flickr**	**Word-Flickr**
Ten Most Frequent Tags	wine (161)	wine (174)	rock (125)	rock (186)	Shoe (168)	shoe (161)
	Wine (39)	chardonnay (80)	music (42)	limestone (172)	shoes (50)	pump (52)
	party (20)	rose (49)	Rock (39)	ocean (28)	foot (33)	shoes (38)
	vino (16)	vin (37)	ROCK (36)	geology (28)	feet (31)	flipflop (31)
	glass (12)	pinot (28)	live (29)	water (27)	Shoe (22)	foot (21)
	2004 (11)	france (27)	metal (26)	cave (25)	legs (11)	feet (20)
	2005 (10)	white (27)	show (23)	coast (23)	Leg (10)	heel (17)
	yen (9)	bottle (22)	punk (22)	climb (22)	toes (10)	flip-flop (14)
	Brokenwood (9)	cabernet (22)	musica (22)	landscape (22)	heel (9)	red (14)
	roadtrip (9)	Wine (22)	york (20)	sea (22)	black (9)	sneakers (13)

Table 1: *Selected results from a comparison between Flickr and WordFlickr queries.*

3.2.1 General Findings

As the clustered tags show (see [Tab. 1]), the use WordFlickr has an impact on the diversity of tags. In some cases, WordFlickr’s overall range of tags is narrower while the diversity in the requested “category” is wider. Flickr’s results for “wine”, for instance, include the tags “glass” and “party”. Hence, the overall range of terms is rather wide including even word that are not directly related to the original query. WordFlickr’s results, on the other hand, include tags such as “chardonnay” and “rose”. These terms refine the term “wine”, and increase the (vertical) variety of terms directly related to the user’s query.

This aspect also has an effect on serendipity. While users can still explore the tags in WordFlickr’s search results and find potentially interesting and new resources, their scope might be limited to a narrower context than in regular Flickr queries.

3.2.2 Tag Inconsistencies

As described elsewhere, tag usage in systems like Flickr and del.icio.us is inconsistent (see [Mathes 2004; Guy and Tonkin 2006]). The results from the comparison between

Flickr and WordFlickr support these findings. The tags in the query results for "shoe" in [Tab. 1] illustrate several discrepancies that may occur. Both singular and plural forms of nouns are present: shoe and shoes, foot and feet, leg and legs. Moreover, both upper and lower case are used: shoe and Shoe.

As mentioned above, FlickrClustr could be easily modified so that inconsistent terms are merged into single clusters. Although this would "correct" the clusters generated by the system, the actual problem would remain unsolved because a query for "shoe", for instance, might still not include photos tagged "shoes".

3.2.3 Tag Usage and Interpretation

This section qualitatively analyses the tags present in the results of the Flickr and WordFlickr queries listed in [Tab. 1] and offers an interpretation. The first finding is that, in some cases, WordFlickr yields less tags that are only relevant to owners of photos than Flickr. When Flickr is queried for "wine", for instance, personal tags such as "2005", "party", and "roadtrip" are frequently used among the resulting images. These tags are often not relevant for people searching images but only for the owners of photos. The WordFlickr query for "wine", on the other hand, is able to eliminate such tags. WordFlickr lists tags such as "chardonnay", "rose", and "pinot" which are vine varieties (i.e., hyponyms). Hence, WordFlickr is capable of offering tags in the results that are semantically closer to the user's query. (However, it cannot be stated if the images provided by WordFlickr are more appropriate for the user.)

The query for "rock" shows an interesting effect. While Flickr includes tags such as "music", "live", "metal", and "show", WordFlickr provides tags like "limestone", "geology", and "climb". This means that Flickr's query relates the term to rock music, whereas WordFlickr interprets the term "rock" as stone. This is particularly astonishing because a manual WordNet query shows that the lexical database retains both meanings of the word "rock". However, in this example WordFlickr also includes tags such as "ocean", "water", and "sea". Since they are not directly connected with the user's query some of WordFlickr's most frequently used tags are semantically even more distant that Flickr's tags.

The last of the three terms analysed is the example used throughout this paper—"shoe". In this case, the results of both Flickr and WordFlickr resemble each other. The influence of the WordNet database can be seen in the tags produced by WordFlickr: they include tags such as "pump", "flip-flop", and "sneakers" (i.e., hyponyms). However, both queries contain significant numbers of tags that are not directly associated with the user's query: "foot", "leg", and "heel". Hence, in this situation WordFlickr cannot provide more appropriate results than Flickr.

4 Conclusion

In this paper, WordFlickr was introduced as a solution to the Vocabulary Problem. The concept of WordFlickr is based on WordNet, a lexical database with semantic relations between words from the English language.

An informal experiment compares search results from the prototype implementation of WordFlickr with results from Flickr. It is hardly possible to quantise the differences between Flickr and WordFlickr. The qualitative results, however, imply that WordFlickr excels Flickr in some cases, where the most frequent tags in WordFlickr's search results are semantically closer to the users' initial query than in Flickr's results. On other occasions, though, the results of Flickr and WordFlickr are alike. Hence, it can be concluded that WordFlickr yields results that are, at worst, as good as Flickr's search results. Therefore the use of WordFlickr's concept can be a valuable addition to tagging systems such as Flickr.

References

[CiteULike 2006] CiteULike, http://www.citeulike.org/.

[del.icio.us 2006] del.icio.us, http://del.icio.us/.

[Flickr 2006] Flickr, http://www.flickr.com/.

[Furnas et al. 1987] Furnas, G. W., Landauer, T. K., Gomez, L. M., Dumais, S. T.: "The Vocabulary Problem in Human-System Communication"; Communications of the ACM, 30, 11 (1987), 964-971.

[Furnas et al. 2006] Furnas, G. W., Fake, C., von Ahn, L., et al.: "Why Do Tagging Systems Work"; Proceedings and Extended Abstracts of the Conference on Human Factors in Computing Systems (CHI 2006), Montréal, QC, Canada (2006), 36-39.

[Golder and Huberman 2005] Golder, S., Huberman, B. A.: "The Structure of Collaborative Tagging Systems"; http://www.hpl.hp.com/research/idl/papers/tags/tags.pdf, Accessed October 11th, 2006.

[Guy and Tonkin 2006] Guy, M., Tonkin, E.: "Folksonomies: Tidying Up Tags?"; D-Lib Magazine, 12, 1 (2006). See also http://www.dlib.org/dlib/january06/guy/01guy.html.

[Hammond et al. 2005] Hammond, T., Hannay, T., Lund, B., Scott, J.: "Social Bookmarking Tools (I): A General Review"; D-Lib Magazine, 11, 4 (2005).

[Kolbitsch 2006a] WordFlickr, http://www.kolbitsch.org/research/wordflickr/.

[Kolbitsch 2006b] FlickrClustr, http://www.kolbitsch.org/research/flickrclustr/.

[Kolbitsch and Maurer 2006] Kolbitsch, J., and Maurer, H.: "The Transformation of the Web: How Emerging Communities Shape the Information we Consume"; Journal of Universal Computer Science, 12, 2 (2006), 187-213.

[Marlow et al. 2006] Marlow, C., Naaman, M., Davis, M., boyd, d. (sic): "Taggin Paper, Taxonomy, Flickr, Academic Article, ToRead" (sic); Proceedings of the Seventeenth Conference on Hypertext and Hypermedia (HT06), Odense, Denmark (2006), 31-40.

[Mathes 2004] Mathes, A.: "Folksonomies – Cooperative Classification and Communication Through Shared Metadata"; http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.pdf, Accessed January 25th, 2006.

[Tonkin 2006] Tonkin, E.: "Folksonomies: The Fall and Rise of Plain-text Tagging"; Ariadne, 47 (2006).

[WordNet 2006] WordNet, http://wordnet.princeton.edu/.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

The Three Pillars of 'Corporate Web 2.0': A Model for Definition

Alexander Stocker, Gisela Dösinger
(Know-Center Graz, Austria
{astocker, gdoes}@know-center.at)

Anwar Us Saaed
(Knowledge Management Institute,
Graz Technical University, Austria
anwar.ussaeed@student.TUGraz.at)

Claudia Wagner
(Institute for Networked Media,
Joanneum Research, Austria
claudia.wagner@joanneum.at)

Abstract*:* There is plenty of hype around the term *Web 2.0*. While the change of the web has gathered increased attention from the Web 2.0 community itself, it seems that the corporate world feels unconfident how to apply the principles of the Web 2.0 to their businesses. Due to the ambiguity and fuzziness of the concepts describing Web 2.0 there is a lot of uncertainness. Highly affected industry branches like the media industry show interest but their fears of loosing their competitiveness because of not knowing how to handle the phenomenon Web 2.0 are evenly strong. Nevertheless, little academic work has been conducted on the implication of Web 2.0 to the business domain. This paper bridges the gap in having a deeper look into the phenomena of Web 2.0 leading to the development of a more graspable model for corporate use of Web 2.0 encapsulating a business focus and hence defining the term *Corporate Web 2.0*. By showing how the introduced model is applied, it helps companies including the media and other businesses to derive a business value from the new Internet.

Keywords: Web, Web 2.0, Corporate Web 2.0, Business Model, Business Processes, Social Aspects of Web, Technological Aspects of Web, Business Aspects of Web
Categories: A.0., A.1., J.4.

1 Introduction

In the year 2004 the term *Web 2.0* was first introduced by Tim O'Reilly to summarize a set of principles and practices attracting attention. Three years later, a vast number of corresponding applications and more than 1400 Web 2.0 start-ups [web2slides, 07], [MapofWeb2.0, 07] considerably signalize that Web 2.0 is throughout the land.

The Web 2.0 – circumscribed by eight principles [O'Reilly, 04], however not concretely defined – has emerged as the new powerful and dynamic user centred web, by default equipped with social features. The majority of Web 2.0 applications including podcasts, blogs and wikis are designed according to the principle of user generated content. These facets of Web 2.0 making the users the masters of the dynamics created hype on the web. In the Web 2.0 world users and their interactions (connect, communicate and collaborate) are the sources of information shaping a kind of collective intelligence. Companies naively tried to copy the success of companies like Amazon (www.amazon.com) and Spreadshirt (www.spreadshirt.com) which have entirely based their business on Web 2.0.

The McKinsey Global Survey How businesses are using Web 2.0 presented a widespread but careful interest in Web 2.0 technologies by companies around the globe [McKinsey, 06]. Companies want to adopt Web 2.0 for being more successful in their business, but the ambiguity surrounding the business perspective as regards Web 2.0 seems to hinder high adoption rates in the corporate environment.

Analysis of the discussion on the web and the results of related surveys [McKinsey 06], [ArthurD.Little, 06], [BoozAllenHamilton, 06] revealed the demand for a clear concept for Web 2.0 in the business context: Three main observations have been levied: (1) People talk about Web 2.0 in the company with different concepts and definitions in mind. (2) The internet showed some tendency towards monopolization: principles and patterns of Web 2.0 have been partly derived from the large web-companies surviving the 2001 dot-com bubble, but nevertheless these companies were founded without having knowledge of these principles. From the perspective of the authors, those principles cannot be simply applied to the large number of SMEs trying to get some business value from the Web 2.0. (3) To the best knowledge of the authors, there are very few scientific publications outlining the organizational impact of Web 2.0.

For that reasons, the authors try to define the term Corporate Web 2.0 by introducing a more general model as an effort to remove the ambiguity of the concepts of Web 2.0 in the business context. The paper proceeds with Section 2 discussing the special demand of clearer concepts to the media industry which is highly affected from the new trends on the internet. Section 3 takes a deeper look into the phenomenon Web 2.0 and tries to spot generic new elements of Web 2.0, finally leading over to the presentation of a model for Corporate Web 2.0. Section 4 addresses the ability of Web 2.0 of reshaping companies with respect to the predefined model. The paper closes with a conclusion and refers to future work in section 6.

2 Relevance of Corporate Web 2.0 to the media industry

There are three ways to look at how a society is informed: The first is that people are trusting and will read, listen to and watch what is provided to them. Second, people seek out for an informed broker to tell them what is important for them. Third, people can, because of their smartness, sort out relevant things for themselves and may find their own vision of the truth [Bowman, 03]. Already in 1995, Nicolas Negroponte [Negroponte, 95] predicted that readers of online news would be given the ability to choose only topics and sources of interest to them. He called it the *Daily Me*. Guardians of traditional journalism were grumbling but nevertheless the portals of online media industry showed a huge tendency towards enabling personalization for their customers. What Negroponte was not able to predict in 1995 is the fact that in the future people will not only refer to the personalization of media-content delivered to them by means of personal filtering. In addition, collective filtering will be done by the internet-audience to separate the subjectively interesting material from junk. So Web 2.0 has introduced a fourth way on how the society is informed – a user driven way. With regard to Web 2.0 the media industry is perpetually losing its traditional role as an intermediary and gatekeeper to the world of content.

As outlined above, the media industry is one of the industries that as a result of the relationship to content and media-channels is highly affected by what is summarized under Web 2.0. The huge participation of the users combined with the enormous quantity of content created by them has sent an alert to almost all the business divisions. Its implications to the content or media related businesses including music companies, news agencies, publishing companies and television companies led to an evolution of innovative business start-ups handling this type of content. Traditional media companies belonged to the forerunners, showing interest in the challenges and opportunities arising from the new internet. Nevertheless, due to the lack of concrete concepts and models they started to fear for their competitiveness because of expecting market segment loss on the battlefield internet.

The new internet era of content creating users (www.flickr.com, www.digg.com, www.youtube.com, www.slashdot.com, ..) together with the visible growth of users favouring the consumption of user created content over the content created by the media industry is deemed to show a huge impact on the traditional media business. The ease of use in transparent, collaborative and participatory environments together with networking constitute the pillars of the new media paradigm. The relevance of this topic for the media business is pictured by a number of activities carried out by media business. A German example is the buy-in of Pro7-Sat1 Consortium, purchasing a 30% share on myvideo (www.myvideo.de interweaving user created content from myvideo with traditional television broadcast.

The explanation of the phenomenon Web 2.0 by referring to a more concrete model will serve as the foundation for media companies gaining benefit from Corporate Web 2.0. Though knowledge of such a model is fundamental, it is not limited to the media industry. The work provided in this paper will remove the ambiguity from Web 2.0 in the corporate context and will be the basis for the development of more conceptual business strategies based on Corporate Web 2.0 by the companies.

3 Towards a definition of Corporate Web 2.0

3.1 A deeper look into the phenomena of Web 2.0

The term Web 2.0 is understood in different ways by different people, depending on the domain of the community they belong to. Depending on the discipline its focus is set to (1) a set of helpful technologies that support in different ways including the support of social interactions on the web (e.g. Computer Science), (2) a social phenomenon resulting in the creation of communities and social-networks by using technology (e.g. Sociology) or (3) a creator of business value derived from technological and social structures of the new web (e.g. Business Administration).

From the perspective of the authors of this paper, Web 2.0 is a natural evolution of the web, not driven by a particular social or technological change. From a more generic view, the new elements of the Web 2.0 compared to what is generally understood by the old web can be reduced to (1) the technical simplification of tools for content creation, social interaction and community building due to the rapid technological evolution of the web and (2) the need of humans for social interaction with like-minded people which they more often locate on the web and (3) the fact, that the web became an important part of knowledge societies. To sum up, the web became more interesting, gaining the attraction of a lot more users, hence moving itself away from the web of the big institutionalized content creators and technocrats towards the web for everybody. To the business perspective, this implies that (1) more valuable data is automatically created by the users bottom-up leading to the opportunity of transforming this data into a business value by using currently developed methods [Fueller et al, 04] and (2) more users on the web imply more consumers since the web is seen as global marketplace from the perspective of the companies.

The adoption of Web 2.0 from the perspective of the companies encapsulates a strong business focus. Companies will grab the social and technological aspects inherited by the Web 2.0 in order to gain more efficiency and effectiveness in reaching the business goals leading to the *business aspects*. From the perspective of the authors of this paper, corporate usage of Web 2.0 (referred to as: Corporate Web 2.0) encompasses three different aspects: *Technological aspects* make up all the well known technologies and tools such as blogs, wikis, content syndication tools, ajax, dhtml and more. *Social aspects* reflect a mind-shift of the users who are encouraged taking part in the development of new social structures and content. This leads to the creation of virtual communities and social networks all over the web making the knowledge of the users explicit and helping a collective intelligence to develop [Kolbitsch, 06].

3.2 An approach for definition

Corporate Web 2.0 can be defined as transformation of the social and technological aspects of the new internet into business, leading to a redesign of existing business processes or even to an evolution of new business models.

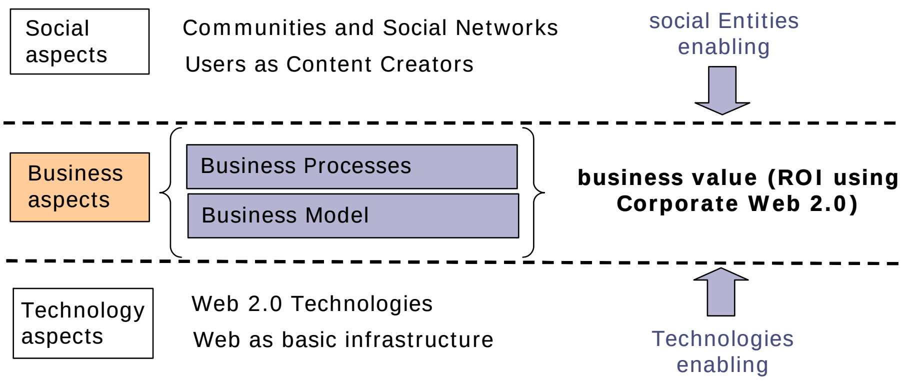


Figure 1: A model for Corporate Web 2.0

The business of a company can be defined by considering (1) the business model and (2) the business process model. While the term business process has achieved quite a good understanding of what it actually describes, the term business model is often interpreted in different ways. Hence the meaning of the term business model is explained by the authors with regards to the relevant literature.

The management of a company naturally derives its business strategy from the business vision in a first step. Second the business model, which can be seen in a nutshell as the logic of a company describing how to generate revenues, is shaped out of the business strategy. According to Timmers [Timmers, 98] the term 'business model' describes *'an architecture for the product, service and information flows, including a description of the various business actors and their roles; and a description of the sources of revenues'*. Osterwalder [Osterwalder 04] introduces a good explanation of what a business model actually does: *'A business model is a conceptual tool that contains a set of elements and their relationships and allows expressing a company's logic of earning money. It is a description of the value a company offers to one or several segments of customers and the architecture of the firm and its network of partners for creating, marketing and delivering this value and relationship capital, in order to generate profitable and sustainable revenue streams.'* Bergholtz [Bergholtz, 02] define a business model as a model concerned with value exchanges among business partners describing the *'what a company does'*.

According to the ISO Standard ISO9000:2000, a process is a *'set of interrelated or interacting activities which transform inputs into outputs'*. A business process model focuses on operational and procedural aspects of business communication, describing *'how the company does its business'* [Bergholtz, 02]. A business process is related to a concrete business activity like innovation, marketing or service and support.

In the traditional web business was quite different: The content available on the web was either created by the companies themselves or commercial content like advertisements. Links to the business processes were very weak or not existent. Many talks were based upon the influence of the technological impact of the internet to consumer to consumer, business to consumer or business to business scenarios.

In brief, technological and social phenomena have to be embedded (1) in the business processes, reflecting a stronger interweavement of business processes with the web or (2) in the business model itself. In a Corporate Web 2.0 environment employees, customers and partners contribute value on the one hand and obtain value on the other hand in a win-win scenario. Going for Corporate Web 2.0 introduces new feasibilities of engaging the stakeholders in a way in which both sides benefit.

4 Opportunities for Corporate Web 2.0

Companies may gain a business advantage by applying the Corporate Web 2.0 at some or even all of its business processes if the business is strongly interwoven with the web. As well, Corporate Web 2.0 enables the conceptual design of new business models. Hence the opportunities for companies in generating a business value from Corporate Web 2.0 logically relate to the dependence of the companies' business on the web. Companies like Google or Yahoo are strongly interwoven with the web and therefore might gain the maximum business value from Corporate Web 2.0.

Production-oriented companies like Audi or Nokia can benefit from reshaping some of the business processes to adopt Corporate Web 2.0 according to their business strategy.

The figure below shows the causality of the opportunities of Corporate Web 2.0 to generate business value to the dependence of a business on the web.

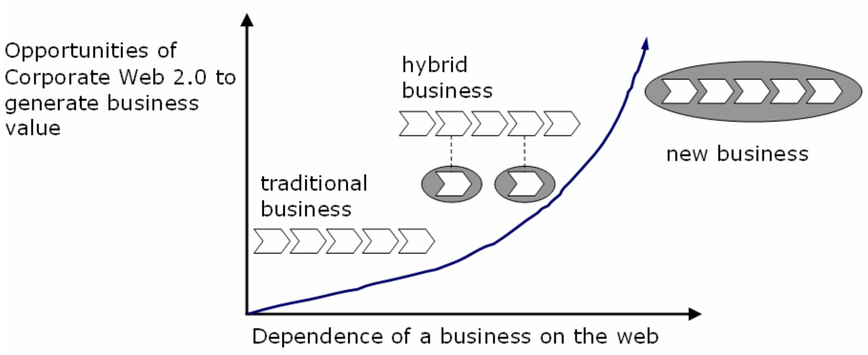


Figure 2: Corporate Web 2.0 and Dependence of Web for the business

First, the dependence of a business on the web can be measured in terms of number of business processes, which are interwoven with the web. The more business processes interwoven, the higher the number of opportunities for the companies in creating value from Corporate Web 2.0. Metcalf's Law and Reed's Law described the relationship of value and utility of a network with its size. Since Corporate Web 2.0 is about users such as employees, partners and consumers the non linear increase according to both laws is mapped to the factors on the two axes.

Five cases of companies considering the three pillars social, technological and business are presented in the next paragraphs. The corporate use of Web 2.0 (Corporate Web 2.0) focuses on the business perspective. Business processes showing more tendencies to the social structures of the web are more adequate to adopt Corporate Web 2.0: Innovation, product-development, marketing or service and support are empowered by the opportunity to benefit from the new knowledge structures on the web.

The development of the new infotainment console for Audi cars (www.audi.com) was driven by the users, visiting the Audi web-site [Fueller et. al, 04]. By the use of an interactive product-configuration-tool, expectations, preferences and trends were spotted with the help of the Audi community. Community based innovation opened up the innovation process to the web leading to a more accepted product-design on behalf of Audi's customers.

Lego (www.lego.com) introduced the concept of LEGO-Ambassadors, a community based program of LEGO hobbyists sharing their experience. Product development was opened to the Virtual LEGO-Communities gathering knowledge about preferences and wants, leading to the development of high price LEGO products such as the Imperial Star-Destroyer from Star-Wars which would not have been developed without the intervention of the community. These products became a huge business success.

Fender (www.fender.com), a manufacturer of guitars and related equipment founded in 1946 adapted its marketing process when launching its new campaign grounded on the sharing for multimedia-content on the web. The primary goals to reach more customers in a more innovative way and to attract more users to the Fender website were definitely achieved.

Beside that, a series of companies with new business models according to Corporate Web 2.0 has emerged. Two examples of companies, with business models based on content created by the users are introduced.

Revver (www.revver.com) provides a web-site for sharing user generated videos. Unlike Youtube Revver does not focus on social-network and community-building, but on marketing and value creation. Revver established a technology for the monetary valuation of user generated content by connecting users with potential companies acting as sponsors in a marketplace. The more views a video gets the more money the user can realize. The most popular video was viewed over 4 million times generating $25.000 for the amateur producers [ZdNet, 07].

Ideawicket (www.ideawicket.com) provides an ‘Open Innovation Portal’ as an instrument for innovators to socialize their concepts and products as well as for companies to source innovation from customers. Users are able to post ideas of products or services of need which are not available on the market yet. The amateur inventors will be connected to companies seeking out for innovative ideas for product development. Innovators and Companies can connect by using this platform. Companies may achieve a fresh perspective on new prototypes saving costs for creativity, while innovators are given a chance to present their ideas and concepts plus realization of value.

5 Conclusion and Future Work

In this paper the authors took a deeper look into the phenomenon Web 2.0. As a main contribution, a model of ‘Corporate Web 2.0’ focusing on the business perspective was developed. The predefined model highlighted the necessity of transferring social and technological aspects into business processes or the business model itself. Companies are enabled to act and find innovative ways using the model presented in the paper.

This paper established a foundation work which will be followed up by future contributions on the ascendancy of Web 2.0 on organizations. Future research will also include standard concepts for stake-holder engagement according to the Corporate Web 2.0 model, addressing the question to which extent social and technological aspects are valuable in this respect.

Acknowledgements

The Know-Center is funded by the Austrian Competence Center program Kplus under the auspices of the Austrian Ministry of Transport, Innovation and Technology (http://www.ffg.at), by the State of Styria and by the City of Graz. Part of this contribution is also funded by the Higher Education Commission of Pakistan (http://www.hec.gov.pk).

References

[ArthurD.Little, 06] Arthur D. Little, Web Reloaded? Driving Convergence in the ‘real world’, 2006.

[Bergholtz, 02]M. Bergholtz, P. Jayaweera, P. Johannesson, P. Wohed, Process Models and Business Models – A Unified Framework, in Proceedings of ER (Workshop), Springer Lecture Notes in Computer Science, 2002.

[BoozAllenHamilton, 06] Booz Allen Hamilton, Web 2.0 – Mythos oder Realität, 2007.

[Bowman, 03] S. Bowman, C. Willis, We Media. How audiences are shaping the future of news and information, The Media Center, 2003.

[Fueller et al, 04] J. Füller, E. Holger, H. Mühlbacher, Community Based Innovation: A Method to Utilize the Innovative Potential of Online Communities, in: Proceedings of the 37th HICSS Conference, Hawaii, 2004

[ISO9000:2000, 00] ISO - International Organisation for Standardization, Qualitätsmanagementsysteme: Grundlagen und Begriffe (ISO 9000:2000), Dezember 2000. http://www.iso.org.

[Kolbitsch, 06] J. Kolbitsch, H. Maurer, The Transformation of the Web: How Emerging Communities Shape the Information We Consume, February 2006.

[MapOfWeb2.0, 07] Web 2.0 Innovation Map, 2007, http://www.fourio.com/web20map/.

[McKinsey, 06] McKinsey, How businesses are using Web 2.0: A McKinsey Global Survey, 2006.

[Negroponte, 95] N. Negroponte, Being Digital, Vintage Books, 1995.

[OReilly, 04] T. O’Reilly, What is Web 2.0. Design Patterns and Business Models for the Next Generation of Software, 2004, http://tim.oreilly.com/lpt/a/6228.

[Osterwalder, 04] A. Osterwalder, The Business Model Ontology: A Proposition in a Design Science Approach, 2004.

[Timmers, 98] P. Timmers, Business Models for Electronic Markets, in Journal on Electronic Markets 8 (2), 1998.

[Web2slides, 07] Web 2.0 Slides, 2007, http://www.web2slides.com.

[ZdNet, 07], ZdNet: Tech News, Blogs and White Papers for IT Professionals, 2007, http://blogs.zdnet.com/micro-markets/?p=290.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

A Theory of Co-Production for User Generated Content – Integrating the User into the Content Value Chain

Tassilo Pellegrini
(Semantic Web School, Austria
t.pellegrini@semantic-web.at)

Abstract: The concept of co-production was originally introduced by political science to explain citizen participation in the provision of public goods. The concept was quickly adopted in business research targeting the question how users could be voluntarily integrated into industrial production settings to improve the development of goods and services on an honorary basis. With the emergence of Social Software and web-based collaborative infrastructures the concept of co-production gains importance as a theoretical framework for the collaborative production of web content and services. Current research in human computation has adopted the concept for the semantic enrichment of web content by collaborative tagging. This article argues that co-production is a powerful concept, which helps to explain the emergence of user generated content and the partial transformation of orthodox business models in the content industries.

Keywords: user generated content, tagging, co-production, collaboration, content value creation, media economics, innovation cycle, business model
Categories: A.0, K.0, K.4, K.5, K.6

1 Introduction

While from its early days the Internet has been a place of social interaction [21, 23] increasing spread of Internet connections and improvements in the usability of (mobile) web applications have lowered the participatory barriers for users to actively engage in virtual communities and the production of web content. The tremendous success of Wikipedia as a communitarian project of high quality content provision, the increasing popularity of blogging as a low cost personalized editorial activity and the rising importance of tagging platforms as referential architecture for the approval of relevance and quality has raised attention of how user generated content (UGC) is affecting and altering orthodox business models in content industries.

Media companies - especially aggregators - have already begun to take advantage of the increasing availability of UGC on the Internet for various reasons, either by setting up their own collaborative platforms or by taking over established service providers. With the growing amount of UGC available increasing attention is being paid to the question how collaborative action on the web can be used for value creation taking into account that users contribute time, skill and resources on a voluntary and honorary basis.

2 Foundations of UGC

Wikipedia.org defines user generated content as “various kinds of media content that are produced or primarily influenced by end-users, as opposed to traditional media producers, licensed broadcasters and production companies.” [24] UGC production embraces activities like programming, publishing, commenting, referencing, reviewing, rating, syndicating, tagging and querying.

UGC emerges out of a micro-content-based, self-organizing infrastructure. It is embedded within a “architecture of participation” based on the principles of bottom up networks, self-service, openness, self-regulation and decentralization. [18] In monetary terms tools and services are cheaply available and cost-effective, being designed not just for the production of content but also for the management of relationships and shared resources.

UGC embraces 1st order editorial content and 2nd order meta-content which makes it applicable for reuse outside the context it has been created in. Tags and folksonomies are examples for meta-content which is being used to generate surplus statements, views and references within a domain. Semantic enrichment through 2nd order content therefore is a crucial, indispensable means in the improvement of content services.

According to media economics goods that are being traded on a market can be characterized by their excludability from use and their rivalry in consumption. [15] By applying these two dimensions to UGC the following matrix can be drawn:

		Rivalry	
		Yes	No
Excludability	Yes	Private Goods (i.e. Universal Resource Identifiers	Club Goods (i.e. Blogs, Tags, Mashups, Folksonomies)
	No	Common Pool Goods (i.e. water, air, space) – *not relevant to UGC*	Public Goods (i.e. Open Access Repositories like Wikipedia)

The matrix shows that although UGC can have several good characteristics it is best described as a club good, characterized by non-rivalry in consumption but excludability in use. The latter aspect might be intriguing to those who think of web content as a public good because of its free availability in terms of monetary costs. But taking into account that the use of web content is bound to production resources like hardware, software, Internet connections, skills and time, and the usage of the content is highly often bound to conditions of use like access fees, referencing, revealing personal data, agreeing to user tracking and profiling, and/or the exposure to advertisements we find an exchange pattern between service provider and content producer that makes the club good paradigm applicable.

3 Web Collaboration as a Means of Co-production

Theories of co-production originate from public policy research. [3] According to Ostrom [19] “co-production is a process through which inputs from individuals who are not in the same organization are transformed into goods and services […] that transform citizens into safer, better educated or healthier persons." Co-production implies that users (i.e. citizens) can play an active role in producing (public) goods and services of consequence to them. According to Incera et al. [14] collaboration can be defined as a state of mutually beneficial relationships between two or more parties who work toward common goals by sharing responsibility, authority and accountability for achieving results. Hence collaboration can be seen as a specific mode of co-production to improve structural deficiencies in the content value chain. Research in media economics has shown that co-production is used as an efficient means to shift the costs of production from the service provider to the user. [11]

Empirical investigation in the development of open source software revealed that co-production amounts for about Euro 1,2 billion for freely available software which leads to 36% in savings in corporate research and development per year compared to conventional ways of software development in Europe. [9]

In the case of proprietary software over 650.000 programmers tested and co-developed Windows 2000 representing as much as $ 500 million worth of effort. [5]

Investigations into the Japanese mobile mail magazine market revealed that it is almost entirely built on the honorary engagement of users in the production and continuous provision of media content. [8]

Recent research in human computing has focussed on the improvement of collaborative infrastructures for value creation in service provision like the semantic enrichment of images through collaborative tagging. [10, 1]

Over the past few years microeconomic theory has paid increasing attention to the role of the user in value creation along the innovation life cycle by the application of cyberinfrastructures. [2, 10, 16, 3, 7] Widespread cyberinfrastructures like the web are a necessary (but not sufficient) precondition to achieve economy of scale on co-production communities, where the users share production resources and facilities with the providers, in a manner of enterprise collaboration. [13] Co-production has to be conceptionalized as a two level infrastructure consisting of (1) a (top down) service provider, who provides the stable running, toolkits and marketing of content services and (2) the (bottom up) content producer who adopts the infrastructure to provide creative work in exchange of improvements in services, reputation, visibility, self-esteem, fun etc.

Service providers profit from the selective integration of users along the product life cycle which can significantly improve the overall production process [6] leading to lower transaction costs in information provision and search, and accumulating critical masses of data for service development, provision and the reduction of cycle times. [10] But to sufficiently enable and govern co-production service providers have to establish adequate incentive and remuneration models to encourage participation, promote trust and secure quality standards.

4 Examples: Collaborative Tagging in ESP, Peekaboom and Phetch

The following section introduces applications that have been developed by Carnegie Mellon University School of Computer Science. The examples illustrate some low level approaches incorporating the idea of co-production for the semantic enrichment of images on the web by collaborative tagging. The empirical set up pays account to the fact that certain improvements in quality of service are hard to achieve by machines, but easy to handle by humans. Hence the human is seen as an extension of the computing system.

According to Heymann [12] collaborative tagging systems are “a good way to leverage large numbers of users to help organize very large, rapidly changing corpora which would be difficult to organize automatically. Often, this works because users are working in their own self interest as they mark an object with a particular tag, and when all of these tags are aggregated together, the system can make assumptions about objects based on the aggregate activities of hundreds of thousands or even millions of users.”

The basic idea behind the following applications was to let users do the semantic enrichment of images in their spare time by playing it as a game. According to Ahn [1] 5,000 people continuously playing the game could assign a label to all images indexed by Google in 31 days. To achieve this two or more anonymous players gather in a team to contest against time and higher score-holders in labelling pictures, locating objects within them and phrasing their content.

4.1 ESP Game

ESP (http://www.espgame.org) is a Java application for the collaborative tagging of images. A team of two randomly chosen anonymous players come together for a limited amount of time to describe images by using tags. For each matching tag both players collect grants which are added to their individual score.

The purpose of the game is to improve image search by collecting annotation data on a lexical level based on a quota-model. This means that just those tags are stored as legitimate annotations of an image that pass certain quota-indicators.

4.2 Peekaboom

Peekaboom (http://www.peekaboom.org) is a Java application to identify the location of an object within an image. Peekaboom is an extension of the ESP Game as it deepens the semantic enrichment of images from a lexical level to a spatial description of objects located within the image.

Two randomly chosen anonymous players take turns either presenting an object within an image or guessing what the object could be. The first player gets an image along with a word related to the image (i.e. cow). By clicking on certain parts of the image the first player reveals little portions of the image to the second player. The object of the game is for the second player to type the word associated to the image from whose perspective, the game consists of a slowly revealing image, which has to be named. Once the second player has guessed the correct word the team moves on to the next image and switches roles.

4.3 Phetch

Phetch (http://www.peekaboom.org/phetch/) addresses accessibility problems of images by visually impaired people. It was designed to produce descriptive captions of an image so that screen readers – programs that convert the text of a webpage into synthesized speech – are capable to process the content of the image and visually impaired people can share a common experience of the web.

Phetch is a 3 - 5 person game to generate explanatory phrases for randomly chosen images from the Web. The players are grouped with others from the Web. One player is chosen as a describer, and the others are the seekers. The describer gets an image and has to help seekers find it in a given image corpus by typing strings related to the captions. Given only text from the describer, the seekers must find the image using a search query which presents the seekers a collection of possibly matching images. The first seeker to find the correct picture wins and becomes the next describer. Each session of the game lasts 5 minutes, during which all players should go through as many images as possible.

5 Success Factors in Co-production Environments

Successful co-production builds on reciprocity between service provider and content producer. The following success factors are crucial to lead co-production to its full potential.

5.1 Non-Monetary Reciprocity Systems

For the service provider co-production engagement aims at fostering customer binding, market penetration and service visibility, providing higher customer value and means of differentiation. It is also a means for market development by nurturing communities that provide the critical mass of data to support product and service development, improvement, diversification and personalisation. The viral, community-based nature of co-production is an efficient test-bed for community driven innovation [20, 5] that is dependent on lead users who are acting as (honorary) trend scouts and opinion leaders in highly risky markets partially cannibalizing orthodox market research.

For the content producer engagement in co-production is slightly less utilitarian. Co-production engagement can be seen as an attitude, a way of life mixing personal interest with mutual progress in a means of self-marketing trying to avoid underutilization of knowledge, skills and time. [19] In the case of gaming users experience fun in competing against time and higher score holders by contributing to the qualitative improvement of the service. Empirical investigations in other co-production of public services have shown that the acquisition of self-esteem, skills, social reach and higher visibility are important drivers to engage in co-production. [4]

5.2 Monetary Reciprocity Systems – The Case of Second Life

Co-production gains are difficult to quantify. Nevertheless some authors argue that financial incentives can play an important role in deepening the relationship between service provider and content producer. [4] Elaborate remuneration models take account of the fact that increases in productivity can be achieved through coupling voluntary engagement with material rewards. In the case of Second Life (http://www.secondlife.com/) the co-production architecture is built in a way that rewards users with the possibility to capitalize their engagement by developing services on their own. The precondition to this is the establishment of a virtual market based on a currency (the Linden Dollar) that enables surplus transactions between service providers and content producers. The coupling of the virtual currency to "real word" financial transactions by stable exchange rates (the LindeX) further deepens the remuneration value users acquire by participating in the co-production setting.

Second Life provides a hierarchical, multi-level remuneration model incorporating the infrastructure providers (Linden Labs) on top with various service providers down the taxonomic functionality of the game. Further on Second Life enables its users to switch between various roles and so take full advantage of either being a service provider or a content producer.

5.3 Bonus Systems for Collaborative Content Production – A Scenario

Bonus systems – here defined as the deliberate rewarding of positive feedback – could be a powerful means to encourage UGC and the participation in co-production settings. In the case of collaborative tagging scores acquired through a game or any other form of mutually positive interaction could be assigned material value for exchange and trade outside the production setting.

A business model could take the following form: A user contributes content to a service by registering at an "accounting server". Any other user who is subscribed to the "accounting server" is able to rate the contributed content for quality, validity, relevance etc. Each feedback is being collected and transformed by an algorithm into a score incorporating indicators like number of tags, downloads, references, reuse etc. taking in consideration the position of each user in the collaboration value chain. The collaboration value chain is built on the principle that all contributors gain rewards measured by the value their contribution adds to a virtual community. To enable market conditions the score acquired by each user in the value chain is transformed into a virtual currency mutually accepted by service providers within a business web to grant access to surplus goods and services, i.e. premium services, personalization, discounts, downloads etc.

Rewarding models based on bonus systems could promote the spread of cyberinfrastructures by giving users the possibility to gain material rewards through collaborative engagement and simultaneously acquire literacy and skills in the use of advanced web technologies. They could work as an incentive to lower the acceptance barriers for paid content and services offered by media companies on the web. Bonus systems could also be seen as a metric for productivity in co-production settings, which is a necessary precondition to allow the evaluation of user participation and centrality within a given domain or community. Nevertheless fraud behaviour has to be taken serious as it might spoil the business model as a whole.

6 Conclusion

Co-production is a powerful concept. Transformations can take place along various dimensions including the modes of production (i.e. from design to co-design), the relationship between producers and consumers (i.e. from consumer to prosumer) and maybe even the socio-economic conditions in industrialized societies (i.e. from desintegrated to reintegrated production). [22]

To secure quality of service knowledge about motivational aspects and desire for privacy preservation are crucial as trust into the service provider and positive feedback are indispensable success factors in any co-production setting. [17] Further on empirical evidence has shown that from the content producer's side the perceived product satisfaction correlates strongly with process satisfaction. [20] Created solutions – previously unmatched by existing offers on the market – must provide the configuration that matches the user's preferences. Data about failed co-production projects revealed that one of the biggest challenges in co-production is to integrate the content producers' data adequately along the product life-cycle and pool the benefits back to the users. [ibid]

With the ever growing penetration of cyberinfrastructures among public information systems co-production settings could be used to foster the integration and balancing of interests between citizens, businesses, policy makers and administrations by improving access to information, consultation and participation. Evidence from the physical world shows that tax benefits work as a strong incentive to co-production.

To achieve this, advances in digital rights management technologies and policy aware computing might lay the technological foundations to develop software-based business models and privacy preservation mechanisms in co-production settings enabling new forms of employment and public services ideally based on a social contract.

References

[1] Von Ahn, Luis; Dabbish, Laura (2004). Labeling Images with a Computer Game. Paper presented at CHI 2004, April 24–29, 2004, Austria. See also: http://www.cs.cmu.edu/~biglou/ESP.pdf, viewed June 20, 2007

[2] Bitran, Gabriel; Gurumurthi, Suri; Sam, Lin Shiou (2006). Emerging Trends in Supply Chain Governance. Cambridge: MIT Sloan School of Management Working Report 227

[3] Bovaird, Tony (2005). Beyond Engagement and Participation. User and Community Co-production in Local Governance. Bristol Business School: Working paper. See also: http://rural.carnegieuktrust.org.uk/files/rural/Beyond%20Engagement%20and%20Participation-%20T%20Bovaird.pdf, viewed: June 15, 2007

[4] Boyle, David; Clark, Sherry; Burns Sarah (2006). Hidden Work. Co-production by People outside Payed Employment. York: Joseph Rowntree Foundation

[5] Crawford, Mike; Rutter, Deborah; Thewall, Sarah (2004). User Involvement in Change Management. Report for the National Coordinating Centre for NHS Service Delivery and Coordination. London: Imperial College. See also: http://palin.lshtm.ac.uk/hsru/sdo/files/project/18-final-report.pdf, viewed June 20, 2007

[6] Fichter, Klaus (2005). Modelle der Nutzerintegration in den Innovationsprozess. Berlin: IZT – Institut für Zukunftsstudien und Technologiebewertung

[7] Franke, Nikolaus; Piller, Frank (2004). Value Creation by Toolkits for User Innovation and Design. The Case of the Watch Market. In: Journal of Product Innovation Management, 21/6, 2004, pp. 401-415

[8] Funk, Jeffrey L. (2004). Neue Technologie, neue Kunden und die disruptive Technologie des mobilen Internet. Erfahrungen aus dem japanischen Markt. In: Zerdick, Axel; et al. (Eds.). E-Merging Media. Kommunikation und Medienwirtschaft der Zukunft. Berlin: Springer Verlag

[9] Ghosh, Rishab Aiyer (2006). Study on the Economic Impact of Open Source Software on Innovation and the Competitiveness of the Information and Communication Technologies (ICT) Sector in the EU. European Commission: Contract ENTR/04/112.

[10] Golder, Scott; Huberman, Bernardo (2006). Usage Patterns of Collaborative Tagging Systems. In: Journal of Information Science, 32/2, 2006, pp. 198-208.

[11] Hess, Thomas (2004). Medienunternehmen im Spannungsfeld zwischen Mehrfachverwertung und Individualisierung. Eine Analyse für statische Inhalte. In: Zerdick, Axel; et al. (Eds.). E-Merging Media. Kommunikation und Medienwirtschaft der Zukunft. Berlin: Springer Verlag

[12] Heymann, Dieter (2006). Final Project: On The Use and Abuse of Collaborative Tagging Data. See also:
www.stanford.edu/class/cs229/proj2006/Heymann-DeterminingTheInformationValueOfTags.pdf, viewed May 10, 2007

[13] Hsu, Cheng (2006). Models of Cyberinfrastructure-Based Enterprise and their Engineering. In: Hsu, Cheng (Eds.). Service Enterprise Integration. An Enterprise Engineering Perspective. Boston: Springer

[14] Incera, Jose; Mejia, Marcelo; Roberts, Keith (2005). Online Collaboration as Emerging Technology. Final Project Paper on INFT 625. See also:

http://collaborationwiki.pbwiki.com/f/Online_Collabloration_Paper.doc, viewed Mai 1, 2007

[15] Kiefer, Marie Luise (2001). Medienökonomik. München: Oldenbourg Verlag

[16] Leadbeater, Charles (2006). The User Innovation Revolution. Report by the National Consumer Council, London, UK

[17] Lieberman, Henry; Smith, Dustin; Teeters, Alea (2007). Common Consensus: a webbased game for collecting commonsense goals. Paper presented at IUI'07, January 28–31, 2007, Hawaii, USA.

[18] O'Reilly, Tim (2005). What is Web 2.0? Design Patterns and Business Models for the Next Generation of Software. See also:

http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html?page=1, viewed April 30, 2007

[19] Ostrom, Elinor (1996). Crossing the Great Divide: Coproduction, Synergy and Developement. In: World Development, 24/6, 1996, pp 1073 – 1087

[20] Reichenwald, Ralf; Piller, Frank (2003). Von Massenproduktion zu Co-Produktion: Kunden als Wertschöpfungspartner. In: Wirtschaftsinformatik,45/5, 2003, S. 515 - 519

[21] Rheingold, Howard (1993). Virtual Communities. New York: Addison Wesley

[22] Weber, Michael; Fröschl, Karl (2006). Semantic Web als Innovation in der ökonomischen Koordination. In: Pellegrini, Tassilo; Blumauer, Andreas (Hg.). Semantic Web. Wege zur vernetzten Wissensgesellschaft. Berlin: Springer Verlag

[23] Weinberger, David (2002). Small Pieces Loosely Joined. A Unified Theory of the Web. New York: Basic Books

[24] Wikipedia (2007). User-generated content. In:
http://en.wikipedia.org/wiki/User_generated_content, viewed April 30, 2007

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Compensation Models for Interactive Advertising

Steffen Zorn
(Business School, University of Western Australia, Australia
szorn@biz.uwa.edu.au)

Astrid Dickinger
(Department of New Media Technology, Modul University Vienna, Austria
astrid.dickinger@modul.ac.at)

Abstract: Due to a shift in the marketing focus from mass to micro markets, the importance of one-to-one communication in advertising has increased. Interactive media provide possible answers to this shift. However, missing standards in payment models for interactive media are a hurdle in the further development. The paper reviews interactive advertising payment models. Furthermore, it adapts the popular FCB grid as a tool for both advertisers and publishers or broadcasters to examine effective interactive payment models.

Keywords: Interactive Advertising, Compensation Model, Classification
Categories: H.3.1, H.3.2, H.3.3, H.3.7, H.5.1

1 Introduction

The focus of marketing is shifting from mass to micro markets, due to changing demography and nuanced product preferences. Digital technology and hypercompetition accelerate this process. This shift has advertising implications as a heterogeneous marketplace requires more one-to-one communication rather than one-to-many communication (Kotler, Jain et al. 2002)

Interactive media, like the Internet, interactive TV or mobile media provide possible answers to this shift, with advantages over traditional media. They extend advertisers' ability to reach and engage viewers (potential customers) through interactive features, which increase customer involvement and could lead to this one-to-one relationship (Brodin, Barwise et al. 2002). Furthermore, the intensity of interaction may associate positively with satisfaction (File and Prince 1993).

Interactive TV, digital TV with a "return path" to enable communication between consumer and broadcaster (Brodin, Barwise et al. 2002), has two broad forms. The major form today, DRTV is limited in that users usually get just an on-screen overlay to exchange personal information for a special offer or other incentives (Mercier and Barwise 2004). One reason for DRTV's popularity could be that it is simple to set up and run. The second form, DAL, offers a wide range of advertising possibilities by taking the viewer in a separate environment showing additional information, yet, permitting a direct-response function as in DRTV (Mercier and Barwise 2004). Advertising figures emphasise the potential of this new medium. While most traditional media lose advertising customers (Hanson and Kalyanam 2007), interactive TV has increased revenues. Peter Birch, head of interactive sales at UK's largest broadcast network, ITV, predicts a 35% rise in interactive advertising revenue for his company in 2007 (Nicholson 2006), after a revenue of £157m in 2006 (ITV plc 2007).

Traditional mass communication such as television often assumes customers are homogenous (one to many). Little interaction between customers and the advertising company occurs (Hoffman and Novak 1997). On "traditional" TV for example, the advertiser usually buys advertising time on a certain programme on a particular TV station. Prices usually depend on the estimated audience, so prime time or spectacular events like January 2007 "Superbowl" in the USA cost US$ 2.6 million for a 30-second spot (Monica 2007).

Interactive media enable consumers to reply to an advertiser's communication. In contrast to mass-media, interactive communication can go in either direction (many to many) (Hoffman and Novak 1996).

Even though interactive media have bright prospects, they also face problems generating advertising revenue: There is for example no standard regarding payments for Internet TV advertisements (Crampton 2007) and regarding advertising on the web, "the complexity of the medium in general hinders the standardization process" (Novak and Hoffman 2000).

Swain (2005) addresses the question of compensation models for interactive marketing predicting that agencies will gain in importance. The interviews with marketing professionals showed that agency compensation has not yet been related to interactive communication success as there is still *"thinking in terms of traditional marketing communication measures"* (Swain 2005). Thus, there is a call for research in the field of developing appropriate *"measurement-based methods for agency compensation"*. (Swain 2005).

This paper therefore reviews interactive advertising payment models as well as advertising models in traditional media. Furthermore, a taxonomy developed in this paper argues which payment model aligns with certain types of goods or services, as not every payment model suits every product. Moreover, it will take the aim of the advertising campaign into account, i.e. consider wheter the aim is an increase in sales or an image campaign, also impacting the choice of compensation model. The FCB grid, a two-dimensional grid differentiating purchasing situations developed by the Foote, Cone & Belding advertising agency, serves as a basis for our discussion (Vaughn 1980).

The remainder of the article proceeds as follows. First, compensation models for advertising on TV and the Internet are presented. Then, our application of the FCB grid and compensation models for interactive TV is presented. The paper closes with a discussion of the results and future research avenues.

2 Compensation Models for Interactive Advertising

Academics have examined different forms of interactive advertising such as Internet advertising (Shen 2002, Spake et al. 1999). Spake et al. (1999) differentiate between behaviour-based compensation and outcome-based compensation. The latter involving some sort of measurement of the outcome such as sales increase, brand share etc. The measurement methods available (audience feedback, attitude change, brand equity, change of behaviour and/or attitude, image) are tied to the final compensation model chosen. Some measurements may require extended market research (e.g. image analysis) and are thus priced on a fee basis.

The most popular interactive advertising medium, the web, offers a large variety of advertising forms like pure text messages, picture or video elements. However the the Internet mostly uses 'banner advertisements' and 'target communications'. Banner advertisements are graphic images and text that try to entice users to click on the banner to learn more. Target communications could range from a single, simple website to a series of linked pages (Novak and Hoffman 2000). Usually a click on a banner leads the viewer to a specific target communication.

Interactive TV offers some more possibilities which can be devided in two types, which are within the broadcast stream or alongside the broadcast stream (Caugerghe and Pelsmacker 2006). Within the stream offers DAL, impulse response, microsite and a contact me function in the commercial. In the content products can be placed and banners can be shown. Alongside the stream the possibilities include a walled garden (logos, banners, games, websites), logos and banners in the electronic program guide, direct mailing and video on demand. Cauberghe and Pelsmacker (2006) compare different forms of interactive advertising including above mentioned in their paper.

Companies need to get further insights into how to compensate for interactive advertising. Therefore we draw on models developed for the Internet to adapt them in the interactive TV context:

- Flat-fee pricing charges the advertiser for their ads on a website in a certain period (e.g. per month). Flat-fee can be without or with traffic guarantees. If accurate traffic information is available, companies can use a pay-per-view model (Novak and Hoffman 2000).
- Pay-per-view (PPV): Usually measured in CPM (price for 1000 impressions), the publisher gets a fee (dependent on the popularity of the page – usually a few dollars or more for the mentioned thousand views) for each ad shown on a publisher's website (Hanson and Kalyanam 2007). Even without clicking on the banner, the mere exposure can increase ad and brand awareness (Briggs and Hollis 1997). However, this payment model can motivate the publisher to attract mass audiences instead of focused audience segments (Novak and Hoffman 2000).
- Pay-per-click (PPC): In contrast to pay-per-view, the publisher is paid only if the visitor clicks on the ad. Google, for example, charges \$0.01 and up per click, depending on the popularity of the keywords (Google 2007). A disadvantage of PPC is ´hit or click fraud´ – the click-through rate is artificially inflated (Anupam, Mayer et al. 1999).
- Pay-per-sale (PPS): The publisher receives a commission for purchases done on the target site. The payments are usually higher because they are more valuable for the target site. Amazon, for example, pays up to 15% depending on the product sold (Amazon 2007).
- Pay-per-lead (PPL): This method requires that visitors take a specific action in response to an ad banner, e.g. registering for an account (Anupam, Mayer et al. 1999). ´Hit-shaving´ is the main threat of both the PPS and PPL pay-

ment models. The advertiser fails to report a lead or a sale to the publisher (Anupam, Mayer et al. 1999).

Other forms of compensation include ´banner swapping´ or ´banner exchange´, where firms exchange ads between each others website without or with a fee (Turban, King et al. 2006). These exchange models are inappropriate for this paper, as they require two advertising companies and no publisher. In addition, there are pricing models, which are hybrids of the above mentioned forms.

3 Choosing the Right Model

Often the negotiating power of the advertiser or the publisher seems crucial to the choice of a model (Hanson and Kalyanam 2007). However, there is no best model and not every model goes with the advertising campaign of a certain product. For example the PPV model seems appropriate for products with universal appeals such as telephone rates or travel services (Mangani 2004). For specialized products, PPV might be unsuitable. Advertisers and publishers negotiating payment models should consider three points:

- aim of the campaign / the ad
- the goal of the user (Hollis 2005)
- the type of product

Planning the target communication starts with specifying the intent of the marketer, that is the aim of the campaign or the ad, and the goal of the user encountering the ad.

The Elaboration Likelihood Model (ELM) model of Petty and Cacioppo (1983) suggests that the degree of involvement of the viewer of an ad is important to predict for these first two points, the campaign's aim and the user's goal.

Customers actively seeking product information usually focus on an advertisement's message. These high involvement consumers follow a central route of processing information. For these consumers, the ad should tend towards being factual (Petty, Cacioppo et al. 1983) and interactive, as people actively searching for information are more likely to click on an interactive ad (Hollis 2005).

In contrast, for low involvement viewers, the focus of the ad should be more on the peripheral cues, for example the ad design, as the consumer follows a peripheral route towards processing information (Petty, Cacioppo et al. 1983). Experience may play an additional role in this discussion, because high experience might lower the necessary involvement (Laczniak and Carlson 1989). Advertising should target towards consumers' information needs and predicted involvement (Hollis 2005).

A basic decision for the advertiser should be to drive an immediate response, e.g. to sell a product, or to increase brand awareness (Hollis 2005). Payment models requiring an active response, such as PPC, may be inappropriate if the advertiser seeks no response or expects the user not to interact due to their low involvement.

Including the type of product in this discussion, the FCB grid offers a basis for negotiating payment models between advertisers and publishers. The two-dimensional

FCB grid has two dimensions, purchase decisions based on thinking or feeling, and high versus low involvement. Combining these two dimensions produces a strategy matrix that isolates product categories and suggests specific marketing considerations. The extended FCB grid in Table 1 shows product categories and serves as a guideline for applicable interactive advertising payment models.

In the first quadrant, high on involvement and thinking, the customers are in the Informative/Thinker stage. They search for information because the product might be complex or at the beginning of the product life-cycle. For the advertiser it is important to supply a potential customer with the information needed to support him in reducing the perceived risk (Vaughn 1980) involved with the purchase decision, leading to clicks on advertisements to request this needed information which suggests a PPC compensation form. An example would be advertisements for a brand new personal computer (high involvement, beginning of the product life cycle) for which a potential buyer would still need a lot of information on the innovations regarding the PC, the programmes installed and the specifics regarding hardware. This information need leads to clicks on the interactive TV advertisement calling for a PPC compensation model.

The second quadrant, high on involvement and feeling, i.e. the Affective/Feeler, also requires high involvement, but the focus is on emotions and feeling not on information (Vaughn 1980). The goal of the ad for a product that requires high involvement and a high level of feeling is to create images, emotions, and arousal in connection with the product. In the long run this may lead to a purchase decision, suggesting a PPL model or flat fee. Car advertisements typically fall into the category of emotion and image creation where the involvement with a product would be high and the focus is on feelings and emotions. The potential customer would not immediately buy the car but could request further information, calling for a PPL compensation. As some sort of image for the brand is created, for example through a video clip, a flat fee could be appropriate too.

In the third and fourth quadrants, the customer involvement is lower. In the third quadrant, Habit formation/Doer, thinking dominates for products like household items or food. The aim of marketers for these products usually is to increase brand awareness and therefore to form habits. Inducing a trial may even trigger subsequent purchase (Vaughn 1980). Vaughn (1986) indicates that exploratory buying may happen in this quadrant, thus, a PPS model appears appropriate depending on the product class. Food may not be bought this way, household goods or cosmetics, however, are possible to be shipped to the customer. Interactive TV advertisements may offer cosmetics like shampoo with a free bottle of conditioner to induce trial of the conditioner. This would work for spontaneous habit purchases as well as trial of new products. The customer has the chance to immediately reply to the advertising stimulus. Thus, a PPS and a PPV model (for perishables) seems sensible (Briggs and Hollis 1997).

In the last quadrant, Self-satisfaction/Reactor, feeling dominates over thinking for self-satisfying products such as cigarettes, liquor or candy. The advertising aim for these products is to increase sales or at least to increase brand awareness. Thus, it is important for those companies to present their products via interactive TV. As such liquor or candy can be promoted via this medium leading to product awareness and even sales. Depending on the possibility to buy products online, a PPV or PPS payment model seems appropriate.

		Thinking	**Feeling**
High involvement	Consumer description	Informative (Thinker)	Affective (Feeler)
	Product Examples	Car, house, furnishings New products	Jewellery, cosmetics, fashion apparel
	Implication for advertiser	Specific information Demonstration	Change of attitude
	Possible payment model	PPC, PPL	PPL, Flat fee
Low involvement	Consumer description	Habit formation (Doer)	Self-satisfaction (Reactor)
	Product examples	Food, household items	Cigarettes, liquor, candy
	Implication for ad-vertiser	Reminder	Attention
	Possible payment model	PPV, (PPS)	PPV, PPS

Table 1: Classification of compensation models for interactive advertising

4 Conclusion

Although it makes sense to use interactive metrics and the derived payment models with interactive media, this paper argues that the appropriate payment model for an interactive advertising campaign could also be one that does not require interactivity (e.g. PPV). The paper adapted the popular FCB grid as a tool for both advertisers and publishers or broadcasters to examine effective interactive payment models.
However, to choose a payment model, advertisers and publishers also have to know the threats of each model and discuss how to avoid them.
It is suggested in the literature, that for example “multi-site” data (on every website involved in the campaign) is one requirement to solve the problems and to obtain further confidence in the interactive media (Novak and Hoffman 2000).
Beyond this confidence discussion, future research should test the classification developed and presented in the grid through both qualitative and quantitative research methods.
Key challenges remain regarding the development of interactive TV and the further diffusion of digital video recorders. Marketing Management predicts that such technologies will destroy advertising effectiveness of traditional TV spots and leading companies will therefore invest in branded entertainment within TV programmes, TV program sponsorship, interactive advertising during TV programs, online video ads and product placement (Marketing Management 2006). These additional forms of TV advertising will require further detailed investigation.

Acknowledgements

Information Diffusion across Interactive Online Media (www.idiom.at) is funded by the Austrian Ministry of Transport, Innovation & Technology and the Austrian Research Promotion Agency within the strategic objective FIT-IT (www.fit-it.at).

References

Amazon. (2007). "How Marketplace Selling Works." from http://www.amazon.com/gp/help/customer/display.html?nodeId=1161238.

Anupam, V., A. Mayer, et al. (1999). "On the Security of Pay-per-click and Other Web Advertising Schemes." *Computer Networks* **31**: 1091-100.

Briggs, R. and N. Hollis (1997). "Advertising on the Web: Is There Response before Click-Through?" *Journal of Advertising Research*(March/April 1997): 33-45.

Brodin, K., P. Barwise, et al. (2002). "UK Consumer Responses to iDTV Report." *Future Media.*

Cauberghe, V. and de Pelsmacker, P. (2006). "Opportunities and Thresholds for Advertising on Interactive Digital TV: A View from Advertising Professionals." *Journal of Interactive Advertising* **7**(1): 25-40.

Crampton, T. (2007, 28.01.2007). "Small screens, new programs." *TECH/MEDIA*, from www.iht.com/articles/2007/01/28/business/tv29.php.

File, K. M. and R. A. Prince (1993). "Evaluating the Effectiveness of Interactive Marketing." *Journal of Services Marketing* **7**(3): 49-58.

Google. (2007). "Account Fees and Payment Options." from https://adwords.google.com/select/AfpoFinder?countryCode=AU.

Hanson, W. A. and K. Kalyanam (2007). *Internet Marketing & e-Commerce, Thomson Higher Education.*

Hoffman, D. L. and T. P. Novak (1996). "Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations." *Journal of Marketing* **60**: 50-68.

Hoffman, D. L. and T. P. Novak (1997). "A New Marketing Paradigm for Electronic Commerce." *The Information Society* **13**(1): 43-54.

Hollis, N. (2005). "Ten Years of Learning on How Online Advertising Builds Brands." *Journal of Advertising Research* **45**(2): 255-268.

ITV plc. (2007). "ITV plc results for year ended 31 December 2006." from http://www.itvplc.com/itv/news/releases/pr2007/2007-03-07/.

Kotler, P., D. C. Jain, et al. (2002). *Marketing Moves: A New Approach to Profits, Growth & Renewal*, Harvard Business School Press

Laczniak, R. N. and L. Carlson (1989). "Examining the Influence of Attitude-Toward-the-Ad on Brand Attitudes." *Journal of Business Research* **19**: 303-311.

Mangani, A. (2004). "Online advertising: Pay-per-view versus pay-per-click." *Journal of Revenue and Pricing Management* **2**(4): 295-302.

Marketing Management (2006). As not seen on TV. *Marketing Management* 15(4): 6.

Mercier, P. and P. Barwise (2004). Digital Television in the UK: Consumer Responses to Interactivity. London, London Business School.

Monica, P. R. L. (2007). "Are Super Bowl Ads Worth the Money?" from http://money.cnn.com/2007/01/24/news/companies/superbowlads/index.htm.

Nicholson, K. (2006, 17.11.2006). "Is iTV a strong commercial medium?" *Campaign. Teddington.*

Novak, T. P. and D. L. Hoffman (2000). Advertising and Pricing Models for the Web. *The Economics of Digital Information and Intellectual Property*. D. Hurley, B. Kahin and H. Varian. Cambridge, MIT Press.

Petty, R. E., J. T. Cacioppo, et al. (1983). "Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement." *The Journal of Consumer Research* **10**(2): 135-146.

Turban, E., D. King, et al. (2006). *Electronic Commerce - A Managerial Perspective*. New Jersey, Pearson Education.

Shen, F. (2002). Banner Advertisement Pricing, Measurement and Pretesting Practices: Perspectives from Interactive Agencies. Journal of Advertising **31**(3): 59-67.

Spake, D.F.D'Souza, G. Crutchfield T.N. Morgan R.M. (1999). Advertising Agency Compensation: An Agency Theory Explanation. Journal of Advertising **28**(3): 53-72.

Swain, W. N. (2005). "Perceptions of Interactivity and Consumer Control in Marketing Communication: An Exploratory Survey of Marketing Communication Professionals." *Journal of Interactive Advertising* **6**(1): 109-124.

Vaughn, R. (1980). "How Advertising Works: A Planning Model." *Journal of Advertising Research* **20**(5): 27-33.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Enhancing Music Learning Experience through a Dedicated Web 2.0 Oriented Service

http://e-guitare.univ-reunion.fr/figs/

Olivier Sébastien
(IREMIA, University of Reunion, France
olivier.sebastien@univ-reunion.fr)

Noël Conruyt
(IREMIA, University of Reunion, France
noel.conruyt@univ-reunion.fr)

Stéphane Calderoni
(IREMIA, University of Reunion, France
stephane.calderoni@univ-reunion.fr)

Abstract: The Web 2.0 philosophy has brought new ways of using the web as a content repository and a sharing platform. Non-computer skilled people can now publish their own text, images, videos and/or sounds and take part in communities created around topics they like.
Our idea is thus to use this new communication mechanism to assist skill learning. Skill is indeed a difficult knowledge to learn on a text form as it is hard to describe movements, gestures or procedures in this way: sometime, a picture or a video is better than a thousand of words.
As a popular field of experience, we have focused our attention on music learning, and more particularly on guitar pieces learning. Music is very representative of skill learning, it is both a physical and an intellectual activity. The "Gloss2U" service we describe in this article takes into account the specificities of this learning process, especially as events are time-related.
To achieve this, we rely on new equipment conditions that are nowadays gathered. User-friendly multimedia tools opened new horizons and broadband networks (ADSL, cable, ...) are becoming more and more common. Almost everybody is a potential content producer with just a webcam or a cell phone. Our system is therefore audio and video based and users interact with each others by submitting contributions, called glosses, in the context of the piece they are learning. A dialogue starts between student users (learners) and experimented ones (professors), to discuss encountered problems and the way to solve them. The other strength of this system is that it can act as a knowledge repository for forthcoming practitioners.

Keywords: Knowledge management, skill-oriented learning, Music learning, Web 2.0-based services, Multimedia, glosses

Categories: H.4.2, H.5.1, H.5.2, H.5.4, H.5.5

1 Introduction

Defining the Web 2.0 philosophy is not easy. Since the first use of the phrase by Tim O'Reilly in 2004 (refer to [O'Reilly, 05]), it has generated a lot of discussion about its being innovative or just a buzz word. Nevertheless, there is no doubt that it has introduced a new vision of the web-based activities and we can depict a few aspects that are particularly important to us.

First of all is the ability for anybody to become a content producer. In the past, putting one's own text, image, sound and/or video online required various computer and network skills. Nowadays, lots of creative user-friendly tools (software, hardware, online services) have appeared allowing a neophyte person to publish his blog in a few hours after having bought a computer.

Then, there is the social aspect that lays behind: Web 2.0 sites allow users to share their production, but to create communities too. The strength of the system thus comes from the aggregation of hundreds or thousands of contributions around specified themes (thanks to tagging), these contributions being cross-linked.

Leaders are often identified in communities, allowing to set a hierarchy. For instance, one can often see a ranking next to an user name. This is an interesting point because the leaders are presented as experts and they tend to spread their knowledge to the others, allowing them to progress in this particular domain. Of course, legitimacy of the leaders can (and sometime is) be discussed but there are various ways to check it. For example, in a video production community, the leaders are the ones whom it is the professional activity and who therefore have a rich experience to share. This example may be transposed to any topic.

Our idea is thus to use these now wide-spread tools, services and users interaction mechanisms to create a web-based system called Gloss2U that facilitate musical knowledge management. In this context, we are particularly interested in tacit knowledge, which is difficult to lay on paper.

2 Musical knowledge management

To describe the concept, we will firstly present and justify the main features we wanted to implement and in a second time, give examples of existing applications or services that have inspired us as far as the user experience is concerned.

2.1 Presentation

2.1.1 What

The idea is to build a web application that uses well known Web 2.0 mechanisms to facilitate the learning of musical tunes. The originality of the system comes from its architecture, that matches a musical structure in addition to a web one. Indeed, an important factor to take into account is time, as music is basically a succession of events that occurs at particular moments. It is thus necessary to discuss these events (how to play them, how to make the transition between them) in their temporal context. As guitar fans, we have narrowed our experimentation field to this instrument because it seems difficult to manage the specificities of several instruments from the beginning. This choice is also reasonable because guitar is popular and easy to access. The experience should of course be derivated to other instruments in a second time.

2.1.2 Why

The best way to learn an art has been the private lesson with a teacher for ages. There are of course lots of alternative media that have been created during the years to materialize this knowledge and facilitate its transmission. But none of them has the same efficiency as the in-situ course.

On the other hand, paper was the only medium for centuries. Then came audio recording and video. Computing capabilities brought another dimension: interactivity. This is the cement between these media and it can be used to add a meaning. Finally, Internet drew the links between users and media, and even from users to users, as in the Web 2.0 philosophy.

This is a new strength that allows to gather knowledge that is spread all over the world. Such a learning system thus contributes to two major achievements in Knowledge Management: facilitating musical knowledge transmission and archiving it in a more natural way than ever, using multimedia. Can we imagine what classical music teaching would have been if we had video and audio recordings of Mozart or Bach performing their own pieces?

2.1.3 How

The system is firstly a communication tool between users. Two classes of users are identified: learners and teachers. The formers encounter problems during their learning. The system allows them to ask for help by submitting questions in the context of the piece they are currently learning. The professors provide answers to these questions in the same way. At this stage, this is more or less like a forum or a chat dialogue.

But dealing about music using text is not natural nor efficient. That is why all the exchanges above are audio and video based. Users are able to record themselves using the device of their choice: webcam of course, but also camcorder, still camera with video mode and even cell phone. Consumer electronic market provides nowadays various easy ways to record events. Quality is not always the best, but is regularly improved. Ease of use is of course a *sine qua non* condition. We would fail if users have to focus more on system handling than on learning itself.

From the conception point of view, each submission is called a gloss. This word refers to linguistics: it is a short contextual explanation of a word or a phrase. All the glosses of a book are usually gathered in a specific section called glossary.

2.2 Sources of inspiration

As the system should be easy to use, we have to rely on handling that people are used to. The system we want to provide should not require to learn new technical skills. We therefore have to start from the actions they already manage to do. Thus, let us see a selection of popular applications and services that are representative of typical user experience. One point must be clarified here: our aim is not to copy these sources of inspiration but to refer to mechanisms used by them that we subscribe to. We may not demand more effort to our public than these examples do.

2.2.1 Blogs/Social networking

Blogs are online private diaries. The author shares articles he writes and from which his guests can react to. Although being initially textual, blogs now embed photos, audio and video media.

The number of blogs dramatically increased thanks to the online user-friendly page editors provided. Updating one's blog is as easy as writing a letter using a word processor.

MySpace is one of the most famous blog service provider. Its success comes from the social aspects that is promoted.

2.2.2 Youtube

The interesting aspect with Youtube is that it has democratized online video broadcasting, an process that used to be very difficult to do in the past for several reasons: file size, format selection, streaming constraints, ... Posting a video to Youtube is not so easy, but the process is now a classic one.

2.2.3 iLife

The iLife software suite is a bundle that comes along with all Apple Macintosh computers. It proposes very simple applications to create music (Garage Band), capture, manage and edit photos and videos (iPhoto and iMovie), use these contents to burn a DVD (iDVD) or share it on the Internet (iWeb). What is interesting here is the exceptional usability of these products along with the natural connection between them. Workflow is simple and can output quality content, for a consumer production.

3 Our Gloss2U Service Implementation

We will firstly present our technical choices and then describe the resulting system.

3.1 Technical concerns

Building an online video-based service immediately refers to specific constraints. Consulting the content is not a problem, trouble starts when it comes to produce and publish it. The question is: what is the easiest way for a user to submit a video contribution to the system? Our answer is to use an attached audio/video capture device (webcam, camcorder, ...) to create a stream that is encoded and uploaded in real time to a server. This way, there is no need for the user to learn how to capture/edit/encode/upload his submission. The ability to upload directly a file still remains for those who want to do it.

3.1.1 Flash technology

The choice of a web technology is here very easy. Flash, from Adobe, is simply the only one that allows to create an homogeneous web application achieving our goal. It is used both on the server and on the client sides. Although being proprietary, we can list four major reasons to select this technology:

- Runtime communality: the Flash Player is free and installed on most of the Internet connected consumer computers: penetration rate is 98% of the computers park on the Internet. It is available for Windows, MacOS and Linux operating systems.

- Flash Video (FLV): this is a video format that appears to be a very efficient one, having one of the best quality/compression ratio. It can be sent from the server via HTTP (progressive download) or via RTMP (streaming) protocols. Moreover FLV encoding can be performed from the Flash player itself to prepare the video broadcasting for the server part.
- ActionScript: this is the programming language of the Flash environment. Years after years, it has been dramatically enhanced and is now object oriented and provide a rich API.
- Flash Media Server: this server application can manage multimedia streams, even in real time (live broadcasting). It allows to do direct stream record from a browser.

3.1.2 Global architecture

Oppositely to traditional web network architecture, communication is here heavily bi-directional because of the video upload process. The architecture lays on four layers:

- HTTP server: it hosts the start page embedding the Flash application. It handles direct file upload, too. Only FLV format is supported as an input for the moment.
- Flash Media Server handles the audio/video/data content: video streaming as well as video capture. Streaming is also interesting because author content is never written to user's hard disk, ensuring copy control.
- MySQL database system management: where are recorded all the data relative to users, pieces and contributions.
- Client station: access the above mentioned servers from a web browser with the Flash plug-in installed.

3.1.3 Server side requirements

Apart from the server applications themselves, there are specific requirements. In fact, the quality of the service depends directly on the ability to react quickly to requests. Thus, video streaming demands a symetric wide band Internet connexion and high disks performances to handle multiple files simultaneous access. Contrary to these points, processing needs are not as high because FLV encoding is normally not achieved on the server.

3.1.4 Client side requirements

As far as the hardware is concerned, user should have an audio/video input device (webcam, etc), a broadband Internet access (ADSL at 512 Kbps is a minimum) and a computer which has processing power enough to complete real time FLV encoding to upload to the Flash Media Server.

As far as the software is concerned, it is rather simpler. A browser with the latest Flash plug-in is sufficient.

We do not mention here the non (so) technical requirements. However, content quality rely on them. Here are examples of these requirements: room lighting, camera sensor quality, background noise, etc.

3.2 Resulting system

Following all the previous recommendations, we built a system having the described main features. It is located at http://e-guitare.univ-reunion.fr/figs/

As far as usability is concerned. we rely on the well-known mechanisms for video handling: traditional remote control and media selection. Online video capture is more unusual. We therefore decided to mimic traditional consumer video editing software like iMovie or Windows Movie Maker, in an even simpler form.

In fact, we applied another principle to facilitate user experience: on-screen elements (windows, controls, indications) that are not useful in a specific context are hidden. They appear when they are needed.

Finally, we added the question/answer communication scheme, that is to say, it is possible to start a thread that embeds different levels of discussion. This part is very similar to what is implemented in traditional forums based on a tree structure.

3.2.1 Musical content specificities

At this point, it is interesting to see what features were added to fulfill our musical learning goal.

We have video performances of our pieces recorded by professional guitarists who are teachers too. We use them as reference for the learners. They are provided with two particularities: they are multi-angles and can be slowed down. One can pick the best angle that show the information he needs at half the speed. This is important as playing guitar is, basically, a question of succeeding to do the good gestures with the left and the right hand. For example: how to play the right hand arpeggio in “Forbidden Game”? We therefore have the possibility to focus on this hand.

Teachers' video are interesting to us for another reason: they are used as time reference. As learning work is based upon pieces, the idea is to create discussions on localized parts of the piece. For instance, a submission dealing with a difficult chord is attached to the moment when this chord is played in the professor's performance.

Two time references are possible. The first one is based upon musical systems (lines of the paper score). If this document is not available, the tune is divided into logical parts: intro, verse, refrain, bridge, outro, ... The time division is set as input thanks to an XML document making a parallel between video time codes and system numbers.

The second time reference is marker based: user places start and end flags (small triangles) with the mouse or keyboard shortcuts. These markers can also be used to loop a part of the piece: music learning rely heavily on repetition. We have to facilitate it.

This is not an obligation for the learner to manage contextualisation, everything is done automatically, unless he wants to be more precise.

3.2.2 Interface

The main screen of the application is shown on , with all the panes displayed. Here are the captions:

- 1: professor reference video
- 2: video gloss
- 3: gloss meta-data. Notice the “Reply” button

- 4: gloss listing for current system
- 5: discussion navigator pane: allow to browse glosses, their replies and to start a new thread
- 6: navigation buttons
- 7: back to piece/tune listing
- 8: Angle selector
- 9: playback (and record) remote control
- 10: markers (position flags that are also used to loop in the piece)
- 11: markers/loop activator
- 12: record controls
- 13: record quality settings

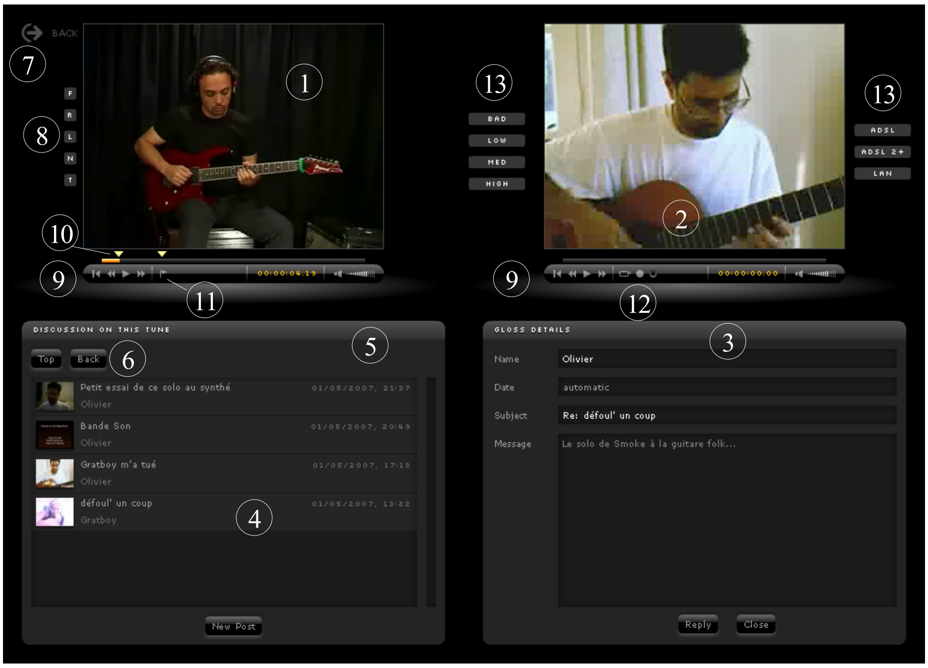


Figure 1: main screen with gloss pane deployed

When arriving for the first time, only the left half of the picture is visible until a gloss has been selected or the “New Post” button is invoked.

In the same way, the record controls appear when needed, that is to say, either when starting a new gloss or when replying to an existing one. Very basic presets are proposed to match user equipment: picture size (Bad/Low/Med/High) and connexion speed (ADSL/ADSL2+/LAN). Default setting (Med/ADSL) will fit for everybody.

3.2.3 Typical use case

Let us now follow the typical navigation path of a learner in our system to illustrate its interest.

- First thing to do is to pick a piece in our collection. Many criteria are proposed to make a choice: difficulty, style, type of guitar, etc.
- User can listen and watch the professor performing the piece. Multi-angles feature allows to focus on the hands from different points of view.
- The musical score is also available to download and print (PDF format).

- With these two medias, the learner can study the piece part by part by using the loop feature.
- During its progression, he can see the previous discussions regarding his current position in the navigator pane.
- Once a difficulty that is not discussed is encountered, he may post a new gloss. If the markers are activated their time-codes are stored. In all cases, the new contribution is indexed to the current system.
- After having recorded the video (live or file upload), user fills in a small meta-data form (his name, the subject and a comment).
- A teacher will post an answer to his question, using the same procedure.
- The process can be repeated as many time as needed.

4 Future work and conclusion

This web-based tool is in fact the new element of an instrumental e-learning project called e-guitare initiated in 2003. The result of this project is a DVD-based interactive learning system. This system is more comfortable to use for the practical guitar learning phase because it provides rich-media content, without network limitations. One of our objective is therefore to couple the online and the offline tools. They are highly complementary: the offline version succeeds in professor to student communication whereas the other manages learner to professor and/or learner to learner communication.

The whole new system will have to be tested. We are attached to the user centred development according to a process exposed in I-Know'2006 [Conruyt, 06]. Validating our vision is a path that will be explored in the forthcoming months.

Acknowledgements

We would like to thank the European Union, the French State and the Regional council of Réunion Island for supporting this research.

References

[O'Reilly, 05] T. O'Reilly, What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software, September 2005.

[Adobe, 06] Flash Environment – Adobe; Link: http://www.adobe.com/products/flashmediaserver/

[Calderoni, 07] Flashblog (French), 2007, Link: http://flashblog.univ-reunion.fr/

[E-guitare, 03-06] Site du projet e-guitare; Link: http://e-guitare.univ-reunion.fr

[Sébastien, 05] Sébastien, Conruyt, Natural music e-learning system development: the e-guitare user-centred approach, ICOOL 2005, South Africa, 2005.

[Conruyt, 06] Conruyt et al, A Methodology for Designing E-services from a Co-Design Platform, i-Know'2006, Austria, 2006

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Making Learning Management Systems Work - Usability as Key Issue for System and Interface Design

Michael Burmester, Frank Thissen, Jörg Westbomke
(Stuttgart Media University , Germany
{burmester, thissen, westbomke}@hdm-stuttgart.de)

Abstract: New media applications like E-Learning or Computer supported collaborative learning are becoming more and more popular and the internet plays a key role in this development. These applications open up new ways in education, which were not thinkable only some years ago. Nowadays the progress in technology like broadband internet access, video on demand, streaming media, etc. allow to realize such applications. But as first approaches show, to realize a good E-Learning application is not only a matter of technology it is also a matter of good interface and system design. Often a good interface/system design, which puts the user and his needs into the focus, is as important for a successful application as the realized functionality.
This papers presents a study investigating the learner perspective on the use of learning management systems by conducting focus groups. Based on the statements we derived requirements and propose design improvements in order to support and facilitate collaborative learning applications.

Keywords: CSCL, usability engineering, learning management system, learner-centered design
Categories: H.5.2, H.5.3, K.3.1

1 Introduction

The European research project Collaboration Across Borders (CAB, funded by EC Minerva 110681-CP-1-2003-1-UK-MINERVA-M) established a network of collaboration between educational institutions across Europe. The project intends to facilitate effective online collaboration between students from all over Europe and abroad. The objectives are to investigate, evaluate and reflect the benefits and optimization potential of peer group collaboration.

We assume that besides personal and contextual factors, technology can support and facilitate – or obstruct – the collaboration and learning process. For collaborative activities in most cases learning management systems (e.g. MOODLE, WebCT, IBT®SERVER) are employed. Analyzing the application of these learning management systems (LMS) it becomes obvious that they follow a uniform teacher-oriented didactical structure and force the learner to maintain this. This is in conflict with theoretical assumptions of collaborative learning, which follow a learner-oriented perspective of learning [Bett 03]. For collaboration activities, LMS offer functions like chat, discussion forums, e-mail, and whiteboard [Schulmeister 03].

Corresponding to the well-known user-centered design approach (UCD), a learner-centered design process (LCD) is needed in order to design LMS that fit the requirements of learners and their context of use [Guzdiul 95]. Following a LCD, the necessity of usability evaluation of e-learning applications is more and more accepted [Ardito 04, Zaharias 04]. However, an important issue is the analysis how students really work with LMS and what they really need for their learning and working process [Convertino 04].

2 Qualitative Research Method

The main objective of the CAB project is to initiate and support collaborative learning across borders. For this purpose, a LMS is employed, which offers tools for collaboration that support the students' teamwork. The following research question guided this study: What aspects of LMS must be improved from a usability point of view including functionality and user interface design issues to support collaborative learning?

2.1 Focus Groups

To answer our research question out of a learner-centered perspective we first had to understand student work and learning during CSCL seminars from the student point of view to derive requirements and design optimizations for LMS. The qualitative data acquisition method focus group as a group interview technique is well suited for that purpose. One central advantage compared to classical interview techniques is that the interviewees not just respond to the interviewer's questions, but also react to the answers of the other focus group participants. Another advantage is that moderation techniques can be used in order to qualify statements and to support participants to get deep into the question focus.

Four focus groups were conducted. Each focus group followed the same set of questions and moderation techniques was based on a common questions and moderation guide. This guide contained all instructions for the moderator, the co-moderator, the protocol recorder and the technology assistant. The focus groups were conducted in the usability lab of the Stuttgart Media University. The participants need to have experience with at least one CSCL course in a university context.

2.2 Qualitative Data Analysis

Focus groups mainly generate verbal data. The adequate way of data analysis is qualitative data analysis. The theoretical bases for that is the grounded theory [Strauss 96]. The analysis structure applied in this study is a combination of practical strategies and the analysis strategy of grounded theory. The analysis procedure has been carried out in five steps:

1. Transcript of all verbal statements (number of statements recorded in the four focus groups ranged from 67 to 642; average of 394 statements).
2. Coding of all statements with inductively defined and iteratively optimized codes (a code represents the underlying concept of a verbal statement)
3. Interpretation of all codes and identification of themes (range of 11 to 29 themes between the four focus groups, average 20 themes)

For all focus groups:

4. Review of all focus group interpretations and re-interpretation of the complete verbal material
5. Integration of code interpretations of all four focus groups and identification of requirements

The first three steps followed a practical approach in order to identify the main themes for each focus group. In step 4 and 5 an analysis according to the grounded theory supported by the software tool ATLASti was applied.

2.3 Action science

In software design and development, the definition of requirements is a central activity to describe the envisioned system. According to Carroll and Rosson [Carroll 04], theory development in human-computer interaction research can be achieved by designing artifacts in combination with analysis of the design rational and investigation of the use of artifacts. Applying action science to our research questions means in a first step to formulate requirements based on the qualitative analysis and in a second step to design prominent LMS aspects identified in the data. Consequently, the design method was scenario-based, which is a proved method in the field of CSCL.

3 Results, relations and requirements

3.1 General attitudes towards CSCL and motivation factors

Students said that CSCL has several benefits over traditional seminars. Major benefits are seen in the extensive communication and collaboration with others. Exchanging information is felt as a strong benefit. Students appreciate the possibility to work independently from time and location. "*For me it is the open time management. That is simply more flexible. For me that is the main benefit*". Working hours during the night and at home are very famous. The work based on CSCL seems to be more goal-directed and concentrated. Students said that learning with students from other countries would be very valuable for them.

As a conclusion we can say that students initially accept and appreciate CSCL. But there are several de-motivating factors, which have the potential to destroy the positive attitude and motivation. Very important de-motivation factors were technical problems with the LMS. "*The platform [LMS] was de-motivating, because it didn't work correctly. We must use it, but I always postponed the work*". Technical problems are an important obstacle for learning and working.

Bad usability is also associated with the technical de-motivation factors. Examples of bad usability are ill-sorted discussion threats, no sorting possibility of document pools, poor readability of chat protocols, no filter possibilities e.g. in link directories. "*the usability of [<LMS name>] was horrible. It took a lot of time. Very often, new windows appeared which I didn't want. The link names were not clear to me.*". Associated with usability problems, students complained about the esthetic design of the LMS interfaces. The design was not appealing for them and they asked for personalization possibilities of the design. Students complained that information material is very often presented as text. Together with text-based synchronous and asynchronous communication means such as chat, e-mail, instant messages, discussion forums, the text overload is a strain for the students.

De-motivation of some participants of a working group can have an effect on other participants because all participants can see if a person provides shorter feedback or smaller inputs to joint work. From these results, we derive the following general requirements:

[**R1**] Technical problems lead to frustration and must be avoided.

[**R2**] The usability standard of LMS should obviously be improved.

[**R3**] The visual presentation of a LMS should be appealing for the target group.

[**R4**] Material should be presented in different medial presentation formats like video, audio or text-to-speech possibilities in order to reduce text reading.

3.2 Introduction in a CSCL course

Students felt the start of a virtual seminar as a very important phase. In this phase, it is necessary to reduce anxiety and insecurity. Nearly all CSCL seminars reported during the focus groups had a presence introduction at the beginning. In general, blended learning elements were quite usual. For a seminar with students distributed over different countries, this is not possible. Therefore, the following requirement is important for pure CSCL seminars.

[**R5**] All CSCL courses must show an introduction of the seminar: the technology and its usage, the other students as well as the expected activities during the CSCL seminar. In order to cope with technical problems (see R1), it is also important to offer communication means independently from the LMS.

3.3 Work organisation

After introducing into a course, the formation of working groups is another critical phase of collaborative learning in groups. The students clearly prefer the formation of working groups by deciding on their own who is in their working group. They prefer to work with students they know very well. For successful work in groups, the students said that it is important to know the working style, the motivation and the knowledge background of the group members. Working groups seem to work more effectively and efficiently, if they well know each other. For collaboration between different countries, presence meeting are normally not possible. Therefore, LMS should support the process to learn more about the other participants.

[**R6**] A LMS should provide means for group forming. A prerequisite for that is that the group members are introduced to each other at the beginning of their joint work. The tutor must initiate and should use support of the LMS.

Group work discussed in the focus groups show some typical group activities, like splitting project tasks in sub-tasks, decision which tasks have to be accomplished in group work and individual work, distributing tasks between group members, agreeing on milestones and deadlines, solving group problems or presenting results. Looking at these group activities we see that a working group involves intense collaboration and communication. The LMS experienced by the students did not sufficiently support these types of activities. Of course, it is possible to communicate and to exchange documents, but e.g. explicit work distribution, basic project management features such as definition of tasks, responsibilities, milestones, deadlines etc. were poorly supported. Furthermore, features are not designed according to the whole task sequence. For example, on many LMS, initiating a chat is mostly just opening a communication channel. However, a chat meeting as synchronous communication needs an appointment. Planning appropriate dates and informing all participants about that date necessary to initiate a chat meeting.

Normally, student working groups are rather small and the projects are rather small as well. Therefore, not highly sophisticated project management and CSCW tools are required, but the working group members need a complete picture about the ongoing collaboration and the activities of the working group (activity awareness).

[**R7**] The LMS should support group activities in project-based working and learning groups. Very important is basic project management.

[**R8**] LMS tools should reflect the whole task flow, i.e. for synchronous communication, it must be possible to manage a meeting date.

3.4 Problems and conflicts in group work

Several problems of group work were reported in the focus groups. For example group meetings are felt as ineffective and inefficient or team members drop out of the group work, obstructing the group work. Looking at these problems of teamwork in student working groups it becomes clear that a main reason for that is a lack of knowledge how to organize and how to manage teamwork. In virtual working groups, insufficient support of group work by the LMS aggravates this situation. Beside these problems, it became clear in the focus group discussions that group work of students is accompanied by conflicts. Some of the reasons for conflicts were: Differences in the interest on the seminar topic, differences in the motivation, differences in previous knowledge, unconstructive discussions and a lack in finding compromises or lacking or unsuccessful decisions on responsibilities.

Conflicts are connected with negative emotions. Conflicts get worse in virtual teams due to the general disadvantage of communication via electronic media in comparison to face-to-face meetings. So electronic communication is very often text-based, therefore it is difficult and exhausting to formulate a message in the right way. And the lack of non-verbal communication often cause new misunderstandings.

Problems and conflicts obstruct the group work. It is of central importance to install features to prevent and solve problems and conflicts. The following requirements contribute to that goal:

[**R9**] The role of a moderator or team leader is important for successful group work. The moderator must be supported by information on that role and provided with tools to fulfill the role.

[**R10**] It is important to support the problem-solving processes in a team. Additional tools for group work would be e.g. a brainstorming tool or other creativity techniques like mind mapping.

[**R11**] Basic project management functions are needed in order to plan tasks, make a time schedule and have a group calendar. This calendar must have a course view, a group view as well as a personal view of the individual user.

[**R12**] LMS should provide communication means that support natural auditory or video-auditory communication. This reduces the text focus and the risk of misunderstandings because of additional information via the non-verbal channel.

[**R13**] For asynchronous communication, the prompt reaction should be supported by e.g. alerting the receiver of a message to answer in time.

[**R14**] The communication between a single student and/or the working group must be supported, e.g. by a virtual consultation hour with the tutor.

[**R15**] Conflict solving must be supported by early detection and systematic support of conflict solving techniques.

4 Consequences and Perspectives

Our focus group results show that the existing LMS do not sufficiently imply solutions to solve the discussed problems. On the one hand, the virtual form brings new difficulties and challenges for group work, but on the other hand, the technology used for collaboration could add value to solve group-work-specific problems. When reading the following design proposals, some might say that such solutions still exists. However, functions for group organization like a calendar can be found in today’s LMS, but very often, these calendars do not have all required functions, are not specific for groups or hard to operate.

4.1 Support of group activities

Three areas of group activities can be derived from the focus groups. (1) Of central importance is a group area: e.g. a virtual team room. This is the area where the team activities can take place, e.g. a team chat room, team forum, team document pool, team calendar for co-ordination of tasks, milestones, appointments. (2) In addition to that, there is a course room. Here all groups can present their results and get feedback from other groups. (3) All users need their own CSCL desktop, which can be personalized with respect to visual presentation as well as important tools and functions for the user, e.g. organization of personal documents, calendar, etc.

4.2 Selected Solutions

According to the idea of action science [Carroll 04], the results of the focus groups and the requirements were transferred to design solutions. The following design solutions were developed as scenario machine prototypes, i.e. a prototype showing the interaction of the user and the user interface along a scenario:

- Introduction phase at the beginning of a CSCL seminar: log into the seminar, get acquainted with the others, forming a working group.
- Get acquainted with the other students by playing games in order to initiate first co-operation of the students and to ensure that the profiles are filled out.
- Group calendars, which support the co-ordination appointments, deadlines and milestones of the tutor, the complete course and the working group as well as the working group members.
- Wrap-up information since the last log-in, so that the working group member sees the new mails, messages, tasks, appointments etc. since the last log-in.
- Chat-based virtual meetings within the working group supporting the moderator role with moderation techniques (right to speak, agreement on

communication rules, meeting agenda, etc.) combined with document or application sharing in order to discuss work results or work together.

- A brainstorming tool should support creative processes and problem solving situations.
- A conflict-solving tool in order to automatically support working groups in coping with group conflicts.

In the following, we present two exemplary design solutions: one game for being acquainted with each other and a conflict-solving tool.

4.2.1 Get to know to each other

It is very important that partners in working groups know each other very well. Very often, the user profiles are not filled out on LMS. Therefore, interactive technology can be used to motivate participants of a CSCL course to get to know each other. One possibility is to start with games. Figure 1 shows a prototype of a game, which works in two steps: First step is a game based on the metaphor of "speed dating". That means that two participants at a time have some time to answer predefined questions and to exchange the answers. Based on that, all participants have the chance to have a short dialogue with all others. When this phase is finished, statements of selected participants are presented and the user has to guess who could have said what. By answering all questions, the user profile is filled out.

Figure 1: Prototype of a game to get to know each other

4.2.2 Conflict solving

The focus group results showed that conflicts are very common in working groups. The idea is to support teams in coping with conflicts. Of course, the tutor is a very important person in supporting teams, but students stated that they are reluctant to contact the tutor when conflicts appear. Some strategies for solving conflicts are known by the students, but they do not apply to them in all cases.

The basic design strategy is to apply moderation techniques to conflict solving. The tool comes up after each log-in and presents the next step in the problem-solving process. When finished, the user can proceed with his work. The next step of the conflict-solving process is always launched, when all participants have finished the previous step. The conflict-solving tool operates in the following phases:

(1) Log-in: After each log-in, the participants are asked to rate their mood concerning the group work. If the mood is neutral or positive, nothing happens.
(2) Decision: If the average mood assessment over all group members is in the negative range for a defined time (say 5 days), then conflict solving is proposed by the system. The group member can decide whether she or he wants a supported conflict solving (see Figure 2). If all participants answer that conflict solving is appropriate, the tool starts with the next phase.
(3) Problem collection: All participants are asked to collect their problems, which could cause the bad group mood.
(4) Importance rating: The importance of the collected problems is rated.
(5) Reason collection: Reasons for the most important problem are collected.
(6) Solution finding: Possible solutions are collected in the working group.
(7) Solution rating and selection: The solutions accepted by the majority of group members are selected.
(8) Application planning: A plan is made how to implement the accepted solution in the group work by defining concrete activities, responsibilities and deadlines.
(9) Feedback: A final feedback in form of a verbal statement is collected from and presented to all group members.
(10) Closing conflict solving: The tool closes and the work continuous.

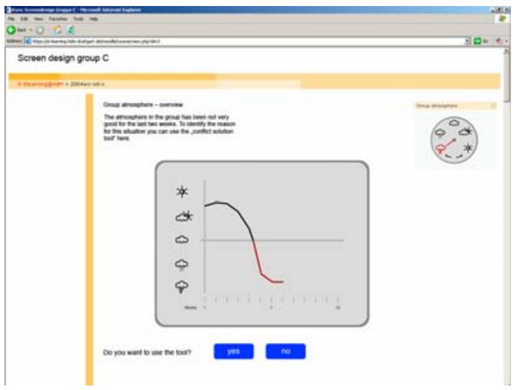


Figure 2: After having rated the mood, the average mood of the working group is indicated and conflict solving is proposed.

5 Conclusion

The objective of our study was to define requirements for LMS to support student collaborations. These requirements were derived from our focus group results. Based on our results, we propose possible design solutions to support collaborative activities, which are especially important if blended learning arrangements are not possible. The next step would be to evaluate these solutions to see if they work in a real university context of collaborative learning. On this basis, a stable and deep understanding on how LMS should be designed can be achieved.

References

[Ardito 04] C. Ardito, M. De Marsico, R. Lanzilotti, S. Levialdi, T. Roselli, V. Rossano, M. Tersigni, Usability of E-Learning Tools. In: Proc. of AVI '04, Italy, May 2004, 80-84.

[Bett 03] L. Bett, U. Rinn, Lernplattformen zwischen Technik und Didaktik in: J. Wedekind/K. Bett (Hrsg.): Lernplattformen in der Praxis, Münster: Waxmann, 2003.

[Carroll 04] J.M. Carroll, M.B. Rosson, Design Rational as Theory. In: J.M. Carroll (Eds.), HCI Models, Theories, and Frameworks. Morgan Kaufmann, Amsterdam, 2004, 431-461.

[Convertino 04] G. Convertino, D. C. Neale, L. Hobby, J. M. Carroll, M. B. Rosson, A Laboratory Method for Studying Activity Awareness. NordiCHI '04, October 23-27, 2004 Tampere, Finland, 2004, 313-322.

[Guzdiul 95] M. Guzdiul, B.K. Yusmin, Learner-Centered System Design: HCI Perspective for the Future. In: Proc of DIS 95 Ann Arbor MI USA, 1995, 143-147.

[Rosson 02] Rosson, M.B. & Carroll, J.M. Usability Engineering – Scenario-based development of human-computer interaction. Morgan Kaufmann, San Francisco, 2002.

[Schulmeister 03] R. Schulmeister, Lernplattformen für das virtuelle Lernen. Evaluation und Didaktik. Oldenburg, München, 2003.

[Strauss 96] A. L. Strauss, J.M. Corbin, Grounded Theory: Grundlagen Qualitativer Sozialforschung. Beltz PVU, 1996.

[Zaharias 04] P. Zaharias, Usability and e-learning: the road towards integration, eLearn, ACM Press New York, NY, USA, Volume 2004, Issue 6.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Personalizing the Web Content on User Perceptual Preferences

Nikos Tsianos, Panagiotis Germanakos, Zacharias Lekkas, Costas Mourlas
(Department of Communication and Mass Media, University of Athens, Greece
{ntsianos,pgerman,mourlas}@media.uoa.gr)

Abstract: This paper introduces a new model of personalized usage of the internet that is based on technologies of user representation, artificial intelligence and semantic augmentation of the content. By taking advantage of internet's unprecedented dynamics, compared to traditional media, this user representation model incorporates cognitive, mainly, psychology theories, combined with parameters that constitute more traditional approaches in user profiling (such as demographics, expertise, etc).
The purpose of this research is to alleviate difficulties that massive approaches impose on areas such as education and information processing, by integrating intelligent adaptive characteristics into web applications; this can lead to a highly adapted to each user's needs content and more effective, in our case, learning.

Keywords: Hypermedia systems, Web-based services, XML, Semantic Web, Multimedia
Categories: E.1, J.4, H.5.2, H.5.4

1 Introduction

It is a fact nowadays that the number of web-pages and information available in the internet is such that it has gradually become very difficult for users to reach information they actually seek out [De Bra, Aroyo & Chepegin, 04]. As a result, designers of intelligent systems have introduced the term of Adaptive Hypermedia, in order to optimize navigation and content provision processes, as a response to the aforementioned innate difficulties.

A system can be classified as an Adaptive Hypermedia System if it is based on hypermedia, has an explicit user-model representing certain characteristics of the user, has a domain model which is a set of relationships between knowledge elements in the information space, and is capable of modifying some visible or functional part of the system based on the information maintained in the user-model [Eklund & Sinclair, 00; Brusilovsky & Nejdl, 04; Brusilovsky, 96b].

In 1996, Brusilovsky identified four user characteristics to which an Adaptive Hypermedia System should adapt [Brusilovsky, 96b; Brusilovsky, 01]. These were user's knowledge, goals, background and hypertext experience, and user's preferences. In 2001, further two sources of adaptation were added to this list, user's interests and individual traits, while a third source of different nature having to deal with the user's environment had also been identified.

The project is co-funded by the European Social Fund and National Resources (EPEAEK II) PYTHAGORAS

Generally, Adaptive Hypermedia Systems can be useful in application areas where the hyperspace is reasonably large and the user population is relatively diverse in terms of the above user characteristics [Brusilovsky, 01; Brusilovsky, 96a; Brusilovsky & Nejdl, 04; Brusilovsky, 96b]. A review by Brusilovsky has identified six specific application areas for adaptive hypermedia systems since 1996 [Brusilovsky, 01]. These are educational hypermedia, on-line information systems, information retrieval systems, institutional hypermedia and systems for managing personalized view in information spaces. Educational hypermedia and on-line information systems are the most popular, accounting for about two thirds of the research efforts in adaptive hypermedia.

Furthermore, speaking in terms of traditional media, generalizing in the level of an Adaptive Web, such architecture makes possible the personalization of each "message"; this is undoubtedly an important step ahead if compared to a general classification of similar user groups that an audience reception theory [Morley, 80] from traditional media studies would perhaps lead to. The main difference is that in the case of personalized content provision, users not only project their personal characteristics on information, but actually receive what they actually are interested in; even more importantly, when more implicit cognitive and emotional processes are involved, users are provided with content and navigational tools that match their individual traits and preferred (consciously or not) ways of receiving and processing information.

The goal of our research is to integrate individual cognitive and emotional characteristics, which we refer to as User Perceptual Preferences Characteristics, as main parameters into the adaptive system we have already developed. Our system, though it concerns web personalization in general, focuses on educational purposes for experimental reasons, and its personalization mechanism relies exactly on mapping the provided content and navigational patterns on each user's preferences and inclinations.

2 User Perceptual Preferences Characteristics

User Perceptual Preferences Characteristics is the new (as compared to traditional approaches) component / dimension of a comprehensive user profile we have built and are in the process of evaluating. This model embraces visual attention and cognitive processes (including emotional parameters) that could be described as user "perceptual and information processing preferences or abilities", aiming to enhance learning efficacy.

User Perceptual Preferences could be described as a continuous mental process, which starts with the perception of an object in the user's attentional visual field, and involves a number of cognitive, learning and emotional processes that lead to the actual response to that stimulus [Germanakos et al, 07] (figure 1).

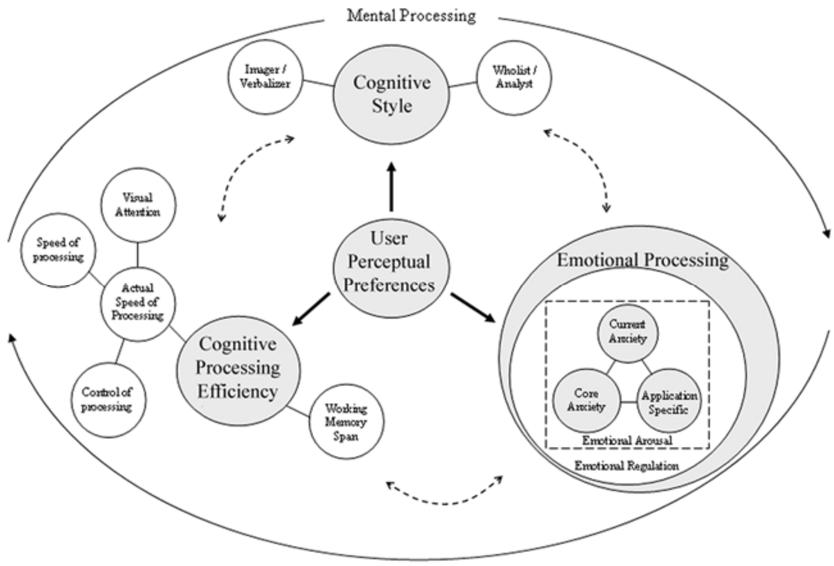


Figure 1: User Perceptual Preference Characteristics – Three-Dimensional Approach

This model's primary parameters formulate a three-dimensional approach. The first dimension investigates the visual and cognitive processing of the user, the second his / her learning style, while the third captures his / her emotionality during the interaction process with the information space.

2.1 Learning Styles

Within the context of educational psychology, theories of learning and cognitive styles have been developed, addressing the issue of individual differences in learning, or more specifically, the perception, processing and retaining of information. Cognitive styles have been defined by Messick as "consistent individual differences in preferred ways of organising and processing information and experience", while by Sternberg and Grigorenko as "a bridge between what might seem to be two fairly distinct areas of psychological investigation: cognition and personality" [Sadler-Smith & Riding, 99]. Learning styles, as a term, are frequently used interchangeably with cognitive styles, but in general are broader concepts that incorporate a greater number of not mutually exclusive characteristics, and focus on learning than cognitive tasks [Cassidy, 04].

Taking into account individual cognitive and learning styles is of high importance, since such an approach "can lead to new insights into the learning process, a greater knowledge of individual differences, and an expanding repertoire of methods for the teacher" [Banner & Rayner, 00].

Regarding the hypermedia information space, amongst the numerous proposed theories of individual style, we consider that Riding's Cognitive Style Analysis [Riding, 01] and Felder / Silverman's Index of Learning Style [Felder & Silverman, 88] implications can be mapped on the information space more precisely, since they are consisted of distinct scales that respond directly to different aspects of the Web space. Still, one must be cautious about reliability and validity issues of such psychometric tools, since many of them often fail to exhibit satisfactory results; that may be true for the case of Felder/Silverman's ILS, that perhaps should be further evaluated [Markham, 04], without of course neglecting the authors' reliability and validity reports [Felder & Spurlin, 05].

On the other hand, learning style theories that define specific types of learners, as Kolb's Experiential Learning Theory [Kolb & Kolb, 05], have far more complex implications, since they relate strongly with personality theories, such as the MBTI [Myers Briggs et al, 98], and therefore cannot be adequately quantified and correlated easily with Web objects and structures. For example, being a "converger" according to Kolb's typology or a "judging" person (from a personality psychometric tool aspect) has implications that are more dominant in a traditional social learning environment than in a hypermedia environment. The case of collaborative learning may provide the basis for assessing the role of such learning styles in performance, but is the beyond the scope of our present research.

As part of our research, we did find significant correlation between academic performance and adaptation on specific learning style [Tsianos, Germanakos & Mourlas, 06], though we now have implemented Riding's typology implications in our current hypermedia application, rather than Felder's that we firstly used. Part of our research is to examine whether such results can be repeatedly found, in order to support the importance of learning and cognitive styles.

We use Riding's CSA since it can be applied on most cognitive informational processing tasks (rather than strictly educational), the implications are quite clear in terms of hypermedia design (visual/verbal content presentation and wholist/analyst pattern of navigation), and is probably one of the most inclusive theories, since it is actually derived from the common axises of a number of previous theories.

2.2 Cognitive Parameters

We refer to the term visual attention in the sense of tracking user's eye movements, and in particular scanning his / her eye gaze on the information environment [Gulliver & Ghinea, 2004]. Our prime concern is identifying the parts of information that are of main interest to the user, depending on his prior knowledge.

Visual attention is composed of two serial phases: the pre-attentive and the limited-capacity stage. The pre-attentive stage of vision subconsciously defines objects from visual primitives, such as lines, curvature, orientation, colour and motion and allows definition of objects in the visual field. When items pass from the pre-attentive stage to the limited-capacity stage, these items are considered as selected. Interpretation of eye movement data is based on the empirically validated assumption that when a person is performing a cognitive task, while watching a display, the location of his / her gaze corresponds to the symbol currently being processed in working memory and, moreover, that the eye naturally focuses on areas that are most likely to be informative. Regarding the role of visual attention, we intend to extent our research to that direction with the use of the eye tracking tool.

Cognitive Processing parameters could be primarily summarised in (i) control of processing (refers to the processes that identify and register goal-relevant information and block out dominant or appealing but actually irrelevant information), (ii) speed of processing (refers to the maximum speed at which a given mental act may be efficiently executed), and (iii) working memory (refers to the processes that enable a person to hold information in an active state while integrating it with other information until the current problem is solved) [Demetriou et al., 93; Demetriou & Kazi, 01].

At this point, we have calibrated the psychometric tools that we use to measure these parameters, and we have preliminary results that show statistically significant correlations with academic performance. As expected by theory, proper personalizing techniques (like reducing the volume of information in case of low memory and processing abilities) have indeed optimized performance.

2.3 Emotional Processing

On the basis of the research conducted by Goleman [95], as well as Salovey & Mayer [90], who have introduced the term, we developed an Emotional Control questionnaire that examines the 3 out of 5 scales that comprise the Emotional Intelligence construct (according to Goleman), since factors that deal with human to human interaction (like empathy) are not present in our Web- applications - at least for the time being.

However, this indirect measurement of a (moderated) person’s emotionality does not address emotions and their effect on performance. In terms of actual emotions that affect academic performance, bibliographic research has shown that anxiety is often correlated with academic performance [Cassady, 04], as well with performance in computer mediated learning procedures [Smith & Caputi, 05; Chang, 05]. Subsequently, different levels of anxiety should have also a significant effect in cognitive functions.

The concept of general anxiety is indicative of a person’s tendency to exhibit lower performance in information processing tasks, but not all circumstances are the same. For that reason, we measure not only one’s general (core) anxiety, but also application specific anxiety and his current self-reported anxiety. In our experimental case, application specific anxiety is measured by a questionnaire that refers to educational test anxiety, while users can self-report their levels of anxiety using an indicative bar that is embedded in the interface.

Thus, we refer to anxiety as the basic component of the generalized concept of Emotional Arousal. However, we do not by any means consider anxiety as the sole predictor of the effects of emotions on academic performance; this is why figure 1 illustrates Emotional Arousal as an area, that even though it is dominated (research wise) by the aforementioned concepts of anxiety, the role of additional emotions is also implied.

Moreover, we believe that combining the level of anxiety of an individual with the moderating role of Emotional Control, it is possible to clarify, at some extent, how affectional responses of the individual hamper or promote learning procedures. By personalizing the educational content that our already developed adaptive system provides on the basis of this concept of emotionality, we can avoid stressful instances and take full advantage of his / hers cognitive capacity at any time.

Emotional Control and Emotional Regulation (generalized concept) maintain the same relationship that characterizes Emotional Arousal and anxiety. Emotional Regulation is not a process that can be completely analyzed with tools derived from EQ theories; consequently, though the concept of Emotional Control offers a practical and useful tool, the area of Emotional Regulation (see figure 1) represents wider and more elaborate processes, that in terms of psychometrics can be indirectly measured through the concept of Emotional Control.

At a practical level, we assume that users with high anxiety levels lacking the moderating role of Emotional Control are in a greater need of enhancing the aesthetic aspects of our system, while users with low anxiety levels focus more on usability issues.

The empirical research we conducted provided us with indications that Emotional Control and Anxiety correlate in a way that supports our assumptions [Lekkas et al, 07]- moreover, unpublished results provide even greater support to our approach, though at this point we are in the phase of analysis and interpretation of the data.

3 User Profiling and Data- Implication Correlation

The construction of users' profiles includes a series of psychometric instruments that reveal their perceptual characteristics. As mentioned above, we use:

- Riding's CSA for the Learning/ Cognitive Styles dimension
- A series of real-time measurements for the Cognitive Parameters (Speed of Processing, Control of Processing, Working Memory and Visual Attention), similar to tests developed on the E-prime platform.
- The Emotional Control 27 item questionnaire we have developed (Cronbach's alpha 0.76), and i) the Test Anxiety Inventory [Spielberger & Vagg, 95] to measure application specific anxiety (educational process in our case) and ii) the State-Trait Anxiety Inventory [Spielberger, 83] to measure general (core) anxiety.

Moreover, while users navigate through our application, they can make use of an anxiety bar, which is part of the interface, in order to self-report feelings of inconvenience and high levels of anxiety that burdens their cognitive effort. This self-report measure will be correlated with general and application specific levels of anxiety in order to clarify the extent of their correlation, and the further optimization of the psychometric process.

The representation of the user model requires the use of meta-data that provide hierarchically structured information about users' characteristics [Germanakos et al, 07]. The tree-structured approach we have adapted is shown in figure 2. It should be mentioned that this diagram focuses on cognitive and emotional characteristics, rather than the traditional ones, that may have a high degree of application specific flexibility and are presented here synoptically.

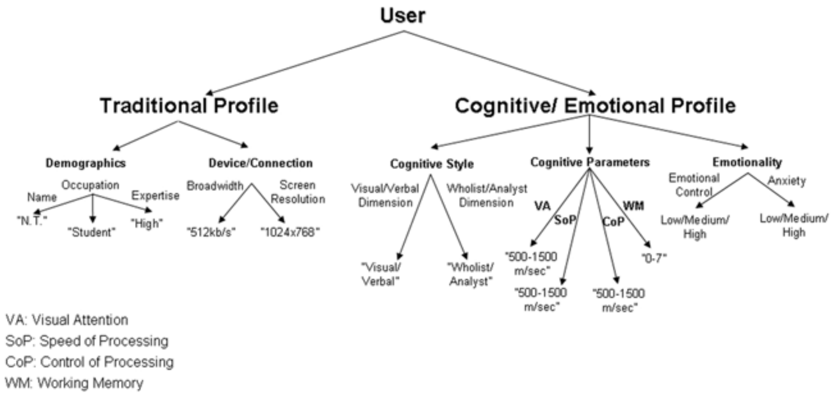


Figure 2: Tree Structure of User Profile

A practical example of data–implications correlation could be as follows, a user might be identified that is: Verbalizer (V) – Wholist (W) with regards to the Learning Style, has an Actual Cognitive Processing Speed Efficiency of 1000 msec, and a fair Working Memory Span (weighting 5/7), with regards to his / her Cognitive Processing Speed Efficiency, and (s)he has low Emotional Processing. The tags affected, according to the rules created and the Data – Implications Correlation, for this particular instance are: *Images* (few images displayed), *Text* (any text could be delivered), *Info Quantity* (less info since his / her cognitive speed is moderate), *Links – Learner Control* (less learner control because (s)he is Wholist), *Additional Navigation Support* (significant because (s)he has low emotional processing), and high *Aesthetics* (to give more structured and well defined information, with more *colors*, *larger fonts*, *more bold text*, since (s)he has low emotional processing).

At this point it should be mentioned that in case of internal correlation conflicts primary implications take over secondary ones. Primary implications are those correlated with high degrees of anxiety and the lack of Emotional Control, that by definition interfere with cognitive processing at all levels, regardless of the aforementioned perceptual inclinations. Additionally, since emotional processing is the most dynamic parameter compared to the others (users may take advantage of the anxiety bar if they wish), any changes occurring at any given time are directly affecting the yielded value of the adaptation and personalization rules and henceforth the format of the content delivered.

4 Adaptive IntelliWeb and Future Work

In support of the aforementioned concepts and user profiling model, we have implemented our theoretical approach into a web application that we refer to as Adaptive IntelliWeb. This is the main component of our AdaptiveWeb System[1,2],

[1] http://www3.cs.ucy.ac.cy/adaptiveweb/
[2] The system will be demostrated in the UM 2007 Conference

responsible for adapting the content according to users' profiles (personalization). Using this framework, users can navigate through all personalized web sites; it also provides extra navigation and learner control tool to help users comprehend the personalized content in the most optimized way.

The Semantic WebEditor is an equally important component of the system, since it provides an interface for building web-sites that offer personalization. At this point, this editor is still under development. The Profile Construction application includes all the psychometric tools mentioned above, plus the traditional profile building procedure, and automatically creates the meta-data enriched user profile, necessary for the mapping process that takes place in the Adaptive IntelliWeb. Our system has been developed on the .Net framework.

As depicted above, we have focused on the educational aspect of personalization, since it provides an excellent basis for examining in depth the role of cognitive and emotional parameters, and the success of the personalizing process can be assessed in terms of learning performance, rather than simply measuring user satisfaction.

The Data- Implications Correlation that we have described above defines in fact the adaptation rules that alter the content presentation and navigational tools and support within the context of our e-learning application.

Consequently, our future work focuses on conducting a number of experiments for further assessing the weight of each factor involved in the learning process, and the actual impact of personalization in optimizing learner performance.

5 Conclusions

User profiling is a basic prerequisite for exploiting internet technologies that, by allowing a new way of distributing information, enhance the effectiveness of this medium. Towards perhaps a new form of the Web, we argue that when information processing (at a considerable extent) is involved, users' intrinsic perceptual preferences should be taken into account, in order not only to alleviate navigational difficulties, but to ensure the effective role of this medium in multiple information-based activities.

In any case, intelligent systems seek out to bridge human to human with human to computer interaction, and educational or learning computer-mediated processes should move to this direction. The one-size-fits-all approach seems rather obsolete when adaptation techniques have been developed at a level that permits the implementation of psychology theories that deal with individual differences and preferences. Consequently, if within the context of a traditional classroom learners are dealt as individuals, an absolutely massive approach in the hyperspace seems less pedagogical viable.

The main challenge, of course, is the quantification of the value each user characteristic takes both in personalization and academic achievement weights; empirical research on the basis of our AdaptiveWeb platform is expected to provide such answers.

References

[Banner & Rayner, 00] Banner, G. & Rayner, S., Learning language and learning style: principles, process and practice. Language Learning Journal, Summer 2000, 21: 37-44.

[Brusilovsky, 01] Brusilovsky P., Adaptive Hypermedia, User Modeling and User-Adapted Interaction 11: 87-110.

[Brusilovsky, 96a] Brusilovsky P., Adaptive Hypermedia: an attempt to analyse and generalize, In P. Brusilovsky, P. Kommers, and Streitz (Eds.), Multimedia, Hypermedia, and Virtual Reality, Berlin: Springer-Verlag, 288-304.

[Brusilovsky, 96b] Brusilovsky P., Methods and techniques of adaptive hypermedia, User Modeling and User Adapted Interaction, 1996, v6, n 2-3: 87-129.

[Brusilovsky & Rayner, 04] Brusilovsky P. & Nejdl W., Adaptive Hypermedia and Adaptive Web, © 2004 CSC Press LLC.

[Cassady, 04] Cassady, C.C., The influence of cognitive test anxiety across the learning–testing cycle, Learning and Instruction 14: 569–592.

[Cassidy, 04] Cassidy, S., Learning Styles: An overview of theories, models, and measures. Educational Psychology, 24 (4): 419-444

[Chang, 05] Chang, S. E., Computer anxiety and perception of task complexity in learning programming-related skills, Computers in Human Behavior 21: 713–728.

[De Bra, Aroyo & Chepegin, 04] De Bra, P., Aroyo, L. & Chepegin, V., The Next Big Thing: Adaptive Web-Based Systems, Journal of Digital Information, 5 (1), article n. 247, http://jodi.tamu.edu/Articles/v05/i01/DeBra.

[Demetriou et al, 93] Demetriou, A., Efklides, A. & Platsidou, M., The architecture and dynamics of developing mind: Experiential structuralism as a frame for unifying cognitive development theories. Monographs of the Society for Research in Child Development, 58 (Serial No. 234), 5-6.

[Demetriou & Kazi, 01] Demetriou, A. & Kazi, S., Unity and modularity in the mind and the self: Studies on the relationships between self-awareness, personality, and intellectual development from childhood to adolescence. London: Routdledge.

[Eklund & Sinclair, 00] Eklund J. & Sinclair K., An empirical appraisal of the effectiveness of adaptive interfaces of instructional systems. Educational Technology and Society 3 (4), ISSN 1436-4522.

[Felder & Silverman, 88] Felder, R.M. & Silverman, L.K., Learning and Teaching Styles in Engineering Education, Engineering Education 78: 674-681.

[Felder & Spurlin, 05] Felder, R.M. & Spurlin, J., Applications, Reliability, and Validity of the Index of Learning Styles, International Journal of Engineering Education, v. 21, n. 1: 103-112.

[Germanakos et al, 07] Germanakos P., Tsianos N., Lekkas Z., Mourlas C., & Samaras G, Capturing Essential Intrinsic User Behaviour Values for the Design of Comprehensive Web-based Personalized Environments, Computers in Human Behavior Journal, Special Issue on Integration of Human Factors in Networked Computing. (to appear)

[Goleman, 95] Goleman, D., Emotional Intelligence: why it can matter more than IQ, New York: Bantam Books.

[Gulliver & Ghinea, 04] Gulliver, S.R. & Ghinea, G., Stars in their Eyes: What Eye-Tracking Reveals about Multimedia Perceptual Quality, IEEE Transactions on Systems, Man and Cybernetics, Part A, 34(4): 472-482.

[Kolb & Kolb, 05] Kolb, A. Y. & Kolb, D. A., The Kolb Learning Style Inventory – Version 3.1 2005 Technical Specifications. Experience Based Learning Systems Inc.

[Lekkas et al, 07] Lekkas Z., Tsianos N., Germanakos P. & Mourlas C., Integrating Cognitive and Emotional Parameters into Designing Adaptive Hypermedia Environments, In Proceedings of the Second European Cognitive Science Conference (EuroCogSci'07), Delphi, Hellas, May 23-27, 2007. (to appear)

[Markham, 04] Markham, S., Learning Styles measurement: a cause for concern, Technical Report, Computing Educational Research Group.

[Morley, 80] Morley, D., The 'Nationwide' Audience: Structure and Decoding, London: BFI.

[Myers Briggs et al, 98] MyersBriggs, I., McCaulley, M. H., Quenk, N. L. & Hammer, A. L., MBTI Manual (A guide to the development and use of the Myers Briggs type indicator). Consulting Psychologists Press.

[Riding, 01] Riding R., Cognitive Style Analysis – Research Administration, Published by Learning and Training Technology.

[Sadler-Smith & Riding, 99] Sadler-Smith, E., Riding, R., Cognitive Style and instructional preferences, Instructional Science, 27, (3): 355–371.

[Salovey & Mayer, 90] Salovey, P. & Mayer, J.D., Emotional intelligence, Imagination, Cognition, and Personality, 9, (3): 185-211.

[Smith & Caputi, 05] Smith, B. & Caputi, P., Cognitive interference model of computer anxiety: Implications for computer-based assessment, Computers in Human Behavior 21, 713-728.

[Spielberger, 83] Spielberger, C. D., Manual for the State-Trait Anxiety Inventory (STAI). PaloAlto, CA: Consulting Psychologists Press.

[Spielberger & Vagg, 95] Spielberger, C. D., & Vagg, P. R., Test anxiety: A transactional process model, In C. D. Spielberger and P. R. Vagg (Eds.), Test anxiety: Theory, assessment, and treatment (pp. 3-14). Washington, DC: Taylor & Francis.

[Tsianos, Germanakos & Mourlas, 06] Tsianos, N., Germanakos, P. & Mourlas, C., Assessing the Importance of Cognitive Learning Styles over Performance in Multimedia Educational Environments, In Proceedings of the 2nd International Conference on Interdisciplinarity in Education (ICIE2006), Athens, May 11-13: 123-130.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Workshop on Multimedia Metadata Applications

Werner Bailer, Peter Schallauer
(JOANNEUM RESEARCH, Graz, Austria
{werner.bailer, peter.schallauer}@joanneum.at)

Harald Kosch
(University of Passau, Germany
Harald.Kosch@Uni-Passau.de)

Ralf Klamma
(RWTH Aachen University, Germany
klamma@i5.informatik.rwth-aachen.de)

Michael Granitzer
(Know Center, Graz, Austria
mgrani@know-center.at)

Different multimedia metadata description standards have been proposed in the recent years, for example MPEG-7 and MPEG-21, as well as a number of formats and standards tailored to specific application domains (e.g. broadcasting). In addition, the application of Semantic Web technologies to multimedia content has brought forward the use of formats such as RDF for the description of multimedia content.

The overall objective of the workshop is to investigate how these formats and standards for multimedia metadata are used in real-world applications. The workshop concentrates on the applications based on MPEG-7 and MPEG-21, but includes also contributions related to other multimedia metadata description standards/formats, especially those integrating semantic information. As many of these standards are relatively new, the lessons learned from using them in real applications can provide valuable feedback for their future development. The integration of these technologies within systems and the interoperability with other (legacy) metadata formats are also relevant topics.

This is the 7$^{\text{th}}$ Workshop of the Multimedia Metadata Community[1], which addresses these topics and aims to bring together experts from research and industry in the area of multimedia metadata and semantics. Building on the tradition of having workshops of the community attached to I-KNOW in 2005 and 2006, this workshop takes place in the context of I-MEDIA 2007. Contri-

[1] http://www.multimedia-metadata.info

butions have gone through a peer review process, using the infrastructure of the TRIPLE-I conference. It has to be noted that about half of the submissions came from organisations that have not been part of the Multimedia Metadata Community before.

Thanassis Perperis, Sofia Tsekeridou and Sergios Theodoridis present an ontology based approach for multimodal detection of violence in video data. Walter Allasia, Fabrizio Falchi and Francesco Gallo propose an approach for searching multimedia data based on rights information using similarity between IPR attributes. Martin Höffernig, Michael Hausenblas and Werner Bailer describe a semantic model of temporal decompositions in content descriptions in order to validate their consistency. Ralf Klamma, Marc Spaniol and Dominik Renzel present an MPEG-7 based and community aware tagging mechanism for multimedia data in order to enable interoperability and semantic enrichment. Wolfgang Jochum analyses the current status of fragment definition formats and defines a set of requirements for developing fragment identification standards.

In addition to presentation sessions, there is for the first time a demonstration session at the workshop that allows presenting prototype applications based on multimedia metadata standards to a larger audience. Accepted demonstrations are documented as short papers in the proceedings. Yiwei Cao et al. present Virtual Campfire, a mobile social software for cross-media communities. Helmut Neuschmied demonstrates a MPEG-7 based video annotation tool supporting time based annotation of content semantics. Martin Höffernig et al. show VAMP, a service for the validation of MPEG-7 documents w.r.t. the semantics of a profile. Peter Schallauer et al. present a system for the description of summarised visualisation of video quality information. Walter Allasia et al. demonstrate the PrestoSpace publication platform, a system for searching and retrieving broadcast archive content.

Acknowledgements

The workshop chairs would like to thank the TRIPLE-I organisation team for their support and the K-Space Network of Excellence[2] for sponsoring the best paper award.

[2] http://www.k-space.eu

Program Commitee

Susanne Boll, University of Oldenburg, Germany
Lionel Brunie, INSA de Lyon, France
Yiwei Cao, RWTH Aachen University, Germany
Vincent Charvillat, ENSEEIHT, France
Richard Chbeir, LE2I Laboratory, Bourgogne University, France
Stamatia Dasiopoulou, ITI Thessaloniki, Greece
Mario Döller, Passau University, Germany
Victor Manuel Garcia-Barrios, University of Technology Graz, Austria
Christian Gütl, University of Technology Graz, Austria
Martin Haller, TU Berlin, Germany
Michael Hausenblas, Joanneum Research, Austria
Werner Klieber, Know-Center Graz, Austria
Klaus Leopold, Klagenfurt University, Austria
Mathias Lux, Klagenfurt University, Austria
Vincent Oria, NJIT, USA
Vedran Sabol, Know-Center Graz, Austria
Ansgar Scherp, UC Irvine, CA, USA
Marc Spaniol, RWTH Aachen, Aachen, Germany
Peter Stanchev, Kettering University, USA
Georg Thallinger, Joanneum Research, Graz, Austria
Raphaël Troncy, CWI, The Netherlands
Christian Timmerer, Klagenfurt University, Austria
Martin Wolpers, KU Leuven, Belgium

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

An Ontological Approach to Semantic Video Analysis for Violence Identification

Thanassis Perperis
(University of Athens, Greece
a.perperis@di.uoa.gr)

Sofia Tsekeridou
(Athens Information Technology, Greece
sots@ait.edu.gr)

Sergios Theodoridis
(University of Athens, Greece
stheodor@di.uoa.gr)

Abstract: Along with the rapid increase of available multimedia data, comes the proliferation of objectionable content such as violence and pornography. We need efficient tools for automatically identifying, classifying and filtering out harmful or undesirable video content for the protection of sensitive user groups (e.g. children). In this paper we present a multimodal approach towards the identification and semantic analysis of violent content in video data. We propose a layered architecture and focus on ontological and knowledge engineering aspects of video analysis. We demonstrate the development of two ontologies defining violent hints hierarchy that low level analysis, in visual and audio modality, respectively should identify. A violence domain ontology, as a reality representation, defines higher-level semantics. Taking under consideration extracted violent hints, spatio-temporal relations and behaviour patterns higher-level semantics automatic inference is possible.

Key Words: Multimedia semantics, violence detection, ontology, inferencing
Category: I.2.4, H.5.1, I.2.10

1 Introduction

Along with the dissemination of internet technologies and digital television comes the exponential growth of available multimedia content. Almost everyone is able to provide content, accessible by large portions of the population, with limited or not at all central control. Common users and industry demand intelligent, human-like methods, to search, classify and filter this huge amount of data. Research towards this direction focuses in enabling machines to automatically extract the meaning, or semantics, of multimedia data and annotate them using MPEG-7, MPEG-21 or other metadata standards. MPEG-7 [Manjunath et al. 2002] defines metadata descriptors for structural and low level aspects of multimedia documents, as well as, high level Multimedia Description

Schemes (MDSs) encapsulating multimedia content semantics. Having the adequate standardised definitions for multimedia content description, we need efficient audio and visual event detectors, ontologies [Chandrasekaran et al. 1999] deploying Semantic Web Languages for high-level formal semantics representation, and tools to optimally handle their interoperation, in an as automatic way as possible.

When we though attempt to deploy Semantic Web trends and ontologies on multimedia data, which are complex in nature, multi-modal, of significant size, requiring extensive and efficient analysis to reduce the data space and extract representative features and descriptions, it is obvious that we are faced with a very challenging task. On the one hand, multimedia data are automatically processed to extract merely low- to medium-level descriptions, as the ones dictated by the Audio and Visual Parts of MPEG-7. The definition model of such descriptions is often referred to as Multimedia Ontology, tackling the data characteristics. On the other hand, ontologies represent domain knowledge in the form of high-level concepts, hierarchies and relations. The interoperation of Multimedia and Domain Ontologies, along with the optimal definition of the latter, to support automated semantic annotation of multimedia data is a major research focus [Hunter 2003] in various application domains.

In this paper, we propose an ontological and knowledge engineering approach for automatic identification and annotation of violent scenes in video data. Previous, mainly low level analysis based approaches are either focused in a limited set of violence actions (kicking, fist fighting, etc.), in order to ease the solution, or are able to only extract simple semantics (e.g. explosions, gunshots). Our approach by focusing on tackling an extensive range of complex violent acts in video data, based on violence domain knowledge representation, using ontologies and reasoning on or inferring from results obtained by multimodal analysis of both visual [Makris et al 2007] and audio [Giannakopoulos et al 2006] modalities, attempts to proceed further. The utmost goal is to automatically detect any violence hidden in video data, annotate them accordingly and enable filtering of content for parental control. A crucial step in the overall methodology is the best possible definition of the underlying ontologies. To optimally combine multimedia descriptions with the violence domain ontology, the knowledge representation process has involved the definition of modality violence ontologies (audio, visual) that essentially map low-level analysis results to simple violence events and objects (medium-level semantics), as well as a violence domain ontology that defines the hierarchy of violence concepts and inherent relationships, irrespective of data, starting from abstract and complex ones to more simple and concrete ones (shared by all ontologies). The latter is used as input to the inference engine that undertakes the fusion of results from medium level semantics to lead to higher level ones and to infer knowledge about existing violence.

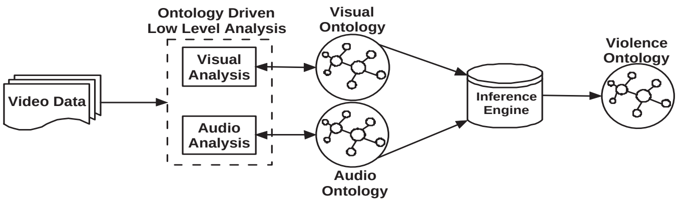


Figure 1: Conceptual Architecture

2 Ontological Methodology for Violence Identification in Video Data

Violence characterization is quite subjective and this creates difficulties in defining and identifying violent content unambiguously. We may define as violence any situation or action that may cause physical or mental harm to one or more persons - violence thus is a very abstract concept. Violent scenes in video data refer to scenes that include such actions. Thus, a very crucial step in violence identification in video is to represent the violence domain knowledge as effectively as possible, in all its complexity, abstractness and hierarchy depth. Previous research towards violence identification in video data is limited and in most cases examines only low level features to extract simple semantics. In [Vasconcelos, Lippman 1997] the design of a simple feature space, for scene categorization based on the degree of action (i.e. degree of violence), is presented. In [Data et al. 2002] the problem of detecting human violence, such as fist fighting, kicking, in video data captured using a stationary camera, is based on motion trajectory information and on orientation information of a person's limbs. Audio data for violence detection is used as an additional feature in TV drama and movies in [Nam, Tewfik 2002], where abrupt changes (i.e. explosions) in energy level of the audio signal are detected using the energy entropy criterion. A formal representation of the violence domain, to drive violent acts detection, has never been attempted before. We make a step forward towards this direction. Our overall goal is to devise a multimodal analysis, fusion and inferencing methodology towards automatic semantic violence extraction and annotation of video scenes, aiming further at content filtering and enabling parental control. The conceptual architecture of our overall methodology is shown in Fig. 1.

We identify as major processes of the system, a visual analysis step, an audio analysis step, each one interconnected with the corresponding data ontology, an inference engine and a domain (violence) ontology. In the following paragraphs, the definition of the different ontologies that drive the low-level processing and the high-level inferencing is described in detail.

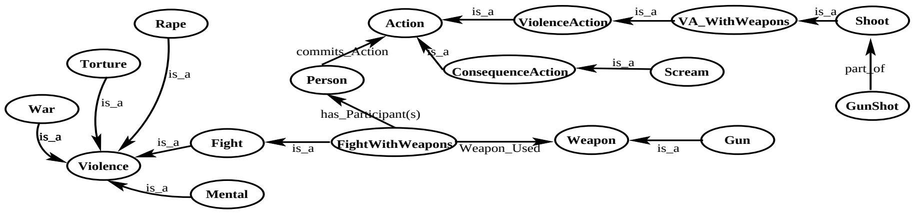


Figure 2: Violence Ontology

2.1 Violence Domain Ontology Definition

As we pinpointed previously a crucial step in violence identification in video data is the effective violence domain knowledge representation. We begin our analysis by presenting a violence domain ontology defining complex semantics of extensive violent acts, also found in movie data, as well as cross-modal relations of medium level semantics. This is the first attempt to conceptualize violence in an organized way. Our violence domain definition appears as a composition of psychologists' view of violence and violent acts and extended investigation through observation of violent acts in video data. Thus the violence domain ontology as a knowledge representation can be further exploited by researchers and organizations investigating violence (i.e. psychologists, pedagogists, police).

Although our ontology comprises a generic representation of violence we will focus our analysis in the movie violence domain. In a movie scene containing violence (e.g. torture, fight, war) a spectator can quickly grasp the form of violence (e.g. fighting without weapons), recognize a sequence of violent (e.g. punching, kicking), of generic (e.g. running, walking) and of consequence (e.g. falling, crawling, scream) actions. The direct application of this process demonstrates how the hierarchy (taxonomy) of violent actions, along with their inter-relations (e.g. a punch is followed by a scream), is constructed, formulating the violence domain ontology. The presented movie violence ontology is implemented in OWL-DL [Smith et al. 2004] using Protégé. In Fig. 2 we demonstrate the higher level concepts of the violence ontology (left part) and an instantiation of the ontology (right part) for the violent action "shoot", demonstrating the way the concept is related with semantically higher (violence) and lower (gun, person) level concepts. The medium level classes (actions) of the ontology are strongly related with the inference engine, since they represent multimodal actions inferred using reasoning by relating to the visual and audio ontology and the single modality analysis and classification results. Furthermore the simplest, more concrete, concepts (e.g. weapon, scream) in the violence domain ontology are further represented along with their low level features in the visual and audio ontologies respectively, thus defining the association mechanisms of the three ontologies.

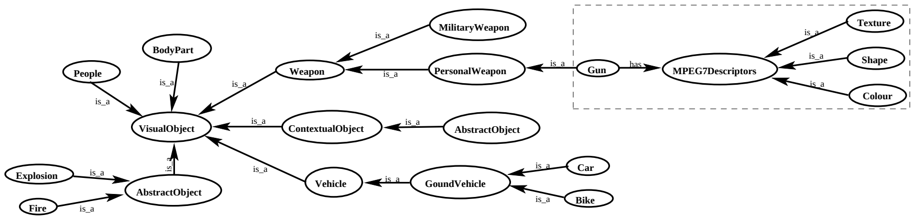


Figure 3: Visual Ontology

2.2 Visual Ontology Definition for Violence

Every action described in the violence domain ontology is usually composed of some primitive actions or events (either visual or auditory) and include a set of objects (e.g. guns, swords). The visual ontology (Fig. 3) for violence defines a taxonomy of moving objects (e.g. people, weapons, body parts, military vehicles), stationary (contextual) objects (e.g. walls, fences, furniture), abstract objects (e.g. explosions, fire, injuries) and primitive actions (e.g. crawling, running, falling). We note that the detection of some of the aforementioned concepts does not directly imply violence (e.g. bottle), but in the context of violence (e.g. hit on the head with bottle) its identification might be very important. The visual ontology further includes the MPEG-7 visual descriptors and MPEG-7 MDSs, which describe visual features such as colour, texture, shape and motion and semantic information of video respectively, associated with the above mentioned taxonomy entries. Furthermore it specifies generic event/concept detetectors' input parameters in order to extract the specified objects and actions from the video data, along with their low level features. Thus the identified concepts and the corresponding features are instantiated based on the ontology. Following the previously reported example of gunshot recognition, in the marked area of Fig. 3 we demonstrate the description instantiation of the visual object "gun" from the low level processing and classification algorithms. This example further demonstrates the linking between the violence and the visual ontology through common terms, from a different viewpoint.

2.3 Audio Ontology Definition for Violence

In violent scenes one can recognize a set of audio events indicative of violence like gunshots, screams, explosions, hit sounds. Moreover indication of violence can be drawn from the background music (e.g. action scenes with gunshots usually are accompanied with intense music) or the emotional speech of an actor (e.g. angry speech might be followed by some sort of fight). Thus, we have further implemented the aforementioned audio events (Fig. 4) in a taxonomic

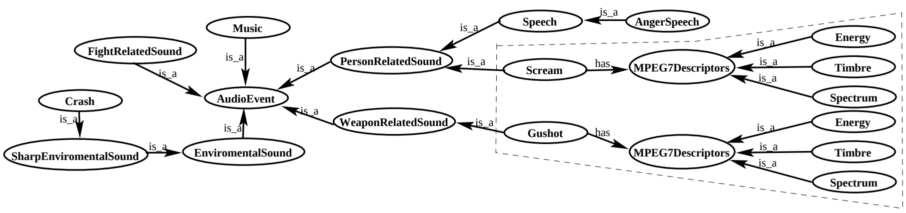


Figure 4: Audio Ontology

way, defining the audio ontology for violence. The ontology is extended with MPEG-7 audio descriptors and MPEG-7 MDSs, to specify low level features and semantically describe the audio data respectively. A set of classification algorithms [Giannakopoulos et al 2006] (e.g. bayesian networks, SVMs) is responsible for instantiating audio descriptions in compliance with the ontology including the corresponding sound segments, their categorization and the values of their low level features (e.g. spectrum, timbre, energy). In the marked area of Fig. 4 we demonstrate the instantiation potentials of "gunshot" and "scream" following the aforementioned example. We note that the high level actions (gunshot, scream) are also represented both in the violence domain and the audio ontology. Thus, as in the case of the visual ontology, the association between the two ontologies is evident.

2.4 Inference Engine Design

Previously we have demonstrated the development of the three ontologies, the ontology driven low level(audio, visual) analysis and the linking of the three ontologies through common terms. At this step we can extract medium level semantics from the single modality analysis, but in order to infer higher level semantics [Pratikakis, Tsekeridou 2005] (e.g. more complex, abstract and extensive violence cases) represented in the violence domain ontology we must take under consideration spatio-temporal relations and behaviour patterns. Let's describe this with the familiar example of "shoot". The "shoot" concept can be automatically extracted based on the initial analysis results and on the sequence or synchronicity of audio or visual automatically detected events such as two persons in visual data, the one holding a gun, while gunshot sound and scream of pain is detected in the audio data. To fulfil such scenarios as the one presented above, we should solve the problem how to fuse and interchange semantics from different modalities.

To support reasoning mechanisms, it is required that apart from the ontological descriptions for each modality, there is a need for a cross-modality ontological description which interconnects all possible relations from each modality and constructs rules that are cross-modality specific. It is not clear, whether this can be achieved by an upper multimedia ontology or a new cross-modality ontology that will strive toward the knowledge representation of all possibilities combining media. It is evident though, that the cross-modality ontology, along with the single modality (audio, visual) ones, greatly relate to the domain (violence) ontology, i.e. to the application at hand.

Furthermore, in this new cross-modality ontology, special attention should be taken for the representation of the priorities/ordering among modalities for any multimodal concept. This translates to sequential rules construction. However there are cases, where simultaneous semantic instances in different modalities may lead to higher level of semantics, that synchronicity is also a relationship to be accounted for. Apart from the spatial, temporal or spatio-temporal relationships that need to be accounted for, there is also the issue of importance of each modality for identifying a concept or semantic event, to solve possible contradictions between unimodal results. This may be represented by means of significance weights. Finally reasoning using partial, imprecise information should be supported in order to tackle uncertainty of extracted medium level semantics and support real time interoperation with data ontologies and concept/event detectors (i.e. in order to identify the aforementioned concept "shoot", if the "gun" is detected, the inference engine must initiate the "gunshot" and "scream" concept detectors through the audio ontology to fullfill the recognition proccess).

3 Experimental Data Setup

We have collected a corpus of 10 movies in MPEG-4 format, containing a variety of violent scenes, composed of both auditory and visual clues. We are in the process of producing manual annotations to form the essential ground truth, as MPEG-7 description instances, based on the violence terms and concepts existent in all defined ontologies. This ground truth data will be used by all processes involved, semantic audio analysis and violent events identification, semantic visual analysis and violent events identification, as well as late fusion methodology and inferencing for complex violent events identification, in order to assess their performance and identification accuracy.

4 Conclusions and Future Work

We have proposed an ontological and knowledge representation approach to define the underlying semantics for violence characterization in video data. This is the first step before providing the inferencing mechanisms in order to automatically identify violent scenes and their context in video data. Thus, this work has to further tackle the question: *how to fuse and interchange semantics from* *different modalities?*. We are in the process of exploring the usage of basic probabilistic inference methods (Bayesian/belief networks, HMMs), probabilistic reasoning (probabilistic logic, PR-OWL) rule construction and the development of low level analysis procedures ([Giannakopoulos et al 2006, Makris et al 2007]). Furthermore we intend to subsequently apply a similar approach to identify and filter out pornographic content.

Acknowledgements

This paper is part of the 03ED375 research project, implemented within the framework of the Reinforcement Programme of Human Research Manpower (PENED) and co-financed by National and Community Funds (75% from E.U.-European Social Fund and 25Greek Ministry of Development-General Secretariat of Research and Technology).

References

[Chandrasekaran et al. 1999] Chandrasekaran, B., Josephson J., Richard Benjamins, V.: "What Are Ontologies, and Why Do We Need Them?"; IEEE Intelligent Systems, 14, 1 (1999), 20 - 26.

[Manjunath et al. 2002] Manjunath, B.S., Salembier, P., Sikora, T.: "Introduction to MPEG-7: Multimedia Content Description Interface"; John Wiley and Sons / England (2002).

[Data et al. 2002] Datta, A., Mubarak, S., Lobo N.V.: "Person-on-Person Violence Detection in Video Data"; Proc. of ICPR2002, Quebec City, Canada, Aug. (2002), 433-438.

[Smith et al. 2004] Smith, M.K., Welty, C., McGuinness, D.L.: "OWL Web Ontology Language Guide"; W3C Recommendation 10 February 2004, http://www.w3.org/TR/owl-guide/.

[Hunter 2003] Hunter J.: "Enhancing the Semantic Interoperability of Multimedia through a Core Ontology"; IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Conceptual and Dynamical Aspects of Multimedia Content Description, 13, 1 (2003), 49-58.

[Nam, Tewfik 2002] Nam, J., Tewfik, A.H.: "Event-driven video abstraction and visualisation"; Multimedia Tools and Applications, 16(1-2), 55-77, 2002.

[Vasconcelos, Lippman 1997] Vasconcelos, N.; Lippman, A., "Towards semantically meaningful feature spaces for the characterization of video content"; Proc. of ICIP1997., Washington, DC, USA, Oct 1997, vol.1, 25 - 28.

[Giannakopoulos et al 2006] Giannakopoulos T., Kosmopoulos D., Aristidou A., Theodoridis S.: "Violence Content Classification Using Audio Features"; Proc, 4th Hellenic Conference on Artificial Intelligence (SETN'06), Heraklion, Crete, Greece, May 18 - 20, 2006.

[Pratikakis, Tsekeridou 2005] Pratikakis I., Tsekeridou S.: "Use Case : Semantic Media Analysis for Intelligent Retrieval"; W3C Multimedia Semantics Incubator Group , http://www.w3.org/2005/Incubator/mmsem/wiki/Semantic_Media_Retrieval_Use_case.

[Makris et al 2007] Makris A., Kosmopoulos D., Perantonis S., Theodoridis S.: "Hierarchical feature fusion for visual tracking"; Accepted to be published in Proceedings of IEEE International Conference on Image Processing 2007 (ICIP2007).

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

A Similarity Approach on Searching for Digital Rights

Walter Allasia
(EURIX, Torino - Italy
allasia@eurixgroup.com)

Fabrizio Falchi
(ISTI-CNR, Pisa - Italy
fabrizio.falchi@isti.cnr.it)

Francesco Gallo
(EURIX, Torino - Italy
gallo@eurixgroup.com)

Abstract: We present an innovative approach that treats the right management metadata as metric objects, enabling similarity search on IPR attributes between digital items. We show how the content base similarity search can help both the user to deal with a huge amount of similar items with different licenses and the content providers to detect fake copies or illegal uses. Our aim is the management of the metadata related to the Digital Rights in centralized systems or networks with indexing capabilities for both text and similarity searches, providing the basic infrastructure enabling the private use and the commercial exploitation as well.
Key Words: rights, information retrieval, multimedia information systems
Category: H.3.3, H.5.1, K.5.1

1 Introduction

Nowadays we are dealing with several devices able to consume and produce lots of digital items, exponentially growing. People from all over the world are creating their own digital contents like images and audio/video files, sharing them by means of electronic mail, Web sites, chats, multimedia messaging services and several distributed systems. The digital contents are mostly provided by the personal use of high tech devices. Our cultural heritage is no longer made up only of videos, images and text documents provided by "institutional" public or private bodies but also of the digital contents provided by every connected digital device as well.

In order to be able to guarantee the preservation and access of these digital items, we have to take into account their Digital Rights management during the creation phase and especially during the search of something provided by someone else. Several approaches have been proposed so far for managing Digital Rights and many standards are available for representing them, but usually open as well as trusted systems provide a simple attribute search on a single specific type of license.

In this paper, we propose a different approach for indexing and searching the information related to the licenses of the digital items, towards a more flexible and open network infrastructure.

2 Backgrounds

2.1 Digital Rights Management

The European project *Networked Audiovisual Systems and Home Platforms* produced an important report [NAVSHP 2005], describing a set of requirements to be satisfied by whatever DRM system. Some of these requirements are harshly criticized, mainly those concerning the analogy with the contract laws and the development and use of free and open source content [Doctorow 2006]. At the same time many initiatives, such as DMP[1], Chillout[2] and MediaLive[3] are trying to provide the basis for a DRM infrastructure. Recently some standards on the expression of the license have been proposed and started to be commonly used, but it is still very difficult today to deal with DRM systems in such an heterogeneous environment as the Web, since at the moment we are very far from having a common agreement on the adoption of a standard DRM system. In this paper we are dealing with the definition of the license and in particular with the search capability of a digital item by its license. We do not want to provide any guideline nor implementation of the software for controlling the respect of use of the license as a DRM system has to guarantee. We want to provide an innovative approach for managing the license attributes in order to be able to search for a similar digital object as query with its related digital license. Many solutions have been adopted so far in the Web for expressing a license which are widely used, such as AdobeContentManager[4], CreativeCommons[5], MPEG-21 REL [MPEG-21 REL 2005], ODRL [Iannella 2002] and PRISM[6].

Since the digital items available on the Web are mainly audio/video files produced by the personal use of digital devices, we focus on the language for expressing the license that are widely adopted for defining this kind of items. We consider as an example three different license formats: CreativeCommons, MPEG-21 REL and ODRL (see Table 1 from [Roberto García González 2006]). Unfortunately the metadata for expressing the license in the three formats described above are quite different from each other and a mapping is required in order to process them in the same query. An example of the mapping for the group of metadata referring to the *Use-type Rights* group is shown in Table 1.

[1] http://www.dmpf.org/
[2] http://chillout.dmpf.org/
[3] http://www.medialive.com/
[4] http://www.adobe.com/products/contentserver/
[5] http://creativecommons.org/
[6] http://www.prismstandard.org/

Use-Type Rights		
CreativeCommons	**MPEG-21 REL**	**ODRL**
reproduction		
		display
	execute	execute
	play	play
	print	print

Table 1: Use-Type Rights group mapping

2.2 Distributed Systems

Many network infrastructures are arising in order to provide the bases for Web sharing and searching functionalities on digital items. Most of them are peer-oriented networks, such as eMule[7] or BitTorrent[8] for images and audio/video files and Joost[9] for video streaming. Furthermore several multimedia platforms enable the automatic audio/video processing for the cataloging and the indexing of digital items [Messina et al. 2006] and in combination with the network infrastructure will provide powerful solutions for digital content management.

2.3 The Metric Space Approach

Although many similarity search approaches have been proposed, the most generic one considers the mathematical metric space as a suitable abstraction of similarity [Zezula et al. 2006]. The simple but powerful concept of the metric space consists of a domain of objects and a distance function that measures the proximity of pairs of objects.

In the metric space $M = (\mathcal{D}, d)$ defined over a domain of objects $\mathcal{D}$ with a total (distance) function $d : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$, the following properties hold $\forall x, y \in \mathcal{D}$:

$$
\begin{array}{ll}
d(x,y) \geq 0 & (\textit{non-negativity}), \\
d(x,y) = 0 \text{ iff } x = y & (\textit{identity}), \\
d(x,y) = d(y,x) & (\textit{symmetry}), \\
d(x,z) \leq d(x,y) + d(y,z) & (\textit{triangle inequality}).
\end{array}
$$

The metric space approach has been proved to be very important for building efficient indexes for similarity searching. A survey of existing approaches for centralized structures can be found in [Zezula et al. 2006] and [Samet 2006]. Two examples of well known centralized structure for indexing metric objects are M-tree [Ciaccia et al. 1997] and D-Index [Dohnal et al. 2003].

[7] http://www.emule.org
[8] http://www.bittorrent.com
[9] http://www.joost.com

Very recently scalable and distributed index structures based on Peer-to-Peer networks have also been proposed for similarity searching in metric spaces, i.e. GHT* [Batko et al. 2004], VPT*, MCAN [Falchi et al. 2005] and M-Chord [Novak and Zezula 2006] (see [Batko et al. 2006] for a comparison of their performances). Currently many research projects are investigating these fields, such as SAPIR[10], a project funded by European Research Area in the 6^{th} Framework Program, that aims to develop cutting-edge technology that will break the barriers and enable search engines to look for large scale audio-visual information by content, using the query by example paradigm. SAPIR intends to propose new solutions for an innovative technological infrastructure for next-generation Multimedia Search Engines. This research effort should lead towards a distributed, P2P based, search engine architecture, as opposed to today parallel search engines within a centralized Web data warehouse.

3 Metric Distance Example for Licenses

We now illustrate an example of a metric distance defined over the IPR information. The main common groups of the expression languages of the licenses can be identified as [Roberto García González 2006]: *Agent Data Element*, *Manage-type Rights*, *Reuse-type Rights*, *Transfer-type Rights*, *Use-type Rights*, *User Constraint*, *Device Constraint*, *Limits Constraint*, *Temporal Constraint*, *Aspect Constraint*, *Target Constraint*, *Payment Constraint*, *Usage Conditions*.

Let $\mathcal{D}$ be the domain of metadata related to the license of any given object. For any $x \in \mathcal{D}$ we define $x_1, x_2, ..., x_n$ as the n main groups and $x_{i,1}, x_{i,2}, ..., x_{i,n}$ as the n_i attributes for the i-th main group. The global distance is defined as the weighted sum of the distances between the main groups, i.e.

$$d(x, y) = \sum_{i=1}^{n} w_i \cdot d_i(x_i, y_i) \ . \tag{1}$$

The distance between the same groups of two distinct licenses can be defined as:

$$d_i(x_i, y_i) = \sum_{j=1}^{n_i} w_{i,j} \cdot d_{i,j}(x_{i,j}, y_{i,j}) \ . \tag{2}$$

The distance $d_i(x_i, y_i)$ between two values $x_{i,j}$ and $y_{i,j}$ of the j-th attribute of the group i must be defined considering the specific attribute type. In case $x_{i,j}, y_{i,j} \in 0, 1$ we can use the distance $|x_{i,j} - y_{i,j}|$. Please note that a specific weight to this distance can be given by setting $w_{i,j}$.

The same distance can be used whenever $x_{i,j}, y_{i,j} \in \mathbb{R}$. However, more sophisticated metric distances could be used for specific numerical attributes. As

[10] http://sysrun.haifa.il.ibm.com/sapir/index.html

	$term_1$	$term_2$	$term_3$	...	$term_m$
$term_1$	0	$\alpha^{i,j}_{2,1}$	$\alpha^{i,j}_{3,1}$	...	$\alpha^{i,j}_{m,3}$
$term_2$	$\alpha^{i,j}_{2,1}$	0		...	$\alpha^{i,j}_{m,2}$
$term_3$	$\alpha^{i,j}_{3,1}$	$\alpha^{i,j}_{3,2}$	0	...	$\alpha^{i,j}_{m,3}$
...	...	...	...	...	...
$term_m$	$\alpha^{i,j}_{m,1}$	$\alpha^{i,j}_{m,2}$	$\alpha^{i,j}_{m,3}$	...	0

	$Notice$	$Attr$	SA	SC
$Notice$	0	0.3	0.6	1
$Attr$	0.3	0	0.3	0.7
SA	0.6	0.3	0	0.4
SC	1	0.7	0.4	0

Table 2: Distance values for attributes taken from terms in a given dictionary (*left*) and proposed values for *CreativeCommons* terms for the *Requirements* (*right*)

an example, for fees we suggest to define the distance as:

$$d_{i,j}(x_{i,j}, y_{i,j}) = |\log(x_{i,j}) - \log(y_{i,j})| = \left|\log\left(\frac{x_{i,j}}{y_{i,j}}\right)\right|,$$

since given a fee as query, the user would be much more interested on the proportion between its query and a given fee. Unfortunately any non 0 fees would be at infinite distance from 0 objects. To avoid this problem we suggest that whenever the fee is 0 the value used for evaluating the distance is 0.01. Please note that the well known distance gap ratio ($(x - y)/y$) would avoid the 0 fee problem but it is not a metric.

For an attribute whose value can be a term in a given vocabulary, we propose a specific approach. If the j-th attribute of the i-th group is a term taken from a specific vocabulary of m terms, we can define the distance $d_{i,j}(x_{i,j}, y_{i,j})$ between the two values according to what reported in Table 2 (left).

It is assumed that the values of α are manually chosen according to the semantic of the given terms. In particular, if all $\alpha^{i,j}_{a,b} = 1$ when $a \neq b$, textual attributes are considered as binary attributes. For $d_{i,j}(x_{i,j}, y_{i,j})$ to be a metric the matrix must be symmetric and all the diagonal values must be 0 and

$$\forall l\ ,\ \alpha^{i,j}_{a,b} \leq \alpha^{i,j}_{a,l} + \alpha^{i,j}_{l,b} = \alpha^{i,j}_{l,a} + \alpha^{i,j}_{l,b}\ .$$

Let x and y be the metadata about the license attributes. Considering for example the CreativeCommons schema for the *Requirements* [11] (the restrictions imposed by the license), we can assign a set of values to the $\alpha^{i,j}_{m,n}$ terms as shown in Table 2 (right), where: *Notice* requires that the copyright and license notices must be kept intact; *Attr* stands for Attribution and requires that credit must be given to copyright holder and/or author; *SA* stands for ShareAlike and requires that derivative works must be licensed under the same terms as the original work; *SC* stands for SourceCode and requires that source code (the preferred form for making modifications) must be provided for all derivative works.

[11] http://creativecommons.org/technology/metadata/implement

Using the *triangle inequality* reported above, it can be shown that if all the distances defined for the attributes in a given group are metric, the proposed distance for the given group is still a metric. Defining the global distance between two license as the weighted sum between the main groups, this global distance is still a metric one. For indexing the licenses metadata using the global distance d in a single index for similarity searching in metric spaces, all the weights w should be fixed in advance. However, if we want to specify at query time the weights w_i for the single groups to be used for searching, we can use distinct indexes for each d_i and then combine the results coming from the various indexes using optimal aggregation algorithms as the ones described in [Fagin 1996]. Moreover, in this case we do not need the global distance function to be metric, but just all the d_i. In this case the aggregation must be monotone. Thus, using separate indexes for each d_i and then combining them using the algorithms described in [Fagin 1996], more aggregation functions could be used and they could even be specified at search time. Obviously there is a price to be paid for that: efficiency. A single global metric distance function can be more efficiently indexed using a single index structure for metric spaces. An extension of the proposed global distance which is still metric is a sort of Minkowski Distance combination:

$$d(x,y) = \sqrt[k]{\sum_{i=1}^{n} w_i \cdot |d_i(x_i,y_i)|^k} \tag{3}$$

The same approach could be used for combining the distance values among the attributes of the same group.

4 Significant Use Cases for Photo Search

Most of the search engines available on the Web provide nothing but the "full text" and/or "attribute" search capabilities. However, many research projects are developing audio and image "similarity" search. According to our proposal, a user will be able to search for an image similar to the one provided considering both the multimedia content (content base) and the related license (provided by the user as well). Furthermore the user can apply for searching similar images regarding the multimedia content and a specific kind of license defined by mean of attributes. Since the user can search for content-based similarity and license similarity independently, we are now focusing on scenarios where they are combined. Two important combination scenarios are:

1. The user is searching for images similar to a given one both considering its visual appearance and license "file"

2. The user is searching for images similar to a given one but with a license similar to a different one (informations from other images can be provided)

In the first case, the user is interested in images similar to a given one both considering its content and its license. This is the typical case in which the user has an image which satisfied his needs both in terms of content and license. The search engine will display as result the ranked list of images similar to that provided according to the content and to the license. In the second case the user has an image which he does like, but that has a license which does not satisfy its needs. The user can search for an image similar to the given one but with a different license. In this case the license the results are requested to be similar to a license that can be either taken from another image or specified using a form.

A special case of this second scenario is searching for copyright violation. Imagine a professional photographers agency that wants to be sure that nobody is making a fake use of their own pictures and/or non authorized use of the associated copyrights. The agency can query the system providing the picture to be searched and can provide the attributes for an open license or something "similar" to an open one. If the system will find a result, it means either that someone has made the same picture or that someone is sharing a non authorized copy of the picture. This use case is innovative because the current search engines are focused on the content sharing and are not addressed to the "control" of the contents themselves, delegating this feature entirely to the DRM systems.

5 Conclusions

We have proposed an innovative approach for managing the attributes and metadata referred to the expression language adopted for describing a license for Digital Rights. The metadata shown are taken as examples and should be changed to fit the needs of the software infrastructure the user has to deal with. This approach considers the IPR attributes as *special features* which a specific distance function can be applied to. For efficiently indexing the data it is important that this distance is a metric.

The Right Management warrantee has been deeply studied in the last few years and lots of solutions are available. However not much has been done concerning the "retrieval" of the license associated to the digital items. Since many standards are available, we will reasonably have many types of license and once we have to deal with thousands of items, the attribute search over the licenses could be not enough to handle the problem. We propose the adoption of the *Similarity Search* for the IPR attributes. In this way the license we are looking for can be easily provided, instead of all the attributes of a specific license format in a complex GUI. Moreover, we can also have a ranked list of results, according to the metric function, by defining the distance between the licenses.

Acknowledgments

This work was partially supported by the SAPIR (Search In Audio Visual Content Using Peer-to-Peer IR) project, funded by the European Commission under IST FP6 (Sixth Framework Programme, Contract no. 45128).

References

[Messina et al. 2006] A. Messina, L. Boch, G. Dimino, W. Bailer, P. Schallauer, W. Allasia, M. Groppo and M. Vigilante: "Creating Rich Metadata in the TV Broadcast Archives Environment: The PrestoSpace Project"; Proc. Axmedis 2006, IEEE Computer Society (2006), 193-200.

[NAVSHP 2005] Networked Audiovisual Systems and Home Platforms Group: "NAVSHP (FP6) DRM Requirements Report". Technical Report, European Community 6^{th} Framework Programme (Sep 2005).

[Doctorow 2006] Cory Doctorow, European Affairs Coordinator: "Critique of NAVSHP (FP6) DRM Requirements Report"; Technical report, Electronic Frontier Foundation (2006).

[MPEG-21 REL 2005] ISO/IEC - Information Technology - Multimedia Framework (MPEG-21), 21000-5 (2005).

[Iannella 2002] R. Iannella: "Open digital rights language (ODRL)"; Version 1.1, World Wide Web Consortium, W3C Note (2002).

[Roberto García González 2006] Roberto García González: "A Semantic Web Approach To Digital Rights Management"; PhD Thesis, Department of Technologies, Universitat Pompeu Fabra, Barcelona, Spain (Apr 2004).

[Zezula et al. 2006] P. Zezula, G. Amato, V. Dohnal and M. Batko: "Similarity Search. The Metric Space Approach"; Volume 32 of Advances in Database Systems. Springer, Heidelberg / New York (2006).

[Ciaccia et al. 1997] P. Ciaccia, M. Patella and P. Zezula: "M-tree: an efficient method for similarity search in metric spaces"; Proc. VLDB '97: 23^{rd}, Morgan Kaufmann, Publishers Inc. (1997), 426-435.

[Dohnal et al. 2003] V. Dohnal, C. Gennaro, P. Savino and P. Zezula: "D-index: Distance searching index for metric data sets"; Multimedia Tools Appl., 21, 1 (2003), 9-33.

[Samet 2006] H. Samet: "Foundations of Multidimensional and Metric Data Structures"; Computer Graphics and Geometric Modeling. Morgan Kaufman Publishers Inc. (2006).

[Batko et al. 2004] M. Batko, C. Gennaro and P. Zezula: "Similarity grid for searching in metric spaces"; In Peer-to-Peer, Grid and Service-Orientation in Digital Library Architecture. 6^{th} Thematic Workshop of the EU Network of Excellence DELOS. LNCS Springer, 3664 (2004), 25-44 .

[Falchi et al. 2005] F. Falchi, C. Gennaro and P. Zezula: "A content-addressable network for similarity search in metric spaces"; Proc. DBISP2P '05, LNCS Springer, 4125 (2005), 98-110.

[Novak and Zezula 2006] D. Novak and P. Zezula: "M-chord: a scalable distributed similarity search structure". Proc. Infoscale'06: 1^{st}, ACM Press (2006), 19.

[Batko et al. 2006] M. Batko, D. Novak, F. Falchi and P. Zezula: "On scalability of the similarity search in world of peers"; Proc. Infoscale '06, ACM Press (2006), 20.

[Fagin 1996] R. Fagin: "Combining Fuzzy Information from Multiple Systems"; Proc. Fifteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, ACM Press (1996), 216-226.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Semantics of Temporal Media Content Descriptions

Martin Höffernig, Michael Hausenblas, Werner Bailer
(JOANNEUM RESEARCH Forschungsgesellschaft mbH
Institute of Information Systems and Information Management
Steyrergasse 17, 8010 Graz, Austria
{firstname.lastname}@joanneum.at)

Abstract: The temporal structure of multimedia content is an important aspect of the description of time-based media and needed in many applications. Expressive content description languages, such as MPEG-7, provide tools for describing the temporal decomposition of content into segments. Although the semantics of temporal decomposition are apparent, the validation of the semantics of the temporal decompositions (e.g. temporal extent of child segments, gaps, overlaps) is not possible on the syntactic level. We propose therefore to model the semantics of temporal decompositions using ontologies and rules. As a proof of concept we apply the formalisation to a validation use case, implemented as a Web application.

Key Words: Metadata, Ontology, Semantic Web, MPEG-7, Temporal Decomposition, Time Interval, Validation
Category: H.5.1

1 Introduction

The description of multimedia content is of growing importance in a number of applications dealing with multimedia content creation, processing and archiving. Media content descriptions can be on a global scope (i.e. describing only metadata related to a complete media item, such as title and production information) or related to spatial, temporal and spatiotemporal segments of the content. An important aspect of a detailed content description of time-based media is the description of the temporal structure of the content, i.e. its decomposition into formal and logical units, such as e.g. shots, scenes or speech segments. There are several use cases where semantic annotations on these segments are relevant.

Imagine a YouTube[1]-like service, which provides summaries of the videos in the database in order to facilitate browsing. To produce such summarisation clips automatically, annotations on temporal segments of the source video are used in order to determine the relevant snippets that are put into the summary. This requires combining semantic descriptions of the content and the temporal structure of the source material.

Now imagine that we use the videos we have found with the help of the summaries to edit new content, like in classic post-production or in a Web 2.0

[1] `http://www.youtube.com`

application such as jumpcut[2]. Instead of just getting the final video as output, it would be great to also get a metadata description of the output. This requires metadata editing, i.e. automatically applying the edit decisions taken on the audiovisual material to the related metadata. There are annotations for each of the segments in each of the source contents and the edit operations creates a new segmentation. The task is to identify which metadata from the source applies to which segments of the target content, and whether there are potentially conflicting descriptions from the two source contents.

Last but not least, one can think of the automatic, semantic validation of temporal decomposition. Systems may produce descriptions of media assets that conform to a certain standard, such as MPEG-7, on a syntactic level—but how about the semantics? We aim at answering this question in this paper by formalising the semantics of temporal decompositions of media content descriptions.

1.1 Existing Work

The description of the temporal structure of the content is one of the most important aspects of a detailed content description of time-based media, and in particular a strength of MPEG-7 over other multimedia content description standards. The flexibility of MPEG-7 is based on allowing descriptions to be associated with arbitrary multimedia segments or regions, at any granularity, using different levels of abstraction. The downside of the breadth targeted by MPEG-7 is its complexity and its fuzziness [Bailer and Schallauer 2006, Ossenbruggen et al. 2004]. For example, very different syntactic variations may be used in multimedia descriptions with the same intended semantics, while remaining valid MPEG-7 descriptions. To reduce this syntax variability, MPEG-7 has introduced the notion of *profiles* that constrain the way multimedia descriptions should be represented for particular applications. Profiles are therefore a way of reducing the complexity of MPEG-7 (i.e. only a subset of the whole standard can be used) and of solving some interoperability issues (i.e. English guidelines are provided on how the descriptors should be used and combined). However, these additional constraints are only represented with XML Schema[3], and, for most of them, cannot be automatically checked for consistency by XML processing tools. In other words, profiles provide only very limited control over the semantics of the MPEG-7 descriptions [Hunter 2001, Nack et al. 2005]. Because of this lack of formal semantics, the resulting interoperability problems prevent an effective use of MPEG-7 as a language for describing multimedia.

In [Troncy et al. 2006] the authors present an approach to formalise a subset of the semantic constraints of the Detailed Audiovisual Profile (DAVP)[4].

[2] `http://www.jumpcut.com`
[3] `http://www.w3.org/XML/Schema`
[4] `http://mpeg-7.joanneum.at`

The formalisation of the semantic constraints can be used to automatically validate *semantically* the conformance of MPEG-7 descriptions to a given profile [Troncy et al. 2007].

In this work we do not focus on the semantics of such a temporal segment in terms of the type of unit it represents (e.g. shot, scene), as this is already modeled in the ontology described in [Troncy et al. 2006]. We concentrate on the semantics of the temporal segmentation [Allen and Ferguson 1994]. A temporal decomposition of a segment is a container for a set of segments, thus defining parent-child relations between the segment to be decomposed and the segments in the set. The temporal extent of a segment is specified by its time point and duration elements, which are pattern-restricted strings in MPEG-7. In addition, attributes of the temporal decomposition specify, whether overlaps of segments or gaps between them are allowed.

1.2 Problem Formulation

The semantics of the temporal decomposition are clearly defined. However, due to the limitations of XML Schema, documents containing one of the following two violations of temporal decomposition semantics are still valid w.r.t. to the profile schema:

Invalid parent-child segment relation A temporal decomposition of a segment into subsegments is only meaningful if the time range filled by each of the subsegments is at most the time range of the segment being decomposed, i.e. a part of a temporal segment cannot start before or end after its parent segment.

Gap and overlap A temporal decomposition can be qualified whether the subsegments in the decomposition overlap or have gaps between them. These properties are specified with the `gap` and `overlap` attributes of the decomposition that have a `true`/`false` value. There is, however, no mechanism to check whether the actual time description of the segments conforms to the value of the attribute or not.

An example of a temporal decomposition of a segment is shown in Figure 1. Segment S_1 is decomposed into Segments S_2, S_3 and S_4. For example S_1 has start point t_1 and a duration d_1. This temporal decomposition contains three gaps (between t_1 and t_2, between t_3 and t_4, and between t_7 and t_8) and one overlap between t_5 and t_6.

Our approach is to model the semantics of temporal decompositions using Semantic Web languages to formalise the semantics, and later inference tools to check the semantic consistency of the segments. Section 2 describes the approach we are proposing and Section 3 its implementation and integration into the

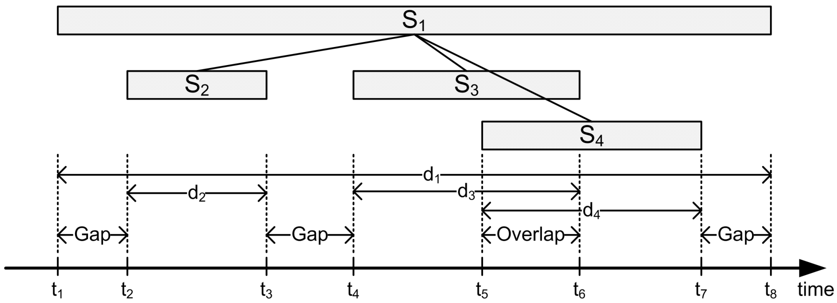


Figure 1: Temporal decomposition of segment S_1 into three segments (S_2, S_3, S_4) with gaps and overlap.

validation service. In Section 4 we conclude the discussion and outline future research.

2 Formal Representation of Temporal Media Descriptions

An ontology is used for the formal representation of temporal segments. A temporal segment is described by a start point and a duration. The ontology contains classes and properties for describing the temporal behavior of a temporal segment and the relations between these temporal segments. Hence the ontology models (i) the time interval of a temporal segment with start point and duration, (ii) the parent-child relation between temporal segments and (iii) the temporal decomposition attributes of temporal segments (overlap and gap).

2.1 An Ontology for Temporal Segments

Several classes and properties are needed to model the required relationships:

Class `Segment` This is the main class in the ontology. Every temporal segment is an instance of class Segment. Every instance of this class has exactly one `hasStartPoint` relation and exactly one `hasDuration` relation.

Class `ParentSegment` This class describes all temporal segments that are decomposed into further temporal segments (using `hasChild`, `hasAssertedGap` and `hasAssertedOverlap`. This class is a subclass of class `Segment`.

The exemplary temporal decomposition in Figure 1 is partially represented as an ontology in Figure 2[5].

---

[5]
```
prefix tsmd: http://mpeg-7.joanneum.at/semantics/temporal#
prefix ex: http://mpeg-7.joanneum.at/semantics/example#
prefix rdfs: http://www.w3.org/2000/01/rdf-schema#
prefix rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#
```

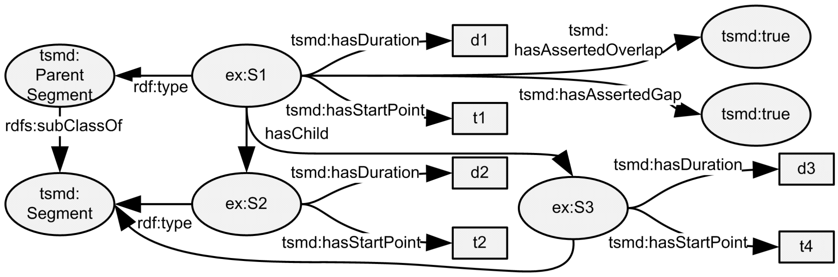


Figure 2: Excerpt of the ontology for describing the temporal decomposition shown in Figure 1.

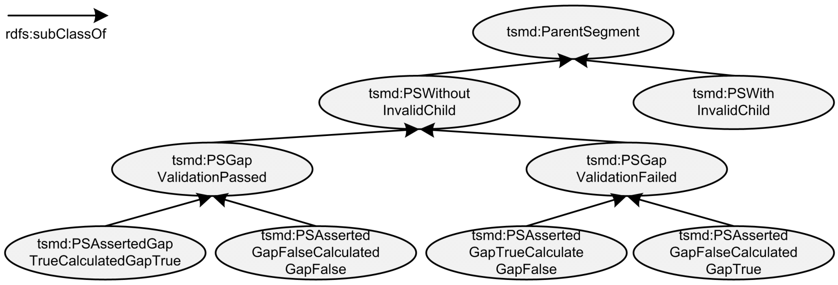


Figure 3: Validation hierarchy (classes for parent-child relation- and gap- verification)

3 Validation of Temporal Decompositions

The purpose of the validation process is to find invalid parent-child segment relations and to verify the asserted gap and overlap relation of a parent segment.

The presented ontology is capable of representing a temporal decomposition of a segment. For the validation purpose the classes depicted in Figure 3 are relevant. Classes needed for overlap verification are not depicted for simplicity.

3.1 Validating Temporal Decompositions using Rules

Rules are used to produce new statements about a temporal segment. The Jena rules syntax[6] is used for defining the rules. First the rule `calculate_end_point` computes the value of the property `hasEndPoint`, which represents the end point of a temporal segment. Additional rules for calculating the property values

[6] `http://jena.sourceforge.net/inference/index.html#rules`

```
[parent_has_invalid_child_true:
 (?parent rdf:type tsmd:ParentSegment),
 noValue(?parent tsmd:hasInvalidChild tsmd:true),
 (?parent tsmd:hasChild ?child),
 (?parent tsmd:hasStartPoint ?parent_sp),
 (?parent tsmd:hasEndPoint ?parent_ep),
 (?child tsmd:hasStartPoint ?child_sp),
 (?child tsmd:hasEndPoint ?child_ep),
 parentHasInvalidChild(?parent_sp, ?parent_ep, ?child_sp, ?child_ep)
 ->
 (?parent tsmd:hasInvalidChild tsmd:true)]

[parent_has_invalid_child_false:
 ?parent tsmd:hasInvalidChild tsmd:false
 <-
 (?parent rdf:type tsmd:ParentSegment),
 noValue(?parent tsmd:hasInvalidChild tsmd:true)]
```

Figure 4: Rules for computing `hasInvalidChild` property value.

(`true` or `false`) of the properties `hasInvalidChild`, `hasCalculatedGap` and `hasCalculatedOverlap` are defined. These properties are needed to define the classes for validation (see Figure 3) in the ontology. The rules for computing the value of the property `hasInvalidChild` are depicted in Figure 4. For this purpose, the forward rule `parent_has_invalid_child_true` and the backward rule `parent_has_invalid_child_false` are needed. New statements generated by the forward rule are the input for the backward rule. If all terms of the forward rule pass, a new statement with property `hasInvalidChild` and value `true` is generated for an instance of class `ParentSegment`. If no invalid parent-child relation has been found the backward rule produces a new statement with property `hasInvalidChild` and value `false`.

3.2 Implementation

As shown in Figure 5 the following steps are performed to semantically validate the temporal decomposition of a segment:

1. Classify parent segments (using a reasoner)
2. Calculate invalid children, gaps and overlaps respectively (using rules)
3. Classify the validation results (using a reasoner)

The approach is integrated into the VAMP service[7] for semantic validation of MPEG-7 descriptions [Troncy et al. 2007] and enables it to validate the description of the temporal media structure in terms of semantics.

[7] `http://vamp.joanneum.at`

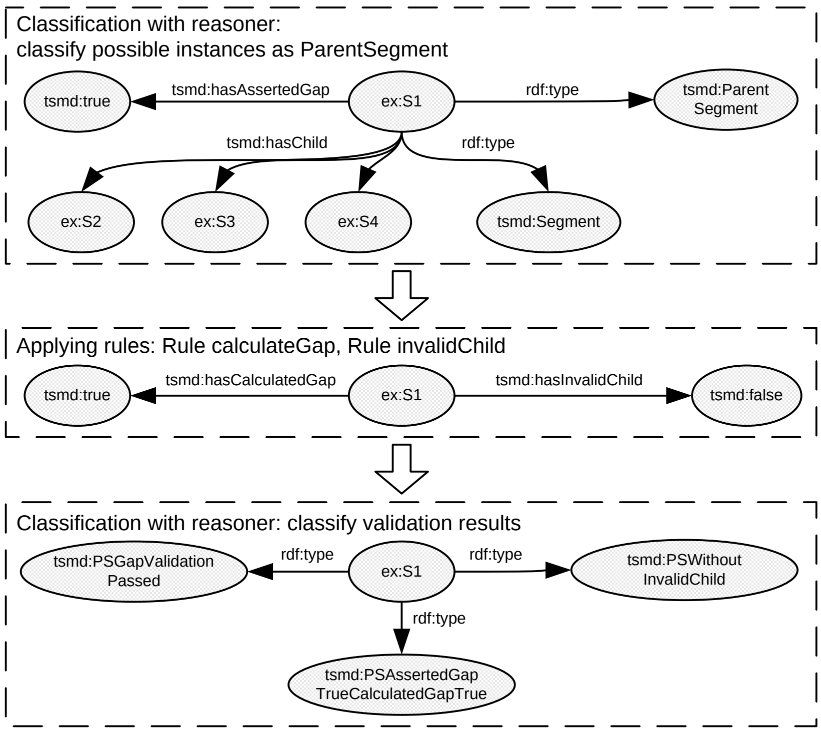


Figure 5: Validation steps for parent-child relation and gap

4 Conclusions and Future Work

We have proposed a formalisation of temporal decompositions in multimedia content description. Rules are subsequently used to check the validity of a the temporal description. The chosen approach is generic, as it models general properties of the description of temporal multimedia content structure, independent of the actual description format or standard. It can thus be applied to the validation of descriptions using other standards that support the concept of temporal segmentation, such as MXF DMS-1.

Another application are for the proposed model is that of search and retrieval applications, that allow querying temporal segments of the content. The metadata describing the content are often related to different segments, so that only subsegments are relevant to the query. Our approach can be used to determine the segment that shall be returned from the segments to which to the metadata description of the content relates.

A possible enhancement for the validation of temporal segmentation is to more precisely report the violation, i.e. the segments and the elements/attributes that caused it. This would improve the quality of the error messages that the user receives and thus the usability of the service.

Acknowledgments

The authors would like to thank Raphaël Troncy, Philip Hofmair and Rudolf Schlatte for their contributions. The research herein was partially supported by the European Commission under the contracts FP6-027026, "Knowledge Space of semantic inference for automatic annotation and retrieval of multimedia content - K-Space", IST-2-511316, "IP-RACINE: Integrated Project - Research Area CINE", and IST-004124, "NM2: New Millennium – New Media".

References

[Allen and Ferguson 1994] James F. Allen and George Ferguson. Events and actions in interval temporal logic. *Journal of Logic and Computation, Special Issue on Actions and Processes*, 4(5):531–579, Oct. 1994.

[Bailer and Schallauer 2006] Werner Bailer and Peter Schallauer. The Detailed Audiovisual Profile: Enabling Interoperability between MPEG-7 based Systems. In *12th International MultiMedia Modelling Conference (MMM'06)*, pages 217–224, Beijing, China, 2006.

[Hunter 2001] Jane Hunter. Adding Multimedia to the Semantic Web - Building an MPEG-7 Ontology. In *First International Semantic Web Working Symposium (SWWS'01)*, Stanford, California, USA, 2001.

[Nack et al. 2005] Frank Nack, Jacco van Ossenbruggen, and Lynda Hardman. That Obscure Object of Desire: Multimedia Metadata on the Web (Part II). *IEEE Multimedia*, 12(1), 2005.

[Ossenbruggen et al. 2004] Jacco van Ossenbruggen, Frank Nack, and Lynda Hardman. That Obscure Object of Desire: Multimedia Metadata on the Web (Part I). *IEEE Multimedia*, 11(4), 2004.

[Troncy et al. 2007] Raphaël Troncy, Werner Bailer, Michael Hausenblas, and Martin Höffernig. VAMP: Semantic Validation for MPEG-7 Profile Descriptions. Technical Report INS-E0705, (CWI) Centrum voor Wiskunde en Informatica, 2007.

[Troncy et al. 2006] Raphaël Troncy, Werner Bailer, Michael Hausenblas, Philip Hofmair, and Rudolf Schlatte. Enabling Multimedia Metadata Interoperability by Defining Formal Semantics of MPEG-7 Profiles. In *1st International Conference on Semantics And digital Media Technology*, pages 41–55, Athens, Greece, 2006.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Community-Aware Semantic Multimedia Tagging – From Folksonomies to Commsonomies

Ralf Klamma, Marc Spaniol and Dominik Renzel
(Lehrstuhl Informatik 5, RWTH Aachen, Germany)
{klamma|spaniol|renzel}@i5.informatik.rwth-aachen.de

Abstract: Tagging is an extremely popular mechanism in many Web 2.0 applications to create metadata supporting search and retrieval of arbitrary multimedia information like digital images, video or audio. However, compared to the syndicated multimedia information itself, the metadata are still "sticky". They cannot be accessed across several Web 2.0 applications, their semantic enrichment is not possible and they cannot be embedded in the local practices of communities of practice. Here, we present a multimedia tagging mechanism based on the international standard MPEG-7 for community-aware, standard compliant tagging of semantically enriched metadata implemented in the M7MT proof-of-concept application.

Keywords: Multimedia Management, MPEG-7, Web-Services, Information Systems
Categories: H.3.3, H.3.4, H.3.5, H.5.1, H.5.4

1 Introduction

Due to its simplicity and intuitiveness tagging has become a globally adopted technique for categorizing and retrieving multimedia on the web. People use web services for sharing and tagging their images with flickr [Flickr 2007], videos with YouTube [YouTube 2007], bookmarks with delicio.us [Delicious 2007], etc. thereby creating so called folksonomies. However, a "semantic gap" between the technical extraction of data and the semantically correct interpretation of contents can be recognized [DelBimbo 1999]. In this aspect, existing multimedia tagging systems have three crucial shortcomings. At first, these systems only offer plain keyword tagging, where tags carry their semantics implicitly only. Despite their potential in improving search and retrieval of multimedia contents, tagging systems face the problem inherent in the implicit semantics of the vocabulary used for tagging [Furnas et al. 1987, Marlow et al. 2006]. Particularly, the semantics are not accessible for further machine processing. Current trends and evolving standards in multimedia technology are intended for enriching multimedia content with semantic metadata leading to more advanced multimedia management and retrieval methods in order to handle the dramatically increasing amount of publicly available multimedia content on the web [Benitez et al. 2002]. Consequently, the tags itself should carry their semantics explicitly in order to make this additional information machine-accessible. A second issue in existing systems is a certain lack of community awareness. Existing systems understand their users as one community having a common interest and practice with regard to a specific medium. However, sub communities evolving within these systems have different terminologies and viewpoints on contents of many different media. Thus, different communities should be able to create different – even contradictory – community-specific terminologies for multimedia contents. Finally, the third shortcoming of existing systems is that they offer only basic community support. That means, users can form groups and restrict access to group specific media. But none of the existing systems is capable of commsonomies: The cross-media and cross-community wide sharing of community-specific folksonomies.

This paper addresses the above mentioned issues and introduces a proof-of-concept implementation of a community-aware semantic tagging system called *M7MT*. *M7MT* incorporates MPEG-7 based semantic multimedia descriptions [ISO 2002, ISO 2003] within a Lightweight Application Server (*LAS*) for MPEG-7 compliant community hosting [Spaniol et al. 2006, Klamma et al. 2006]. The next chapter therefore compares related tagging systems and describes their capabilities the processing the implicit semantics of multimedia as well as community-awareness. Then, we introduce our conceptual approach towards community-aware semantic tagging by commsonomies. After that, we present *M7MT*, our proof-of-concept implementation of a community-aware semantic tagging system. The paper ends with conclusions and gives an outlook on future research.

2 Related Tagging Systems

There exist several systems for tagging. However, most of them only support a single media type instead of providing cross-media tagging support. Even more, these systems are basically incapable of distinguishing between the different community contexts a user currently is member of. We will now briefly introduce the most prominent systems for different kinds of media types and explain their central features.

Flickr is a typical representative of a Web 2.0 application [Flickr 2007]. It provides its users with functionalities to describe, tag and arrange images in web-based collections. Similar features are also provided by *flickr*'s parent company *Yahoo! Photos* [Yahoo 2007a]. However, *flickr* recently introduced some elementary community support, which will lead to integrated version of *Yahoo! Photos* into *flickr*. Comparable with the systems described before, YouTube is being used for the community wide-sharing of videos [YouTube 2007]. In the music domain, *last.fm* offers its users possibilities to share tags about mp3-songs [Last.fm 2007]. Again, similar features are also available by the competitor system *Odeo* [Odeo 2007]. The tagging of information about web-sites is possible with *delicio.us* [Delicious 2007]. However, again only a single medium is being supported, namely bookmarks. Thus, *delicio.us* does not support any specific mean to distinguish tagging of web-sites different from blogs, e.g. by semantic concepts. That is the place where *Technorati* comes into play. *Technorati*, for instance, is a system dedicated to tagging of blogs, only [Technorati 2007]. Likewise, *Yahoo! Podcasts* is a tool dedicated for the annotion of podcasts and vodcasts [Yahoo 2007b]. However, there is even no combination of tagging features between *Yahoo! Photos* and *Yahoo! Podcasts*.

What can be seen from the brief introduction of the related tagging systems are basically three things:

- Tagging support is mono-medial only
- There are no high-level concepts for the typification of tags

- No distinctions are made based on the user's community context

In order to overcome the three problems mentioned above, we will now introduce commsonomies in order to allow the community-aware tagging of multimedia.

3 Commsonomies: Community-Aware Semantic Multimedia Tagging

In this section we first present community-aware semantic multimedia tagging on the conceptual level with a focus on extending classic keyword tagging by semantic and community-awareness concepts.

3.1 Semantic Extensions

In our previous work [Spaniol et al. 2006] we already presented semantic tagging as an extension of plain keyword tagging by additional metadata. Based on the MPEG-7 semantic description scheme we assigned semantic information to tags. While plain keyword tags are represented by their name exclusively, semantic tags consist of a name, an optional definition, a mandatory type and optional type specific information. Following the MPEG-7 standard, semantic tags are classified into the seven semantic tag types *Agent*, *Object*, *Place*, *Time*, *Event*, *Concept* and *State*. Each of these seven types allows the specification of additional type-specific information such as geographic coordinates for locations, time points resp. intervals for time, parameter name/value pairs for states, etc. One prominent problem of plain keyword tagging that is additionally overcome by semantic tags is the potential risk of semantical ambiguities. As one example consider the word "Portrait" being a polysemy of different meanings: a certain kind of painting or a dedicated camera angle. While plain keyword tagging users would assign the identical keyword tag to two media, semantic tagging reflects this difference in semantic meanings by assigning two different semantic tags. The ambiguity problem also occurs in the context of different communities, possibly having agreed upon different definitions of the same term.

3.2 Community-Awareness Extensions

Existing plain keyword tagging systems allow users to assign tags to media without reflecting any community memberships. Every user has access to all tags assigned by all users, possibly within the contexts of different communities. However, it is not possible to specify, in which community context a tag assignment has been defined. We intend to gap this shortcoming by modelling community-specific tag assignments using the concept of community forests, i.e. a set of hierarchies along with a special notion of community membership semantics. If a user is explicit member of a community, he is considered member of all ancestor communities within the same community hierarchy. This extended notion implies that tag accessibility has to be controlled by the system. A user should only be able to access a tag, if he is member of the community in whose context the tag has been assigned. To illustrate the above ideas, the following example provides a possible scenario to demonstrate community-aware tagging of one specific multimedia content item. Figure 1 below shows a theoretical scenario. Each tree node represents a specific community and is annotated with the set of semantic tags assigned to the considered multimedia content item in the context of the corresponding community. Semantic tags *s1* and *s2* have been assigned to the multimedia content item in the context of community *c1*, *s3* in the context of *c2* and *s4* in the context of *c3*. No tags have been assigned in the context of *c4* and *c5*.

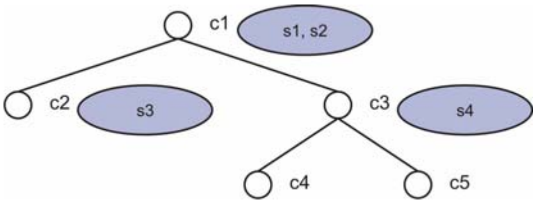


Figure 1: View of a commsonomy for a single multimedia content item

Now consider three users *u1*, *u2* and *u3* being members of different communities. *u1* is explicit member of community *c2*, *u2* is member of *c1* and *u3* is member of *c2* and *c5*. Let us now recall community membership semantics. If a user is explicit member of a group *g*, he is implicitly considered member of all ancestor groups of *g*. Accordingly, *u1* is member *g1* and *g2* and thus has access to semantic tags *s1*, *s2* and *s3*. Analogously, *u3* is member of *g1*, *g2*, *g3* and *g5* and thus has access to all semantic tags *s1*,...,*s5* while *u2* has access to *s1* and *s2* only. Figure 2 demonstrates user-specific tag accessibility for users *u1*, *u2* and *u3*.

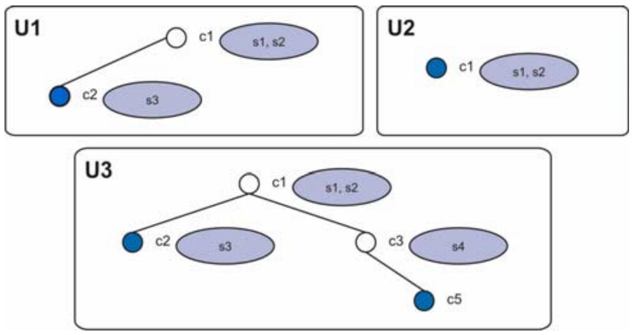


Figure 2: User-specific commsonomy tags depending on community affiliations

4 *M7MT*: Multimedia Commsonomies

In this section we present our proof-of-concept implementation of a community-aware semantic multimedia tagging system. On the server side we employ our MPEG-7 compliant Lightweight Application Server (*LAS*) for MPEG-7 Services in Community Engines (cf. [SKJR06] for details). Next, we briefly explain the key concepts for community management in *M7MT*. After that we explain the MPEG-7 Commsonomy-Services of *M7MT* in more detail. Finally, the user interface of *M7MT* is introduced and its community features are highlighted.

4.1 Community Management in *M7MT*

LAS provides a set of built-in core services offering community management functionality. The *LAS* usermanager maintains users and communities (groups in *LAS* terminology) as well as their general and group-specific access rights modelled as roles. The *LAS* object manager provides the access to security objects. In the following paragraphs basic *LAS* community management concepts are explained in detail.

Managing Users & Groups

For each user there is a list of roles, that can be assigned to him either as global permissions or prohibitions. These roles define, which service methods a user is allowed resp. forbidden to invoke. In addition to users, *LAS* maintains a hierarchical group structure being built of a number of group trees, i.e. a group forest. Groups are defined by a unique id, a unique name, some arbitrary XML structure for the optional storage of additional group information and a list of members. *LAS* group memberships carry special semantics. If a user *u* is member of a group *g*, then he is implicitly also considered member of all ancestor groups of *g* within the same group tree. Special group roles can be assigned to members in order to define the particular rights they have within this particular group.

Managing Permissions and Roles

A permission in *LAS* defines access rights to services and their methods. *LAS* offers four levels of granularity for the definition of *LAS* permissions:

- *Root Permission:* all services including all methods
- *Service Permission:* one specific service including all methods
- *Service Method Permission:* one specific method of a specific service
- *Service Method Signature Permission:* one specific method of a specific service carrying a specific signature

The granularity levels define an implication relation. The root permission implies service permissions for all services, a service permission implies service method permissions for all methods of this service, and so on.

Managing Security Object Access

The *LAS* objectmanager maintains an access control list (ACL) for each security object. Similar to the UNIX filesystem an ACL defines access rights on three different axes: users, groups and all others. An ACL contains an arbitrary number of ACL permission- resp. prohibition collections for users, groups and others in order to control the access to a security object in a specific service method context. The content of an ACL permission collection is interpreted as permissions. The content of an ACL prohibition collection is interpreted as prohibitions.

4.2 MPEG-7 Commsonomy-Services

In our previous work [SKJR06] we introduced a set of two services involved in the process of semantic multimedia tagging: a semantic service for the management of MPEG-7 semantic basetype descriptions and a multimedia content service for the management of MPEG-7-based multimedia content descriptions. Both services used a built-in *LAS* component for the interaction with a native XML database (e.g. eXist [Exist 2007, Meier 2002] or Oracle 10g [Cyran 2005]) storing the MPEG-7 descriptions. Semantic tagging is realized in the multimedia content service by adding semantic basetype references to the semantics descriptor of a multimedia content descriptor. In order to create support for community-aware semantic multimedia tagging, we introduced an additional custom *LAS* security object type for controlling access to semantic basetype references within a multimedia content description. Notice the difference between controlling access to a semantic basetype description and to instances of semantic basetype references.

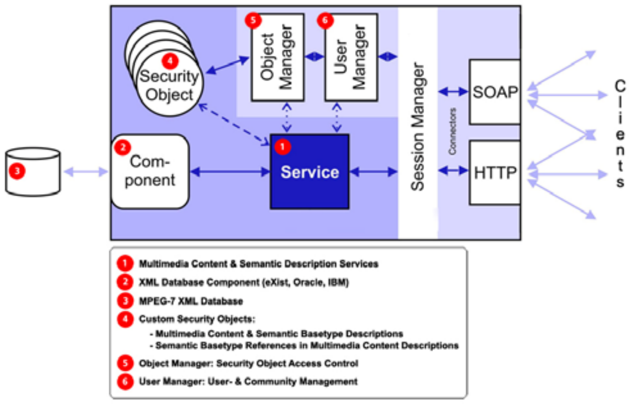


Figure 3: Usage & combination of LAS concepts for multimedia commsonomies

Community-awareness is now realized by controlling the ACL of such a semantic basetype reference security object, especially the group-axis. If a user intends to tag a multimedia content item, he will first use the semantic basetype service to check, if the set of semantic tags he wants to assign completely exists in the system already. If this is not the case, he can use the semantic service to create the missing semantic basetype descriptions. In the next step the respective semantic basetype references are assigned to the multimedia content description in a given community context, i.e. *LAS* groupcontext. If a particular semantic basetype has already been assigned to the same medium from another community context, the corresponding semantic reference security object is adjusted by adding an appropriate permission in the group section of its ACL. If such a security object does not exist, it is created with the appropriate ACL. Removal of a semantic tag within a given community context is achieved by either removing the corresponding group permission from the ACL, if the semantic tag has been assigned in more than one community context or even removing the whole semantic basetype reference, if the tag has been assigned in one single community context. On retrieval of a multimedia content description by a user, the multimedia content service checks the access rights to all assigned semantic basetype references and only returns those tags that are accessible within a community context the calling user is member of, either explicitly or implicitly. Figure 3 shows an excerpt of the *LAS* overall architecture including annotations to illustrate how basic *LAS* concepts are used and combined in order to achieve community-aware semantic multimedia tagging.

4.3 Community-Aware Multimedia Contents in *M7MT*

The user interface of *M7MT* allows users to obtain community-awareness depending on the context they are currently involved. Here, users create semantic tags and assign them to multimedia content items in a specific community context. Visibility of a semantic tag depends on the user's particular community memberships. In order to demonstrate community dependent tag visibility on client side, figure 4 shows three different user views on an image depending on the users' individual community memberships in *M7MT*. If a semantic tag has been added in a specific community context and the viewing user is member of this community, the semantic tag is rendered as a thumbnail being part of a multimedia information overlay. Tags from communities a user is not member of are invisible. The previously introduced example (cf. figure 3) has been mapped to one possible real world example of a UNESCO community and its sub communities tagging a picture of a Buddha painting in Bamiyan Valley, Afghanistan during a fieldwork. The lower part of figure 4 shows the semantic tag thumbnails for each of the users *u1*, *u2* and *u3*.

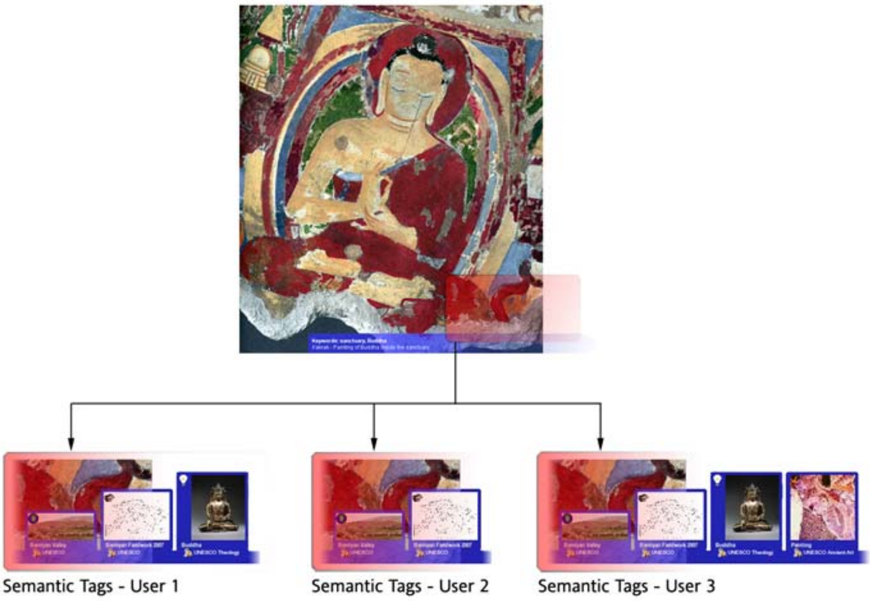


Figure 4: Commsonomy tag visibility for members of different communities in M7MT

5 Conclusions & Outlook

In this paper we identified several crucial shortcomings in existing multimedia tagging systems in the Web 2.0. Metadata as tags are still “sticky” in the current generation of Web 2.0 applications and not accessible across applications borders even if we can syndicate the multimedia information itself. Plain keyword tags only carry semantics implicitly. Additionally, existing systems do not exhibit community-aware tagging. Thus, we proposed a community-aware semantic multimedia tagging system overcoming these gaps. Support for semantic multimedia tagging is realized as LAS services using MPEG-7 semantic basetype descriptions as semantic tags which are assigned to multimedia content by adding semantic basetype references to the corresponding MPEG-7 multimedia descriptions. The *LAS* built-in community support, especially the concept of security objects and their ACLs is used to create a community-aware semantic tagging. Community-aware tagging services are essential for next generation mobile multimedia information systems where search and retrieval will be supported by context-aware services using location, time and community information in parallel to offer best possible results to mobile users.

References

[Benitez et al. 2002] A. B. Benitez, H. Rising, C. Jörgerisen et al.: "Semantics of Multimedia in MPEG-7", 2002.

[Cyran 2005] M. Cyran: "Oracle Database Concepts, 10g Release 2 (10.2)", B14220-02, Oracle, October, 2005.

[DelBimbo 1999] A. Del Bimbo: Visual Information Retrieval. Morgan Kaufmann, 1999.

[Delicious 2007] Delicio.us: "Social Bookmarking", delicio.us, 2007 [last access: 1.6.2007].

[Exist 2007] eXist: "An Open Source Native XML Database", www.exist-db.org, 2007 [last access: 1.6.2007].

[Furnas et al. 1987] G. W. Furnas, T. K. Landauer, L. M. Gomez, and S. T. Dumais: "The vocabulary problem in human-system communication", 1987.

[Flickr 2007] flickr™: "Photo Sharing", www.flickr.com, 2007 [last access: 1.6.2007].

[ISO 2002] ISO ISO/IEC: Information Technology – Multimedia Content Description Interface – Part 3: Visual. ISO/IEC 15938-3:2002, Intl. Organization for Standardization, 2002.

[ISO 2003] ISO/IEC: Information technology – Multimedia Content Description Interface – Part 5: Multimedia description schemes. 15938-5:2003, ISO, 2003.

[Klamma et al. 2006] R. Klamma, M. Spaniol, and Y. Cao: "MPEG-7 Compliant Community Hosting", M. Lux, M. Jarke, H. Kosch (Eds.): MPEG and Multimedia Metadata Community Workshop Results 2005, J.UKM Special Issue (Journal of Universal Knowledge Management), Springer, Vol. 1, No. 1, 2006, pp. 36-44.

[Last.fm 2007] Last.fm: "The Social Music Revolution", www.last.fm, 2007 [last access: 1.6.2007].

[Meier 2002] W. Meier: "eXist: An Open Source Native XML Database", In: Web, Web-Services, and Database Systems, NODe 2002 Web and Database-Related Workshops, Erfurt, Germany, October 7-10, Revised Papers, volume 2593 of LNCS, Springer-Verlag, Berlin Heidelberg, pages 169 – 183, 2002.

[Marlow et al. 2006] C. Marlow, M. Naaman, D. Boyd, M. Davis: "HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, To Read", 2006.

[Odeo 2007] Odeo: "Millions of FREE MP3s, Podcasts, and More", www.odeo.com, 2007 [last access: 1.6.2007].

[Spaniol et al. 2006] M. Spaniol, R. Klamma, H. Janßen and D. Renzel: “LAS: A Lightweight Application Server for MPEG-7 Services in Community Engines”, K. Tochtermann, H. Maurer (Eds.): Proceedings of I-KNOW '06, 6th International Conference on Knowledge Management, Graz, Austria, September 6 - 8, 2006, J.UCS (Journal of Universal Computer Science) Proceedings, Springer, pp. 592-599, 2006.

[Technorati 2007] Technorati, www.technorati.com, 2007 [last access: 1.6.2007].

[Yahoo 2007a] Yahoo! Photos, http://photos.yahoo.com, 2007 [last access: 1.6.2007].

[Yahoo 2007b] Yahoo! Podcasts: “Discover and enjoy all the best podcasts and vidcasts”, http://podcasts.yahoo.com, 2007 [last access: 1.6.2007].

[YouTube07] YouTube: “Broadcast Yourself”, www.youtube.com, 2007 [last access: 1.6.2007].

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Requirements of Fragment Identification

Wolfgang Jochum
(University of Vienna, Distributed and Multimedia Systems, Austria
wolfgang.jochum@univie.ac.at)

Abstract: The task of creating specific references rests on specifications that qualify how parts of resources can be addressed. The lack of standards for fragment identifiers has lead to the problem that links, metadata and references merely point to whole resources. Although it is suggested that fragment identification is specified with a media type's MIME type registration, there are few formats that provide a fragment identification scheme. Furthermore formats that specify fragment identification schemes have not agreed on a common set of requirements.

In this paper we present an overview of the current status of interoperable fragment definitions, point out promising activities that promote interoperable fragment definitions and suggest strategies to promote uniform fragment identifiers. Additionally a set of requirements is defined and described to ease the development of fragment identification standards.

Key Words: fragment identifier, URI, media access, metadata
Category: H.3.3, H.3.5, H.5.1, H.5.4

1 Introduction

Hypertext links are recognized as one of the primary driving forces of the Web and the simplicity of creating links is one aspect for the success of the hypertext Web [Jacobs 04]. Following a link, we are used to navigate directly to the linked information resource, or even to a specific part of the resource. But not all media types define how to address a specific part of a resource with a fragment identifier - FID. For instance, if one would like to link to a selected part of a movie, there is no approved, application and format independent way to do so. But Navigating directly to a specific part of a resource is only one possible application of fragment identifiers. Fragment identifiers can also be used to include a specific part of a resource in another document or to add links or metadata to parts of resources that do not support this internally. The vocabulary required to define the fragments is thereby moved from the metadata or annotation schema to the language that defines the fragment identifier [Geurts 05]. This improves interoperability and usability. Navigation, transclusion and external references are different applications of FID and all have their own specific requirements. By giving an overview on the current status and presenting a set of requirements, this paper promotes uniform fragment identification.

1.1 Fragment Identification on the Web

Fragments can be used to refer to parts of resources on the client-side. This may either be a secondary resource identified by reference to the primary resource or a defined view on the primary resource [Berners-Lee 05]. Fragment identifiers allow authors to reference aspects of existing resources independent of the resource provider. This is possible, because the fragment identifier is separated from the URI - Uniform Resource Identifier before the resource is requested from the server. After the resource has been retrieved the identified fragment is processed by the user agent.

A Web browser or user agent that follows a link to a HTML - Hypertext Markup Language document identified by a URI knows how to interpret the optional fragment identifier, because it is declared in the MIME[1] media type 'text/html'[2]. So if a fragment is declared, the browser will render the whole page and scroll to the identified element. Although it is recommended that a MIME media type registration should contain information on fragment identifiers [Freed 05], there are only few media types that provide fragment identifier specifications. Web formats like HTML, SMIL - Synchronized Multimedia Integration Language and SVG - Scalable Vector Graphics use anchors or named elements to specify link targets that can be used by FID. Although this is sufficient for navigating, it is insufficient for transclusion and external references. Especially multimedia document types are missing clear semantic and syntactic description of fragment identifiers [Ossenbruggen 04, Nack 05, Arndt 07]. The recently published standard MPEG-21 FID[ISO 06] and proposals like temporal URI[Pfeiffer 05] and text/plain FID [Wilde 05] prove that developers need a unified way to address fragments in diverse formats.

A lively discussion is held upon the fact, that content negotiation of the http protocol and fragments don't go well together. In section 3.5 of [Berners-Lee 05] a fragment identifier is described as a component of a URI that allows *indirect* identification of a secondary resource by reference to a primary resource. The secondary resource is the fragment and the primary resource is identified by the URI without the fragment identifier. That is where a possible problem arises when using content negotiation of the http protocol [Fielding 99]. This mechanism selects an appropriate response format of an URI. If a fragment identifier is used with an URI using content negotiation, the fragment identifier has to be consistent across all formats that may be retrieved from that URI. This is well known [Berners-Lee 97] and no agreement has been found how to overcome this potential pitfall. The most simple solution is to avoid combining fragment identifiers and content negotiation.

1 RFC 2046: Multipurpose Internet Mail Extensions Part Two: Media Types [Freed 96]

2 RFC 2854: The 'text/html' Media Type [Connolly 00]

After an overview of the status of fragment identification, a set of requirements are presented. Based on these requirements, suggestions how to add fragment identifications to a wide range of common multimedia types are given.

2 Current Status

This part gives an overview of current standards and projects that define fragment identifiers.

HTML, SMIL and SVG support named fragment identifiers. Additionally the fragment identifier of SVG allows to specify a desired view of the document.

XML - Extensible Markup Language files can use the XPointer Framework [Grosso 03] as a basis for fragment identifiers. XML based formats should define their own fragment identifiers. Although one might assume that fragment identification is similar to HTML this is incorrect [Jacobs 04].

Adobe has defined a set of open parameters that can be used as fragment identifiers [Taft 04]. The parameters can be used to define a highlight, to jump to a named destination or page, to search in the document and to define view options. Also it is possible to enable and disable specific controlls for the user like the tool bar, status bar, message bar, navigation panes and scroll bar.

Fragment Identification of MPEG Resources - MPEG-21 FID - is defined in Part 17 of the MPEG-21 framework. It supports all MPEG resources and can be used to address parts of any resource whose MIME type is one of: audio/mpeg, video/mpeg, video/mp4, audio/mp4, application/mp4, application/mp21 and video/MPEG4-visual. It is based on the XPointer Framework and adds temporal, spatial and spatio-temporal axis, logical units, byte ranges, masks for videos and items and tracks of ISO Base Media Files [ISO 06, WG 05].

The two following specifications for fragment identifiers are not standards, but represent the ongoing effort to establish interoperable fragment identifiers for various media types.

The Internet-Draft *Specifying time intervals in URI queries and fragments of time-based Web resources* - Temporal URI - addresses the problem of fragment identification for temporal offsets and time intervals in time-based Web resources. Although it was originally developed to support specific resources (Annodex, CMML2) it can be used with all information resources that relate to a timeline of events [Pfeiffer 05] .

With fragment identifiers from the Internet-Draft *URI Fragment Identifiers for the text/plain Media Type* - text/plain FID - positions, ranges and query results can be addressed. Additionally a hash value can be used to check if a fragment is still valid [Wilde 05].

3 Requirements of Fragment Identifiers

Based on our research on fragment identification and previous work[Rutledge 01, Wilde 05, Pfeiffer 05, ISO 06] a comprehensive set of requirements for fragment identification is presented.

3.1 Source of Fragment Definition

A fragment can either be defined in the destination resource, in a separate location or inside the fragment identifier. In the first case, the author of the resource has to specify the fragments before they can be used; a prominent example is HTML. Other standards like XPointer support addressing into the internal structures of documents without having to modify it.

3.2 Fragment Identification Type

Three main fragment identification types can be distinguished: measured, nominal and structured.

Measured fragments provide dimension specific metrics to identify fragments. Usual dimensions are spatial, temporal and spatio-temporal. The metrics rely on semantics of the dimension and are inherent to the resource. In most cases measured fragment identification is coding format independent.

Nominal fragments use given names or ids within the destination document. The semantic of the fragment is defined by the media type. While HTML and SMIL only support fragments to be used as link targets, SVG allows to define views.

Structured fragments use the physical or logical structure of a resource type to identify a fragment. Queries are another way to identify structural fragments. The specification of text/plain FID [Wilde 05] shows how to use regular expressions to identify fragments in text files. The XPointer Framework has been defined as a basis to identify fragments in XML documents based on various properties, such as element types, attribute values, character content and relative position.

Depending on the resource type and its semantics, fragment identification can be measured, nominal, structured or any meaningful combination of these fragment types. A prominent example of a fragment identifier using structured, nominal and measured fragments is the MPEG-21 FID. MPEG-21 FID is an ISO standard that is based on the XPointer framework and defines fragment identifiers with respect to media semantics of MPEG resources.

3.3 Fragment Presentation

If easy distinction between fragment and context has to be provided to the user, a reasonable solution is highlighting the fragment with a given style. In order to have coherent presentations of the same fragment across user agents, presentation behavior has to be specified with the fragment identifier. Style attributes of the highlight can be defined as part of the fragment or within the destination resource. In CSS3 - Cascading Style Sheets Level 3 the pseudo class target can be used to define style of the link destination [Celik 04].

Another presentation dependent fragment that can to be defined is the view of a resource. Formats like SVG and PDF define attributes in their fragment identifiers that allow adjustments of the region that is displayed.

3.4 Fragment Context Removal

In order to improve reuse of existing resources, an author may crop or clip resources thereby creating a fragment. Using a fragment identifier one can also create a portion of the original resource. This implies that the fragment is separated from the context. The context of a fragment is a portion of a resource, that is not a fragment [Rutledge 01]. As discussed in [Rutledge 01] a unified fragment identification would introduce unification, consistency and simplicity to Web fragmenting.

The author of the FID should be able to control whether context has to be removed. By definition, context removal of URI fragments is done by the client. Although it is obvious that server side context removal is more efficient in terms of network traffic, this behavior guarantees, that fragment identification is independent from the provider of the resource.

3.5 Fragment Robustness

Since resources may change, means to improve robustness should be added to fragment identification. An example how this can be done by adding a hash sum is available in [Wilde 05]. The hash sums are used to check if a resource has changed.

4 Evaluation of the Current Status

All fragment identification formats presented in this paper except XPointer are evaluated with respect to the identified requirements. XPointer does not specify appropriate semantics for fragment identification of specific XML-based data formats [Jacobs 04].

4.1 Fragment Definition and Identification

MPEG-21 FID is the most expressive language and can be extended to support other pointer schemes. In contrast to the open framework of MPEG-21 FID all other formats are focused on specific formats. The text/plain FID format provides a complete set of identifiers for resources with the media type text/plain. Temporal URI defines ways to address temporal fragments similar to MPEG-21 FID. HMTL, SMIL and SVG use nominal fragments for identification, with the limitation, that only elements that have been given names can be used as fragments.

4.2 Fragment Presentation

None of the fragment identification formats support the definition of style to be used by user agents that present a fragment. HTML and SVG documents may use the CSS3 Hyperlink Presentation Module to specify the presentation properties of hyperlinks, but currently CSS3 is still a working draft.

4.3 Fragment Context Removal

HTML, SMIL, PDF, MPEG-21 FID and text/plain FID do not specify context removal behavior. SVG allows to define client side context removal using view parameters.

Temporal URI uses a query (?) instead of a fragment (#) to support server side context removal. A temporal query supports server-side context removal with the limitation that the server has to be capable of resolving a Temporal URI query. Temporal queries have the same addressing scheme as fragments which allows the author to easily choose between client and server side processing of the fragment identification.

4.4 Fragment Robustness

Fragment robustness is only considered by text/plain FID.

5 Recommendations

In the authors opinion MPEG-21 FID is a promising step towards unified fragment identification for multimedia resources. MPEG-21 FID is a comprehensive standard for fragment identification and has the potential to act as basis for unified fragment identifier specifications, particularly for multimedia resources. It supports all MPEG resources, can be used for almost all audiovisual multimedia resources and has means to add support for other resources.

Just because MPEG-21 FID is very expressive and provides universal schemes for a whole domain, it may have problems becoming widely-used. Without openly available fragment identification processors it is far easier to define proprietary schemes. Feasible tools and support for developers must be the first step towards unified fragment identification. In addition a core profile of fragment identifiers for common use cases and defining mappings for unsupported identifiers outside the core profile can simplify adaptation. Having syntax and semantics that are not directly compatible with other Web technologies may be another disadvantage within the Web community. Furthermore, special requirements like robustness and presentation have not been specified in the MPEG-21 FID. Although it is possible to add them, it is still necessary to agree on a specification to become a standard that ensures interoperability.

6 Conclusions

The need for unification, consistency and simplicity of fragment identifications is obvious, but during the evolvement of the Web little effort has been taken to provide useful fragment identifiers for commonly used formats. This is especially true for formats that were originally not designed for the Web or were not meant to be accessed in a flexible way by referring to parts of the resource.

The requirements identified in this paper will help to improve future fragment identifier specifications in terms of extend, interoperability and expressiveness. Together with efforts that have been taken in several projects and standards [WG 05, Pfeiffer 05, Wilde 05] to provide uniform fragment identifiers it forms the basis for a brisk step towards unified fragment identification.

References

[Arndt 07] Richard Arndt, Raphal Troncy, Steffen Staab & Lynda Hardman. *Adding Formal Semantics to MPEG7: Designing a Well-Founded Multimedia Ontology for the Web.* Rapport technique, KU and CWI technical report KU-N0407, January 2007.

[Berners-Lee 97] Tim Berners-Lee. *URI References: Fragment Identifiers on URIs.* Axioms of web architecture, http://www.w3.org/DesignIssues/Fragment.html, April 1997. available at: `http://www.w3.org/DesignIssues/Fragment.html`.

[Berners-Lee 05] T. Berners-Lee, R. Fielding & L. Masinter. *Uniform Resource Identifier (URI): Generic Syntax.* RFC 3986, Internet Engineering Task Force, January 2005.

[Celik 04] Tantek Celik, Bert Bos & Daniel Glazman. *CSS3 Hyperlink Presentation Module.* W3C Working Draft WD-css3-hyperlinks-20040224, W3C, February 2004. http://www.w3.org/TR/2004/WD-css3-hyperlinks-20040224.

[Connolly 00] D. Connolly & L. Masinter. *The 'text/html' Media Type.* RFC 2854, June 2000.

[Fielding 99] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach & T. Berners-Lee. *Hypertext Transfer Protocol – HTTP/1.1.* RFC 2616, Internet Engineering Task Force, June 1999.

[Freed 96] N. Freed & N. Borenstein. *Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types.* RFC 2046, Internet Engineering Task Force, 1996. http://www.ietf.org/rfc/rfc2046.txt.

[Freed 05] N. Freed & J. Klensin. *Media Type Registration.* Best current practice, Internet Engineering Task Force, http://www.ietf.org/rfc/rfc4288.txt?number=4288, 2005 2005. `http://www.ietf.org/rfc/rfc4288.txt?number=4288`.

[Geurts 05] J. Geurts, J. van Ossenbrugen & L. Hardman. *Requirements for practical multimedia annotation.* In Multimedia and the Semantic Web, 2nd European Semantic Web Conference, 2005.

[Grosso 03] Paul Grosso, Eve Maler, Jonathan Marsh & Norman Walsh. *XPointer Framework.* W3C Recommendation, 25 March 2003. `http://www.w3.org/TR/xptr-framework/`.

[ISO 06] (International Organization for Standardization) ISO. *Multimedia framework (MPEG-21) – Part 17: Fragment Identification of MPEG Resources.* ISO Standard ISO/IEC 21000-17:2006, International Organization for Standardization, Geneva, Switzerland, December 2006. DRAFT: ISO/IEC FDIS 21000-17:2006(E).

[Jacobs 04] Ian Jacobs & Norman Walsh. *Architecture of the World Wide Web, Volume One.* W3c recommendation, World Wide Web Consortium, December 2004.

[Nack 05] Frank Nack, Jacco van Ossenbruggen & Lynda Hardman. *That Obscure Object of Desire: Multimedia Metadata on the Web, Part 2.* IEEE MultiMedia, vol. 12, no. 1, pages 54–63, 2005.

[Ossenbruggen 04] Jacco van Ossenbruggen, Frank Nack & Lynda Hardman. *That Obscure Object of Desire: Multimedia Metadata on the Web, Part 1.* IEEE MultiMedia, vol. 11, no. 4, pages 38–48, 2004.

[Pfeiffer 05] S. Pfeiffer, C. Parker & A. Pang. *Specifying time intervals in URI queries and fragments of time-based Web resources.* Network Working Group, Internet-Draft, March 19 2005. `http://www.annodex.net/TR/URI_fragments.html`.

[Rutledge 01] Lloyd Rutledge & Patrick Schmitz. *Improving Media Fragment Integration in Emerging Web Formats.* In Proceedings of the International Conference on Multimedia Modeling 2001 (MMM01), pages 147–166, CWI, Amsterdam, The Netherlands, November 5-7 2001.

[Taft 04] E. Taft, J. Pravetz, S. Zilles & L. Masinter. *The application/pdf Media Type.* Informational 3778, Internet Engineering Task Force, May 2004.

[WG 05] MPEG WG. *Introducing MPEG-21 Part 17 an Overview.* Overview, ISO/IEC JTC 1/SC 29/WG 11/N7221, http://www.chiariglione.org/mpeg/technologies/mp21-fid/index.htm.htm, April 2005. available at: `http://www.chiariglione.org/mpeg/technologies/mp21-fid/index.htm.htm`.

[Wilde 05] Erik Wilde & Marcel Baschnagel. *Fragment identifiers for plain text files.* In HYPERTEXT '05: Proceedings of the sixteenth ACM conference on Hypertext and hypermedia, pages 211–213, New York, NY, USA, 2005. ACM Press.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

MPEG-7 for Video Quality Description and Summarisation

Peter Schallauer, Hermann Fürntratt, Werner Bailer
(JOANNEUM RESEARCH, Graz , Austria
{firstname.lastname}@joanneum.at)

Abstract: Manual quality control of audiovisual content in the different steps of the media production, delivery and archiving process causes significant costs. Semi-automatic quality control requires automatisation of quality analysis, quality metadata interoperability and efficient visualisation tools. In this paper we propose the use of MPEG-7 for standard compliant description of media quality metadata and a quality summary visualisation tool which facilitates efficient exploration of visually impaired content by the user.

Keywords: MPEG-7, Quality, Descriptor, Description Scheme, Multimedia
Category: H.5.1, H.3.1, H.5.2

1 Introduction

Quality control of audiovisual content is important in several steps of the media production, delivery and archiving process. Broadcasters are checking audio and video quality within the ingest process, after editing, after encoding and before play-out for terrestrial, satellite and cable broadcast or for delivery to internet and video-on-demand services. Archives are checking for content integrity at archive ingest and delivery. Content providers are checking post production content for correct encoding and conformance to the required quality and format standard. Quality metadata interoperability is a prerequisite for integrating these tasks in the overall process.

2 Video Quality Description

There exist different tasks in the media production process dealing with quality related metadata. In order to facilitate interoperability and exchange of defect and quality descriptions between these tasks, a standardised way of description must be used. The description must be able to represent all the results obtained by automatic quality analysis tools as well as additional annotation made by operators. It must both support gaining a quick overview of the overall quality, type and severity of the defects present in the material, as well as describing the detailed measures returned by the tools when applicable.

MPEG-7 [MPEG-7 01] is a standard for the description of multimedia content, including structuring the content as well as describing a number of low-, mid- and high-level features for each of the segments in the structure. The defect and quality description is based on the MPEG-7 Detailed Audiovisual Profile (DAVP) [Bailer 06]. The original version of the standard provides very simple means for describing a quality rating and listing defects present in a segment, but without the capability to specify more in depth information. An amendment to the MPEG-7 audio part [MPEG-7/4 04] defines a more detailed description of audio signal quality, allowing describing a set of measures per segment as well as a list of error events with different temporal scope and further properties.

Based on the existing work in the audio part, we have defined a similar description framework for the visual domain with even more capabilities for describing details of defects. A list of defects and quality measures can be described for each segment: quality measure descriptors contain the statistics for a segment, while defect descriptors describe an occurrence of a defect in more detail.

There is a generic visual descriptor for defects which specifies general properties and references the defect in a comprehensive impairment classification scheme (based on the BRAVA broadcast archive programme impairments dictionary[1]). This is the minimum description of a defect, specifying the type of defect and the segment of its occurrence. In addition, specific descriptors for a number of defects and quality measures have been defined[2], which allow to describe their respective properties.

3 Visualisation of Defect and Quality Analysis Results

The visualisation of defect and quality analysis results must support the user in quickly getting an overview of the condition of the material. For that purpose, we have implemented the defect and quality summary viewer shown in Figure 1.

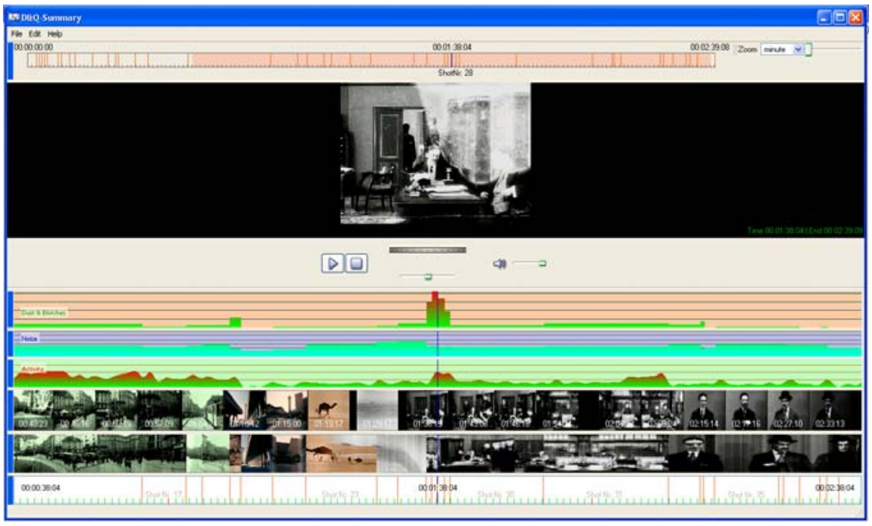


Figure 1: The quality summary viewer.

The tool supports the user in efficiently navigating the content by providing a timeline representation of a number of views. All views are synchronised with the video player. The temporal resolution can be changed so that the user can freely change the level of detail shown. The timeline views show the shot structure of the material, selected representative key frames, stripe images created from the central columns of

---

[1] http://brava.ina.fr/brava_public_impairments_list.en.html

[2] The extension schema can be downloaded from http://mpeg7.joanneum.at

the images in the sequence and a number of graphs visualising defects and quality measures. In the screenshot one of the graphs shows the visual activity, which is not a quality measure, but a helpful indicator in the context of restoration. High visual activity indicates either large scale defects (e.g. blotches) or a high amount of motion, which often complicates the restoration process. The other graphs show the shot-wise dust level as the median fraction of the image area covered by dust and grain noise as the image to grain noise ratio.

The temporally condensed overview allows the user to quickly grasp the frequency and strengths of the impairments in the material. From the statistical measures for the individual defects, especially dust and noise level, the needed restoration steps and tools can be planned. Together with the severity of the defects and the user's knowledge about the capabilities of the restoration tools in use the required manual effort and the restoration costs can be estimated. For those defects which are described as defect events, such as big distortions or missing frames, the user can estimate the restoration effort directly from the number of events. Defect event information can also be used for direct examination and restoration of these time intervals in the movie, without having to view the rest of the material.

Acknowledgements

The authors would like to thank Werner Haas, Georg Thallinger, Hannes Fassold, Harald Stiegler as well as several other colleagues at JOANNEUM RESEARCH, who contributed valuable input to the work. This work has been funded partially under the 6th Framework Programme of the European Union within the IST project "PrestoSpace" (IST FP6-507336, http://www.prestospace.org).

References

[Bailer 06] Bailer, W., Schallauer, P.: "The Detailed Audiovisual Profile: Enabling Interoperability between MPEG-7 Based Systems," *Proc. of 12th International Multi-Media Modeling Conference*, Beijing, CN, Jan. 2006.

[MPEG-7/4 04] Information Technology – Multimedia Content Description Interface, Part 4: Audio, ISO/IEC 15938-4:2002/Amd 1:2004.

[MPEG-7 01] Information Technology – Multimedia Content Description Interface, ISO/IEC 15938, 2001.

[Schallauer 07] Schallauer, P., Bailer, W., Mörzinger, R., Fürntratt, H., Thallinger, G.: "Automatic Quality Analysis for Film and Video Restoration", *Proc. of IEEE International Conference on Image Processing*, San Antonio, TX, USA, Sept. 2007, to appear.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

MPEG-7 Video Annotation Tool

Helmut Neuschmied
(JOANNEUM RESEARCH, Austria
helmut.neuschmied@joanneum.at)

Abstract: This paper presents a video annotation system which enables efficiently to annotate video footage. It provides automatic feature extraction methods which support the user in navigation through and structuring of the video content. An annotation at image region level is supported by object redetection and tracking functionalities. The result content description of the footage is saved in MPEG-7.

Keywords: MPEG-7, Video Annotation Tool, Automatic Feature Extraction, Metadata
Categories: H.3, I.4

1 Introduction

A detailed description of the video content can be used for different purposes. It is especially required for content retrieval, film analysis, and as input data for generating interactive video content. The annotation of video footage is a time-consuming task. We have developed annotation tools which enables the efficient annotation in combination with a set of automatic annotation plug-ins. All metadata of the tools is stored in the ISO standard MPEG-7 [ISO/IEC, 01] [MPEG-7 Lib, 07]. The annotation workflow can be seen in ***Figure 1***. The pre-processing tool extracts metadata automatically from the video. After that this data can be edited and extended semi-automatically by the MPEG-7 Video Annotation Tool.

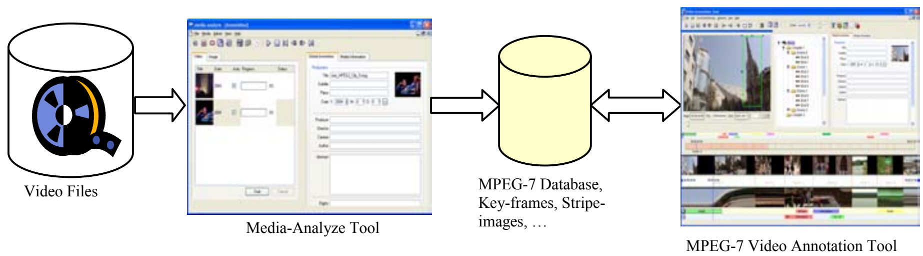


Figure 1: Annotation workflow

2 Media Analyze – Pre-processing Tool

In a first preparation step of the annotation workflow the global description of the source video data has to be specified and an automatic analysis process has to be started. For this a separate tool, the Media Analyze Tool, has been developed. During the automatic content analysis the characteristic camera motion, shot boundaries including dissolves, relevant key-frames, several image similarity features, and the visual features which are required for searching for similar image regions in the video are extracted. The result is a first metadata description of the video which is stored in MPEG-7. The automatic analysis can be started for several videos and then the process can run over night.

The core part of this program is a module framework. The framework serves as an execution environment for analysis modules. The modules are interconnected by the framework to constitute a so-called module graph (see Figure 2). The module graph is defined by a XML file.

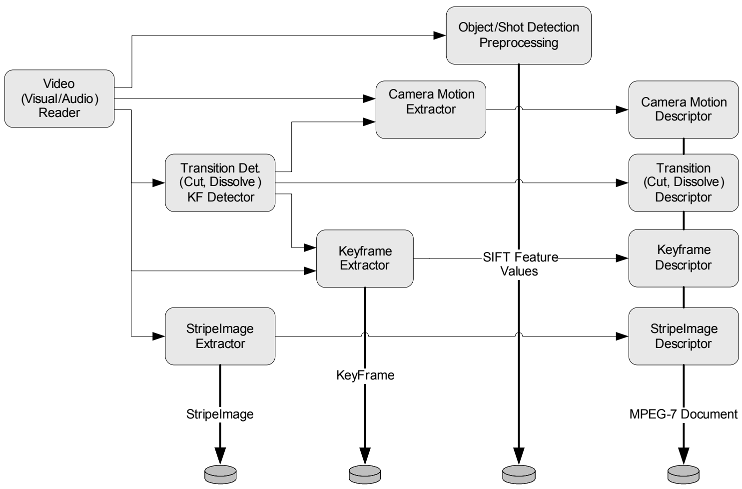


Figure 2: Analyses module graph of the MediaAnalyze Tool

3 MPEG-7 Video Annotation Tool

As soon as the automatic analysis of a video is completed the produced metadata description can be displayed, edited and extended by the MPEG-7 Video Annotation Tool (see Figure 3). This tool has a number of views which enables fast and easy navigation in the video. Through the key frames and the stripe image one gets a quick overview of the video content. There are two time lines, one for the whole video time and one which shows only a selected time period (time zoom). In the time lines the shot boundaries and the dissolves are displayed and they can also be edited. There is also the possibility to structure the video depending on the video content. For example shots can be grouped to scenes; scenes can be combined to build chapters and so on. This structure yields a kind of table of contents and is displayed by a separate view. Depending on the selected structural element different textual annotations are possible. These are for example the title of the structural element, content description, remarks, and specifications about time, location, and persons. At the shot level shooting settings like camera motion, camera angle, or view size can be documented.

The integrated video player has drawing functionalities for the annotation of regions (objects). Image regions can be specified by drawing a rectangle or a polygon. Once an image region is specified (see green rectangle in Figure 3) it is possible to start an automatic search for a similar region in other shots of the video. The result of the object search is displayed in a separate key-frame view. The object redetection functionality is very fast because it uses the pre-processed visual feature values [OBER, 05]. By using the object search result textual annotations can be assigned simultaneously to several objects or shots.

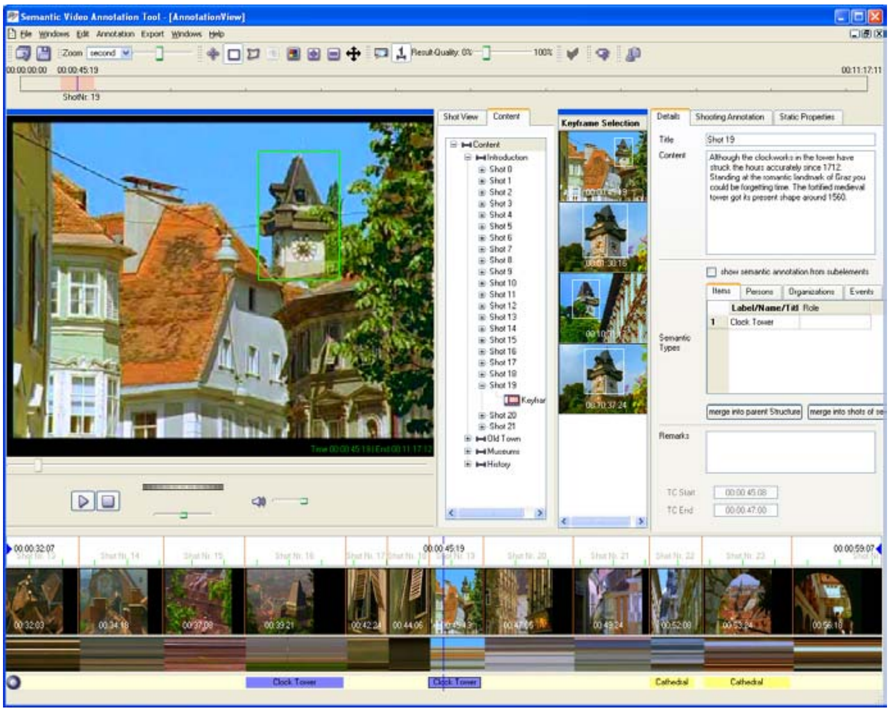


Figure 3: User interface of the MPEG-7 Video Annotation Tool

4 Conclusions

A video annotation system has been presented which allows creating a MPEG-7 content description in different levels of detail (e.g. scene, shot, key-frame, object).

References

[MPEG-7 Lib, 07] MPEG-7 library: http://mpeg-7.joanneum.at

[ISO/IEC, 01] ISO/IEC 15938:2001

[OBER, 05] Ober Sandra, Neuschmied Helmut, Bischof Horst, Object Recognition and Video Indexing for Media Analysis, OEGAI Journal, 24 (1), pp. 4-10, 2005

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

PrestoSpace Publication Platform: A System for Searching and Retrieving Enriched Audiovisual Materials

Walter Allasia, Massimiliano Porzio, Michele Vigilante
(EURIX - Torino, Italy
{allasia, porzio, vigilante}@eurixgroup.com)

Laurent Boch, Giorgio Dimino, Alberto Messina
(RAI-CRIT - Torino,Italy
{l.boch, g.dimino, a.messina}@rai.it)

Abstract: We present the Publication Platform, a component of the PrestoSpace[1] project, which provides retrieval and browsing functionalities of enriched audio-visual material. The Prestospace Factory is a system for enriching audio-visual documents in order to provide automated content and semantic analysis.The Publication Platform provides a user interface for semantic queries and produces a Web page with the results of the AV analysis and additional information about related external documents.

Key Words: content analysis, information retrieval, digital libraries, multimedia information systems

Category: H.3.1, H.3.3, H.3.7, H.5.1

1 Introduction

In the past years the value of audiovisual archives has been rediscovered by broadcasters and new researches have shown that approaches meant to the recovery and availability of archived materials may produce consistent cost savings in the overall programme production processes. In order to achieve this goal, it is essential to adopt metadata and find which kind of information can be used by archive users to retrieve audiovisual items or information concerning the content.

The aim of the Prestospace project [Messina et al. 2006] is to provide technical devices and systems for digital preservation of all types of audio-visual collections. In this paper, we describe the Publication Platform, a component of the PrestoSpace project having the objective of offering suitable mechanisms for retrieving and accessing audiovisual contents based on metadata.

2 System Architecture

The Publication Platform exploits the *EDOB* (EDitorial OBject) format, which is based on MPEG-7[2] and P_META[3] and was specifically created in order to

---
[1] http://www.prestospace.org
[2] ISO/IEC 15938:2001
[3] http://www.ebu.ch/metadata/pmeta/v0102/html

handle the metadata available within the PrestoSpace project. This platform is used to browse enriched digital items coming from the archives of broadcasting companies, although it is suitable for browsing any kind of AV material. It is made up of three different main subcomponents: a Web application, namely the user interface (described in Section 3); a relational DBMS that stores information related to the available programmes; a text search and indexing engine, including a semantic engine for processing natural language queries.

3 The User Interface

The user interface of the Publication Platform allows for several searching approaches. The semantic queries are performed using the KIM (Knowledge and Information Management) platform, described in Section 3.1. The user can submit a keyword or a sentence and perform the search by contribution, title, publication date, publication service, topic and named entities. The submitted queries can be also translated to a different language using the CLIR server (see Section 3.2). Finally the results of the query are shown in the form of a list which can be used by the user to select a document to be browsed in a Web page using the EDOB viewer, described in Section 3.3.

3.1 The KIM Platform

The KIM[4][Popov et al. 2003] platform provides a novel Knowledge and Information Management infrastructure and services for automatic semantic annotation, indexing, and retrieval of unstructured and semi-structured content. As a base line, KIM analyzes texts and is able to recognize references to named entities (like persons, organizations, locations). It should be noted that also the indexing engine mentioned in Section 2 is based on the KIM platform, which provides an engine with indexing capability for the EDOB metadata and therefore it is used in the retrieval phase for full text or semantic queries.

3.2 The CLIR Server

The user can enable a feature to translate the submitted query from the Italian language to the English language and vice versa. Source language of queries can be different from the target language (characterizing metadata). Cross-language Information Retrieval (CLIR) is supported in the Publication Platform by a specific server which adopts the Sense Disambiguation and Translation technique. The description of the Sense Disambiguation algorithm which has been adopted within Prestospace can be found in [Basili et al. 2006]. The system provides two

[4] http://www.ontotext.com/kim/

services: the analysis of the query and extraction of the named entities within the query itself; the translation of common nouns present in the query using the mentioned algorithm. For the best usage, the system requires the submission of a sentence with known sense instead of single keywords.

3.3 The Edob Viewer

The Web page produced as a result of the query process contains a section with video previews and a tree structure showing the segmentation of the programme in editorial parts. This segmentation is also shown as a timeline and it is obtained by an automated content and semantic analysis performed during enrichment phase of the process. The remaining part of the Web page provides several tabs showing the identification information (such as titles, publications, contributions and identifiers), the text of the extracted speech obtained by an Automatic Speech Recognizer (ASR), the results of the semantic analysis, the analysis of the contents analysis and further information about related sources, such as links to Web pages containing external documents whose subject is related to that of the editorial parts.

4 Conclusions

The Publication Platform, with the aim of providing services for semantic analysis and cross language information retrieval, is a valid solution for making more accessible the content of audiovisual digital libraries.

Acknowledgments

The work reported in this paper has been partially supported by the PrestoSpace (www.prestospace.org) project, as part of the Information Society Technologies (IST) Program of the European Commission (Contract IST-507336).

References

[Messina et al. 2006] A. Messina, L. Boch, G. Dimino, W. Bailer,P. Schallauer,W. Allasia, M. Groppo,M. Vigilante, R. Basili , Creating Rich Metadata in the TV Broadcast Archives Environment: the PrestoSpace Project, Proceedings of 2nd International Conference on Automated Production of Cross Media Content for Mulitchannel Distribution, Leeds, (UK), 2006.

[Popov et al. 2003] Borislav Popov, Atanas Kiryakov, Angel Kirilov, Dimitar Manov, Damyan Ognyanoff, Miroslav Goranov, KIM Semantic Annotation Platform, 2nd International Semantic Web Conference (ISWC2003), 20-23 October 2003, Florida, USA. LNAI Vol. 2870, pp. 834-849, Springer-Verlag Berlin Heidelberg 2003.

[Basili et al. 2006] R. Basili M. Cammisa, A. Gliozzo, Integrating Domain and Paradigmatic Similarity for Unsupervised Sense Tagging, Proceedings of the European Conference on Artificial Intelligence, Riva del Garda, (Italy), 2006.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

VAMP: Semantic Validation of MPEG-7 Profiles

Martin Höffernig, Michael Hausenblas, Werner Bailer
JOANNEUM RESEARCH Forschungsgesellschaft mbH
Institute of Information Systems and Information Management
Steyrergasse 17, 8010 Graz, Austria
{firstname.lastname}@joanneum.at

Raphaël Troncy
CWI Amsterdam, The Netherlands
raphael.troncy@cwi.nl

Abstract: This paper describes the VAMP web application for the validation of MPEG-7 descriptions with respect to semantic constraints defined in a profile. The semantic constraints are formalised using an ontology and a set of rules.
Key Words: Metadata, Ontology, Semantic Web, MPEG-7, Validation
Category: H.5.1

1 Introduction

MPEG-7 [MPEG7 2001], formally named Multimedia Content Description Interface, can be used to create complex and comprehensive metadata descriptions of multimedia content. To reduce the syntax variability, MPEG-7 has introduced the notion of *profiles* that constrain the way multimedia descriptions should be represented for particular applications. Since MPEG-7 and profiles are defined in terms of an XML schema, the semantics of its elements has no formal grounding. An approach for expressing this semantics explicitly by formalising the constraints of a profile using ontologies and logical rules is presented in [Troncy et al. 2006]. The use of the MPEG-7 descriptors in a particular context can thus be specified and validated.

VAMP[1] is a Semantic Web Application for validating the conformance of MPEG-7 documents to the semantics of a given profile. The idea and the implementation of VAMP is described in [Troncy et al. 2007].

2 General Workflow

Given a MPEG-7 document, VAMP validates whether it conforms to a selected profile or not. First, the MPEG-7 input document is checked for syntactic validity against the Profile XML Schema. Second, the MPEG-7 description is converted

[1] http://vamp.joanneum.at

into RDF with respect to an ontology capturing the semantics of the selected profile. Since not all of the semantic constraints can be described by an (OWL) ontology, logical rules are also used for representing them. Finally, these RDF triples are the input data for the semantic consistency check of the knowledge base containing the ontology and the logical rules.

In contrast to [Garcia and Celma 2005, Tsinaraki et al. 2004, Hunter 2001, Arndt et al. 2007], we do not intend to completely map the MPEG-7 description tools onto an OWL ontology, but rather use Semantic Web technologies to represent those MPEG-7 semantic constraints defined in natural language that cannot be expressed using XML Schema. Our approach is therefore complementary to these other attempts for formalizing MPEG-7.

3 The VAMP Web Application

VAMP is available as a web interface for humans (Figure 1), and as a REST-style[2] Web service for agents.

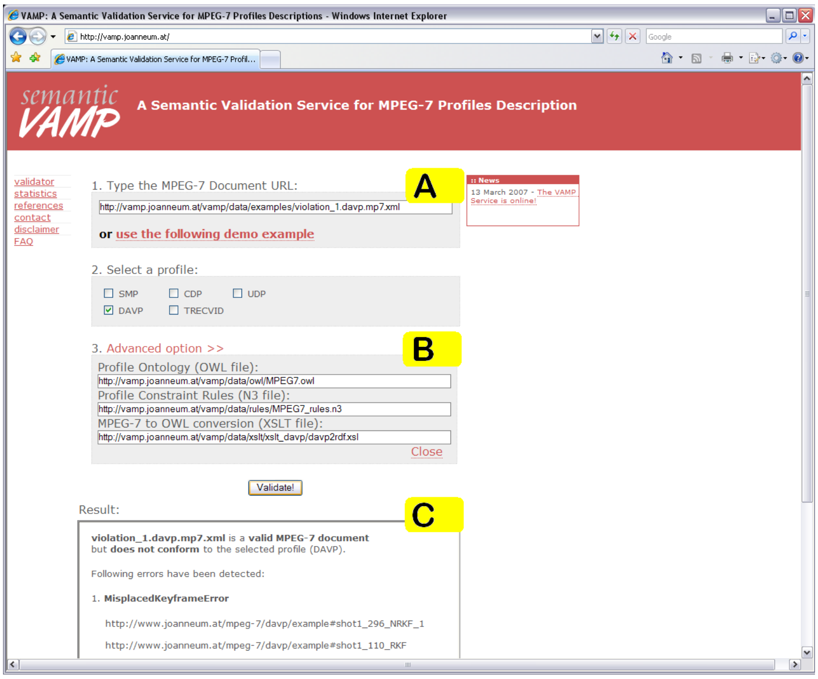


Figure 1: The VAMP Web interface.

[2] `http://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm`

The user enters the URI of the description to be validated (A). In an advanced mode, optional parameters corresponding to an alternative formalisation of the semantic constraints can be entered (B). The `Validate` button provides a meaningful explanation of the errors detected in the description (C).

At the time of writing parts of the Detailed Audiovisual Profile (DAVP) [Bailer and Schallauer 2006] are formalised and available in VAMP. Formalisations of other profiles (SMP, CDP, UDP) and de-facto profiles (such as the TRECVID format) are planned in the near future.

Acknowledgments

The research leading to this paper was partially supported by the European Commission under the contracts FP6-027026, "Knowledge Space of semantic inference for automatic annotation and retrieval of multimedia content - K-Space" and IST-2-511316, "IP-RACINE: Integrated Project - Research Area CINE".

References

[Arndt et al. 2007] Richard Arndt, Raphaël Troncy, Steffen Staab, and Lynda Hardman. Adding Formal Semantics to MPEG7: Designing a Well-Founded Multimedia Ontology for the Web. Technical Report KU-N0407, University of Koblenz-Landau, 2007.

[Bailer and Schallauer 2006] Werner Bailer and Peter Schallauer. The Detailed Audiovisual Profile: Enabling Interoperability between MPEG-7 based Systems. In *12th International MultiMedia Modelling Conference (MMM'06)*, pages 217–224, Beijing, China, 2006.

[Garcia and Celma 2005] Roberto Garcia and Oscar Celma. Semantic Integration and Retrieval of Multimedia Metadata. In *5th International Workshop on Knowledge Markup and Semantic Annotation (SemAnnot'05)*, Galway, Ireland, 2005.

[Hunter 2001] Jane Hunter. Adding Multimedia to the Semantic Web - Building an MPEG-7 Ontology. In *First International Semantic Web Working Symposium (SWWS'01)*, Stanford, California, USA, 2001.

[MPEG7 2001] MPEG-7. Multimedia Content Description Interface. ISO/IEC 15938, 2001.

[Troncy et al. 2007] Raphaël Troncy, Werner Bailer, Michael Hausenblas, and Martin Höffernig. VAMP: Semantic Validation for MPEG-7 Profile Descriptions. Technical Report INS-E0705, Centrum voor Wiskunde en Informatica (CWI), 2007.

[Troncy et al. 2006] Raphaël Troncy, Werner Bailer, Michael Hausenblas, Philip Hofmair, and Rudolf Schlatte. Enabling Multimedia Metadata Interoperability by Defining Formal Semantics of MPEG-7 Profiles. In *1st International Conference on Semantics And digital Media Technology*, pages 41–55, Athens, Greece, 2006.

[Tsinaraki et al. 2004] Chrisa Tsinaraki, Panagiotis Polydoros, and Stavros Christodoulakis. Interoperability support for Ontology-based Video Retrieval Applications. In *3rd International Conference on Image and Video Retrieval (CIVR'04)*, Dublin, Ireland, 2004.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Virtual Campfire – A Mobile Social Software for Cross-Media Communities

Yiwei Cao, Marc Spaniol, Ralf Klamma and Dominik Renzel
(Lehrstuhl Informatik 5, RWTH Aachen, Germany
{klamma|spaniol|cao|renzel}@i5.informatik.rwth-aachen.de)

Abstract: Multimedia creation, annotation and sharing are challenging tasks especially of interdisciplinary, intercultural and intergenerational communities. We present the mobile social software *Virtual Campfire* to provide cross-media and cross-community support for de- and re-contextualization of multimedia content, employing Web Services, the MPEG-7 standard and Web 2.0 technologies etc. Virtual Campfire can enable communities to set up and maintain multimedia community information systems quickly and easily.

Key Words: Information system, Multimedia, MPEG-7, Web Service, Social Software, Web 2.0, Mobile computing, Community of Practice.
Categories: H5.1 Multimedia Information Systems, H2.4 Systems.

1 Introduction

Virtual Campfire is an approach to providing cross-media and cross-community support for the de- and re-contextualization of multimedia content. Virtual Campfire serves as a framework for various web services enabling communities to share knowledge on the Web 2.0. It employs advanced multimedia technologies in order to professionally support interdisciplinary, intercultural as well as intergenerational communities. We present a selection of services on the basis of the Lightweight Application Server (LAS) serving as the backbone of the Virtual Campfire framework to show its applicability in various application scenarios.

2 Objectives

Our research is framed by three distinct projects covering the many facets of multimedia processing for communities of practice:

- The Aachen based research cluster established under the excellence initiative of the German government "Ultra High-Speed Mobile Information and Communication (UMIC)" (www.umic.rwth-aachen.de)
- The interdisciplinary collaborative research center of the German Science Foundation on "Media and Cultural Communication" (www.fk-427.de)
- The EU IST Network of Excellence in Professional learning focusing on technology enhanced learning PROLEARN (www.prolearn-project.org).

In Virtual Campfire, we focus on two specific aspects in multimedia processing. First, we incorporate multimedia content interoperability by employing the MPEG-7 multimedia metadata standard. Second, we provide a framework for multimedia-centric services that help professionals create community information systems in order to foster multimedia enhanced knowledge sharing in communities.

3 What Will Visitors See?

We present the following MPEG-7 compliant multimedia services of Virtual Campfire based on LAS [cf. Figure 1]:

- Interdisciplinary, intercultural and intergenerational learning support by geographical hypermedia services [Klamma et al. 06].
- MPEG-7 compliant tagging fostering the management and access of multimedia content by semantic tagging [Spaniol et al. 06a].
- Media integrated storytelling to support learning from digital stories in communities of practice [Spaniol et al. 06b].

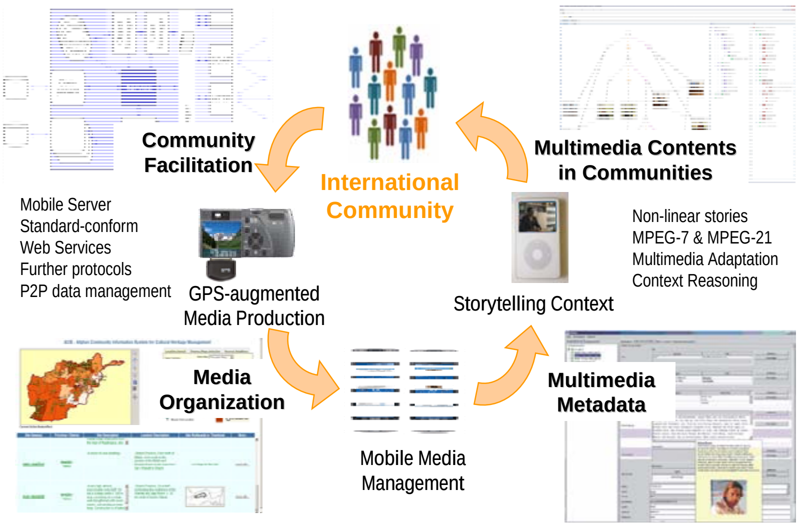


Figure 1: Scenario of Multimedia Service usage in Virtual Campfire

In detail, Virtual Campfire enables communities like cultural heritage communities, multimedia communities, learning communities, and geospatial communities to create, annotate and share multimedia content. Virtual Campfire provides a set of services including mainly the user and role access management service, the FTP service, the geo-tagging service, the feature extraction service, the multiple database connector service, the MPEG-7 service and the story service.

All these services are deployed on LAS, which is a platform-independent lightweight middleware server implemented in Java. It has been developed for the purpose of providing web services which can be shared among various tools in order to support the work of communities in practice. The LAS architecture is based on the

LAS Connectors using HTTP and SOAP, the LAS Components as the minimal functionality units, and the LAS Services easily and directly used by the client side. The LAS Java API and its concepts can be easily used to extend the server functionality.

To summarize, within Virtual Campfire communities are able to produce images, audios and videos by mobile devices with the relevant location information. The multimedia content could be annotated by MPEG-7 compliant tagging services. Stories can be created by selected pre-annotated multimedia content. These stories are suitable and interesting learning contents for the cross-media communities on various platforms e.g. the Apple iPod.

4 What Is Innovative About This Exhibit?

Virtual Campfire offers an open architecture that helps professionals flexibly create information systems in versatile application domains. On the basis of LAS it combines advanced multimedia standards and database technologies like spatial databases and XML databases facilitating the creation and maintenance of state-of-the-art (mobile) information systems for MPEG-7 compliant multimedia support. Furthermore, Virtual Campfire shows a new way to bridge between Web 2.0 tagging as well as folksonomy-based approaches and advanced multimedia standards such as MPEG-7 and Dublin Core approaches leading to location-based semantic multimedia annotations. Meanwhile, it provides communities more opportunities to create, access, share and even reuse multimedia content in the Web 2.0 era.

Acknowledgements

This work was supported by the excellence initiative of the German government within the research cluster “Ultra High-Speed Mobile Information and Communication (UMIC)”, by the German National Science Foundation (DFG) within the collaborative research center SFB/FK 427 “Media and Cultural Communication”, and by the 6th Framework IST programme of the EC through the NoE on Professional Learning (PROLEARN) IST-2003-507310. We thank our students Dimitrios Andrikopoulos, Pham Manh Cuong, Andreas Hahne, Holger Janßen, and Monika Pienkos for the system development.

References

[Klamma et al. 06] R. Klamma, M. Spaniol, M. Jarke, Y. Cao, M. Jansen and G. Toubekis: Standards for Geographic Hypermedia: MPEG, OGC and co. E. Stefanakis, M.P. Peterson, C. Armenakis, V. Delis (Eds.): Geographic Hypermedia - Concepts and Systems, LNG&C, ISBN 3-540-34237-0, Springer-Verlag, pp. 233-256, 2006.

[Spaniol et al. 06a] M. Spaniol, R. Klamma, H. Janßen and D. Renzel: LAS: A Lightweight Application Server for MPEG-7 Services in Community Engines. K. Tochtermann, H. Maurer (Eds.): Proceedings of I-KNOW '06, Graz, Austria, September 6 - 8, 2006, J.UCS (Journal of Universal Computer Science) Proceedings, Springer, pp. 592-599.

[Spaniol et al. 06b] M. Spaniol, R. Klamma, N. Sharda and M. Jarke: Web-Based Learning with Non-linear Multimedia Stories. W. Liu, Q. Li, R. W. H. Lau (Eds.): Advances in Web-Based Learning, Proceedings of ICWL 2006, Penang, Malaysia, July 19-21, Springer-Verlag, Berlin Heidelberg, LNCS 4181, pp. 249-263, 2006.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Spatial Applications as Social Knowledge Provider. An Approach for Regional Virtual Communities

Daniela Kuka, Günther Kolar, Michael Badics, Christa Schneebauer
Eva Kühn, Stefan Eibelwimmer, Fadil Kujundzic, Nicolas Naveau
Clemens Häckel
(Ars Electronica Linz GmbH, Austria
firstname.secondname@aec.at)

Abstract: Existing applications in visualization of blogs and online conversations have largely focused on the discovery of social networks and knowledge flows in the online space. At the same time, another class of visual spatial applications allows us to uncover the relationships between people and real space. This can be used to infer how individual and collective behaviour is not only a function of social contexts but also spatially situated in real physical environments. This paper presents a socio-technical approach and a prototype public tool integrating these two aspects in a synergetic way: (1) using social software to facilitate community forming in specific regional, spatial application scenarios and (2) eliciting new knowledge from community-based interaction processes in spatial settings. The prototypical implementation in a case-study in Upper Austria demonstrates how the developed socio-technical tool enables new forms of cross-municipal communication and interaction which blend together physical, informational and social dimension of space.

Keywords: Social Software, Spatial Applications, Digital Communities, Conversation Visualization, Digital Cartography, Collaborative Mapping, Cross-Municipal Communication
Categories: H.5, H.5.2, H.5.3, J.5, K.4.1, K.4.2

1 Introduction

This paper presents a socio-technical approach and a prototype public tool integrating two aspects in a synergetic way: the discovery of social networks and knowledge flows by the visualization of blogs and online conversations and the usage of spatial applications to uncover relationships between people and real space.

The project places emphasis on (1) using social software to facilitate community forming in specific regional, spatial application scenarios and (2) eliciting new knowledge from community-based interaction processes in spatial settings. Thus, we are demonstrating the work on a public tool that enables new forms of cross-municipal communication and interaction.

2 Approach

To improve the value of online-platforms in regional contexts, we merged a specific *cartographic mapping* and a specific *conversation visualization* approach. Synthesizing online networking, geographical localization and visualization methods [see e.g. Donath, 99 and Lima, 05] should create intuitive access to the interaction patterns and the user-generated collective knowledge of a regional community.

2.1 Collaborative Mapping and Online Conversation

The method of resolution for the prototype is to connect the elaborated functions of a conventional internet forum with an additional module to embed a geographical (regional) dimension as reference and orientation framework for the community. *The connection of a cartographic map (static and closed) with online conversation (dynamic and ongoing) is conceptualized and designed to generate additional information of the relationship between geographic frames of reference and the occurring communication and interactions processes of individual user and user groups.* Enabling the interrelation between spatial information and user generated content (collaborative mapping) reflects the localization and structure of the community and its representation as *network of people* with specific backgrounds. Geographical frames are related to specific discourses that are relevant especially for the people in this area.

2.2 Visual Studies and Online Conversation

The combination of the usage of the cartographic material and the visualization patterns, here realized as a visualization of the communication activities on the map, allow "*to make the complex accessible*" and "*to make the hidden visible*" – the geographical distribution of members and their degree of involvement into specific subject-matters concerning the social issues of a region. Mapping, thus, is a designer's task to invent strategies for information visualization enabling new interpretations [see Abram et al., 06]. Using a map as tool for supporting online conversation necessitates implementing a dynamic quality that is realized by dynamic visualization of the user's activity. Thus, the difference to actual forms of visualizing online communication or blog activity is that it is based on a synthesis of cartographic mapping and visual analysis.

The self reflexive practices activated by a visualization on a map are essential for the collective identity of the community and, moreover, the dynamic expression of social activity as well as its interpretation. Referring to the categories of [Judelman, 04], both complexity (list / all articles of a thread) and interrelationships (location, article, user) are visualized.

3 Visualization Patterns in the Prototype

The data base for the visualization is automatically generated from both the user profile (registration data) and the user's activities in the forum (number of posts per thread). The user profile contains the localization of single participants as well as special group belonging(s) relevant for the forum. Two types of visualization, the visualization of the discussion and of the community, are realized: The user can switch between a text forum with common functionalities and a visualization mode that is displaying the discussion in a thread on the scalable map (figure 1). Items with different coordinates and sizes are representing the local position as well as the intensity of participation of a participant. This display format is reflecting the geographical dissemination of the subject-matters and allows judging the spreading and relevance of a discussion with regard to the intensity of participation.

The visualization of the community allows intuitively identifying the geographical distribution of the members as well as the density of participation. A more differentiated image of the digital community is available by several filter functions selected by individual parameters. By changing the maps (e.g. political/physical and statistics maps), additional knowledge about a specific community frame can be gained due to e.g. a comparison of the community's activity on a specific topic and the appropriate statistic data for the section. The generated image of the spatiotemporal progression of the online-discussion allows to cross-read the contents in the context of municipal specific statistical maps.

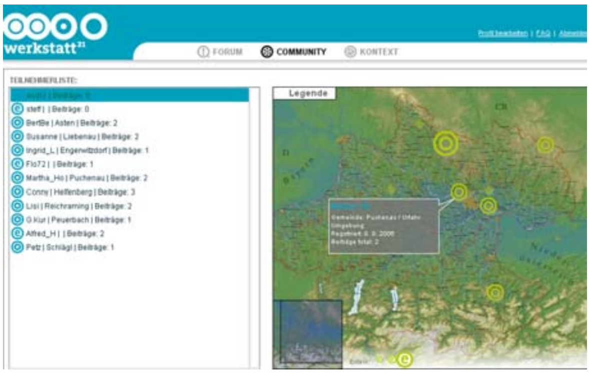


Figure 1: Snapshot of the Discussion Visualization in the PROvisionS Prototype

4 Synthesis and Conclusions

The result is a spatial application [see Hockenberry, 06] that enables, as public tool, new forms of cross-municipal communication and interaction and elicits new knowledge from this community setting both for the users themselves and for external research. Starting from scientific research and project evaluations, we've chosen the approach to combine a specific *collaborative mapping approach based on cartographic material* and a specific *visualization approach representing asynchronous online discussions both in structure and content.* Due to the cartographic mapping approach, it was possible to create a close link between the real life experience and the digital space of the community.

Due to the visualization approach it was possible to redesign linear-chronological text based communication supported by the most conventional online platforms and suffering from insufficient transparency, incoherent accumulation of huge masses of data etc. Users can access the hidden knowledge reservoirs of the community (structure, history and localization of the online discussion) at the first glance. The combination of both approaches opens up a democratic *place* facilitating self-reflection and identification within specific regional context. Knowing about the origin of posts and annotations leads to a coherent image of the real social environment and the focal point of topic discussion. Due to democratic principles and easier identification of collective interests and fields of action, online communication within regional contexts facilitates to quickly shift from knowledge to activity and vice versa. The visualization functions also serve as ‘diagnostic tool’ for different branches, scenes and milieus.

5 Future Prospects

The visualization of interaction patterns shall be enhanced to facilitate network analysis and graphic/visual representations of the contents of a discussion. In this context, the Ars Electronica Futurelab started collaborations to develop tools synthesizing the current results with semantic and self-learning functions.

Acknowledgements

The R&D of two prototypic applications of 'Social Software as tool for cross-municipal communication processes: a community network for Upper Austrian citizens” was funded within proVISION, a programme of BMWF Ministry of Science and Technology, Austria.

References

[Abram et al., 06] Abram, J., Hall, P.: “ELSE/WHERE: MAPPING. New Cartographies of networks and territories”; University of Minnesota Design Institute, Minneapolis (2006), 12

[Donath, 99] Donath, J. et al.: “Visualizing conversations”; Proc. of HICSS-32, reprinted in the Journal of Computer Mediated Communication 4, 4 (1999), also appeared as electronic version, http://smg.media.mit.edu/papers/VisualConv/VisualizeConv.pdf

[Hockenberry, 06] Hockenberry, M.: „Grounding for a Computational Model of Place”; MIT MS Thesis, Massachusetts Institute of Technology, Massachusetts (2006), http://pubs.media.mit.edu/pubs/papers/hock-ms.pdf

[Judelman, 04] Judelman, G.B.: "Knowledge Visualization. Problems and Principles for Mapping the Knowledge Space"; M.Sc. Thesis (2004), http://www.gregjudelman.com/thesis.html

[Lima, 05] Lima, M.: “blogviz - Mapping the dynamics of Information Diffusion in Blogspace”; Master Thesis (May 2005), http://www.blogviz.com/blogviz/

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

I-SEMANTICS '07

International Conference on Semantic Systems

Editors

Tassilo Pellegrini
Sebastian Schaffert

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

The NEPOMUK Project - On the Way to the Social Semantic Desktop

Tudor Groza, Siegfried Handschuh, Knud Möller
(DERI, National University of Ireland, Galway, Ireland
tudor.groza@deri.org, siegfried.handschuh@deri.org, knud.moeller@deri.org)

Gunnar Grimnes, Leo Sauermann
(DFKI Kaiserslautern, Germany
gunnar.grimnes@dfki.de, leo.sauermann@dfki.de)

Enrico Minack
(L3S Hannover, Germany, minack@l3s.de)

Mehdi Jazayeri, Cédric Mesnage
(Faculty of Informatics, University of Lugano, Switzerland
mehdi.jazayeri@unisi.ch, cedric.mesnage@lu.unisi.ch)

Gerald Reif
(University of Zurich, Switzerland, reif@ifi.unizh.ch)

Rósa Gudjónsdóttir
(Royal Institute of Technology, KTH, Sweden, rosag@kth.se)

Abstract: This paper introduces the NEPOMUK project which aims to create a standard and reference implementation for the Social Semantic Desktop. We outline the requirements and functionalities that were identified for a useful Semantic Desktop system and present an architecture that fulfills these requirements which was acquired by incremental refinement of the architecture of existing Semantic Desktop prototypes. The NEPOMUK project is primarily motivated by three real-life industrial use-cases, we briefly outline these and the processes used to extract required functionalities from the people working in these areas today, and we present a selection of typical tasks where the Semantic Desktop could be of benefit.
Key Words: Semantic Desktop, Personal Information Management, Semantic Middleware
Category: H.3.7, H.5.4

1 Introduction

In traditional desktop architectures, applications are isolated islands of data – each application has its own data, unaware of related and relevant data in other applications. Individual vendors may decide to allow their applications to interoperate, so that e.g. the email client knows about the address book. However, today there is no consistent approach for allowing interoperation and a system-wide exchange of data between applications. In a similar way, the desktops of different users are also isolated islands - there is no standardized architecture for interoperation and data exchange between desktops. Users may exchange data by sending emails or upload it to a server, but so far there is no way of seamless communication from an application used by one person on their desktop to an application used by another person on another desktop.

The problem on the desktop is similar to that on the Web. On the Web we are faced with isolated data islands, and also as on the desktop there is not yet a standardized approach for finding and interacting between applications.

The Social Semantic Desktop (SSD) paradigm adopts the ideas of the Semantic Web, which offers a solution for the Web. Formal ontologies capture both a shared conceptualization of desktop data and personal mental models. RDF serves as common data representation. Web Services - applications on the Web - describe their capabilities and interfaces in a standardized way and thus become Semantic Web Services. On the desktop, applications (or rather: their interfaces) will therefore be modeled in a similar fashion. Together, these technologies provide a means to build the semantic bridges necessary for data exchange and application integration. The SSD will transform the conventional desktop into a seamless, networked working environment, by loosening the borders between individual applications and the physical workspace of different users.

The aim of the NEPOMUK project[1], described in this paper, is to provide a standardized description of a SSD architecture, independent of any particular operating system or programming language. Reference implementations will show the feasibility of the standard. The paper is structured as follows: we start with Section 2 by describing the engineering cycle we follow in the project. Then we detail in Section 3 scenarios captured from real-world case-studies and in Section 4 a list of functionalities extracted from these scenarios. Section 5 depicts the current version of the NEPOMUK SSD Architecture, while Section 6 shows related approaches for building the SSD. In Section 7 we state our conclusions.

2 NEPOMUK **Engineering Cycle**

The NEPOMUK project relies heavily on existing software developed by the partners. On the other hand, usability research is being held with the case studies partners by interviewing potential users of the SSD. This specific set up of the project led us to develop our engineering cycle (Figure 1). This cycle represents the way we intend to merge the existing technologies and the needs from users.

Clockwise, Figure 1 shows the forward engineering cycle. We analyzed the end-user's intended *usage* of the SSD, studied the different use cases and formulated them into *scenarios*. We generalized the individual scenarios and extracted the common *functionalities* that make up the SSD. These functionalities formed

[1] http://nepomuk.semanticdesktop.org/

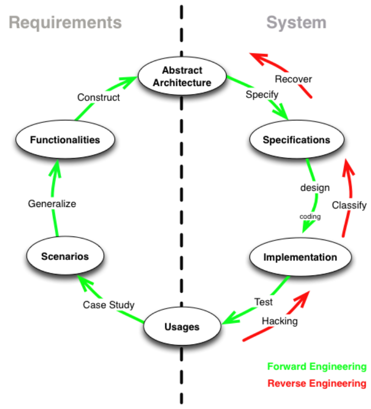


Figure 1: NEPOMUK Engineering Cycle

the basis to define the reference *architecture* which in turn lead to the service *specification* and *implementation* that is tested by the end-users. On the other hand, partners already started to hack components that are likely to be needed by the SSD or had component developed before the NEPOMUK project started. Therefore, we reverse engineered these components to get their specifications and used the gained experience when defining the architecture.

The construction of the architecture of the SSD is therefore the combination of different parts, (1) the requirements and objectives from the vision of the SSD driving the NEPOMUK project, (2) the functionalities from user studies (forward engineering), and (3) the service specifications of existing implementations (reverse engineering). The overlaps between these areas give us confidence in the needs. Combining these three sources results in a complete architecture. Thus the architecture represents a shared understanding of all partners involved in the project and we see it as a roadmap towards the realization of the SSD.

In the next sections we present some of the scenarios which we considered as being particularly representative for the SSD paradigm, then we show the list of functionalities abstracted from the user study material.

3 Scenarios

The study of user needs regarding collaboration on the SSD is a major goal of the NEPOMUK project. User studies were carried out in the project at the case study partner sites, which are companies and research labs working in the area of business software, biomedical research, Linux development, and management consulting. The type of work performed varies between the case study partners, but what they have in common is the fact that the employees are knowledge workers, receiving, interpreting and structuring information on a daily basis. The purpose of the user studies was to understand the work environment in order to develop a SSD that meets the knowledge workers' needs and requirements. 40 contextual interviews [4] and seven video brainstorming workshops [9] were performed with employees at the different partner sites. To document the resulting user requirements 14 personas and 40 usage scenarios [6, 7] were created, illustrating the user needs, desires and expectations on the SSD. Personas are fictitious persons that represent different user groups and are always based data collected in user studies. A persona and a related scenario where the persona uses the SSD that we will develop, is an effective way to illustrate how the users want the SSD to operate. In this section we summarize a collection of the usage scenarios with the help of our primary personas.

Dirk gets task from Claudia. Claudia is working on a project deliverable and she identifies tasks to be done. She adds the tasks to the project and assigns them to Dirk. Dirk is notified of his new tasks and he accepts the responsibility for some of the tasks and Claudia is notified. Dirk realises that some tasks require more specific knowledge so he declines them and suggests allocating them to Martin. Claudia reassigns the tasks to Martin.

Josephine follows-up the project plan. Josephine is following up on an active project she is administrating. It involves Karen and a few other trainers. She accesses the project plan and browses to see if everything is on schedule. The project plan is connected to the trainers calendars and all changes in their calendar get fed directly into the project plan.

Karen edits a document with another person. Karen is to give a presentation for an existing client in a couple of days. The presentation is new and the purpose is to sum up a series of training programs she has performed for the client. When she is working on a slide she can see a new graphical layout suggestion by Josephine who is concurrently working on the presentations graphic form. The system allows them to collaborate, make changes, discuss and explain their intentions and thoughts.

Karen shares experience. Karen finds the time to take care of some administration issues. She just finished a project successfully and feels that the experiences should be shared with her colleagues. She opens the course material, marks it Shared, and adds a few keywords to make sure that people interested find the material.

4 Functionalities

In order to integrate the requirements expressed in the scenarios and other materials produced in the case studies we need to use a more formal approach. All the material was processed by a group of members of the project coming from different areas: developers, case study partners, architects and usability designers. The results of this workshop is an homogeneous list of functionalities required to satisfy the scenarios. For each functionality, we provide a name, a short textual description, inputs, outputs and the relevant material in which this functionality was discovered. We grouped these into five cluster:

Search enables users to *search* for resources amongst different sources (either locally or on the network). Users also need to find relevant resources by querying by example.

Desktop. On their desktop, users *manage resources*, they use legacy applications to either create or edit documents therefore NEPOMUK needs to *integrate these applications*. NEPOMUK should provide a *notification management* system for the user to receive informations regarding shared resources and configure the ways she is notified. Even when *offline*, users should be able to access relevant resources transparently. We see *desktop sharing* as the ability to share applications or windows.

Profiling by *logging* the user's activity, NEPOMUK should be *trained* to behave according to the specific user's needs. This automatic behaviours must be *tailorable* and include *annotations* and information regarding *trust* with other users or sources(*i.e.,* if a user do not trust an information source, he should not receive results from this source).

Data Analysis. To ease semantic annotation of unstructured documents, such as text, users can use *keyword extraction*. Search results might need to be rearranged using *sorting and grouping*. The use of *reasoning* provides with new information.

Social. At the social level, the *management of groups and users* enhances *social interaction* and ease *resource sharing*. *Access rights management* tackles with the security needs. Users can *publish and subscribe* to relevant stream of information, such as the modifications made to a particular resource or the results of a search.

The discussion around these functionalities lead to the architecture which integrates the user requirements and the SSD vision. This architecture is discussed in the following section.

5 Architecture

In this section we present an overview of the NEPOMUK architecture. The architecture, as show in Figure 2, is organized in three layers. The NEPOMUK SSD is made up by the user's individual desktops which are organized in a peer-to-peer (P2P) fashion. To support the communication between the peers, the lowest layer is the *Network Communication* layer. This layer provides an *Event-based System*, which is responsible for the distribution of the events on between

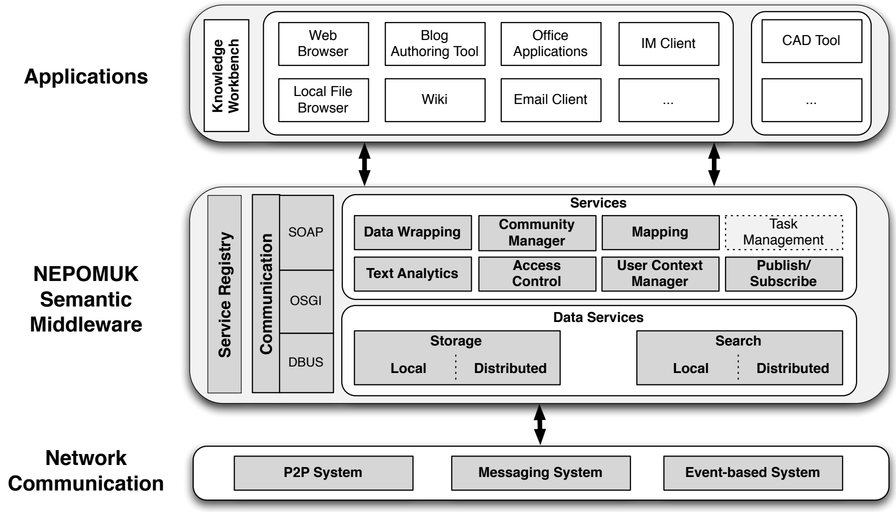


Figure 2: Layered NEPOMUK Architecture.

the NEPOMUK peers. The events carry an RDF graph as payload describing the cause of the event. The *Messaging System* routes the messages to receiver. The *Peer-to-Peer File Sharing System* enables the shared information space. It will be based on GridVine [2]. GridVine is based on P-Grid [1] and provides a distributed index which supports RDQL search queries.

On top of the Network Communication Layer, the NEPOMUK Semantic Middleware provides the core services the NEPOMUK SSD is made up from. The goal of the NEPOMUK project is to propose a reference architecture for the SSD that can be implemented on top of different operating systems such as MS Windows, MacOS, and Linux. Hence, different communication techniques such as SOAP over HTTP, OSGI[2], or D-Bus[3] can be used for interaction between the NEPOMUK services depending on the platform. Therefore, we decided to use WSDL as communication technique and programming language independent interface definition language to specify the service interfaces. The services have to register at the *Service Registry*.

The *Data Services* are responsible to control the insertion, modification, and deletion of resources on the NEPOMUK desktop. A resource can be a user, a document, a calendar entry, an email, and so on. It provides a service to store the RDF meta-data in the *Local Storage*. A resource and their RDF description

[2] OSGi Alliance – `http://www.osgi.org/`
[3] D-Bus – `http://www.freedesktop.org/wiki/Software/dbus`

can either be manually added to the NEPOMUK desktop or the *Data Wrapper* or the *Text Analysis* service extracts the information form desktop applications such as email clients of calendar applications. The Data Wrapper will be used to extract meta-data form structured data sources (*e.g.,* email headers, calendar entries, *etc.*) and will be implemented based on Aperture [3]. The Text Analysis service will be used to extract meta-data form unformatted text (*e.g.,* email bodies, text processor documents, *etc.*). For local queries and for offline working the RDF meta-data is stored in the Local Storage. If a resource is shared with other users in an information space, the meta-data is also uploaded to the distributed index of the P2P file sharing system. The *Search* service can either issue a *local* search in the local storage or a *distributed* search in the underlying P2P system.

Ideally only one ontology exists for a domain of interest such as contact data, calendar events. In reality, however, we are faced with many ontologies of (partly) overlapping domains (*e.g.,* foaf and vCard for contact data). Therefore, the NEPOMUK middleware provides a *Mapping Service* that can be used by other middleware services and services in higher layers to translate RDF graphs from a source ontology to a target ontology.

Actions a user performs on the shared information space have to be approved by the *Access Control* System. Depending on the group membership of a user, maintained in the User/Group Management, the *Community Management* grands the privileges to perform the action. The access rights, the user, and the group data are stored as RDF graphs in the distributed index of the peer-to-peer system. This data is encoded using the access right ontology and the user/group ontology, which belong to the NEPOMUK core ontologies.

The NEPOMUK middleware logs the actions a user performs on the resources on his desktop. The logged data is stored in the Local Storage and is analyzed by the *User Context Manager* to capture the current working context of the user. The working context of the user is used to suggest meaningful actions to the user depending on the task a user is currently working on.

The *Publish/Subscribe System* allows users to subscribe to events in the NEPOMUK system. The subscriptions are stored as SPARQL queries [10] which are matched against the RDF payload of the events. When the subscription, *i.e.,* the SPARQL query, matches the event, the *Messaging System* looks up the preferred notification media (*e.g.,* email, instant messaging, SMS) and delivers the messages. The Messaging System is further used for synchronous and asynchronous communication between NEPOMUK users.

The NEPOMUK Middleware provides the core services of the NEPOMUK architecture. These services can be accessed via the NEPOMUK API. An application programmer can build usage specific services on top of the NEPOMUK API. By using the functionality provided by the API, the programmer can implement new functionality according to the end-users' business requirements.

Hence, the basic set of services provided by the NEPOMUK API can be customized and extended by businesses and organizations. For example, a company might be interested in integrating *Task Management* system whereas another might be interested in having document versioning support for resources. The end-user specific services are shown in dashed boxes in Figure 2.

The top layer of the architecture is the presentation layer. It provides a user interface to the services provided by the NEPOMUK desktop. The presentation layer is built using the NEPOMUK API. Many desktop applications are possible sources for resources that should be managed by NEPOMUK. Therefore, each desktop application should integrate support for the NEPOMUK Middleware. Since this assumption does not hold for most applications, we developed plug-ins and add-ons to enable a seamless integration for popular applications such as the MS Office Suite, which for example extract email or calendar data and adds them as resources to the NEPOMUK desktop. However, with in the NEPOMUK project we develop applications such as *Wikis* or *Blog Tools* that have generic support for the SSD and build directly on tho of the NEPOMUK API.

In addition, the *Knowledge Workbench* is the central place to browse, query, view, and edit resources and their meta-data. This way the Knowledge Workbench aims to replace current file management tools such as the File Explorer. If the SSD is extended by usage specific services, the application programmer has also to provide the corresponding user interface in the Presentation Layer.

6 Related Work

In the following we review the most important projects related to establishing a SSD. These projects are coming from the research, business, as well as the open-source community. After a brief general overview of each project, we want to learn from the related work as the conclusion of this section.

Gnowsis Semanic Desktop. The first research project targeting a Semantic Desktop system is the *Gnowsis Semantic Desktop* [12]. Its goal is to complement established desktop applications and the desktop operating system with Semantic Web features, rather than replacing them, while primarily focusing on *Personal Information Management* (PIM). The thesis addresses the problems of how to identify and represent desktop resources in an unified RDF graph.

Haystack. A major research project concerning an integrated approach in our field is *Haystack* [11]. Application-created barriers of information representation and accessibility are removed by simply replacing these applications with Haystack's word-processors, email client, image manipulation, instant messaging and other functionality. Haystack was ground-breaking in terms of the dynamic creation of user interfaces, but ended before establishing any standards.

Semex. Another relevant Personal Information Management tool is *Semex* (SEMantic EXplorer) [8]. Semex concentrates on the problem of Reference Reconciliation, meshing objects and relations seamlessly together. They combine three measures for this approach being evaluated on one of the author's private dataset. In contrast, NEPOMUK will add more reconciliation algorithms from the Semantic Web and evaluate the data integration in industry scenarios.

IRIS. The idea of the PIM system *IRIS* [5] is to have an integrated environment, similar to Haystack, but based on standard software, which is integrated into one coherent interface, allowing to classify and display related information. By today, the project lists only one publication introducing their approach.

Apogée. The *Apogée* project aims at building a framework to create Enterprise Development Process-oriented desktop applications, independent from vendor or technologies. Probably due to its status of an industrial project, it aims at implementing state-of-the-art features, but not beyond.

All integrated Semantic Desktop systems faced similar problems. First problem is evaluation and verification of the ideas in industry settings. Most systems like IRIS or Semex are not evaluated yet, they are only used by the developers in self-experiments. With its case studies, NEPOMUK will provide a testbed to show the implications of the whole Semantic Desktop in both, industrial environments and open-source communities.

Second problem is that the projects do not consider collaborative work and the interconnection of Semantic Desktops at all. They concentrate on a single user scenario, whereas NEPOMUK also tackles collaborative knowledge work.

Last and probably most significant problem is integration. While for example DBin shows the aspect of collaborative work, it does not connect to desktop applications. Though Haystack provides a well evaluated user interface, it does not re-use established Desktop applications users are used to, thus faces the user with a new environment. Further, none of the projects established standards which would increase interoperability and reusability.

Each system accommodates singular beneficial features, but also suffers from flaws like usability problems, bad performance, or missing functionality. These projects are designed as an integrated system, and despite the fact that most prototypes are open-source, it is not straightforward to reuse components of one for the other in order to amplify their features and extinguish their weaknesses. In contrast, NEPOMUK will establish a framework and standards so that components can be reused and are interoperable, creating a better whole.

7 Conclusion

This paper has given a very brief overview of the motivations, goals and progress of the NEPOMUK project. We have described the features and functionalities that our vision of a Social Semantic Desktop requires, based on observation of real knowledge-workers and their struggle with information integration using today's technology. Using an engineering process where we worked backwards from the desired functionalities and requirements, while simultanously refined a collection of existing Semantic Desktop research prototypes, we devised an architecture for the Semantic Desktop. This architecture enabled us to build a prototype featuring some of the required functionalities, and it is released as open-source and is available for download from the NEPOMUK Web-site[4].

The core aim of the NEPOMUK project is to specify an standard for Semantic Desktop communication and processing. We are basing our work on well-established standards of Web and Semantic Web technologies, and we hope that our Semantic Desktop standards in turn will provide a fertile ground for future projects. By having a flexible and easily extendible architecture we hope that over the next years any developer looking to solve information integration problems on the desktop will look to NEPOMUK as a framework for their projects, thus by the time the project is ending, NEPOMUK will have become a useful entity in it's own right, with an active community and untold possibilities.

References

1. K. Aberer, P. Cudré-Mauroux, A. Datta, Z. Despotovic, M. Hauswirth, M. Punceva, and R. Schmidt. P-grid: a self-organizing structured p2p system. *SIGMOD Record*, 32(3):29–33, 2003.
2. K. Aberer, P. Cudré-Mauroux, M. Hauswirth, and T. V. Pelt. Gridvine: Building internet-scale semantic overlay networks. In *3th International Semantic Web Conference ISWC 2004*, pages 107–121. Springer Verlag, 2004.
3. Aperture a java framework for getting data and metadata, Last visited March 2007. `http://aperture.sourceforge.net/`.
4. H. Beyer and K. Holtzblatt. *Contextual Design ? Defining Customer-Centered Systems*. Academic Press, San Diego.
5. A. Cheyer, J. Park, and R. Giuli. Iris: Integrate. relate. infer. share. In S. Decker, J. Park, D. Quan, and L. Sauermann, editors, *Proc. of Semantic Desktop Workshop at the ISWC, Galway, Ireland, November 6*, volume 175, November 2005.
6. A. Cooper. *The Inmates are Running the Asylum: Why High-Tech Products Drive Us Crazy and How to Restore the Sanity*. SAMS, Indianapolis, 1999.
7. A. Cooper and R. Reinman. *About Face 2.0: The Essentials of Interaction Design*. John Wiley & Sons, 2003.
8. X. Dong and A. Y. Halevy. A platform for personal information management and integration. In *CIDR*, pages 119–130, 2005.
9. E. Mackay, A. Ratzer, and P. Janecek. Video artifacts for design: bridging the gap between abstraction and detail. In *Designing interactive systems: processes, practices, methods, and techniques, DIS '00*. ACM Pres, 2000.
10. E. Prud'hommeaux and A. S. eds. SPARQL query language for RDF. W3C Working Draft, 4. Octoberl 2006. `http://www.w3.org/TR/rdf-sparql-query/`.
11. D. Quan, D. Huynh, and D. R. Karger. Haystack: A platform for authoring end user semantic web applications. In *International Semantic Web Conference*, pages 738–753, 2003.
12. L. Sauermann. The gnowsis-using semantic web technologies to build a semantic desktop. Diploma thesis, Technical University of Vienna, 2003.

[4] `http://dev.nepomuk.semanticdesktop.org`

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

From Philosophy and Mental-Models to Semantic Desktop Research: Theoretical Overview

Danish Nadeem
(Institute of Cognitive Science, University Of Osnabrück, Germany
dnadeem@uos.de)

Leo Sauermann
(Knowledge Management Lab, DFKI GmbH, Kaiserslautern, Germany
sauermann@dfki.uni-kl.de)

Abstract: This paper examines issues on Philosophy of Artificial Intelligence (AI), Cognitive Science and Mental Models. The paper provides a philosophical grounding for the researchers in Personal Information Management (PIM). An overview is given on various philosophical aspects of computer-based activities. Discussions on the theories relevant to understand the goals for the Semantic Desktop community are elicited. Philosophical theories are not immediately transparent to the programmers, but the ideas discussed here are intended to emphasize a theoretical foundation, with respect to Semantic Desktop long term goals. The goal of this paper is to examine the theories of Philosophy and to provide a conceptual idea to design user-intuitive Semantic Desktops. The paper tries to induce scientific curiosity among the Semantic Desktop researchers as well as to develop the future Semantic Desktops to realize *Weak AI*.

Key Words: mental model, knowledge management, semantic desktop
Category: H, H.1.0, H.1.2, H.5.0

1 The Semantic Desktop idea

In [Sauermann et al.2005] Sauermann gives the possible definition of a Semantic Desktop:

> *A Semantic Desktop*[1] *is a device in which an individual stores all her digital information like documents, multimedia and messages. These are interpreted as Semantic Web resources, each is identified by a Uniform Resource Identifier (URI) and all data is accessible and queryable as RDF graph. Resources from the web can be stored and authored content can be shared with others. Ontologies allow the user to express personal Mental Models and form the semantic glue interconnecting information and systems. Applications respect this store, read and communicate via ontologies and Semantic Web protocols. The Semantic Desktop is an enlarged supplement to the user's memory.*

[1] http://semanticdesktop.org

Semantic Desktop is a research approach to provide Personal Information Management [Boardman2004] to the user in an elegant way. It is a challenging issue for Semantic Desktop community to develop a computer application that recognizes the Mental Model of its user and maps Mental Model to the desktop - integrating (or substituting) all other applications [Heim2006]. The theories of AI and Cognitive Science should provide an insight to tackle this challenge and provide a theoretical understanding of design methods to be adopted to develop a user-oriented Semantic Desktop framework. There have been several prototypes that aim towards the goal of Semantic Desktop such as Gnowsis [Sauermann et al.2005], Haystack [D.Quan et al.2003], IRIS [Cheyer et al.2005]. The goal of this paper is to elicit holistic view of Semantic Desktop research from philosophical point of view. The philosophical theories are significant to any applied research towards Artificial Intelligence (AI). Research in Semantic Desktop is a novel attempt to address the issues of AI. Thus, the following section mainly discusses philosophy and theories of Cognitive Science to induce scientific curiosity among the Semantic Desktop researchers.

2 Philosophy, AI and Cognitive Science

The main philosophy behind Artificial Intelligence is based on the question *Can machines think*? Although the question itself doesn't clearly define what is implied by *think*. Within the AI community there has always been a debate *whether* it is possible or not to build intelligent thinking machines which are as good as human in terms of decision making, problem solving, conscious thinking to the extent of making creative discovery or showing sophisticated behavior like understanding the concept of *emotional love*. According to Dietrich [Dietrich2002], the most important of the *whether* problems lie at the intersection of theories of the semantic *content of thought* and the nature of computation. Human thinking is the manipulation of contentful thoughts, involved in cognitive processes as making inferences, recognizing patterns, planning and executing activities. During such processes our thoughts are manipulated to refer to the various things and concepts in our real world. The *content of thought* would possibly mean an expression or summary in our mind about a particular concept. But when this view is translated to computers a real hard problem emerges, because machine cognition is generally based on an algorithmic manipulation and computation of well defined data structures. The role of philosophical problems come to a play at this moment, thoughts have contents or internal states, now the question is when computers do addition of 1 to 2, how internal states are actually denoted to the number 1 or 2 ? Cognitive Science raises many of such interesting questions that are worthy of investigation by Philosophers. For example, if a computer infers *Fido will chase the cats* from the given facts *Dogs chase the cats* and *Fido is a dog*, do any of its internal states refer to dogs, cats, fido and the act of chasing.

> *If computations are radically different from thoughts in that they cannot have semantic content, then it is unlikely that computers can think* [Dietrich2002].

These problems are addressed in AI and Cognitive Science as *the problem of mental content or representational content*. Dietrich further identified that the second set of whether-possible problems of AI surrounds *the nature of rationality*. He claims humans constantly evaluate ways of achieving goals and rank them according to various measures such as the probability of success, efficiency and consequences. They evaluate the goals themselves and constantly gauge the relevance of one piece of information to another, the relevance of one goal to another and the relevance of evidence to achieving a goal. Humans are often successful in such evaluation however computers are not, thus human level rationality are not obtainable in an intelligent machine. The third set of whether-possible problems of AI are the issues of addressing the reasoning powers of human mind. In 1931, Kurt Gödel demonstrated by his famous *incompleteness theorem*[2] that within any given branch of mathematics, there would always be some propositions that cannot be proven for true or false using rules and axioms. Gödel proved that all *consistent* (i.e. there is no contradiction in a statement and its negation is also true), logical systems are *incomplete* (in the sense that every statement in the language of Number Theory cannot be either proved or disproved). The logical problems surrounding the self-reference and the incompleteness of certain axiomatic systems (logical systems) seem to be a barrier in building an intelligent machine.

Gödel's Theorem has been used to argue that a computer can never be as smart as a human being because the extent of its knowledge is limited by a fixed set of axioms, whereas people can discover unexpected truths. Philosophers made AI conceivable by considering the ideas that the mind is in some ways like a machine, that it operates on knowledge encoded in some internal language and that thought can be used to help arrive at the right actions to take [Russell and Norvig2003]. Even though Philosophy existed a long before computers came into existence, there has always been a question of how minds work, how do human minds work and can non-human have minds? These are real hard problems that often resulted in heated debates among the Philosophers, AI researchers, Cognitive Scientists. Many have now chosen a computational perspective, because of the various tools being available to study the intelligent behavior in detail. At current state of research, AI is still in its infancy to exhibit the higher-level human cognitive abilities and thought processes in computers. The AI research is actively viewed with respect to two schools of thought to achieve machine intelligence. They are termed as *Strong AI* and *Weak AI* by John Searle [Searle1980]:

---

[2] http://mathworld.wolfram.com/GoedelsIncompletenessTheorem.html

– **Strong AI** is the belief that artificial intelligence can truly reason and solve problems, strong AI supposes that it is possible for machines to become sapient, or self-aware, but may or may not exhibit human-like thought processes. As Searle claimed:

  > *..according to strong AI, the computer is not merely a tool in the study of the mind; rather, the appropriately programmed computer really is a mind.*

– **Weak AI** refers to the use of software to study or accomplish specific problem solving or reasoning tasks that do not encompass the full range of human cognitive abilities, unlike strong AI, a weak AI does not achieve self-awareness or demonstrate a wide range of human-level cognitive abilities, and is merely an (arguably) intelligent, more specific problem-solver. The current softwares are Expert systems that are used commonly for specific purposes. For example, there are expert systems that can diagnose human illnesses, make financial forecasts, and schedule routes for delivery vehicles. Some expert systems are designed to take the place of human experts in specific knowledge domains, while others are designed to support them. Expert systems and also the Chess program are a part of Weak AIs.

In the following sections an outline is given about the research in Mental Models and its practical relevance to Semantic Desktop is examined.

3 Mental Models

Mental Models have been studied by cognitive scientists as part of efforts to understand how humans know, perceive, make decisions, and construct behavior in a variety of environments. From the Vannevar Bush ideas of how human mind operates *As we may think* [Bush1968] to the Xanadu Project by Ted Nelson, Philosophers have always been interested in *how we think about the things in our world and how it is represented in our minds*? The term *Mental Model* was first mentioned by Craik in his 1943 book, *The Nature of Explanation* [Craik1943]. It said that humans make use of internal models of external reality, which enable them to better understand and react to situations in their environment. In his view people operate on mental representations to simulate real world behavior and produce predictions. In other words this implies humans are not just physically situated in its environment, but they also have their own internal model of it, which allows them to deal with that external reality of world.

After Craik, literatures on Mental Model appeared in three theoretical approaches, viz. Johnson-Laird's (1983) [Johnson-Laird1983] theory of Mental Models, a collection of work on Mental Models of natural phenomena and devices by Genter and Stevens (1983) [Gentner and Stevens1983], and Paivio's (1986) [Paivio1986] dual coding approach for classification of mental representations.

The Johnson-Laird volume proposed Mental Models as a way of describing the process which humans go through to solve deductive reasoning problems. His theory included the use of a set of diagrams to describe the various combinations of premises and possible conclusions [Johnson-Laird1983]. Johnson-Laird proposed three types of mental representations: (1) Propositional representations: which are pieces of information resembling natural language. (2) Mental Models: which are structural analogies of the world. (3) Mental imagery: which are perceptual correlates of models from a particular point of view. Another book appeared in the same year by Gentner and Stevens. They proposed that Mental Models provide humans with information on how physical systems work. This approach could be generalized to a number of situations that humans face, including the behavior of objects according to laws of physics [Gentner and Stevens1983].

The fundamental philosophical issue addressed within the context of Mental Model is that *things are not the way it is represented in our thoughts*, for instance, thoughts about computer are not computer itself, rather probably a conceptualization of the features of computer. The question is then how the abstract thoughts manage to represent the things in such a way, even though there are missing information. More interestingly, our thoughts are not always restricted to represent existing things, but there are things that cannot possibly exist (rectangular basketball), things that do no exist (unicorn), things which are not perceivable (limit of universe). How all these representations are possible without the things existing itself or the way it exist. Human thoughts have *semantic content*, which is missing in computers. For instance, when a number is added to itself it is twice the number, this is general principle for computation which an algorithm computes for computers, but for humans we have more intuitive information processing mechanism and deeper understanding, humans can rank the worth of computation based on thought content. According to Dietrich [Dietrich2002], when a computer does addition there occurs a cascade of causal processes which implements an algorithm that in turn, if followed exactly guarantees that two numbers will be added. We can make an analogy to wonder why computers and its processing is different than humans, lets take a coffee machine example. A coffee machine doesn't know at all about the type of coffee, nor does it represent coffee and also it has no knowledge about the coffee it is preparing. It is given rather a configuration to use different ingredients to produce specific coffee types. Can we encode semantic content to improving this situation in an intelligent machine? If yes, then we can hypothesize of thinking computer or at least attempting towards a human type information processing rather than simple procedural computations without any thought content involved.

One way out of this dilemma is to attempt to develop a philosophical the-

> *ory of mental content that clearly explains how thoughts get the content that they do. Then we could just check to see whether computations could get content in the same way. If they can, then AI is on firm ground, if they cannot then AI is without hope [Dietrich2002].*

According to Dietrich comment, the nature of semantics are viewed on two perspectives namely: (1) *world-mind relations*: It saw semantics as essentially associated with the truth, causation and getting along in the world. (2) *mind-mind relations*: It saw semantics as essentially associated with being able to draw certain inferences, construct plans and in general determine how one thought and representation relates to another.

The two views mentioned above are necessary to understand and develop the theory of representational content. Within the Semantic Desktop framework a computer can be causally connected to the environment and its representations can be implemented. Semantic Desktop could be a proxy to user's world view to represent and relate concepts a person keeps in her mind. Although a thought content would still remain a problem for computers to comprehend.

The Semantic Desktop is a relatively new research approach motivated to tackle the challenges for Personal Information Management to help the users making sense of their ever increasing personal information. Cognitive Science research perspectives help to develop a theoretical base here, thus the Mental Model would surely become the vocabulary for Semantic Desktop community for representing complex and ever changing world of information. Also for the Human-Computer Interaction (HCI) practitioners a Mental Model provides a set of beliefs about system functionality. Human interactions with system is based on these beliefs [Norman1990].

Usability issues of Semantic Desktop is also tightly connected to the user's Mental Model. But it is quite challenging to meet the needs of an expert as well as a novice users' Mental Model, further discussion explains the reason. A design of Semantic Desktop should be consistent with person's natural Mental Model about the concepts, ideas and everyday objects encountered in environment. For instance, the way people organize their paper-work while doing a specific task like writing a thesis or working on a project should be reflected in Semantic Desktop, all the relevant objects should be related and tagged in a sensible way to efficiently organize the work. Semantic Desktop should provide an interface to model the physical paper-way of organizing with which a user is most familiar in daily activities.

Many existing systems put too many demands on the users that use them, users are often required to adjust the way a system works. A system with inaccurate Mental Model leads to frustration and inconvenience. Moreover with an increasing demand for usability in technology products and the people's dependence on computers, we have to expect the non-experts interacting with the system for Personal Information Management. Such user could be lacking technical expertise and tolerance. The role of technology products such as Semantic Desktop must accommodate the needs of the users of future generation who are more diverse, less technical, very explorative and quite impatient [Gribbons1999]. The system designer bores the responsibility of capturing user's expectations and hypothesize more about user-adaptiveness to deliver intuitive and predictable system consistent with the user's Mental Model.

4 Realizing Theories in Semantic Desktop

This section describes how Mental Models described in [Section 3] could be realized as design methods for Semantic Desktop development. Semantic Desktop should be designed to help users build productive Mental Models of system functionalities. As argued in this paper, the usability issues of Semantic Desktop are closely connected with user's Mental Model. A user should be informed through interfaces in a such a way so that she could exploit the maximum benefit of Semantic Desktop goals. There have been excellent efforts to address the issues of Mental Models [Davidson et al.1999, Sasse1997] to have a better understanding of user-interaction with the systems. According to McDaniel [McDaniel2003] Mental Models are the conceptual representation within the person's mind that helps her to understand the world [Kurtz2003]. Mental Models may be (1) An Image, (2) A script, (3) A set of related Mental Models, (4) A controlled vocabulary, or (5) A set of assumptions.

> *We should create these Mental Model descriptions during user analysis to document users' current understanding. Then, during a design phase, we should create the target model to show the Mental Model we want users adopt.* - McDaniel [McDaniel2003]

Many research from the area of Cognitive Science give some generic ideas that human mind is quite flexible, and our learning abilities, memorizing and conceptualization is quite often based on analogy of relationships among the real-world objects. In digital ambience in relevance to HCI we have distributed and different information but they are often connected with multiple relationships, explicitly or implicitly, which means either it is visible from a given information model (explicit) or it could be inferred (implicit). Based on our discussions we identified some of the general features and criteria for Semantic Desktop to support user in making productive Mental Models. They are illustrated as follows.

- **User Expressivity**: The Semantic Desktop should enable the user to express her personal concepts. Users should be allowed to make their real-world concepts in an intuitive way. User tend to categorize their information according to their own sense, thus any strict categorization feature should be avoided. This would support user's information filing behavior.

- **Reducing Cognitive Overload**: The Semantic Desktop is meant to reduce cognitive overload of information. In [Kirsh2000], David Kirsh points out that too often the information falls between the cracks of our classifying scheme and we are faced to go through the challenging process of creating new indices and categories or painfully stretching the old ones. Whenever we create a new category or stretch an old one there is the danger that where we place the information will be forgotten next time we look for it. This is all stressful specially because the less one has system for dealing with invaded information (i.e. information in the form of emails, newsletters etc.) and the more one must make ad-hoc decisions for each incoming piece of information.

  > *The psychological effort of making hard decisions about invaded information is the first cause of cognitive overload* [Kirsh2000]

- **Preserving Association (user context)**: We have known from the vision of Bush [Bush1968] that human mind operates by association. With one item in its grasp it quickly moves to the next item suggested by the association of thoughts stored as a web of trails. Semantic Desktop should be able to provide users with necessary priming needed to preserve the context in which a user is currently in and a possible trails to move other related contexts. For example, a user reading a paper might also like to know the authors homepage, other co-authors and related publications.
- **Improving accuracy in Mental Model**: Accuracy of Mental Model is proportional to Usability of the system. On one side we have a user with a "Mental Model" who wants to use the system for a specific task, on the other side there is a computer system which follows user behavior and knows user preference and goals (i.e. the system keeps a user model). The task of Semantic Desktop would be to map the Mental Model to the user model for the person to use the system more effectively. Normally the Mental Models people create of computer systems are inaccurate [Norman1983]. By designing the systems that help people to create a more accurate Mental Model of the system, usability would significantly improve [Norman1983]. This means that if the designer creates the correct design model and communicates the model successfully through the system image. Then users interacting with the system will develop an appropriate user model, which will allow them to interact with the system successfully.
- **Personalization**: Personalization of Semantic Desktop systems would enhance user satisfaction and productivity. Each person has her own mental representation of the concepts from the real-world. Semantic Desktop is a way to express these mental concepts with respect to each user. Users prefer to make their own customization according to their interests, motivation and expertise. By giving the user freedom to personalize their concepts would

support their memory and familiarity. This would encourage the user to use Semantic Desktops in an intuitive way.

Mental Models are meant to simplify the reality. They are promising human factors in consideration of design, but the real difficulty exist in methodology to design system that would help the user to create best Mental Model [Preece et al.1994]. Therefore the efforts should be focused on providing an accurate initial Mental Model and capturing user behavior over time. This approach has been taken in PIMO- an Ontology to support initial Mental Model of user, see [Sauermann2006] for details.

5 Conclusions

Cognitive Science gives a better insight to understand Mental Models but to capture and validate users Mental Models poses difficulties. The potential for rewards of improved design and increased usability based on correct Mental Models compensate for the effort, but they are still an open area of challenging research. The ideas discussed in this paper provides a holistic view from the perspective of Philosophy and Cognitive Science theories, which could help in building Semantic Desktop to meet the users requirements. The theories of Mental Models discussed in the paper is an important knowledge needed for any steps in building Semantic Desktop, although not detailed but there is a hint given in this work to emphasize a theoretical foundation as well as practical applications of Mental Model. This has to be kept in mind while designing specification for Semantic Desktop long term goals. In Gnowsis context, we can identify success stories by individual users about their experiences on how Semantic Desktop should ideally work. Comparing their answers would reflect different Mental Models and expectations. This should be used for redesigning prototype to exploit system features and improve usability. We still lack the proper evaluation of the Semantic Desktop systems, specially with the non-expert users to judge how well the systems adopts to user's Mental Model. Moreover, it would be interesting to investigate if Semantic Desktop is an effort realizing *Weak AI*. One such research direction would be to investigate experimentally, how Semantic Desktop systems provide an aid to human memory.

Acknowledgements

This research is conducted at German Research Center for Artificial Intelligence within the Knowledge Management Lab to bring the theories of Cognitive Sciences to the Semantic Desktop research. Special thanks to Dr.Thomas Roth-Berghofer, for his critical review and suggestions. This research is supported by the European Commission under contracts IST-2006-027705 NEPOMUK. The

expressed content is the view of the authors but not necessarily the view of the project consortia.

References

[Boardman2004] Boardman, R. (2004). *Improving Tool Support for Personal Information Management.* PhD thesis, Imperial College London.

[Bush1968] Bush, V. (1968). As we may think. *Conversational Computers*, pages 10–12.

[Cheyer et al.2005] Cheyer, A., Park, J., and Giuli, R. (2005). Semantic desktop workshop 2005 paper — openiris. *ISWC 2005.*

[Craik1943] Craik, K. ((1943)). *The Nature of Explanation.* Cambridge University Press, London, New York.

[Davidson et al.1999] Davidson, M. J., Dove, L., and Weltz, J. (November 15, 1999). Mental models and usability. Technical report, Depaul University, Chicago.

[Dietrich2002] Dietrich, E. (2002). Philosophy of artificial intelligence. In *The Encyclopedia of Cognitive Science (pp.203-208).*

[D.Quan et al.2003] D.Quan, D.Huynh, and Karger, D. (2003). Haystack: A platform for authoring end user semantic web applications. *ISWC*, pages 738–753.

[Gentner and Stevens1983] Gentner, D. and Stevens, A. (1983). *Mental Models.* Hillsdale, New Jersey, Lawrence Erlbaum Associates.

[Gribbons1999] Gribbons, D. W. M. (November 1999). Knowledge-infused design, the "ultimate solution" to product usability. *Help 99 Proceedings*, pages 153–156.

[Heim2006] Heim, D. (2006). Semantic wikis in knowledge management. Master's thesis, Fachhochschule Kaiserslautern.

[Johnson-Laird1983] Johnson-Laird, P. N. (1983). *Mental models:Towards a cognitive science of language, inference, and consciousness.* Cambridge, MA: Harvard University Press.

[Kirsh2000] Kirsh, D. (2000). A few thoughts on cognitive overload. *Intellectica*, 1:19–51.

[Kurtz2003] Kurtz, A. (2003). Mental models - a theory critique.

[McDaniel2003] McDaniel, S. (2003). What's your idea of a mental model? Technical report.

[Norman1983] Norman, D. (1983). Some observations on mental models. *In D. Gentner and A.L. Stevens, Eds. Hillsdale, NJ: Lawrence Erlbaum*, pages 7–14.

[Norman1990] Norman, D. (1990). *The Design of Everyday Things.* New York: Doubleday/Currency.

[Paivio1986] Paivio, A. (1986). *Mental representations: A dual coding approach.* New York: Oxford University Press.

[Preece et al.1994] Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., and Carey, T. (1994). *Human-Computer Interaction.* Wokingham, UK: Addison-Wesley.

[Russell and Norvig2003] Russell, S. J. and Norvig, P. (2003). *Artificial Intelligence: A Modern Approach.* Pearson Education.

[Sasse1997] Sasse, M. A. (April 1997). *Eliciting and Describing Users' Models of Computer Systems.* PhD thesis, School of Computer Science,The University of Birmingham,B15 2TT, England.

[Sauermann2006] Sauermann, L. (2006). Pimo-a pim ontology for the semantic desktop (draft). Draft, DFKI.

[Sauermann et al.2005] Sauermann, L., Bernardi, A., and Dengel, A. (2005). Overview and outlook on the semantic desktop. In *Proc. of Semantic Desktop Workshop at the ISWC, Galway, Ireland.*

[Searle1980] Searle, J. (1980). *Minds, Brains, and Programs.* The Behavioral and Brain Sciences, vol. 3. Copyright 1980 Cambridge University Press.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Improving Search on the Semantic Desktop using Associative Retrieval Techniques

Peter Scheir
(Graz University of Technology, Austria
peter.scheir@tugraz.at)

Chiara Ghidini
(Fondazione Bruno Kessler, Italy
ghidini@itc.it)

Stefanie N. Lindstaedt
(Know-Center Graz, Austria
slind@know-center.at)

Abstract: While it is agreed that semantic enrichment of resources would lead to better search results, at present the low coverage of resources on the web with semantic information presents a major hurdle in realizing the vision of search on the Semantic Web. To address this problem we investigate how to improve retrieval performance in a setting where resources are sparsely annotated with semantic information. We suggest employing techniques from associative information retrieval to find relevant material, which was not originally annotated with the concepts used in a query. We present an associative retrieval system for the Semantic Desktop and show how the use of associative retrieval increased retrieval performance.

Key Words: semantic desktop, associative information retrieval
Category: H.3.3, I.2.4, I.2.6, I.2.11

1 Introduction

On the one hand it is largely [6] [17] agreed that semantic enrichment of resources on the web or desktop provides for more information to be used during search and that this can lead to higher effectiveness of a retrieval system. One the other hand critics [10] as well as advocates [13] of the Semantic Web agree on the low coverage of resources on the current web with semantic information. The sparse annotation of resources with semantic information presents a major obstacle in realizing search applications for the Semantic Web or the Semantic Desktop, which operate on semantically enriched resources. For this reason we propose the use of techniques from associative information retrieval to find additional relevant material, even if no semantic information is provided for those resources.

We describe our approach to information retrieval on the Semantic Desktop and present a retrieval component developed during the first year of the APOSDLE[1] project. The rest of this paper is organized as follows: in section 2 we introduce the concept of associative information retrieval, in section 3 we present statistical information about the knowledge base used in APOSDLE to

[1] http://www.aposdle.org/

demonstrate why the use of associative retrieval techniques suits our needs. Section 4 describes our retrieval system and section 5 its evaluation. We present related work in section 6 and our conclusion in section 7.

2 Associative Information Retrieval

Crestani [5] understands *associative retrieval* as a form of information retrieval which tries to find relevant information by retrieving information that is by some means *associated* with information that is already known to be relevant. Information items which are associated can be documents, parts of documents, extracted terms, concepts, etc. The idea of associative retrieval dates back to the 1960s, when researches [14, 15] in the field of information retrieval tried to increase retrieval performance using associations between documents or index terms, which were determined in advance.

Association of information is frequently modeled as graph, which is referred to as *associative network* [5]. Nodes in this network represent information items such as documents, terms or concepts. Edges represent associations between information items and can be weighted and / or labeled, expressing the degree and type of association between two information items, respectively.

3 Statistics about the Knowledge Base used

We operate on the knowledge base created for the first prototype of the APOSDLE system. The goal of the present system is to help knowledge-workers understand the field of requirements engineering. Therefore this domain is modeled using an ontology. Documents containing learning material (definitions, examples, tutorials, etc.) about requirements engineering are partly annotated with concepts from the domain ontology. The ontology consists of 70 concepts, 21 concepts are used to annotate documents. The document base consists of 1016 documents, 496 documents are annotated with one or more concepts from the knowledge base. We experience a typical scenario here: only parts of the ontology are used for annotation and only parts of the documents are annotated. We see this setting corresponding to the coverage problematic presented in section 1 and employing associative retrieval techniques appropriate to finding relevant material that was not originally annotated with concepts from the domain ontology.

4 The Prototype

The system presented here is based on an associative network consisting of two interconnected layers, one for concepts and one for documents. Nodes in the concepts layer correspond to concepts in the domain ontology. Nodes in the document layer correspond to documents in the system. Concept nodes are associated by means of semantic similarity (cf. section 4.1)), document nodes are associated by means of textual similarity (cf. section 4.2). Concept nodes are associated with document nodes if the concept is used to annotate the document (cf. sections 4.3 and 4.4). The network is searched using a spreading activation algorithm (cf. section 4.5).

4.1 Calculating Semantic Similarity of Concepts

For calculating the similarity of two ontological concepts a symmetric semantic similarity measure is used. The method was presented in [18] and requires two concepts belonging to the same ontology as input. It calculates the semantic similarity between these two concepts according to equation 1. This similarity measure builds on the path length to the root node from the least common subsumer (lcs) of the two concepts, which is the most specific concept they share as an ancestor. This value is scaled by the sum of the path lengths from the individual concepts to the root.

$$sim(c_1, c_2) = \frac{2 \cdot lcs(c_1, c_2)}{depth(c_1) + depth(c_2)} \tag{1}$$

With:

- c_1 ... first concept
- c_2 ... second concept
- lcs ... least common subsumer of two concepts
- $depth$ ... depth of concept in the class hierarchy

Depending on the features present in an ontology different similarity measures qualify to be applied. We chose the measure presented in [18], as a prominent feature of our ontology are taxonomic relations between concepts. An advantage of the used measure is that it tries to address one of the typical problems of taxonomy-based approaches to similarity: relations in the taxonomy do not always represent a uniform (semantic) distance. The more specific the hierarchy becomes, the more similar a child node is to its father node in the taxonomy.

4.2 Calculating Text-based Similarity of Documents

As similarity measure for text-documents we use an asymmetric measure based on the vector space model implemented in the open-source search-engine Lucene[2]. The similarity between two documents is calculated as shown in equation 2.

$$sim(d1, d2) = score(d1_{25}, d2) \tag{2}$$

With:

- $d1$ ... document vector of the first document
- $d2$ ... document vector of the second document
- $d1_{25}$ ... document vector of the first document with all term weights removed except the 25 highest terms weights

$d1_{25}$ is used as query vector for the *score*-measure of Lucene. For extracting the 25 terms with the highest weights, both the document content and the document title are taken into account. The calculation of Lucene's score is depicted in equation **??**. A detailed explanation of the various parameters that can be used to adapt the behavior of Lucene can be found in the Javadoc of the `org.apache.lucene.search.Similarity` class.

[2] http://lucene.apache.org/

$$score(q,d) = coord(q,d) \cdot queryNorm(q) \cdot \sum_{t\_in\_q} (tf(t\_in\_d) \cdot idf(t)^2 \cdot t.getBoost() \cdot norm(t,d)) \qquad (3)$$

With:

- q ... query vector
- d ... document vector
- $coord(q,d) = numberOfMatchingTerms/numberOfQueryTerms$
- $numberOfMatchingTerms$ ... number of terms in document matching query
- $numberOfQueryTerms$ ... number of terms in the query
- $queryNorm(q)$ ... normalization of the query vector, Lucene default used
- $tf(t\_in\_d)$ ... term frequency of current term in document, Lucene default used
- $idf(t)$ ... inverse document frequency of current term in the document collection, Lucene default used
- $t.getBoost() = tf(t\_in\_q) \cdot idf(t)$
- $tf(t\_in\_q)$ ... term frequency of current term in query
- $norm(t,d) = 1/sqrt(numberOfDocumentTerms)$
- $numberOfDocumentTerms$ ... number of terms in the current document

Finding similar documents to a document based on the vector space model is a well researched topic. Equally, Lucene is a frequently used text search engine. Therefore we are confident of the applicability of both, the similarity measure as well as the search engine to our scenario.

4.3 Semantic Annotation of Documents

Despite to other approaches, where fine-grained annotation of the words present in a document with concepts from the ontology is carried out (c.f. [3] or [8]), we follow a more pragmatic approach. We adopt the tagging metaphor and blend this approach with the controlled vocabulary of an ontology. This means that we tag whole documents with a set of concepts the document *deals with*.

We follow this approach two reasons: (1) Although the complete semantics of a sentence contained in a document are not recognized using this approach, the additional information added to the document still provides opportunities to be used at a later time in retrieving material [17] by a limited amount of human involvement. (2) We think that for the near future it makes sense to work on making the Semantic Web a reality, by focusing on bringing little semantics [7] into the current web and taking small steps. We follow this pragmatic approach and try to apply it to the Semantic Desktop in the context of our work.

In APOSDLE, annotation of documents is supported by a plug-in for the ontology editor Protégé[3] and supported by a classification algorithm. This means that we have implemented functionality that suggests a set of concepts for documents to be annotated based on the set of documents already annotated. A detailed description of our approach and an early realization of it can be found in [16].

[3] http://protege.stanford.edu/

4.4 Weighting the Annotations

In our (and other) approach(es) to semantic annotation, a document is either annotated with certain concepts or it is not. From a retrieval point of view this means that a document is either retrieved, if it is annotated with a concept present in the query, or it is not retrieved, if none of the concepts in the query are assigned to the document. Ranking the retrieved document set is impossible.

To allow for ranking the result set and increase the performance of our system we weight the annotations between documents and concepts using a tf-idf-based weighting scheme. This is a standard instrument in information retrieval to improve retrieval results [11]. Our weighing approach is related to the one presented by [3], who are also weighting semantic annotations using a tf-idf-based measure.

$$weight(c,d) = tf(c,d) \cdot idf(c) = tf(c,d) \cdot \log \frac{D}{a(c)} \quad (4)$$

With:

- c ... a concept
- d ... a document
- $tf(c,d)$ ... 1 if d is annotated with c, 0 otherwise
- $idf(c)$ ... inverse document frequency of concept c
- D ... total number of documents
- $a(c)$ ... number of documents annotated with concept c

4.5 Searching the Network

The network structure underlying the system is searched by spreading activation. Starting from a set of initially activated nodes in the network, activation spreads over the network and activates nodes associated with the initial set of nodes. Originally stemming from the field of cognitive psychology, where it serves as a model for operations in the human mind, spreading activation found its way over applications in both neural and semantic networks to information retrieval [5]. It is comparable to other retrieval techniques regarding its performance [9].

Besides systems that use spreading activation for finding similarities between text documents or search terms and text documents, approaches exist, which employ spreading activation for finding similar concepts in knowledge representations [1] [12]. The novelty of our approach lies in combing spreading activation search in a document collection with spreading activation search in a knowledge representation. The formula we use to calculate the spread of activation in our network is depicted in equation 5.

$$A(n_j) = \frac{\sum_{t=1}^{t} A(n_i) \cdot w_{i,j}}{\sum_{t=1}^{t} w_{i,j}} \quad (5)$$

With:

- $A(n_j)$ ... activation of node n_j
- $A(n_i)$ ... activation of node n_i
- t ... number of nodes adjacent to node n_j
- $w_{i,j}$ ... weight of edge between node n_i and node n_j

Search in our network is performed as follows:

1. Search starts with a set of concepts, representing the information need of the knowledge-worker. The concept nodes representing these concepts are activated.
[2. *Optionally*, activation spreads from the set of initially activated concepts over the edges created by semantic similarity to other concepts nodes in the network.]
3. Activation spreads from the currently activated set of concept nodes to the document nodes over the edges created by semantic annotation to find documents that deal with the concepts representing the information need.
[4. *Optionally*, activation spreads from the documents nodes currently activated to document nodes that are related by means of textual similarity and are therefore associated with the document nodes.]
5. Those documents corresponding to the finally activated set of document nodes are returned as search result to the user.

5 Evaluation

The present approach to retrieval on the Semantic Desktop is different from current attempts to retrieval in a desktop environment: (1) the semantic information present in an ontology is taken into account for retrieval purpose; (2) the query to the retrieval system is formulated by a set of concepts stemming from an ontology as opposed to a set of terms (words) as typically used in the context of desktop search. As we are not aware of any standard test corpora for the evaluation of an information retrieval system for the Semantic Desktop we have created our own evaluation environment.

We have evaluated six different configurations of our system using a set of 22 queries. For every query we relevance-judged the first 30 search results. Afterwards we calculated precision at rank 10 (P(10) cf. [2]), precision at rank 20 (P(20))[4] and inferred average precision (infAP[5]). All three evaluation measures rank the tested system configurations in the same order.

Table 5 shows the ranking of the different system configurations. The columns *SemSim*, *TxtSim* indicate whether semantic similarity or text-based similarity was used for the search. The last line (configuration 6) of table 5 is the baseline configuration of our system. The results delivered by this configuration are comparable to the use of a query language as SPARQL combined with an idf-based ranking and no associative retrieval techniques used. Exactly those documents are retrieved that are annotated with the concepts present in the query. All associative search approaches employing semantic similarity (configurations 1, 2, and 5), text-based similarity (configurations 1, 2 and 3) or both (configurations 1, 2, 3, 4 and 5) increase retrieval performance compared to the baseline config-

[4] As we judged 30 documents for every query it would also have been possible the calculate P(30) but as we are aiming on presenting our search results using a sidebar-based interface and we have limited space for our search results there, we are not considering to present 30 results.

[5] infAP was proposed by [19] and performs a random sampling approach to all *judged* results (relevant and not relevant) for a query. The measures P(10) or P(20) only consider judged, relevant results. As infAP takes more information into account than P(10) or P(20) it is considered as a more stable measure.

uration. Additional relevant documents are found, which are not annotated with the concepts used to query the system.

Configuration	SemSim	TxtSim	P(10)	P(20)	infAP
1	Yes (> 0.5)	Yes	0.7	0.6523	0.5728
2	Yes (> 0.7)	Yes	0.6909	0.6477	0.5706
3	No	Yes	0.6636	0.6227	0.5431
4	Yes (> 0.5)	No	0.6545	0.5818	0.4971
5	Yes (> 0.7)	No	0.6364	0.5727	0.46
6	No	No	0.6045	0.5545	0.4176

Table 1: Ranking of system configurations using P(10), P(20) and infAP

For calculating the evaluation scores we have used the trec_eval[6] package, which origins from the Text REtrieval Conference (TREC) and allows for calculating a large number of standard measures for information retrieval system evaluation.

6 Related Work

Beagle++ [4] is a search engine for the Semantic Desktop and indexes RDF-metadata together with document content. Both [3] and [8] present an extension of the vector space model. Together with document content they index semantic annotations of documents and use this information for search. All three are very promising approaches that extend the vector space model using semantic information. None of them employs measures of semantic association.

[12] present a hybrid approach for searching the (semantic) web, they combine keyword based search and spreading activation search in an ontology for search on websites. Ontocopi [1] identifies communities of practice in an ontology using spreading activation based clustering. Both are prospective approaches employing ontology-based measures of association and evaluating them using spreading activation. They do not integrate text-based measures of association into their systems.

7 Conclusions and Future Work

Our experiments encourage us, that the application of associative retrieval techniques to information retrieval on the Semantic Desktop is an adequate strategy. Following recent works [2] [19] in information retrieval system evaluation we are confirmed that the amount of relevance judgments we have used should be increased to have a higher confidence in our retrieval system evaluation. We tend to conclude that text-based methods for associative retrieval result in a higher increase in retrieval performance, therefore we want to explore the approach of attaching a set of terms to every concepts in our domain ontology during

[6] http://trec.nist.gov/trec_eval/

modeling time to provide search results even for concept that are not used for annotation. In addition we want to extend our research towards evaluating different semantic similarity measures.

Acknowledgments

We thank the anonymous reviewers of our submission for their constructive feedback.

This work has been partially funded under grant 027023 in the IST work programme of the European Community. The Know-Center is funded by the Austrian Competence Center program Kplus under the auspices of the Austrian Ministry of Transport, Innovation and Technology (www.ffg.at/index.php?cid=95) and by the State of Styria.

References

1. H. Alani, S. Dasmahapatra, K. O'Hara, and N. Shadbolt. Identifying communities of practice through ontology network analysis. *IEEE Intell. Syst.*, 18(2):18–25, 2003.
2. C. Buckley and E. M. Voorhees. Retrieval evaluation with incomplete information. In *SIGIR '04: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, 2004.
3. P. Castells, M. Fernndez, and D. Vallet. An adaptation of the vector-space model for ontology-based information retrieval. *IEEE Trans. Knowl. Data Eng.*, 19:261–272, 2007.
4. P.-A. Chirita, S. Costache, W. Nejdl, and R. Paiu. Beagle++: Semantically enhanced searching and ranking on the desktop. In *The Semantic Web: Research and Applications*, 2006.
5. F. Crestani. Application of spreading activation techniques in information retrieval. *Artif. Intell. Rev.*, 11:453–482, 1997.
6. J. Heflin and J. Hendler. Searching the web with SHOE. In *Artificial Intelligence for Web Search. Papers from the AAAI Workshop. WS-00-01.*, 2000.
7. J. Hendler. The dark side of the semantic web. *IEEE Intell. Syst.*, 22:2–4, 2007.
8. A. Kiryakov, B. Popov, I. Terziev, D. Manov, and D. Ognyanoff. Semantic annotation, indexing, and retrieval. *Journal of Web Semantics*, 2:49–79, 2004.
9. T. Mandl. *Tolerantes Information Retrieval. Neuronale Netze zur Erhöhung der Adaptivität und Flexibilität bei der Informationssuche.* PhD thesis, University Of Hildesheim, 2001.
10. R. McCool. Rethinking the semantic web, part 1. *IEEE Internet Comput.*, 9(6):88–87, 2005.
11. S.E. Robertson and K. Spärck Jones. Simple, proven approaches to text retrieval. Technical report, University of Cambridge, Computer Laboratory, 1994.
12. C. Rocha, D. Schwabe, and M. P. de Aragão. A hybrid approach for searching in the semantic web. In *Proceedings of the 13th international conference on World Wide Web, WWW 2004*, 2004.
13. M. Sabou, M. d'Aquin, and E. Motta. Using the semantic web as background knowledge for ontology mapping. In *International Workshop on Ontology Matching (OM-2006)*, 2006.
14. G. Salton. Associative document retrieval techniques using bibliographic information. *JACM*, 10:440–457, 1963.
15. G. Salton. *Automatic Information Organization and Retrieval.* McGraw Hill, 1968.
16. P. Scheir, P. Hofmair, M. Granitzer, and S. N. Lindstaedt. The ontologymapper plug-in: Supporting semantic annotation of text-documents by classification. In *Proceedings of the SEMANTICS 2006*, 2006.
17. K. Spärck Jones. What's new about the semantic web?: some questions. *SIGIR Forum*, 38:18–23, 2004.
18. Z. Wu and M. S. Palmer. Verb semantics and lexical selection. In *Meeting of the Association for Computational Linguistics (ACL)*, pages 133–138, 1994.
19. E. Yilmaz and J. A. Aslam. Estimating average precision with incomplete and imperfect judgments. In *CIKM '06: Proceedings of the 15th ACM international conference on Information and knowledge management*, 2006.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Collaborative Wiki Tagging

Valentina Milićević
(University of Nis, Faculty of Electronic Engineering, Serbia
valentina@elfak.ni.ac.yu)

Milorad Tošić
(University of Nis, Faculty of Electronic Engineering, Serbia
mbtosic@yahoo.com)

Abstract: Wikis as well as collaborative tagging have been subject of very intense research last years and active discussion topics in the so-called blogsphere as well. In this paper, we propose collaborative Wiki tagging based on the idea to exploit inherent semantics of the concept of link in Wiki. The low-level integration of Wiki and collaborative tagging of web resources is expected to be effective in enterprise particularly in the personal and group knowledge management application area. We first introduce conceptualisation of collaborative Wiki tagging. Then, we propose a simple scheme for using one of the existing native Wiki syntax to represent tagging data. Collaborative Wiki Tagging Portal Prototype, developed as a proof of concept, is used to give illustrative practical examples of the proposed approach and illustration of the user interface.

Keywords: Knowledge Management, Wiki, Collaborative Tagging
Categories: H.3.3, H.3.5, H.5.3, H.5.4

1 Introduction

Last years we have witnessed a growing interest in enterprise world for the process of generating, managing and sharing knowledge. This knowledge management process is recognised as being of crucial importance for enterprise to be able to effectively manage innovation. The growing need for continuous innovation results in a pressure to exploit all potential of a single person, group, as well as community to generate new knowledge through collaborative means. Also, traditional knowledge management technologies have not delivered as promised. In the same time, new developments are becoming more attractive, such as Semantic Web, social computing, open systems, emergent semantics, etc. As a result, collaborative systems, that provide technological capabilities for collaborative interaction among multiple participants with shared goals and interests across time and place, have recently gained considerable attention [Accenture, 03].

Collaborative tagging systems appear as a new trend in collaboration, gaining growing popularity on the Web. Purpose of those systems is to organize web pages and objects, as a set of resources, representing a new paradigm of organizing information and knowledge on the Web. The collaborative tagging is expected to take a leading role in knowledge work related fields such as information storage, organization and retrieval [Macgregor, 06]. Discussion within blogsphere on the concept of tagging, tagging applications, problems that tagging process retrieves, social and cognitive analysis of tagging [Sinha, 05], tagging formats, tagclouds, hierarchy versus tagging for information classification etc., is very active. Results of the more thorough scientific research have been published also.

In this paper, we propose low-level integration of collaborative tagging and Wiki, in the sense that tagging data is stored using native Wiki syntax. In this way we are able to apply all functionalities available in the Wiki on the embedded tagging data. This includes collaborative editing and full-text search of tagging data, Ajax-like interface components, etc. Resources to be tagged are not limited on the internal Wiki pages only, but can be any Web resource. The idea of deep-integration approach to collaborative Wiki tagging, particularly relation between the concept of a link in Wiki and the semantics of tagging act in tagging systems, is in accordance with the latest research results about integration of semantics, Wiki, and social web. We believe that the concept of collaborative Wiki tagging removes some of the inherent bottlenecks related to group and personal knowledge management.

Rest of the paper is organized in the following way. In Section 2, we introduce concept of collaborative Wiki tagging. Syntax and Wiki text formatting rules for tagging data is discussed in Section 3. Section 4 gives overview of the *Collaborative Tagging Wiki Portal Prototype*. Section 5 discusses some of the recent research results that related to our work, while Section 6 concludes the paper and gives some pointer for future work.

2 Concept of collaborative Wiki tagging

In this paper, we introduce concept of collaborative tagging as an interaction procedure between two (or more) resources. By the interaction, link of a tagged resource is memorised together with some concomitant information about the resource. The link is memorized by the tagging resource (the tagging resource is also interchangeably called agent). The link is reification of identity of the tagged resource in the sense of information that is sufficient to initiate and conduct interaction protocol between participating resources*. The agent with memorizing capability is implemented as a Wiki page. Note that the agent's memorizing capability does not necessary mean intrinsic ability of the agent to activate the memorizing process. Instead, the process may be initiated by some other agent (for example, a human agent in the case of a wiki page). More formal definition of agents, resources, interaction protocol, an agent's knowledge, and addressing have been previously given in [Tosic, 06], and is out of scope of this paper.

In the system of wiki pages, information and knowledge is reified and stored within wiki pages. Traditionally, content of wiki page is interpreted as an informational content primarily used by humans as a document or a Web page.

---

* Identity management of resources on the Web, as well as Semantic Web, is very important while still open problem. Here, we adopt definition of identity relative to the interaction protocol wich is intended to be used. Informally, for an agent, that have intention to interact to some other resource using given interaction protocol, reification of identity of the peer resource is interpreted as information suficient for establishing this interaction protocol.

However, our approach to collaborative wiki tagging is based on assignment of explicit semantics to the content of wiki pages. We identify *1) presentation*, *2) tag* and *3) statement* semantic concepts that we will assign to wiki pages for the purpose of our target application. 1) Wiki page with assigned presentational semantics is a traditional wiki page that can be created and collaboratively edited by several users. The page stores plain text using wiki syntax and text formatting rules. 2) Tag is also wiki page but with assigned tag semantic. The tag semantics is codified within wiki page in a form of a link to a unique meta-semantics page titled "Tag". In other words, every wiki page that contains a link to specific, pre-defined page "Tag"[*] is treated as a tag. 3) Statement is a Wiki page that reifies semantic relation between any two Wiki pages (including tag as well as content pages). Subset of Wiki pages with assigned tags semantics is called TagCloud System Repository, as shown in Figure 1. Tag contains one or more links to external or internal resources. The action of tagging a resource is implemented by appending the tagging data to Wiki page of the every tag used. A Wiki page may belong to one or several tags (Figure 2).

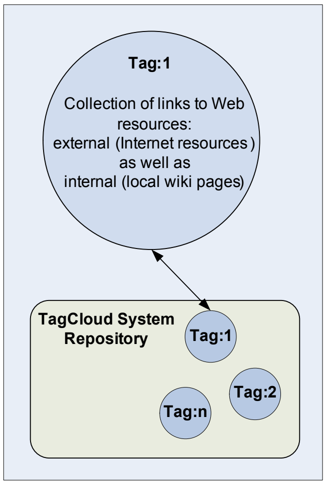


Figure 1: TagCloud System repository

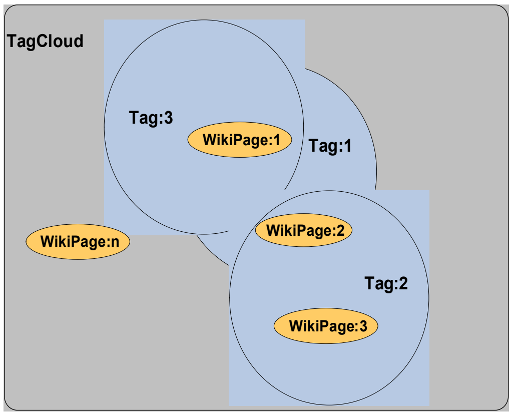


Figure 2: Wiki pages as subsets of tags using concept of tagging

For example, if we tag a presentation wiki page (or any other resource on the web) called A with a tag called B then link to the "Tag" meta-data wiki page and link to the resource A are appended to the content body of the tag page B. Note that there is no difference in format of the stored tagging data when we tag internal wiki page and when we tag any other external web resource. The same resource can be tagged with several tags.

[*] Note that we say nothing about content of the "Tag" page, so it may be empty or may contain some additional informational content.

3 Collaborative tagging Wiki syntax

In this paper, our aim is to store tagging data in Wiki using native text formatting syntax of the used Wiki. Consequently, we have to address two issues: 1) An ontology for tagging, and 2) Syntax for tagging data representation. We use a simple ontology for tagging based on the conceptualization developed elsewhere [Tosic, 06] and described in the previous paragraph. The adopted ontology is in compliance with existing ontologies for tagging [Story, 07]. Nested bulleted lists are used as a basic element of Wiki syntax for tagging data representation. Table 1. shows some of the JspWiki text formatting rules[*], that we use in tagging format. Note that even the syntax is Wiki engine dependent, the concept of the nested lists is not[**]. Therefore, the proposed tagging format can be generally applied in any Wiki engine.

Rule	Description
[link]	Create hyperlink to "link", where "link" can be either an internal WikiPage or an external link (http://)
[text|link]	Create a hyperlink where the link text is different from the actual hyperlink link.
*	Make a bulleted list. For deeper indentations more (**,***) is used.

Table 1: Some of the JspWiki text formatting rules

Using the JspWiki text formatting rules, a tagging data format is:

```
* [ResourceTitle|Link][|Tag] \\ Tags: [tag1][tag2][tag3]
** Clipping: content_of_clipping_if_exist
** Comment: link_to_comment_if_exist
** Posted on date&time, by [username|username_userprofile]
```

HTML preview of the tagging data is as follows:

```
<ul>
  <li>
    <a class="wikipage" href="Wiki.jsp?page=Link"> ResourceTitle </a>
    <a class="wikipage" href="Wiki.jsp?page=Tag"> </a>
    <br />
    Tags:
    <a class="wikipage" href="Wiki.jsp?page=tag1">tag1</a>
    <a class="wikipage" href="Wiki.jsp?page=tag2">tag2</a>
    <a class="wikipage" href="Wiki.jsp?page=tag3">tag3</a>
    <ul>
      <li>Clipping: content_of_clipping_if_exist</li>
      <li>Comment: link_to_comment_if_exist</li>
      <li>Posted on date&time, by <a class="wikipage"
href="Wiki.jsp?page= username_userprofile"> Username </a> </li>
    </ul>
  </li>
</ul>
```

Link may be a wiki-internal page link (page name reference) or an external resource link (URL address is written explicitly in the text, including the protocol prefix). In third line of the HTML code snippet, link to internal wiki page is assigned the

[*] http://www.jspwiki.org
[**] WikiCreole.org is developing an universal Wiki syntax for interwiki compatibility.

wikipage CSS class. In the case of tagging an external resource, the link is assigned the *external* CSS class.

Following the proposed syntax of the concept of semantics reification in Wiki, that we used to implement tagging, we are able to make statements about wiki pages. For that purpose, we introduce `Statement`, `Relation`, and `Category` meta-pages. Format of the Statement is as follows:

```
* [WikiPage_Title|WikiPage][|Statement]
**[Relation] [Category] [Tag]
```

`WikiPage` is the page that we make the statement about. The statement data is written in the wiki page called `WikiPageStatement`. The link `[|Statement]` means that the `WikiPageStatement` page is a statement saying that the `WikiPage` resource is in the `Relation` relation with the resource `Category`

For example, when we make a statement about wiki page `wikipageX`, statement page `wikipageXStatement` is created (if does not exist). Let the relation be `is_same_as`, category be `tagPage`, and tag be `tagX`. Then, format of the statement is:

```
* [WikiPage_Title|WikiPage][|Statement]
**[is_same_as] [tagPage] [tagX]
```

An uunlimited number of statements can be made for a single wiki page. All statements will be written within same statement page for that wiki page. For example, the statement says that `WikiPage` resource is also in relation: `belong_to` with the category `project`, while tag `projectName` assignes name to the project:

```
* [WikiPage_Title|WikiPage][|Statement]
**[belong_to] [project] [projectName]
**[is_same_as] [tagPage] [tagX]
```

4 Collaborative Wiki Tagging Portal Prototype

We have developed a testing prototype of the described collaborative Wiki tagging system, *Collaborative Wiki Tagging Portal Prototype (CWTP)*[*]. The CWTP is aimed to support personal knowledge management, inter- and intra-community collaboration, workflow and process management, interaction, knowledge sharing and dissemination, and heterogeneous information integration (Figure 3.). The prototype supports collaborative tagging, but it is a Wiki site in the same time. It means that every page can be edited, including pages that contain tagging data as well as meta-pages. Edit rights are not publicly available but are instead regulated by an authentication and authorization mechanism at the page level.

---

[*] http://infosys-work.elfak.ni.ac.yu/InfosysWiki-v2-1/Wiki.jsp?page=TagClouds

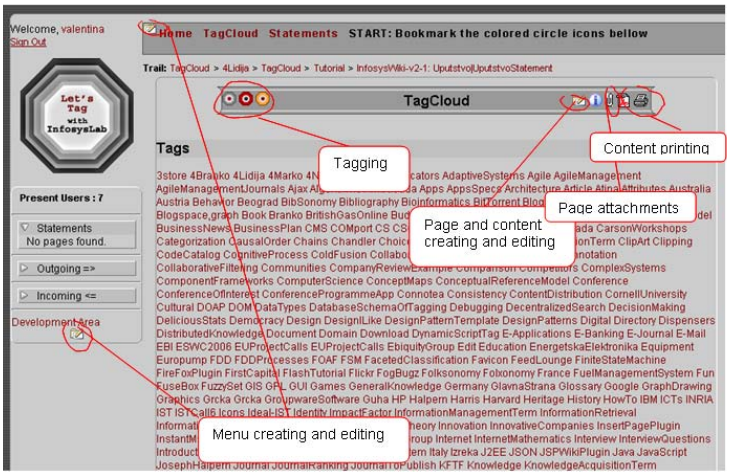


Figure 3: Collaborative Tagging Wiki Portal Prototype (CTWP)

Interaction over structure is represented by means of automatic set of page neighborhood links (links pointing to the page and links pointing from the page) and useful drop-down menus, as well as page-specific menus. Page neighbourhood links are useful for content but even more for semantic navigation. Interaction over structure is augmented with a primitive version of TagCloud that is useful for navigation over presentation content. Using TagCloud, users can navigate through content arranged over tagging hierarchy. Tagging presentation wiki pages and other internal resources allows systematic (re)arrangement of internal structure. Figure 4. shows tagging window for internal resource tagging, while Figure 5. shows external resource tagging.

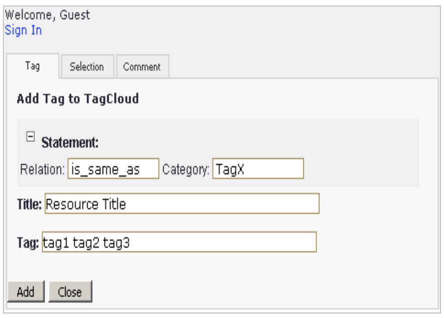


Figure 4: Internal resource tagging

Figure 5: External resource tagging

The differences in tagging windows are in statement and link. When we tag internal resource, we can make statement in the same time. In Tag field users can adduce several tags to which the resource belongs.

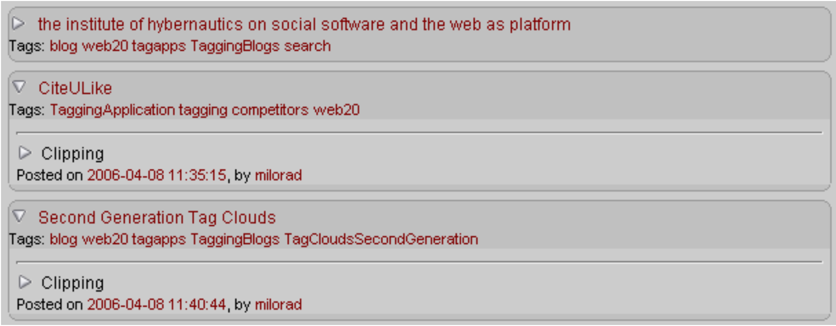


Figure 6: Screenshot of tagging data

5 Discussion and related work

Probably the most known approach to organizing information within a Wiki is Wikipedia categories. Collaborative Wiki tagging presented in this paper associate tags to wiki pages in a similar way. However, we are able to tag any web resource not a wiki page only. Also, cognitive investment made by user and his/her level of attention is much higher in the case of categorization then in the case of free tagging. With respect to low-level syntax, relevant research includes microformats: simple convention for embedding semantics in HTML to enable decentralized development[*], for web resources tagging called tag-rel[**]. Similarity between proposed wiki tagging syntax and .the microformats is inheritance of the existing legacy mechanism (HTML syntax vs. wiki text formatting syntax) and simplicity of the formats. Microformatcs can not be directly edited by end user while wiki tagging page can. However, the difference is simply matter of syntax, so transformation at the level of user interface is possible and is the subject of future research.

Second relevant stream of existing research includes Semantics Web related work, particularly Semantic Wikis [Oren, 06]. Among whole family of different Semantic Wikis, IkeWiki may be the closest to our approach [Schaffert, 06]. The proposed solution is low-level and in this way complimentary approach to Semantic Wikis. Also, semantic collaborative tagging system, as proposed in [Marchetti, 07], is based on semantic assertions that are very close to our Statements.

6 Conclusions

In this paper, we propose concept of collaborative tagging for organization of knowledge stored within a wiki system. We first introduced semantics of collaborative tagging that is implemented in a wiki fashion. Then, we discussed syntax and wiki text formatting rules that we use to store tagging data in a wiki system. Proposed concepts and syntax is used for implementation of our

[*] http://microformats.org/wiki/Main_Page
[**] http://microformats.org/wiki/reltag

Collaborative Wiki Tagging Portal Prototype. We use the prototype extensively for personal knowledge management, group knowledge interaction and project management. Recently, we have started testing the prototype in our undergraduate teaching practice, and we experienced very promising results: improved student-teacher communication, students being more actively involved into learning process, and management of the course being more interactive.

Acknowledgements

Authors want to tank anonymous reviewers for their very useful comments that helped us to improve readability and overall quality of the paper. The research presented in this paper is partially funded by the Serbian Ministry of Science and Environmental Protection.

References

[Accenture, 03] Dynamic Collaboration: Enabling Virtual Teamwork in Real Time, 2003, http://www.accenture.com/xdoc/en/services/technology/research/tech_collaboration.pdf

[Golder, 06] Scott Golder and Bernardo A. Huberman. (2006). "Usage Patterns of Collaborative Tagging Systems." *Journal of Information Science*, 32(2). 198-208.

[Macgregor, 06] George Macgregor and Emma McCulloch (2006), "Collaborative Tagging as a Knowledge Organisation and Resource Discovery Tool", *Library Review*, Vol. 55, No. 5.

[Marchetti, 07] A. Marchetti, et al. (2007), "SemKey: A Semantic Collaborative Tagging System", *Proc. WWW 2007 Workshop on Tagging and Metadata for Social Information Organization,* Banff, Canada

[Oren, 06] E. Oren, M. Völkel, J.G. Breslin, and S. Decker (2006), "Semantic wikis for personal knowledge management." In *Proceedings of the Interntional Conference on Database and Expert Systems Applications (DEXA)*, pp. 509-518.

[Schaffert, 06] Sebastian Schaffert, "IkeWiki: A Semantic Wiki for Collaborative Knowledge Management", In: *1st International Workshop on Semantic Technologies in Collaborative Applications STICA 06*, Manchester, UK, June 2006

[Sinha, 05] Rashmi Sinha, "A cognitive analysis of tagging," Weblog entry, *Rashmi Sinha's weblog,* http://www.rashmisinha.com/archives/05_09/tagging-cognitive.html

[Story, 07] Henry Story, "Search, Tagging and Wikis", Weblog entry, *The Sun BabelFish Blog*, 06. Feb. 2007., http://blogs.sun.com/bblfish/entry/search_tagging_and_wikis (12. July 2007)

[Tosic, 06] Milorad Tosic, Valentina Milicevic (2006), "Semantics of The Collaborative Tagging Systems", *"3th European Semantic Web Conference",* "2nd Workshop on Scripting for the Semantic Web", Budva, Serbia & Montenegro.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Social and Semantic Web Technologies: Semantic Navigation, Digesting and Summarization

Daniel Hladky
(Ontos International AG, Nidau, Switzerland
daniel.hladky@ontos.com)

Christian Ehrlich
(Ontos International AG, Nidau, Switzerland
christian.ehrlich@ontos.com)

Vladimir Khoroshevsky
(Ontos AG, Moscow, Russia
vladimir.khoroshevsky@ontos.com)

Abstract: The evolution of the semantic web brings new possibilities to handle the information overload. This paper presents a solution on how to navigate the web in a semantically meaningful way and to create reports of items of interest on demand. The proposed solution is based on semantic navigation, digesting and summarization of web-content. In the context of the proposed solution text mining is carried out by special agents, which provide an ontology-driven information extraction. Further we demonstrate the possibility of plug-in technologies for Internet browsers that allow digesting and summarization of web-content. This platform allows the creation of mashup services like integrating information from other sources or invoking an online function for purchasing. The paper also highlights the area of social bookmarking.

Keywords: Semantic Web, Multi-Agent System, Ontology-Driven Content Extraction, Semantic Navigation, Digesting, Mashups, Summarization

Categories: H.3.1, H.3.2, H.3.3, H.4.m, H.5.2, H.5.4

1 Introduction

It is generally known that, at the present stage, the worldwide process of accumulating and distributing information has a snowballing nature. Thereby the existing "traditional" information retrieval, storage and analysis methods are continuously loosing their efficiency, while their accuracy decreases and more and more redundant results are created.

According to a common expert estimation, about 80% of all reasonably accessible information is contained in natural-language texts. Further 70% of that information is repeated and creates no additional value to the customer. Meanwhile, the main shortcoming of "traditional" document processing methods consists in their inadequacy which becomes apparent when attempts are made to solve a problem of understanding the meaning of texts and performing, on this basis, semantic search and analysis of information.

To overcome the lack of meaning, the IT community has started several years ago the World Wide Web Consortium (W3C), which, in particularly, proposed the "Semantic Web" concept [Berners-Lee, et al., 01]. Several W3C standards like RDF and OWL have been created in order to maintain metadata [Deborah, et al., 04]. The OWL-like knowledge representation tools do also ensure a technological baseline for generalizing data, retrieving logical links between data elements, drawing inferences, and even finding solutions based on these inferences [Shadbolt,, et al., 06; , Zou,, et al, 04a; Zou,, et al, 04b].

Within this semantic wave stream Ontos AG develops and implements knowledge management technologies for building next generation information systems. The main innovations of Ontos AG are in the field of development and implementation of:

- intelligent systems for information-to-knowledge transformation based on the newest linguistic technologies allowing to extract information from texts and text collections in different natural languages under the control of domain ontologies;
- methods of knowledge handling that employ visual representation of the meaning of documents and/or collections of documents in the form of cognitive maps (special kind of semantic networks); methods of dynamic generation of textual digests from document collections; methods of automatic generation of textual summaries of documents and/or document collections in a specified target language;
- semantic navigation through web-content.

The main goal of this paper is to present the solution for semantic navigation through web-content and some related services. The paper also provides an outlook on future work.

2 Current Situation and Challenges

Today an Internet user who is interested in up-to-date information and news is using different portals or news provider. A common process of search is based on current search engines like Google, Yahoo and newer approaches like Powerset or Searchmash. In the last years it turned out that it becomes increasingly difficult to search with a meaningful focus while avoiding redundancy of results. Nevertheless, each content source might have additional information that would fill the complete picture of the information the user was looking for. There are many projects, academic and commercial, narrowed to overcome the above depicted problems in Europe [ESTC, 2007], USA [SemTech, 2007], and in other countries. Among other, impressive results were received within SEKT [Benjamins, et al., 04] and NEPOMUK [Völkel, 07] projects that bring together researchers, industrial software developers, and representative industrial users, to develop comprehensive solutions for Semantic Web. Interesting intelligent system for automatic extracting and searching knowledge was developed by the GATE team within the SEKT project [Lanfranchi, et al., 07]. An alternative approach presented by Gnowsis [Sauermann, 05] – semantic desktop environment published by the Knowledge Management Lab

of the DFKI and initially developed within the NEPOMUK project, supports semantic annotation of documents by the users with special knowledge management tools.
The next sections will describe our approach of how to tackle such challenges by using semantic web technologies.

3 Ontos Semantic Navigation and Analytics

The Ontos technologies and product line were designed for carrying out tasks pertaining to the processing of information from unstructured or semi-structured sources. The backbone of the solution is the processing and storing of the information retrieved from various content sources. In a first step we describe the main architecture and key components before demonstrating the user interface that allows navigating, digesting and summarization of the content.

3.1 Creation of metadata with multi-agent system

Data can be extracted and annotated from different content sources automatically. A peculiarity of this process is the usage of grid computing technology [Parastatidis, et al., 03] which provides for the system scalability by means of connecting additional computational power whenever the flow of documents is being processed requires so, and also for the higher reliability and fault tolerance of the system as a whole.

3.2 Resource Crawling

"Resource Crawling" means the monitoring of certain content sources (e. g. Internet pages), as well as the acquisition of text documents presented on these pages for processing. By default the resource crawling is performed automatically. In some cases, however, when the information extraction accuracy is of higher importance, the involvement of the system administrator in this process is possible. In such cases the administrator specifies the exact location of documents to be processed, and the conditions of extracting relevant text components from the pages.

The system allows setting up several types of agents responsible for the acquisition of input documents. Such agents scan data sources and retrieve documents by themselves. At the moment their tasks include:

- Crawling pages of Internet sites;
- Crawling file system folders;
- Processing RSS feeds.

Alongside with such agents, the Ontos technologies allow external applications to integrate into the Ontos Annotation Server using different APIs.

3.3 Text Mining – Ontos Annotation Server

The text mining processors are the intellectual core of the Ontos technologies. The tasks of this multi agent service include the automatic linguistic analysis of texts in accordance with specified domain ontologies and on the basis of special sets of linguistic and expert rules. Detailed discussion of an Ontos solution for a text mining component is presented in [Khoroshevsky, 04; Efimenko, et al., 05]. The extracted metadata will be stored in the Expert Knowledge Base (EKB).

3.4 Expert Knowledge Base on RDF-store

The Expert Knowledge Base (EKB) is built up automatically by importing the results of the text mining processes and performing merging and clean-up tasks based on configurable rules. It ensures the storage of entities (objects) and relationships between them extracted from documents, i.e. meta information relevant to a specified domain model, as well as of references to the processed documents. The general flow of information processing at EKB level is illustrated by Figure 1.

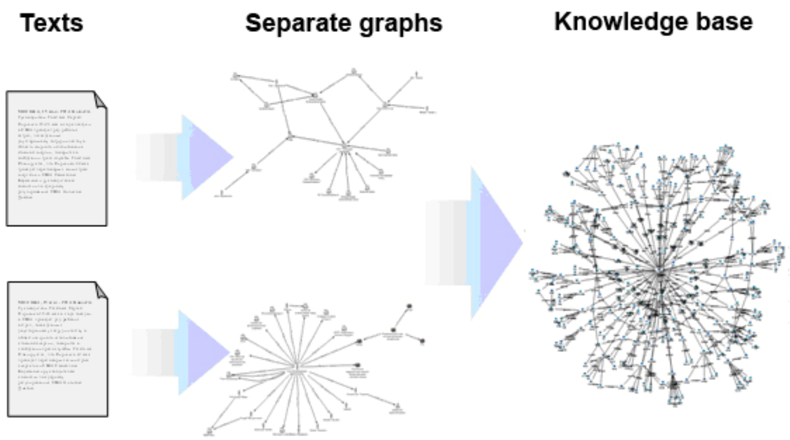


Figure 1: Flow of information at EKB level

4 The Semantic Navigation and Analytics Solution

The current Ontos solution for the Semantic Web includes:

- Semantic navigation
- Digesting and summarization
- Social bookmarking, e. g. manual annotation
- Mashups, e. g. purchasing from online stores

All the above depicted services and components are discussed in the next sections of the paper.

4.1 Semantic Navigation

Semantic navigation allows the user to navigate between the knowledge base objects along the relations, acquire information on attributes of these objects, and view the source documents.

From the user’s point of view, the service operation looks like the appearance of additional hyperlinks on a web page presented in a navigation card (Figure 2).

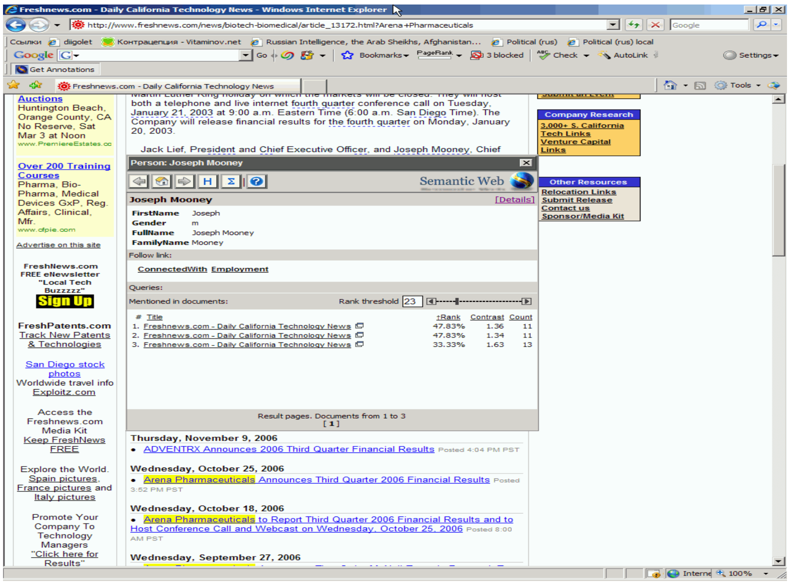


Figure 2: Annotated page with navigation card

After having been extracted by the system during the process of semantic analysis of the current page content, the entities and relations are “superimposed” on the document pages displayed in the browser window, thus establishing the correspondence between a given text fragment on the page and an object in the EKB.

When a mouse pointer is moved over a semantic link, the user can activate a navigation card, this card being a starting point for the semantic navigation. The navigation card contains:

- Object attributes.
- Object relations (navigational routes in the semantic web).
- Documents in which the object was found.

The navigation card allows navigating along semantic links between the objects. During this process, all links to the pages having occurrences of the objects are “drawn up”. The links to the pages are ranked in the card as per their semantic relevance. Therefore, the semantic navigation system integrates web sites from different locations covering similar or neighboring subject areas into one thematic portal.

The semantic relevance is calculated taking into consideration the importance of a subject within a document. The importance is not given by the number of occurrences only, but also by the semantic quality of references from other subjects within the document. Thus it is a very reliable measure to recognize the significance of a page.

The semantic navigation facilitates the "surfing" through the meaning of the content. It also reduces redundancy as duplicated entries have been merged during the mining and merging process. The navigation card allows additional functions to be executed. The following section describes some of them.

4.2 Digesting and Summarization

The semantic digesting service makes it possible to create an on-demand report (digest) consisting of text fragments taken from one document or collection of documents that contain information relevant to a given query. A digest is generated when navigating across the EKB. It includes document fragments (sentences) with entities and relations collected en route (Figure 3).

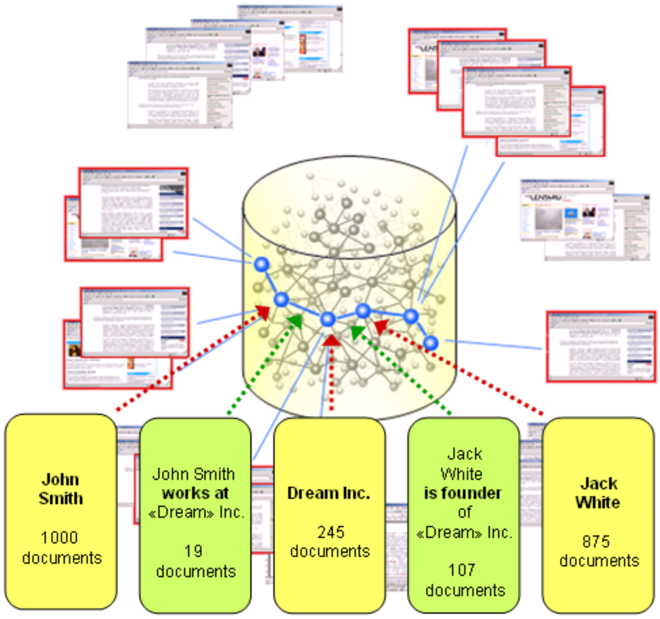


Figure 3: Source chain for digesting

The system stores navigating routes from their starting points up to the current ones, so a "navigation report", or semantic digest, can be generated at any moment.

When a digest is being generated, combinations of sentences with relevant objects and relations between them from evaluated documents are used. A document publication date and hyperlink to the source are indicated for each fragment of the digest (Figure 4).

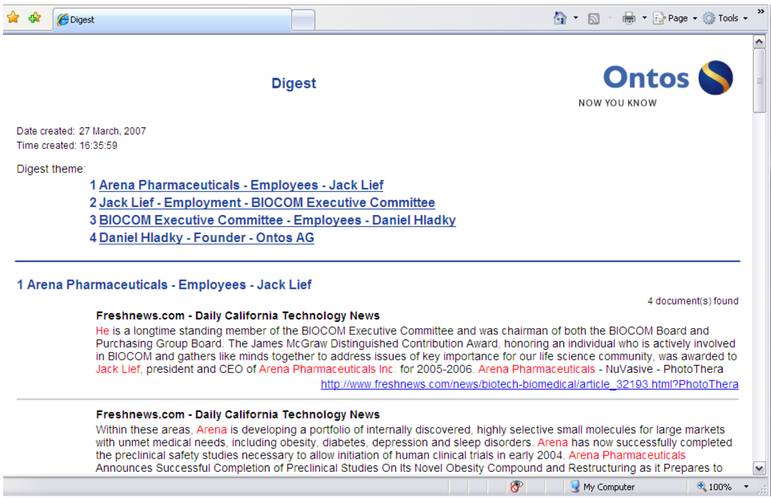


Figure 4: Digest report created on demand

In contrast to the digest, the semantic summarization is a process of creating a textual summary for one dedicated object (e. g. a specific person). The summary takes into consideration all available information in the expert knowledge base.

The summary can hold information of different relation depths, starting from the central object, considering the stored semantic network.

The semantic summarization service is designed for automating the routine work on the initial acquisition and analysis of relevant objects. It can include information extracted from documents of different languages, even languages the user is not able to read personally.

4.3 Social Bookmarking

The current version of the Ontos solution allows social bookmarking in a controlled way. The tool allows a knowledge engineer to annotate or bookmark the content within an Internet browser (Figure 5).

In order to avoid a metadata chaos the system restricts this process in two ways: Firstly, it requires the user to authenticate before being able to modify the data stored in the EKB. Secondly, it uses an ontology driven approach. The user can mark a text fragment and connect it to a named object by selecting the corresponding concept from the ontology.

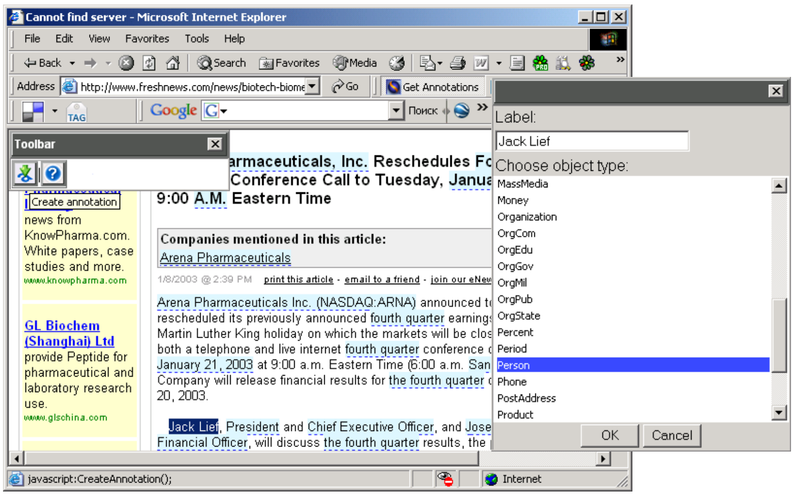


Figure 5: Manual annotation

4.4 Mashups Functions

The navigation card can be extended by Web services in order to start content sensitive processes or to read additional data for a given object and relation. An example is the possibility to place an online order for a specific product. One implemented scenario allows the user to semantically analyze symptoms in relation to drugs. Once a suitable drug is found, the navigation card allows the user to purchase the drug from an online drug store.

Further Web service functions are being explored and tested. Possible scenarios are:

- integrating an online map provider in order to show the location of an address,
- showing pictures or videos of a person or organisation,
- cross-linking to articles of well-known sources like wikipedia.org.

5 Conclusions and Future Work

The development of Ontos AG solutions for Semantic Web is based on text mining systems, which process multilingual text collections within the context of domain models represented by ontologies. In order to make the extracted information readable by human beings, appropriate intelligent services were discussed in this paper.

Ontos AG will enhance the presented solutions continuously by extending existing and adding new domain ontologies. In addition, the functionality of the services and the possibilities to interact with the system will be advanced massively.

References

[Benjamins, et al., 04] V. Richard Benjamins, Jesús Contreras, Mercedes Blázquez, Luis Rodrigo, Pompeu Casanovas, Marta Poblet, The SEKT use case components: ontology and architecture", In: Legal Knowledge and Information Systems, Tom Gordon (Ed.), IOS Press, Amsterdam, pp. 69-77, 2004

[Berners-Lee, et al., 01] Tim Berners-Lee, James Hendler, Ora Lassila, The Semantic Web, Scientific American, May 2001.

[Deborah, et al., 04] Deborah L. McGuinness, Frank van Harmelen, (eds.), OWL Web Ontology Language Overview, W3C Recommendation, February 10, 2004, http://www.w3.org/TR/2004/REC-owl-features-20040210

[Efimenko, et al., 05] Efimenko I.V., Khoroshevsky V.F., Klintsov V.P., Multilingual Content Extraction Systems: Ontology-Driven Semantic Analysis, In the Proceedings of International Conference SPECOM-2005, Greece, 2005.

[Khoroshevsky, 04] Khoroshevsky V. OntosMiner: Multilingual Information Extraction Systems Family. Proceedings of the CAI'2004, Tver, Russia, 2004 (In Russian).

[Lanfranchi, et al., 07] Vitaveska Lanfranchi, Ravish Bhagdev, Sam Chapman, Fabio Ciravegna and Daniela Petrelli, Extracting and Searching Knowledge for the Aerospace Industry, In: Proc. Of ESTC-2007, Vienna, 2007

[Parastatidis, et al., 03] Savas Parastatidis, Jim Webber, Paul Watson, Thomas Rischbeck, A Grid Application Framework based on Web services Specifications and Practices, North East Regional e-Science Centre School of Computing Science, University of Newcastle, Newcastle-upon-Tyne, NE1 7RU, United Kingdom, 2003.

[Sauermann, 05] Leo Sauermann. The semantic desktop - a basis for personal knowledge management, In: Proc. of the I-KNOW 2005, p.p. 294 – 301, 2005.

[Shadbolt,, et al., 06] Nigel Shadbolt, Tim Berners-Lee and Wendy Hall, The Semantic Web Revisited, IEEE Intelligent Systems 21(3) pp. 96-101, May/June 2006.

[Völkel, 07] Max Völkel, From Documents to Knowledge Models, In: Proc. of the 4th Conference on Professional Knowledge Management, p.p. 209-216, 2007

[Zou,, et al, 04a] Youyong Zou, Agent-Based Services for the Semantic Web, Ph.D. Dissertation, University of Maryland, Baltimore County, August, 2004.

[Zou,, et al, 04b] Youyong Zou, Tim Finin and Harry Chen, F-OWL: an Inference Engine for the Semantic Web, http://ebiquity.umbc.edu/get/a/publication/128.pdf

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

User-Driven Semantic Wiki-based Business Service Description

Heiko Paoli, Andreas Schmidt, Peter C. Lockemann
(FZI - Research Center for Information Technologies, Karlsruhe, GERMANY
{Heiko.Paoli, Andreas.Schmidt, Peter.Lockemann}@fzi.de)

Abstract: A key factor for success of companies operating in a globalized market environment is a modern SOA-based infrastructure. An essential component of a SOA infrastructure is the central service registry. Current standards for organizing service registries and their implementations are driven by the technical aspects of the infrastructure. When using such technically organized service registries, business users often fail to find the needed information. With the concepts of Web 2.0 in mind, we present a new approach to the organization and implementation of the business registries that are driven by the needs of business users. The paper discusses the problems of the current technically driven approaches, presents an architecture for a business user-driven service registry and introduces an implementation of the architecture using UDDI and Semantic MediaWiki.

Key Words: Service Oriented Architecture, Service Registry, UDDI, Web 2.0, Semantic MediaWiki, Ontology Engineering, Work Integration, Collaboration
Category: H.1.m, H.3.3, H.3.5, H.5.3, I.2.4, K.4.3, K.6.4

1 Introduction

A key factor for the success of companies operating in a global market environment is a flexible communication and information infrastructure that can be quickly and easily adapted to changing needs. Lately, service orientation has evolved as one of the more promising concepts for providing this flexibility. Information infrastructures that follow the paradigm of Service-Oriented Architecture (SOA) allow information processes to be defined conveniently and with minimal effort as a succession of calls on available services [Cearley et al. 2005, He 2003].

Judging from the many trade journals, service orientation does not yet live up to these expectations. We claim as our thesis that the failure is due to service descriptions that are of little help to the business users. Current descriptions have been written by service developers and just cover technical aspects such as service interface, formal parameters, or supported protocols. But this is not the world of the business users who initiate and control the business processes and react to numerous events in them. They need to know which services are available for which business purpose, which services have to be replaced when a business process has to be changed or whether new services are needed in order to adapt to new requirements [Huhns and Singh 2005].

As part of the solution we propose differentiating between different stakeholders. The design of information processes should be the responsibility of personnel that understands both, information systems and the business processes (we refer to them as *business analysts*). They need to know what the services have to offer to those executing the business (the *business users*). How these services have been technically implemented should be of little concern to them. The implementation of the services, and their connection to information processes, is the domain of *service developers*.

Service registries should address all stakeholders. Current service descriptions, though, concentrate on the service developers. To include the business aspects would be the task of the business analysts. The objective of this paper is to discuss how the analysts can be supported effectively to carry out this task. Any solution should keep in mind that in an environment subject to frequent change, service description cannot be a one-time affair but rather a continuous and collaborative effort among business analysts and service developers [Stojanovic and Dahanayake 2005].

Web 2.0 seems to be an appropriate interaction paradigm in which all stakeholders can be given an active part in service description. This paper presents a new collaborative and lightweight approach to describing services, and shows how business users can take an active part in it, so that a service registry would be able to cover their needs as well.

2 Problem Analysis

As discussed before, service discovery has technical and business ("semantic") facets. The technical part of a service description has always been formulated in a way to make algorithmic processing possible. For the purpose of computer-assisted service discovery the same should hold for the semantic part. Consequently, the business analyst must build a formal model of his or her conceptualization of the business domain, and relate the services to this model.

Technical descriptions specify how services can, and must be used within a computational environment. Consequently, technical descriptions should only concern the service developers. Likewise, semantic descriptions should be solely of interest to the business analysts and users. Moreover, being an abstraction the same service implementation may be applicable in different business situations and, hence, may have more than one semantic description. Consequently, both for technical and application reasons the technical and business aspects of the service description should be kept separate, something that has been known in software engineering as *separation of concerns*[1].

[1] Progr. for Separation of Concerns, `http://www.dmi.unict.it/~tramonta/PSC07/`

Modern business is not a static affair. Continuous change to the business descriptions in the registry is, therefore, a constant challenge. In today's interlinked world the flexibility of SOA should be complemented by a flexible approach where the organization of the business registry should be turned into a collaborative and continuous task along the lines of, say, the Web 2.0 concept.

To summarize, a business-oriented service registry should meet three requirements:

R1 Capture the semantics of business aspects to make services more accessible to business analysts

R2 Keep technical and business aspects of the service description separate for optimal support of the different user groups

R3 Support the collaborative and dynamic evolution of the service description to accommodate changing needs

3 A Comprehensive Approach to Business Service Description

3.1 Basic architecture and workflow

UDDI is practically the only standard for publishing technical descriptions, well suited for this purpose, but lacking the capability to describe business aspects appropriately. To meet requirement R2 Uddi is used a basis. To fulfill requirement R3, a Semantic MediaWiki (SMW) is taken as a collaborative front end of the business registry. Finally, to satisfy requirement R1 we employ ontologies to capture the network of related terms. In particular, our aim is a lightweight ontology that can be easily handled by business experts without extensive training in ontology engineering.

Figure 1 shows the system architecture. It consists of four main components: a UDDI-based technical registry, a SMW-based business registry, an ontology server and an ontology engineering component. The figure also indicates the basic workflow within the architecture. A software developer can use any UDDI-compatible client to publish a technical description of a new service. The developer may add some keywords based on the ontology which are used as an initial categorization for the service. The content of the UDDI Registry is dynamically embedded into the content of the SMW. From now on business users can search or navigate along the content of the SMW and provide additional business information. A SMW is chosen to make the content machine-understandable and to add implicit facts with the help of an ontology server. The ontology engineering component allows the business users to adapt the used business ontology to their needs in a lightweight and collaborative way.

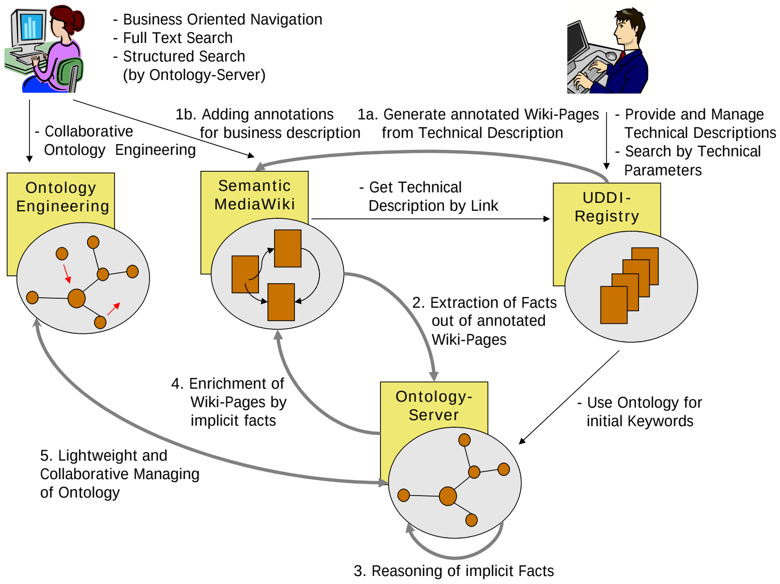


Figure 1: Combining UDDI with a Semantic MediaWiki

3.2 Ontology and Lightweight Ontology Engineering

The left-hand side of Figure 2 gives an example of the organization of our ontology. The top level part provides the domain-independent concepts such as the terms *Concept*, *Business Object* and *Service*. These are refined to a network of concepts of the business domain. Figure 2 shows just three examples of such concepts: the terms *water*, *water level* and *water gage information*. In a collaborative environment the presentation of ontologies is particularly important for effective and efficient use by the business analysts. The business registry is presented in the form of Wiki pages, with relations between concepts mapped to semantic links (right-hand side of Figure 2). In the example, *water* is a top level (business) concept while *water gage* is a business object concept, and *water gage information* is of type *service* and stands for a published service in the SOA infrastructure which will return a *water gage*. We use OWL-Lite as the ontology description language. Currently we use KAON2 as a reasoner, but any other compatible reasoner is also possible [Motik and Sattler 2006].

In the dynamic business environment the ontology itself is also bound to change frequently (see Section 2). Rather than entrusting a central authority with modifying the ontology we rely on the combined and distributed competency of all business analysts. Accordingly, we let the ontology evolve in collaboration of the business experts whenever one sees the need. Since we cannot expect the analysts to be experts in building ontologies, the engineering of the ontology should be made as simple as possible. We ease the task in two ways. For one the ontology is visualized as a graph, and all modifications can be easily done by dragging and dropping the nodes of the visual presentation rather than in some formal language. Second, the range of possible modifications is restricted (hence the name "lightweight engineering"), e.g., concepts can only be connected via *broader-narrower* and *related* relations. By using Wiki pages all modifications become immediately visible to other analysts.

3.3 Annotation of Wiki pages and service discovery

A service such as *water gage information* is initially entered into the system by its developer, who publishes it to the UDDI registry and is encouraged to augment it by intuitive keywords found in the ontology. After the publication a Wiki page is generated for the service, and automatically annotated with the keywords as well as semantic links that are obtained from the relations of the general UDDI data model. Subsequently, the business analyst may carry out the annotation of Wiki pages by means of such SMW features as semantic links, semantic attributes, and inline queries (to embed dynamic content). Many annotations can be obtained from the ontology by navigating through it and extracting further facts, or by using the reasoner to derive implicit facts. For example, on the left-hand side of Figure 2 the solid arrows represent relations that are explicitly available from the ontology (*hasType*, *belongsTo*, *provides*), while the dashed arrows represent relations that are implicitly available because of reasoning through the ontology server. Not only does our approach satisfy requirements R2 and R3, but it clearly does so with great benefit to both, business analyst and service developer. A business analyst can concentrate on the business description and organize and annotate the Wiki pages freely. For example he or she may express the business context of a service, e.g., business use cases, business value etc. The business description is limited neither by the data model of UDDI nor the facilities of WSDL. On the other hand the UDDI registry remains compatible to current SOA implementations.

We do not foresee automatic service discovery. This explains the emphasis we give to the presentation via Wiki pages. Take again the right-hand side of Figure 2. Note that much of the page contents for all terms is automatically generated. In particular, business object pages list all relevant services. Consequently, our approach satisfies requirement R1 as well. The proposed organization of the business registry and the use of a domain ontology well known to the business analysts and users provides a familiar and easy-to-use environment for them.

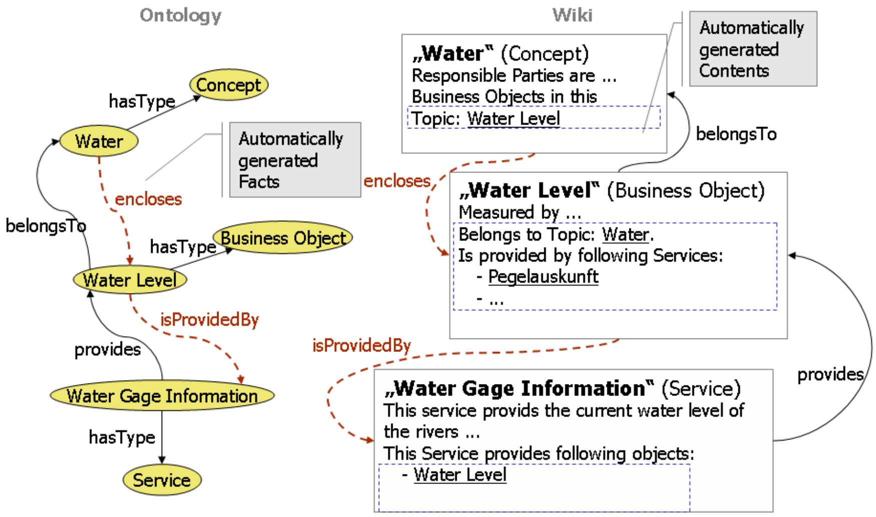


Figure 2: Organization and presentation of the business registry

4 Implementation

The implementation of our service registry consists of a central relational database, which holds the UDDI entries, the SMW pages and the ontology. On top of the relational database we have a J2EE application server and an HTTP server with PHP support. The J2EE application server represents the technical UDDI-compatible registry, which is realized through a standard UDDI framework. This way our implementation is fully compatible to the standard and explicitly allows publishers to use their own UDDI browser if they wish to. The HTTP Server with PHP support represents the business-oriented registry realized through an extended SMW component – the extension is necessary to support the automatic generation of content from the UDDI registry. For ontology engineering we use the existing tool SOBOLEO, a Web-based implementation of a Simple Knowledge Organisation System [Zacharias and Braun 2007].

5 Related Work

Automated discovery of Web services in a UDDI environment using WSDL descriptions is the subject in [Sivashanmugam et al. 2003]. The authors propose the use of an ontology to annotate WSDL message parts in order to add the necessary semantics. The semantics is entirely embedded in WSDL and thus cannot be separated from the technical description. While this may appear acceptable for automatic discovery, it burdens the business analyst, not the least because the semantics is expressed in notations unnatural to the business user. The approach thus fails to meet requirement R2. A bit earlier, [Paolucci et al. 2002] took a similar approach. The authors present in greater detail an algorithm for matching service requests to advertised services based on semantic descriptions. From the point of view of R2 their approach seems somewhat more advanced, since the DAML-S semantics is kept on a semantic layer. A DAML-S/UDDI translator is used to connect the semantic layer to the UDDI registry. But it seems doubtful that business users would feel comfortable with the semantic description or would consider the approach transparent enough to evaluate the outcome of their search.

Consequently, separation of technical and business concerns has not been a pressing issue in the past. [Bergmans et al. 2001] introduce a general model for examining whether and when it makes sense to compose systems from multiple concerns. The authors define a category of composability problems inherent in given composition models and provide criteria for using the separation of concerns paradigm. We conclude that our approach does not fall into the category of composition anomalies so that requirement R2 is indeed justified.

Wiki as a standard for collaborative authoring has been proposed in the past. The authors of [Krötzsch et al. 2006] extend the Wiki concept so that the content of a Wiki becomes machine-processable, and provides an embedded query language. Citing a number of reasons, they seem to confirm that a SMW is ideal as a front end for business analysts. More specifically, for the collaborative task of continued engineering of lightweight ontologies the SOBOLEO system employs a tagging mechanisms: interesting information is shared within a community and tagged by the latter to categorize it [Zacharias and Braun 2007]. Concepts of a lightweight ontology can then be derived from the used tags. The ontology is constructed and changed in a collaborative and Web 2.0-like way.

6 Experiences and Conclusions

The work presented in this paper has its origin in a project that was financed by the Ministry of Environment of Baden-Wuerttemberg. The environmental administration of Baden-Wuerttemberg has a long experience with environmental information systems in service oriented architectures. At the moment a redesign to a modern SOA-based infrastructure is planned by the State Institute for Environment, Measurements and Nature Conservation on behalf of the Ministry of Environment. The main objective is to provide all relevant parts of the system as services by a registry, and it should be possible to add a wide though unknown range of the services in the future. The system should be capable of handling hundreds of business users and service developers. To avoid duplicate work and to make all published services transparent to all business users a business oriented service registry seemed essential. The initial ontology we have used is based on an already existing and widely used taxonomy developed for the environmental information system of Baden-Wuerttemberg. The technical infrastructure as described above was developed in close communication with more than 10 representatives of business analysts and 5 representatives of developers, and was rolled out for a first testing period in April of 2007. First feedback by users sounds encouraging.

The thesis underlying our work is that service orientation will become widespread only if services can be discovered and employed with ease not just by service developers but also by business analysts. We have translated the needs to three requirements, the separation of technical and semantic descriptions, natural use of the semantic descriptions by business people, and a collaborative approach to dealing with the business dynamics. First experiences seem to support our thesis for the narrow scope of environmental information systems. What is definitely needed is more systematic and wider ranging empirical studies before we can be sure that our approach is an important step in overcoming the still existing doubts on the effectiveness of service-oriented architectures.

References

[Bergmans et al. 2001] Bergmans, L., Tekinerdogan, B., Glandrup, M., Aksit, M.: "Composing Software from Multiple Concerns: Composability and Composition Anomalies"; in Proc. of Int. Conf. on Softw. Engineering 2001, Toronto, Can., 2001

[Cearley et al. 2005] Cearley, D., Fenn, J., Plummer, D.: "Gartner's Positions on the Five Hottest IT Topics and Trends in 2005"; Gartner Web Site, 2005. http://www.gartner.com/DisplayDocument?doc_cd=125868 as seen on 2006-09-20

[He 2003] He, H.: "What Is Service-Oriented Architecture"; O'Reilly, 2003. http://webservices.xml.com/pub/a/ws/2003/09/30/soa.html as seen on 2006-09-20

[Huhns and Singh 2005] Huhns, M., Singh, M.: "Service-Oriented Computing: Key Concepts and Principles"; in: IEEE Internet Comp., vol. 09, no. 1, pp. 75-81, 2005.

[Krötzsch et al. 2006] Krötzsch, M., Vrandecic, D., Völkel, M.: "Semantic MediaWiki"; In Proc. of ISWC, Lecture Notes in Comp. Sc., vol. 4273, pp. 935-942, Springer, 2006.

[Motik and Sattler 2006] Motik, B., Sattler, U.:"A Comparison of Reasoning Techniques for Querying Large Description Logic ABoxes"; In Proc. of the 13th Int. Conf. on Logic for Progr. AI and Reas. (LPAR 2006), Phnom Penh, Cambodia, 2006

[Nickull et al. 2006] Nickull, D., McCabe, F., MacKenzi, M.: "SOA Reference Model TC"; OASIS Web Site, http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=soa-rm, as seen on 2006-09-20.

[Paolucci et al. 2002] Paolucci, M., Kawmura, T., Payne, T., Sycara, K.: "Semantic Matching of Web Services Capabilities"; In First Int. Semantic Web Conf., 2002.

[Sivashanmugam et al. 2003] Sivashanmugam, K., Verma, K., Sheth, A., Miller, J.: "Adding semantics to web services standards"; In Proc. of the 1st ICWS, 2003.

[Shen 2004] Shen, Z.: "UDDI v3.0 (Universal Description, Discovery and Integration)"; OASIS Web Site, 2004. http://www.oasis-open.org/committees/uddi-spec/doc/spec/v3/uddi-v3.0.2-20041019.pdf

[Stojanovic and Dahanayake 2005] Stojanovic, Z., Dahanayake, A.: "Service-Oriented Software System Engineering, Challenges and Practices"; Idea Group Publish., 2005.

[Zacharias and Braun 2007] Zacharias, V., Braun, S.: "SOBOLEO - Social Bookmarking and Lighweight Engineering of Ontologies"; in Proc. of WWW '07, Workshop on Soc. and Collabor. Constr. of Struct. Knowl. (CKC), Banff, Canada, 2007.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

A Semantic Web Content Model and Repository[1]

Max Völkel
(FZI Forschungszentrum Informatik Karlsruhe, Germany
voelkel@fzi.de)

Abstract: There is currently no model that is capable of representing the actual *content* of web resources *together* with their semantic web meta-data. This paper presents requirements for semantic content management, a unified human-browsable and human-editable *semantic web content model* (SWCM), and its implementation *swecr*.

Key Words: meta-data, semantic web, content management, RDF
Category: H.3.7, H.5.4

1 Unifying Web and Semantic Web

This paper presents a content management meta-model combining the usability of the web with the expressivity and flexibility of the semantic web.

Although the semantic web is typically characterised as an extension of the existing web [1], there is no formal model describing the resulting mix of content *and* meta-data. The web is targeted for direct human usage e. g. by browsing web pages, but the semantic web is not. Therefore a simple merge of the two models does not result yet in a very usable model for content management.

Requirements for content management are discussed and presented in Sec. 2. The resulting model is called the *Semantic Web Content Model* (SWCM) and described in Sec. 4. Sec. 5 shows the architecture and implementation of swecr – a semantic web content repository.

1.1 The Web

REST [2] is the architectural style used for most parts of the world wide web. REST describes a set of addressable resources which are manipulated by sending self-describing representations to them (c. f. Fig. 1). One of the REST constraints is "hypertext is the engine of application state", which means each representation should contain the URIs of related resources. There is no defined way to model *typed* relations between resources, as RDF allows.

In in the WWW resources are addressed by URIs and representations are character streams plus metadata describing the encoding, type of content and

1 Part of this work has been funded by the European Commission in the context of the IST NEPOMUK IP - The Social Semantic Desktop, FP6-027705. Special thanks to Tim Romberg, Heiko Haller and Daniel Clemente for fruitful discussions.

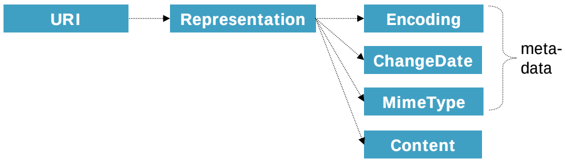


Figure 1: The Web Model: REST

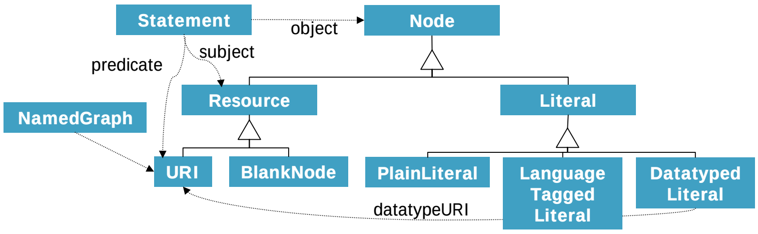


Figure 2: The Semantic Web Model: RDF

other metadata such as the last modification date. In practice, there are many more meta-data fields e. g. to control caching or compression of content.

1.2 The Semantic Web

The Resource Description Framework (RDF) [3] is the basic representation format for knowledge on the semantic web. It was originally defined as a format to describe meta-data about resources on the web. Fig. 2 shows the RDF data model together with the notion of NamedGraphs [4] as used in SPARQL [5]. As such it was never intended to *contain* the actual content of web resources.

Although RDF can conceptually contain binary data, stored in an `xsd:base-64Binary` datatyped literal, there is no defined way to relate URIs with content. Current RDF triple stores are not meant to store larger binary chunks either. Programming libraries for RDF lack ways to describe, access or change the content of web resources themselves.

A second problem with RDF is its lack of authoring tools. These can be divided into two classes: (1) generic: the user can change the schema at runtime, and (2) fixed-schema: the schema is pre-defined. An example of a fixed-schema tool is an address book editor which outputs its data in a fixed RDF format.

Authoring generic RDF without a pre-defined schema is very flexible, but has usability issues: e. g. each RDF resource can have none, one or multiple labels. It is an application level task to decide how to handle this. RDF can be called an *assembly language for data*, that can represent almost everything but lacks higher-order features to make it efficient for direct interaction with humans.

2 Requirements for a Semantic Content Management

Granularity (1). A Semantic Web Content Model must allow to describe web content.

The web began with small personal home pages and grew up with huge search and shopping portals. Since a few years there is a tendency for smaller content granularity, especially on collaborative websites. The term *micro-content* emerged for this set of addressable content consisting of tags (single terms), comments (often not more than a single paragraph), blog posts (often about half a page), images (including meta-data and a title) or video snippets. For most of these micro-content items, the author and the time of creation or last change are automatically logged and used for searching and browsing.

Expressivity (2). The model should be able to offer the same flexibility and expressivity as RDF to describe and relate content resources. Existing popular schema-free authoring tools such as mind-mapping or outlining tools lack the expressivity and data integration abilities of RDF, e. g. in many mind-mapping tools the user may in fact edit only strict trees. In short, existing (micro-)content management applications have low expressivity.

Compatibility (3). Furthermore, A clear path how to use the SWCM together with existing frameworks is desirable. Especially the re-use of existing background-knowledge expressed in RDF should be possible together with SWCM.

Naming (4). Names allow a user to fetch a unit of information in O(1). This is similar to know e. g. the URL of a certain web page or the file name and path of an office file. Human-usable naming is probably an overlooked area of content management. E. g. wikis allow users to use easy-to-remember names to quickly navigate or link to known pages. The semantic web is fundamentally built on URIs, which are unique names for resources. Unfortunately, they are hard to read and use for humans.

Search (5). Any content model should allow to retrieve content conveniently. Queries are usually convenient ways to retrieve a number of items fulfilling certain criteria. A SWCM needs also the ability to query the content, preferably by building on existing query languages.

Renderable representations (6). The model should be usable by end-users, hence some requirements are imposed for meta-data structures: All meta-data items should have a meaningful human-readable representation.

Mandatory inverse relations (7). In order to allow browsing semantic links in a knowledge model, links must be traversable in both directions. Therefore, it is desirable that link types have labels for both directions, e. g. "works for" and "employs". Note: In OWL, inverse relations are allowed but not mandatory

Freedom of formalisation (8). The user needs a way to express content in an informal way, e. g. as plain text, formatted text or box-and-arrow diagrams. Then the user should be able to migrate the knowledge into more formal structures, if desired (c. f. [6]).

Access rights (9). In any system used by multiple persons, access rights soon become a necessity. Fine grained access rights management can become very complex. If there are too many resources or rights to manage, the system becomes unusable.

Versioning (10). A model supporting versioning can be used better for collaborative settings, because users do not have to be afraid of applying changes. Changes not accepted by other community members can be rolled back.

3 Related Work

A number of related content models exist. This section describes them briefly and evaluates them with respect to the requirements.

JCR [7] defines the *Java Content Repository* API (JCR), which has quickly gained much industry attention. To date, there are at least four independent implementations of this standard . JCR handles granularity well, even a mix of large binaries and small single-term words has reasonable performance (1). The expressivity of JCR is also rather high, JCR has a concept of node typing and allows to add relations between nodes (2). However, JCR allows only XPath-style [8] queries and does not allow for graph-like queries (as they are supported by e. g. SPARQL), so (3) is not met. Requirements (4), (6), (7) and (8) are not met at all. (5) is met rather well, as JCR allows to use the familiar query languages SQL and XPath. (9) and (10) are well addressed.

Subversion [9] is an open-source versioning system (10) with access rights (9) and a number of interesting properties. It can handle small text files or larger binaries, but single terms are not in the focus. Therefore (1) is only partly met. Subversion allows to attach key-value pairs to resources, but no relations to other resources, so (2) also only partly met. Subversion repositories can be browsed (6) as trees and have meaningful names (4). There are even some best-practices for naming resources in a Subversion repository[2]. Subversion offers no search (5), and does not address (3), (7) and (8).

[2] http://svnbook.red-bean.com/en/1.0/ch05s04.html#svn-ch-5-sect-6.1

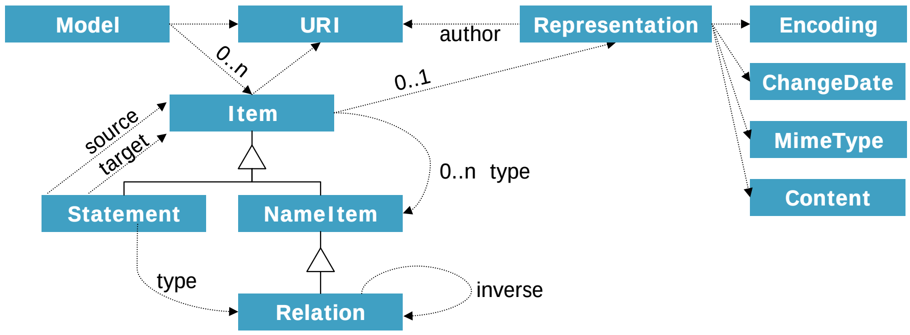


Figure 3: Semantic Web Content Model

4 A Semantic Web Content Model (SWCM)

A Semantic Web content model (SWCM, c. f. Fig. 3) consist of a set of *items*. Item is a central concept which bridges the RDF and the content world as it is both addressable via a URI and can refer to web content. All other elements in SWCM are special kinds of items. Therefore all SWCM entities are addressable and can refer to content of all sizes (Req. 1).

A *NameItem* is a sub-type of Item. It has a content snippet with the mime-type “text/plain”. NameItems represent human-readable and -write-able names. No two NameItems may have the same content within a single SWCM.

A *Relation* is a sub-type of NameItem. It has always exactly one *inverse relation* defined. This makes the model much easier to browse and visualise (Req. 7). E. g. in most semantic GUIs incoming links are rendered different from outgoing links. Therefore it makes a difference for browsing whether a user stated (“FZI” “employs” “Max”) or (“Max” “works for” “FZI”). For this user, this is often an artificial distinction.

A *Statement* is also modelled as a sub-type of Item. This makes it addressable and allows a user to attach content to it. A statement represents a relation from one item (the source) to another item (the target) and always has a relation type. In other words, a statement is like a typed link (Req. 2). Different from RDF, SWCM statements can be addressed themselves.

SWCM has at least the same expressivity as RDF. Each RDF statement (s,p,o) can be represented as two items (s and o) and a relation (p) with an inverse (-p). However, SWCM has some features, that RDF has not: (1) All SWCM content is addressable – this is not true for RDF literals. (2) SWCM statements are addressable, too – RDF has only a rather unclear concept of reification. The similarity between SWCM and RDF allows to convert RDF to SWCM (creating new URIs for literals and lifting them to items).

5 A Semantic Web Content Repository

This section describes the architecture of *swecr*, a *semantic web content repository* (swecr) available under BSD license from `http://swecr.org`. Swecr implements the SWCM.

5.1 Architecture

Swecr is implemented in two layers: The core layer models the state and the repository layer offers an item-centric view on the state.

The **core layer** consists of a simple binary store (BinStore), an RDF Named Graph repository and a text index.

The BinStore allows to access content in a stream and random access fashion. The BinStore interface also takes care of concurrency and allows either multiple reads or single writes. Note that many existing binary storage APIs offer no random access which makes it impossible to use such store for implementing projects such as Semantic File Systems [10].

The RDF repository is modelled as an RDF2Go ModelSet. RDF2Go is an abstraction layer over RDF triple- and quad stores which relives the programmer from choosing a single RDF store for all times. Currently OpenRDF is used as the underlying implementation.

In order to allow queries over the binary content and to speed up queries on RDF literals, Apache Lucene[3] is used as a full text index. The index stores inverse mappings for (item URI, content) and (URI, property URI, RDF literal content).

For indexing binary content, all binary content is indexed after it has been written to the BinStore. In a similar fashion all RDF literals are indexed after they have been written. The proxy pattern is used here to separate the API from the indexing. Removed resources have to be reflected in the index as well.

SPARQL queries should allow queries which access both the RDF and the fulltext index. Similar systems have been developed for Jena (LARQ [4]) and Sesame (LuceneSAIL [5]) already. The general idea is to split the query in two parts and execute them individually: One query part is delegated to the RDF store, the other part is delegated to the full-text index. Then a join is performed by the item URIs. Depending on the result set size, other join strategies should be favoured, i. e. first performing the full-text query and then binding the item URI in the RDF query to the resulting URIs. We are currently implementing this on top of RDF2Go to be independent of the triple store. Note that neither LARQ nor LuceneSAIL currently provide any API for handling binary content.

[3] http://lucene.apache.org/
[4] http://seaborne.blogspot.com/2006/11/larq-lucene-arq.html
[5] http://gnowsis.opendfki.de/wiki/LuceneSail

The core layer has no other obligations than starting, stopping and running the three core components: RDF store, text index and BinStore.

The **repository layer** has no persistent state on its own, instead all updates and queries are delegated to the core layer. This simplifies debugging and will make sharing of state easier.

The repository layer implements 1:1 the SWCM as outlined in Sec. 4. It allows to create, delete and manipulate Items, NameItems, Relations and Statements. Additionally, queries on models are provided.

The repository, however, maintains the runtime state of items. Each item can be locked. After locking, the item cannot be edited by other users. Locks time out automatically if they are now renewed by the requesting application. The requesting application may read who is currently editing a resources to be able to initiate communication processes. Such communication is outside the scope of *swecr*.

5.2 Implementation

This section describes how SWCM structures are stored in RDF. Two RDF models are used for each SWCM, one for the actual data as modelled explicitly by the user (user model) and one for the resulting plain RDF statements that can be used for queries and inferencing (index model).

An **Item** with URI x is simply stored as

```
<x> a swcm:Item.
```

The optional content of the item is stored in the BinStore. A **NameItem** with URI x and and the content (name) "My Thesis" is represented in RDF as

```
<x> a swcm:NameItem;  swcm:hasContent ''My Thesis''.
```

That is, the content of NameItems is currently stored in RDF. In the future, it might instead be stored in the BinStore - this matters only for performance, ease of debugging and the implementation of the query engine. A **Relation** p (e. g. "works for") with its inverse q (e. g. "employs") is represented as

```
<p> a swcm:Relation; swcm:hasContent "works for"; swcm:hasInverse <q>.
<q> a swcm:Relation; swcm:hasContent "employs";   swcm:hasInverse <p>.
```

A **Statement** s from a to b with the relation p is represented as

```
<s> a swcm:Statement;
    swcm:hasSource <a>; swcm:hasTarget <b>; swcm:hasRelation <p> .
```

For this statement, however, some data is written also to the RDF index model:

`<a> <p> <b>.` and `<b> <p-inverse> <a>.`

This means that we currently materialise the inverse triples into the index model. In the future, a rule engine or reasoner might be used instead.

6 Summary

This paper presented the Semantic Web Content Model (SWCM) for content management. The model was obtained by combining the features of the expressive model of the semantic web (RDF) with the human-targeted model of the web (REST). The resulting content model has been refined in a number of ways in order to fulfill the requirements for content management.

Although the resulting model has the same expressivity as RDF, the SWCM is targeted for browsing and authoring by humans. This comes at the cost of additional constraints, e. g. all content has a URI, each relation has an inverse, there can not be two relations having the same "label", etc. The SWCM and its implementation swecr are used in the NEPOMUK-project[6] to realize two end-user modeling tools: iMapping [11] and a Conceptual Data Structures (CDS) [6] authoring tool.

The paper also presented existing approaches and its shortcomings, and an implementation of the SWCM.

Future work includes: Împlementing SPARQL queries combining text index and triple store, access control, versioning, and carry out performance tests.

References

1. Decker, S., et al.: The semantic web: The roles of XML and RDF. IEEE Internet Computing **4** (2000) 63–74
2. Fielding, R.T.: Architectural Styles and the Design of NetworkBased Software Architectures. PhD thesis, U. Mass. (2000)
3. Hayes, P.: Rdf semantics. Recommendation, W3C (2004)
4. Carroll, J.J., Bizer, C., Hayes, P., Stickler, P.: Named graphs, provenance and trust. Technical report, HP (2004)
5. Prud'Hommeaux, E., Seaborne, A.: Sparql. W3C TR working draft (2005)
6. Völkel, M., Haller, H.: Conceptual data structures (cds) – towards an ontology for semi-formal articulation of personal knowledge. In: Proc. of the 14th International Conference on Conceptual Structures 2006, Aalborg University - Denmark (2006)
7. Nuescheler, D.: Content repository api for java technology specification. Technical Report Java Specification Request 170, Day Management AG, Switzerland (2005)
8. Clark, J., DeRose, S.: Xml path language (xpath) version 1.0. Technical report, W3C (1999)
9. Pilato, C.M., Collins-Sussman, B., Fitzpatrick, B.W.: Version Control with Subversion. O'Reilly Media, Inc (2004)
10. Bloehdorn, S., Görlitz, O., Schenk, S., Völkel, M.: Tagfs - tag semantics for hierarchical file systems. In: Proceedings of the 6th International Conference on Knowledge Management (I-KNOW 06), Graz, Austria, September 6-8, 2006. (2006)
11. Haller, H., Völkel, M., Kugel, F.: imapping wiks - towards a graphical environment for semantic knowledge management. In Schaffert, S., Völkel, M., Decker, S., eds.: Proceedings of the First Workshop on Semantic Wikis – From Wiki To Semantics. (2006)

---

[6] http://nepomuk.semanticdesktop.org

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

APOSDLE: learn@work with Semantic Web Technology

Chiara Ghidini
(FBK-irst. Via Sommarive, 18. 38100 Trento, Italy
ghidini@itc.it)

Viktoria Pammer
(Know-Center and Graz University of Technology, Knowledge Management Institute
Inffeldgasse 21a, 8010 Graz, Austria
vpammer@know-center.at)

Peter Scheir
(Know-Center and Graz University of Technology, Knowledge Management Institute
Inffeldgasse 21a, 8010 Graz, Austria
pscheir@know-center.at)

Luciano Serafini
(FBK-irst. Via Sommarive, 18. 38100 Trento, Italy
serafini@itc.it)

Stefanie Lindstaedt
(Know-Center, Inffeldgasse 21a, 8010 Graz, Austria
slind@know-center.at)

Abstract: The EU project APOSDLE focuses on work-integrated learning. Among the several challenges of the project, a crucial role is played by the system's ability to start from the context of the immediate work of a user, establish her missing competencies and learning needs and suggest on-the-fly and appropriate learning stimuli. These learning stimuli are created from a variety of resources (documents, videos, expert profiles, and so on) already stored in the workplace and may be in the form of learning material or suggestions to contact experts and / or colleagues. To address this challenge requires the capability of building a system which is able find, choose, share, and combine a variety of knowledge, evolving content and resources in an automatic and effective manner. The implementation of this capability requires technology which goes beyond traditional query-answering and keyword based search engines, and Semantic Web technology was chosen by the consortium as the most appropriate technology to make information search and data integration more efficient. The aim of this paper is to give an overview of the broad spectrum of Semantic Web technologies that are needed for a complex application like APOSDLE, and the challenges for the Semantic Web community that have appeared along the way.
Key Words: work integrated learning, semantic web technology, APOSDLE project
Category: J.4, I.2.4, I.2.6, I.2.11

1 Introduction

The EU project APOSDLE[1] focuses on work-integrated learning. In a nutshell, the project aim is to develop a software platform and tools to support the process of learn-

[1] `http://www.aposdle.org`

ing@work, that is learning within the context of the immediate work and current work environment of a, so called, knowledge worker.

To deliver a knowledge worker with context-sensitive learning material, tailored to her specific competences, work situation and learning needs, the APOSDLE system needs to know and manipulate, not only the specific domain of knowledge in which the knowledge worker is acting, but a set of other different aspects, spanning from processes, competences, learning needs and methods, user profiles, to the knowledge capital (documents, videos, expert profiles, and so on) available in the work environment and used by the APOSDLE system to build the learning material. Thus, to fulfil its goals, the ASPODLE system must be able to find, choose, share, and combine a variety of knowledge and knowledge artefacts in an automatic and effective manner. The implementation of this capability requires technology which goes beyond traditional query-answering and keyword based search engines, and Semantic Web technology was chosen by the consortium as the most appropriate technology to make information search and data integration more efficient. The goal of this paper is to give an overview of the Semantic Web technologies that are needed for a complex application like APOSDLE, and the challenges for the Semantic Web community that have appeared along the way. In particular we focus on a core set of functionalities that were realised as part of the first prototype of the APOSDLE system, and which constitute the starting point for future enhancements. This core set of functionalities allows starting from a description of the current profile of a knowledge worker, and from her current working task, to determine her missing competences on specific domain elements, to select appropriate material from the knowledge capital of the company and compose it in appropriate learning material. This learning material is finally presented to the knowledge worker to help her acquire the (missing) competences required to fulfil her task. The realisation of this core set of functionalities has required the capability to:

1. store an integrated representation of different domains of knowledge. This integrated representation is what we call the *APOSDLE knowledge base*;
2. connect the knowledge capital with the APOSDLE knowledge base; and
3. retrieve suitable knowledge capital in order to compose learning material.

In the remaining sections we illustrate the Semantic Web technology and challenges related to these three steps. We end the paper with a brief overview of the new technology that we aim at developing for the next version of APOSDLE.

2 The APOSDLE Knowledge Base

The APOSDLE knowledge base is composed of an integrated representation of the following models:

- *Domain Model.* The Domain Model is used to provide a conceptualization of the domain-dependent knowledge with which the knowledge worker (learner) is concerned, and to provide a vocabulary form the annotation of the documents which for

the knowledge capital of the organization. Our choice is to represent the Domain Model as an OWL ontology.
- *Task Model.* The Task Model represents the description of processes with which the knowledge worker is concerned. Since the main emphasis of the Task Model is to represent processes, and their dynamic aspects, our choice is to model processes using the workflow language YAWL (Yet Another Workflow Language) and its modelling tools [van der Aalst and ter Hofstede, 2005].
- *Competence Performance Model.* The Competence Performance Model is concerned with the description of the competences needed to perform tasks of a certain domain (see [Ley et al., 2007]). The competence performance model can be structured in a hierarchy from the more general competence to the more specific competence, and a natural implementation of such a hierarchy and the relation with tasks is via OWL.
- *Instructional Model.* The Instructional Model provides a conceptualization of the main concepts used to characterise the features of documents and competences from a learning perspective. It is mainly domain-independent and its main concepts are taken from the IMAT ontology [Barnard et al., 1999], which is represented in OWL in the APOSDLE system.
- *Knowledge Capital.* The knowledge capital we considered in the initial phase of the project are textual documents. In the following we adopt the APOSDLE terminology and use the more generic term of knowledge artefact instead of document (see [Consortium, 2006a]).

The definition of the APOSDLE knowledge base presents two important challenges, which are illustrated in the remaining part of the section.

Building the APOSDLE Knowledge Base. The Domain Model, the Task Models, and the Competence Performance Model, are domain dependent models which need to be created every time the APOSDLE system is configured and deployed for a new organization. If we consider the typical application environment of APOSDLE, we can safely assume that the organizations will not be interested in ontology engineering and workflow based process representation. Their main interest lies in setting up a tool that enhances the productivity of their work environment and offers to transform it into an integrated work-, learn- and collaboration environment.

This fact provides the APOSDLE developers with the first important challenge: to be able to provide tools that offer as much support as possible to automate and simplify the tasks of (domain) ontology engineering, workflow based process acquisition, and competence performance analysis and representation. This both in terms of graphic and easy-to-use tool interfaces, and automatic knowledge acquisition. During the first stages of the project we have focused mainly on the problem of supporting the Domain Model development, and in the current version of APOSDLE we provide a domain modelling tool implemented as a Protégé plugin. This tool allows us to use all the state-of-the-art facilities of Protégé, and also includes additional features similar to our earlier work described in [Scheir et al., 2006]. Among these features are e.g. relevant term extraction and document clustering which are starting points for helping the APOSDLE ontology engineer to elicit knowledge relevant for ontology creation.

Current and future work goes into two directions. On the one hand, usability-related issues must be addressed. To this purpose, the existing tool is being evaluated by application partners. On the more technical side, advanced ways of supporting users in the task of ontology creation are being researched. Among them are enhanced preparation of relevant terms (e.g. semantic grouping of terms using WordNet senses) and improvement of document clusters that are presented to the user. Concerning the preparation of relevant terms we are investigating the integration of the current Protégé plugin with tools that allows the automatic construction of ontology out of semi-ontological structures, as for instance: concept hierarchies, classifications, file system structure, database schemata. A detailed description of this approach is contained in [Bouquet et al., 2006]. Roughly speaking terms from semi-ontological structures are associated with WordNet senses, and general knowledge extracted from the hierarchical structure of senses in WordNet is used, together with possibly existing specific domain knowledge to support the phase of ontology construction. Other features may encompass guidelines to ontology creation, guided tours for ontology creation, support for selection of suitable ontologies from publicly available sources and embedded ontology evaluation. Our first attempt to embedded ontology evaluation is described in [Pammer et al., 2006].

Diverse sources of knowledge working together. The creation of the APSODLE knowledge base means that the models listed at the beginning of the section need to be integrated. This poses us with two problems: the first one is how to integrate such diverse models, especially OWL models and YAWL models, into a single structure. The second one is how to support the (semi) automatic integration of the different models.

During the first stages of the project we have focused mainly on the first problem. In particular we have decided to store the models in APOSDLE in their original format; this to exploit the peculiarities of each representational approach. In addition we have decided to integrate them by wrapping the YAWL model in OWL and expressing mappings between models in OWL itself. A graphical representation of the resulting structure is given in Figure 1. An additional advantage of this approach is that we can also access (query) the integrated models via query languages as SPARQL as if it was a single ontology.

A next step in APOSDLE is to support the automatic semantic integration of heterogeneous ontology / models. We plan to achieve this by using and extending semantic matchers like CTXMATCH [Bouquet et al., 2003].

3 Including Knowledge Artefacts into the Knowledge Base

Knowledge artefacts are connected with the APOSDLE knowledge base by means of annotations. We have chosen to annotate them with elements of the domain ontology and of the instructional ontology. This in order to indicate the topic(s) of a particular knowledge artefact and its instructional value (e.g., being an Example rather than an Introduction).

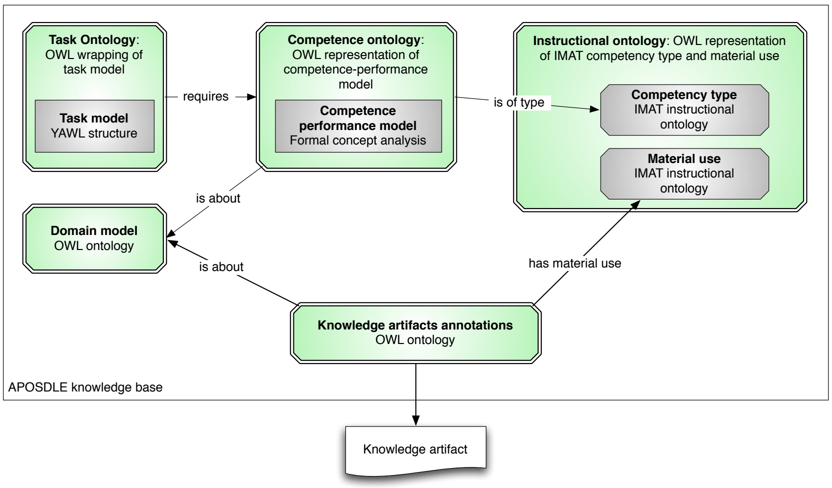


Figure 1: The integrated view

Most of the state-of-the-art semantic annotation tools ([Reeve and Han, 2005] for a survey) share the view that parts of documents are annotated. Within APOSDLE however, we do not need to annotate single names or small sentences of the documents. We rather need to annotate at a paragraph or page level, since we want to use these segments of text, here called knowledge artifacts, as *basic learning elements*, that is elements that can be used to teach a concept, a procedure, and so on. Simple examples of basic learning elements are: a self containing paragraph that provides a definition of a concept, or a graph that describes a particular sequence of actions to achieve a goal.

It is easy to see that a typical document produced in an organization can contain more than one basic learning element. Therefore the annotation is realized in APOSDLE in two main phases:

– *Document segmentation.* In this phase the text is segmented in parts, the knowledge artifacts that can be used for learning purposes.
– *Semantic annotation.* In this phase the knowledge artefacts are labelled with a number of concepts from the APOSDLE knowledge base. A simple example: a definition of ”Use Case” can be labelled with the concept ”Use Case” taken from the domain ontology and with the tag “definition” taken from the instructional model.

Similarly to the domain modelling tool, we have implemented an annotation tool as a Protégé plugin. Currently it supports manual annotation: Given an ontology, documents can be loaded into the annotation tool and semantic metadata can be assigned manually. Additionally, the current version offers content-based classification using the manually annotated documents as training set.

There are multiple challenges associated with the annotation part that must be addressed in APOSDLE. First, annotation must be performed w.r.t. different models. In the current version of APOSDLE, elements from the domain ontology *and* the instructional ontology need to be related to documents. As a low-tech workaround, the user could first annotate documents using the domain model and after that proceed to annotating using the instructional ontology. In the long run, this solution is hardly satisfying. So, annotation using multiple models is necessary. The challenge hereby lies as well in the usage of multiple models in Protégé as in a clear presentation to the user. Also note that above we stated “at least one domain model”. This reflects the fact that we do not suppose that we will always necessarily have exactly a single domain ontology. There might be multiple domain ontologies that represent multiple aspects of the domain to be learned, and this should not be a problem for annotation.

Another challenge lies in classification of content. Currently, the integrated classification is content based (see [Scheir et al., 2006]). This is fine for automatically assigning domain model elements to documents. However, instructional concepts like “Explanation”, “Introduction” or “Example” may not only be defined by the content but also by the structure of a document. Finally, it will also be desirable that annotation may be supported in the process of working with the APOSDLE system - files created in a certain context may be automatically annotated with corresponding concepts.

4 Enhancing Text-Based Information Retrieval with Knowledge Representations

In terms of information retrieval we see APOSDLE as application for the semantic desktop [Sauermann et al., 2005]. We use technology for the semantic web to build a desktop application to support the knowledge worker. In our approach we aim at combining database-like queries to a knowledge representation as found in semantic web technology with classical information retrieval approaches, i.e. the statistical analysis of document content. We expect to be able increase the performance of our system in terms of recall in the work support scenario and precision in the learner support scenario with this approach.

When defining a model of the context of a knowledge worker (see work described in [Ulbrich et al., 2006]) we noticed that there are three classes of objects that can be used to describe the current situation of the knowledge worker:

- A *set of concepts* of a knowledge representation that describes the situation of the knowledge worker, for example the current actions a person performs or the competencies he or she acquires.
- A *set of documents* that are related to the current situation of the knowledge worker, for example the document template he or she is currently interacting with, or the process documentation the person is reading.
- A *set of terms* which are related to his or her current situation, examples for such terms would be parts of documents the person currently views or a text he or she currently types.

To increase the chances of successfully supporting the worker with resources, i.e. to increase recall during a situation the person needs information to perform a certain task we defined an network model taking all three classes of objects (concepts, terms, documents) into account as query items [Scheir and Lindstaedt, 2006]. This model forms the basis of our information retrieval system and allows for searching documents based on content *and* semantic metadata (stemming from the Domain Model). Therefore we are implementing an associative information retrieval [Crestani, 1997] system based on the network model: in a nutshell, we create a large network (or graph) in which we combine concepts, terms and documents. Starting from a given set of documents, terms and concepts we traverse this network returning those documents most closely associated with the given query set. As measures of associativeness we are researching semantic and content-based similarly.

When suggesting learning material to a knowledge worker we aim at increasing precision by again taking semantic annotations into account. This approach is different from the one described previously but operates on top of it. The difference in the two approaches is that when suggesting learning material we have to assure that the person we retrieve resources for is able to learn the retrieved information. We have to assure that the information retrieved builds up a certain competency (i.e. fosters learning) and the person is able the execute the task without our help in the future. Therefore the presented material has to fulfil certain prerequisites, formally defined in the Instructional Model. For example, depending in the subject to teach we only present information of a certain type, as certain information is learned better by giving an example, other by providing a definition. The realize this approach, we take semantic metadata of documents (stemming from the Instructional Model) into account and filter the result set according to the current learning situation.

5 Conclusions

In this paper we have illustrated the main Semantic Web technologies that we have adopted in the first stages of the APOSDLE project to support the tasks of ontology engineering, ontology mapping, semantic annotation and information retrieval. This technology provides a starting point for the development of the APOSDLE platform and is included in the description of the first reference architecture [Consortium, 2006b]. New technology that must be developed to make the APOSDLE platform a success concern: (i) the ability to support user-friendly ontology engineering, (ii) the integration of different forms of ontologies (like the domain ontology and the competence performance ontology) and different forms of models (like the domain ontology and the task model). This requires matching algorithms that go beyond the identification of equivalent concepts in different ontologies; (iii) the automatic annotation of knowledge artefacts that can be used for learning needs, and (iv) advanced algorithms for information retrieval.

Acknowledgments

This work has been partially funded under grant 027023 in the IST work programme of the European Community. The Know-Center is funded by the Austrian Competence Center program Kplus under the auspices of the Austrian Ministry of Transport, Innovation and Technology (www.ffg.at/index.php?cid=95) and by the State of Styria.

References

[Barnard et al., 1999] Barnard, Y., de Hoog, R., and Wielinga, B. (1999). Imat: Re-using multimedia electronic technical documentation for training. In *EMMSEC Conference*, Stockholm.

[Bouquet et al., 2003] Bouquet, P., Serafini, L., and Zanobini, S. (2003). Semantic coordination: a new approach and an application. In *Proceedings of the Second International Semantic Web Conference*, volume 2870 of *LNAI*, pages 130–145.

[Bouquet et al., 2006] Bouquet, P., Serafini, L., Zanobini, S., and Sceffer, S. (2006). Bootstrapping semantics on the web: meaning elicitation from schemas. In *WWW '06: Proceedings of the 15th international conference on World Wide Web*, pages 505–512. ACM Press.

[Consortium, 2006a] Consortium, A. (2006a). APOSDLE Scope and Boundaries - Deliverable 6.1. EU-IST Integrated Project Deliverable 6.1.

[Consortium, 2006b] Consortium, A. (2006b). APOSDLE Software Design and Development Plan 1 - Deliverable 4.1. EU-IST Integrated Project Deliverable 4.1.

[Crestani, 1997] Crestani, F. (1997). Application of spreading activation techniques in information retrieval. *Artif. Intell. Rev.*, 11(6):453–482.

[Ley et al., 2007] Ley, T., Albert, D., and Lindstaedt, S. (2007). Competency management using the competence performance approach: Modeling, assessment, validation and use. In *Competencies in Organizational E-Learning: Concepts and Tools*. Idea Group, Hershey, PA.

[Pammer et al., 2006] Pammer, V., Scheir, P., and Lindstaedt, S. N. (2006). Ontology coverage check: support for evaluation in ontology engineering. In *FOMI 2006. The 2nd workshop: Formal Ontologies Meet Industry.*

[Reeve and Han, 2005] Reeve, L. and Han, H. (2005). Survey of semantic annotation platforms. In *SAC '05: Proceedings of the 2005 ACM symposium on Applied computing*, pages 1634–1638, New York, NY, USA. ACM Press.

[Sauermann et al., 2005] Sauermann, L., Bernardi, A., and Dengel, A. (2005). Overview and outlook on the semantic desktop. In *Proceedings of the 1st Workshop on The Semantic Desktop at the ISWC 2005 Conference*.

[Scheir et al., 2006] Scheir, P., Hofmair, P., Granitzer, M., and Lindstaedt, S. N. (2006). The ontologymapper plug-in: Supporting semantic annotation of text-documents by classification. In *Semantic Systems From Vision to Applications - Proceedings of the SEMANTICS 2006*, number 212 in books@ocg.at, pages 291–301. OCG.

[Scheir and Lindstaedt, 2006] Scheir, P. and Lindstaedt, S. N. (2006). A network model approach to document retrieval taking into account domain knowledge. In *LWA 2006, Lernen - Wissensentdeckung - Adaptivität, 9.-11.10.2006 in Hildesheim*, number 1/2006 in Hildesheimer Informatik-Berichte, pages 154–158. Universität Hildesheim.

[Ulbrich et al., 2006] Ulbrich, A., Scheir, P., Lindstaedt, S. N., and Görtz, M. (2006). A context-model for supporting work-integrated learning. In *Innovative Approaches for Learning and Knowledge Sharing - First European Conference on Technology Enhanced Learning, EC-TEL 2006*, volume 4227 of *LNCS*, pages 525–530. Springer.

[van der Aalst and ter Hofstede, 2005] van der Aalst, W. and ter Hofstede, A. (2005). YAWL: Yet Another Workflow Language. *Information Systems*, 30(4):245–275.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

PIMO - a Framework for Representing *Personal Information Models*

Leo Sauermann, Ludger van Elst, Andreas Dengel
Knowledge Management Department
German Research Center for Artificial Intelligence DFKI GmbH,
{firstname.surname}@dfki.de

Abstract: This paper presents the concept and realization of a *P*ersonal *I*nformation *Mo*del (*PIMO*). A PIMO is used to represent a single users' concepts, such as projects, tasks, contacts, organizations, allowing files, e-mails, and other resources of interest to the user to be categorized. This categorization using multiple criteria was used to integrate information across different applications and file formats. Based on RDF/S, multiple layers were defined: an upper-layer for a minimal set of generic concepts, a mid-layer for refinements, and a user-layer for concepts of the individual user. Our approach was deployed and used in several research projects. The PIMO helps users to categorize resources for Personal Information Management (PIM), it is intended to be the integrative part in personalized systems, such as Social Semantic Desktops.[1]

Key Words: personal information management, ontologies, personalization
Category: H.1.1, H.3.3

1 Motivation

In the EPOS project [6] the use of ontologies was suggested on all levels of the organization, starting with the desktops of individual knowledge workers. Their personal knowledge workspace encompass e-mails, files, contacts, projects and resources from the corporate intranet [2]. Today, products exist to create order on top of these structures, for example mind mapping tools, project management tools, or *Personal Information Management* (PIM) tools. However, the semantics of these structures is typically buried in the individual application.

The core of the EPOS approach is the *Personal Information Model* (*PIMO*). It is a formal representation of the structures and concepts an individual knowledge worker needs, according to her or his personal mental model. It is an application-independent and domain-independent representation. Concepts used to categorise elements in one application will also appear in other applications. Based on studies about file management we know about the importance of these structures for finding and reminding information [3]. The value of the existing

[1] This work was supported by the German Federal Ministry of Education, Science, Research and Technology (bmb+f), (Grant 01 IW C01, Project EPOS, and Grant 01 IW F01, Project Mymory) and by the European Union IST fund (Grant FP6-027705, Project NEPOMUK). The authors thank the EPOS project team for input and Laura Zilles for performing the study which led to the DFKI-mid ontology.

structures should be kept and augmented now. We have already developed the model of using multi-perspective classification [7] and how such structures can be used for personalization [13]. There has been research in synchronizing different applications to each other in a n:n approach [4], the PIMO model lowers the costs here because it allows each application to integrate with the PIMO and not to all other applications. For the *Semantic Desktop*, where Semantic Web technologies are already used, the PIMO is a cornerstone for data integration.

2 Definition of a PIMO

Formal structures such as OWL and RDFS ontologies have been used for Personal Information Management before [17], our approach includes a suggestion for a domain model for knowledge work, coined *PIMO-Upper*. Before taking a closer look, we give a definition of terms which are partly based on the definition of TopicMaps [11, 12]. The **Personal Knowledge Workspace** [9] (or "Personal Information Space" [17]) embraces all data "*needed by an individual to perform knowledge work*". It is (1) independent from the way the user accesses the data, (2) independent from the source, format, and author of the data. **Native Resources** are part of the personal knowledge workspace, personal files of the user, e-mails, and other PIM related resources, such as appointments or contacts. In Topic Maps, this is an occurrence that is categorized. **Native Structures** are categorization schemes for Native Resources such as file-system folders, bookmark folders, e-mail folders, tags. In the (common) case that a user operates in a document-centered way his internal representations of these concepts are already largely reflected in content and structuring of his information elements (see, *e. g.*, [7]). There are file folders called "projects", e-mail folders named with costumers' names, product names in file designators, etc. The **Mental Model** is part of the cognitive system of the person. Subjective to the person, the mental model is individual and cannot be externalized thoroughly. The PIMO aims to represent parts of the Mental Model necessary for knowledge work.
Now, a definition for a Personal Information Model can be given.

Definition 1. A PIMO is a Personal Information Model of one person. It is a formal representation of parts of the users Mental Model. Each concept in the Mental Model can be represented using a Thing or a subclass of this class in RDF. Native Resources found in the Personal Knowledge Workspace can be categorized, then they are occurrences of a Thing.

The vision is that a *Personal Information Model* reflects and captures a user's personal knowledge, *e. g.*, about people and their roles, about organizations, processes, things, and so forth, by *providing the vocabulary* (concepts and their relationships) for required expressing it as well as concrete instances. In other words, the domain of a *PIMO* is meant to be "all things and native resources that are in the attention of the user when doing knowledge work". Though "native" information models and structures are widely used, there is still much potential for a more effective and efficient exploitation of the underlying knowledge. We think that, compared to the cognitive representations humans build, there are mainly two shortcomings in native structures:

– *Richness of models*: Current state of cognitive psychology assumes that humans build very rich models, encoding not only detailed factual aspects, but also episodic and situational information. Native structures are mostly taxonomy- or keyword-oriented.

– *Coherence of models*: Though nowadays (business) life is very fragmented humans tend to interpret situations as a coherent whole and have representations of concepts that are comprehensive across contexts. Native structures, on the other hand, often reflect the fragmentation of multiple contexts. They tend to be redundant (i.e., the same concepts at multiple places in multiple native structures). Frequently, inconsistencies are the consequence.

The *PIMO* shall mitigate the shortcomings of native structures by providing a *comprehensive model* on a *sound formal basis*. In the following, we describe how the concept of a*PIMO* has been realized within the EPOS project.

3 Realization of the PIMO

When building concrete *PIMO*s, we now have the problem of two, potentially conflicting demands: On the one hand, we want to give the user the opportunity to span his information space largely in the way *he* wants. The *PIMO* should model *his* mental models. In consequence, we cannot prescribe much of this structure. On the other hand, "empty" systems often suffer from the cold start problem, being not accepted by user when not already equipped with some initial content. Using a multi-layer approach (see also [14]), we try to find a balance through providing the presentational basis as given, which users *can* incorporate or extend:

– PIMO-Basic: defines the basic language constructs. The class pimo-basic:Thing represents a super-class of other classes.

– PIMO-Upper: A domain-independent ontology defining abstract sub-classes of Thing. Such abstract classes are PersonConcept, OrganizationalConcept, LocationConcept, Document, etc.

– PIMO-Mid: More concrete sub-classes of upper-classes. The mid-level ontology serves to integrate various domain ontologies and provides classes for Person, Project, Company, etc.

- Domain ontologies: A set of domain ontologies where each describes a concrete domain of interest of the user. The user's company and its organizational structure may be such a domain, or a shared public ontology. Classes are refinements of PIMO-Mid and PIMO-Upper, allowing an integration of various domain ontologies via the upper layers.
- PIMO-User: the extensions of above models created by an individual for personal use. Classes, properties and things are created by the user.

3.1 The Representational Assumptions: PIMO-Basic

To represent things and their relations, different standards are available. First, RDFS which can represent classes and subclass relationships, properties and subproperty, and resources that are instances of classes and can be described with properties. Second is OWL which integrates reasoning capabilites and description logic. Third, there are standards to describe mind maps, such as the XTM standard. Using rules or description logic is not in the requirements for personal information models, therefore we have selected RDFS as our representation language instead of OWL.

In PIMO-Basic, two key concepts are introduced: the classes *Thing* and *ResourceManifestation*. *Thing* is a superclass of abstract concepts and physical objects, with the aim of representing them on a conceptual level. *ResourceManifestation* is a class to represent the documents in a computer system. The *native structures and resources* can be transformed to RDF as presented in [5],[15]. They are represented using subclasses of *ResourceManifestation*. The separation of Things from ResourceManifestations was missing in OWL and RDFS. Although SKOS models it as separation between concepts and resources, it doesn't reuse RDFS subclass-relations and therefore domain/range restrictions or any typed properties are not usable.

The idea is that a `Thing` can now occur in one or many resources. This is represented by a `occurrence` relation. For example, the city Rome (as a concept) can occur in a website about business in Rome (a document). This relation allows the annotation of documents according to the ontology. Certain occurrences are more tightly bound to a concept, when the topic of the document to describe exactly this concept, we model these as `groundingOccurrence`. A grounding occurrence of the concept of the City of Rome could be the wikipedia page about it. For people, the grounding occurrence of the Person "Paul" could be the address book entry with the contact information about "Paul", for a comany the website of the company. Independent of the application domain, grounding occurrences provide user-readable descriptions of the concept in question, and can be used to automatically map PIMOs of multiple users, when two concepts from different users have the same grounding, chances increase that the concepts are the same. This is comparable to XTM occurrence-references of non-addressable topics.

Besides implicit mapping using occurrences, it is also possible to explicitly map things. For this, the `hasOtherRepresentation` relation is defined. When two things are formalisations of the same concept, they can be mapped using this property. Note that this should not happen inside one user's PIMO but rather when domain ontologies or multiple PIMOs are mapped. Additional to instances, also classes can be mapped. For this, the meta-class `PimoClass` was created as subclasses of RDFS-Class. Using a meta-class allows adding annotations about classes in a clean way. Mapping classes is realized with the `hasOtherConceptualization` relation.

3.2 DFKI-KM-Mid: Acquisition of an Exemplary PIMO Mid-Level

The upper level of a *PIMO* just makes a few, basic ontological statements about things which exist on a Semantic Desktop, *i. e.*, things which are essential in a knowledge worker's mental model: *Information elements*, *people-*, *organization-* and *process-related things*, but of course also basic ontological categories like *space* and *time* concepts well-known (and imported) from other typical upper-level ontologies. Obviously, the commitment in this statement is very fundamental for the concept of a Semantic Desktop, but also very abstract. In order to avoid a *cold start problem*[2] with PIMO-based applications, we pre-modeled a PIMO-Mid-Level as a refinement of the upper level which serves two purposes: Firstly, the concepts of the mid level serve as *anchor points* for a user's personal incremental extensions of his PIMO. For example, having already a couple of *project types* as examples in his PIMO (instead of just having projects as abstract organizational concepts) makes it probably much easier for him to classify already existing projects or to model new project types. Moreover, offering a common mid level layer to a *group of people* can also be seen as a *seed for a shared conzeptualization* between these people, facilitating information exchange on the basis of these shared parts of their PIMOs. So, conceptually, the *scope* of a PIMO mid-level is a *group of user's* who potentially share many concepts on their Semantic Desktop (*e. g.*, people in the same department), while the control with respect to extensions or modifications is intended to be at the individual user.

In our prototype, we modeled an exemplary PIMO mid-level using the following methodology, consisting of the three phases *seeding*, *reality match*, and *evolution*: In the seeding phase, a couple of exemplary *native structures* (file and email folders) of members of DFKI's Knowledge Management Department were manually analyzed and so laid the basis for an *initial DFKI-KM-Mid model*.

[2] The problem of cold starts is very well known in knowledge-based systems: In the beginning a system, like a shell, just has little of no information and therefore seems not to be useful to a new user. Consequently, he is not motivated to invest in using and feeding the system with new information which would be a prerequisite to be *more* useful.

DFKI-KM-mid mainly consisted of concepts without deeper modeling, like attached slots etc. In the second phase, this initial model was checked by a detailed survey. 23 members of the department were interviewed whether the initial model fit their individual native structures, which concepts were missing in the model or not occurring in their native structures The results from the reality match were used for evolving and extending the DFKI-KM-Mid model. Further extensions have been made by a more detailed modeling of slots and by the integration of third-party ontologies like FOAF and specially tailored domain ontologies like the "*Organizational Repository*", formalizing the employees and projects of the DFKI KM lab.

Here the idea is that when bringing the *PIMO* idea into a specific environment the mid level should be re-modeled in a similar way as described above. [16] shows an example for that in a concrete business scenario.

4 Applied use of the PIMO

Using above prerequisites, the *Personal Information Model* of a user can now be created by assembling the different parts. We will use the example user *Paul* and **Paul's PIMO**[3]. The following steps are necessary: Firstly, PIMO-Basic, PIMO-Upper, PIMO-Mid are imported unchanged Then, one or more domain ontologies are imported, *e. g.*, the "Organizational Repository" of a company. The personal mental model of the user is represented in the user's own domain ontology, called PIMO-User. The user works within his own namespace, abbreviated using `paul:`. The first element is the user himself, `paul:Paul`. He is represented as instance of the class `pimo:Person` and annotated as owner of his PIMO, which is represented as `paul:PaulsPim`. The user can refine existing classes by creating subclasses and instances. Finally, the native resources on the desktop of the user (files, e-mails, address-book, etc) are converted to data vocabularies using adapters. They are matched to the personal mental model and to domain ontologies.

Hence, the *Personal Information Model* (PIMO) of a user can be defined as the sum of imported upper and mid-level ontologies, domain ontologies, one personal mental model of the user (PIMO-User), and the native resources found in heterogenous data sources.

As an example for a project managed by Paul, we assume he is planning to open a branch office of his company in Rome, Italy. This project is represented as `paul:BranchOfficeRome`, an instance of class `pimo:Project`. To express that the co-worker Tim is part of the project, `paul:Tim` was created and related to the project via the `pimo:hasPart` relation. Tim has a grounding occurrence in the address book of Paul, the address book entry is a resource manifestation.

[3] The models can be retrieved from: http://ontologies.opendfki.de/repos/ontologies/pim/pauls-pimo.pprj

The example goes on to create a custom class (`paul:BusinessPlan`) and custom properties (`paul:manager`).

The PIMO was used as a basis for the EPOS project, and the possibilities to expand it and customize it to certain scenarios is decribed in [13]. It is also the basis for data representation in the *Gnowsis* project, which was decribed in [14].

Norberto Fernandez created an approach to populate a PIMO while the user is doing search tasks. The user interface of his SQAPS search engine automatically creates PIMO concepts in the background, annotating them with Wikipedia pages [8] as grounding resources.

5 Related Work

A similar approach was used by Huiyong Xiao and Isabel F. Cruz in their paper on "A Multi-Ontology Approach for Personal Information Management", where they differentiate between *Application Layer, Domain Layer and Resource Layer*. Alexakos et al. described "A Multilayer Ontology Scheme for Integrated Searching in Distributed Hypermedia" in [1]. There, the layers consist of an *upper search ontology layer, domain description ontologies layer, and a semantic metadata layer*.

PIMO is different from XML Topic Maps (XTM) as it allows to use inference and RDFS definitions, also enabling an efficient way to store the data in RDF databases (whereas XTM is based on XML). The main difference to RDF is that Topic Maps Associations are by definition n-ary relations, whereas in RDF the relations are typically binary. In RDF, a similar approach as to XTM is the SKOS vocabulary [10]. It represents all Things using the class *Concept*, this blocks reusing inference and typed properties of concepts (like the "first name" property of a person cannot be modelled in SKOS).

6 Summary and Outlook

In this paper we presented the *Personal Information Model* — PIMO. It is a framework of multiple ontologies to represent concepts and documents that are in the attention of the user when doing knowledge work. Basic concepts such as time, place, people, organizations, and tasks are pre-modelled in a mid-level ontology that can be extended by the user at will, to express their mental model. Items can be assigned to multiple concepts, extending the limitations of current hierarchical file system. The PIMO was used in the EPOS, Gnowsis, and SQAPS research projects, *e. g.*, for personalization [13] and semantic retrieval services.

Future challenges are in refining the upper and mid-level models, based on experiences gained through evaluations within the NEPOMUK project. There, the PIMO will be used as a means for file and e-mail annotation in various software applications, for example in the Linux KDE desktop. More research needs to be directed towards automatically creating PIMO structures based on analysing native resources and structures.

References

1. C. Alexakos, B. Vassiliadis, K. Votis, and S. Likothanassis. A multilayer ontology scheme for integrated searching in distributed hypermedia. In S. Sirmakessis, editor, *Adaptive and Personalized Semantic Web*, number 14 in Studies in Computational Intelligence. Springer, 2006.
2. J.-T. Bähr, L. van Elst, A. Lauer, H. Maus, L. Sauermann, and S. Schwarz. EPOS – Guiding Example. internal report, DFKI, 2004.
3. D. Barreau and B. A. Nardi. Finding and reminding: File organization from the desktop. 1995.
4. R. Boardman. *Improving Tool Support for Personal InformationManagement.* PhD thesis, Department of Electrical and Electronic Engineering Imperial College London University of London, July 13 2004.
5. A. S. C. Bizer. D2rq-treating non-rdf databases as virtual rdf graphs. In *Proceedings of the 3rd International Semantic Web Conference (ISWC2004)*, 2004.
6. A. Dengel, A. Abecker, J.-T. Bähr, A. Bernardi, P. Dannenmann, L. Elst, S. Klink, H. Maus, S. Schwarz, and M. Sintek. EPOS – Evolving Personal to Organizational Knowledge Spaces, 2002.
7. A. R. Dengel. Six thousand words about multi-perspective personal document management. In *Proc. EDM IEEE Workshop.* IEEE, Oct 2006.
8. N. Fernandez-Garcia, L. Sauermann, L. Sanchez, and A. Bernardi. Pimo population and semantic annotation for the gnowsis semantic desktop. In *Proceedings of the Semantic Desktop and Social Semantic Collaboration Workshop at the ISWC*, volume 202 of *CEUR-WS*, 2006.
9. H. Holz, H. Maus, A. Bernardi, and O. Rostanin. From Lightweight, Proactive Information Delivery to Business Process-Oriented Knowledge Management. volume 0, pages 101–127, 2005.
10. A. Miles (Ed.). Simple knowledge organisation system (skos). Technical report, Feb 2004.
11. S. Pepper and G. Moore (Eds.). Xml topic maps (xtm) 1.0. Specification, TopicMaps.Org, 2001.
12. H. Rath. The topic maps handbook detailed description of the standard and practical guidelines for using it in knowledge management. empolis white paper, empolis GmbH, 2003.
13. L. Sauermann, A. Dengel, L. van Elst, A. Lauer, H. Maus, and S. Schwarz. Personalization in the epos project. In *Proceedings of the Semantic Web Personalization Workshop at the ESWC 2006 Conference*, pages 42 – 52, 2006.
14. L. Sauermann, G. A. Grimnes, M. Kiesel, C. Fluit, H. Maus, D. Heim, D. Nadeem, B. Horak, and A. Dengel. Semantic desktop 2.0: The gnowsis experience. In *Proc. of the ISWC Conference*, pages 887–900, Nov 2006.
15. L. Sauermann and S. Schwarz. Gnowsis adapter framework: Treating structured data sources as virtual rdf graphs. In *Proceedings of the ISWC 2005*, 2005.
16. M. Siebert, P. Smits, L. Sauermann, and A. Dengel. Increasing search quality with the semantic desktop in proposal development. In *Proceedings of the Practical Aspects of Knowledge Management PAKM conference*, volume 4333/2006 of *Lecture Notes in Computer Science*, pages 279–290. Springer Berlin / Heidelberg, 2006.
17. H. Xiao and I. F. Cruz. A multi-ontology approach for personal information management. In S. Decker, J. Park, D. Quan, and L. Sauermann, editors, *Proc. of Semantic Desktop Workshop at the ISWC, Galway, Ireland, November 6*, volume 175, November 2005.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Increasing the Versatility of Java Documentation with RDF

Axel Rauschmayer
(LMU München, Germany
Axel.Rauschmayer@ifi.lmu.de)

Anita Andonova
(LMU München, Germany
andani2@web.de)

Patrick Nepper
(CDTM München, Germany
patrick.nepper@cdtm.de)

Abstract: Writing Java documentation is difficult, because there are many redundancies and implicit links. Stonefly extends the RDF editor Hyena to bring more reuse and integration to Java documentation: It manages small-grained units of content in an RDF database and enables various ways of collating and interlinking these units. By mapping the source code structure to RDF, source code references in the documentation become very flexible and can even be returned by a query. Furthermore, Stonefly also plugs into the Eclipse IDE, to keep RDF and source code in sync and to offer several ways of navigating between the two.
Category: D.2.7
Key Words: Java, JavaDoc, documentation, Semantic Web, RDF

1 Introduction

Writing good documentation is a constant challenge as a software engineer. Documentation is a separate layer of conceptual modeling on top of the source code. That is, it reflects the structure of the code while having constructs that are completely independent of it. Stonefly has been created to improve the "modeling abilities" for documentation: On one hand, it allows one to combine small-grained units of information so that there is no duplication and that it can be most efficiently accessed. On the other hand, the source code itself is brought closer to the documentation and can be easily referenced from it. The RDF query language SPARQL can be used to explore the structure of the documentation and the source code.

Stonefly achieves these goals by storing both pure documentation and longer code-related comments in the same RDF[1]. The source code is integrated

[1] For this paper, we assume familiarity with the *Resource Description Format* (RDF, [1]). But even without it, one should be able to follow our exposition by thinking of RDF as an object-oriented database with typed nodes and labeled edges.

by special source code references. STONEFLY builds on top of RDF's nodes, relations and its query language SPARQL to construct documentation by linking, embedding and querying fine-grained units of documentation. STONEFLY plugs into the Eclipse IDE so that references are constantly updated and that authoring and accessing documentation is easy while coding.

This paper is structured as follows: First, we show how STONEFLY is practically used, then we explain its abstract model. Finally, we motivate future research and conclude this paper.

An extended version of this paper includes the source code of the running example, more details on how source code and RDF are synchronized and a section on related work. It is available at `http://www.pst.ifi.lmu.de/~rauschma/bib/`.

2 Stonefly in use

STONEFLY is a plugin for two frameworks: On the one hand, it extends the RDF editor HYENA [2] which is based on the Eclipse IDE. On the other hand, it extends Eclipse's Java tools. Fig. 1 shows what STONEFLY looks like. In this section, we would like to show the reader what using STONEFLY feels like, by going through a series of small motivational examples. We contrast STONEFLY's modus operandi with the traditional way of documenting Java via JavaDoc and separate HTML pages. In this section, we call the separate (more conceptual) HTML pages *manual*, as opposed to *JavaDoc* which refers to the HTML pages that are generated via the JavaDoc tool.

2.1 Structuring and accessing content

JavaDoc and the manual are managed separately which makes links between them brittle. STONEFLY's solution is to store both code-related and more conceptual documentation in a single RDF database and to enable flexible links between the two. All manual information and every JavaDoc comment is stored in an RDF node. STONEFLY encourages small units of content and offers several ways to combine and reuse them. The content itself is written as a wiki [3] page, because wikis have always supported quick authoring and linking which are two attributes we desire for documentation creation.

The most common way to reuse content in JavaDoc or the manual is to link to it. `{@inheritDoc}` can be used to embed content, but it has to be JavaDoc and can only embed information from the overridden Java construct. STONEFLY wiki nodes can embed any other RDF node. The simplest case is obviously to embed wiki pages. But the embedded RDF node can also be a query that returns the set of RDF nodes to be embedded. Queries are stated in the standard RDF query language SPARQL [4] and have access to any meta-data that might be

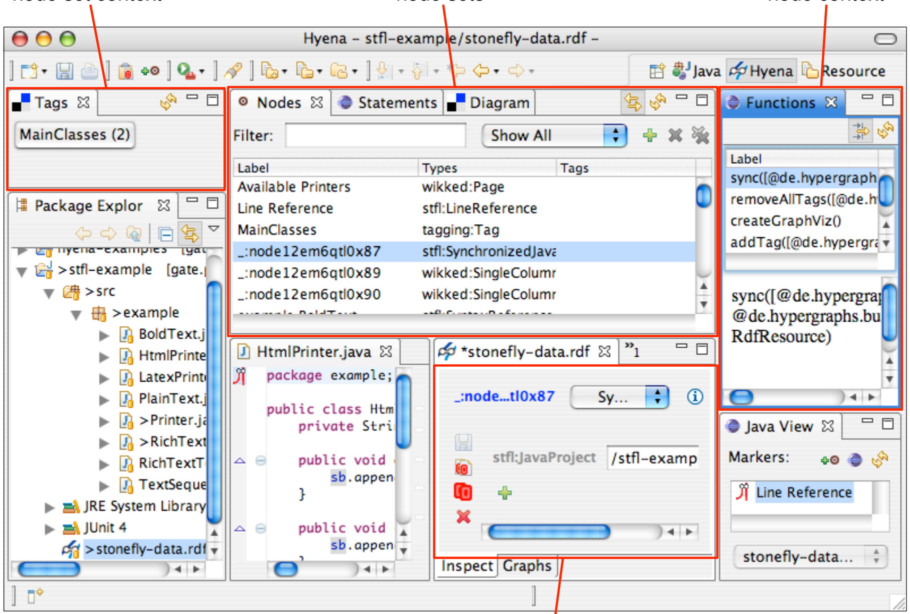


Figure 1: This screenshot shows a typical STONEFLY environment. Clockwise, starting in the top left corner: `Tags` lists the tags that are in use among the currently listed RDF nodes and allows one to filter by tag. In this case, two nodes are tagged with `MainClasses`. `Nodes` and other views show sets of nodes that can be selected. `Functions` is a list of functions that apply to the currently selected node. The `Java View` provides STONEFLY functionality, such as creating a line reference, when a Java editor is active. The editor for `stonefly-data.rdf` contains a graphical *inspector* for the currently selected node. `HtmlPrinter.java` is one class that is referenced from the RDF file; in the top left corner, one can see the marker for a line reference.

attached to the nodes that one is looking for. Thus, content reuse goes beyond simple embedding by dynamically accessing meta-data to determine the relevant content. Fig. 2 shows how this works: The tag node `:tag1` is meta-data for the wiki pages `:page2` and `:page3`. The query node `:query1` retrieves all nodes that are tagged with `:tag1`. The result of this query is embedded inside `:page4`. If one should decide to tag more pages with `:tag1`, those will show up in `:page4` afterwards. `:page3` and `:page1` are examples of linking and embedding single pages.

Note that `:tag1` is meta-data that provides us with context-specificity:

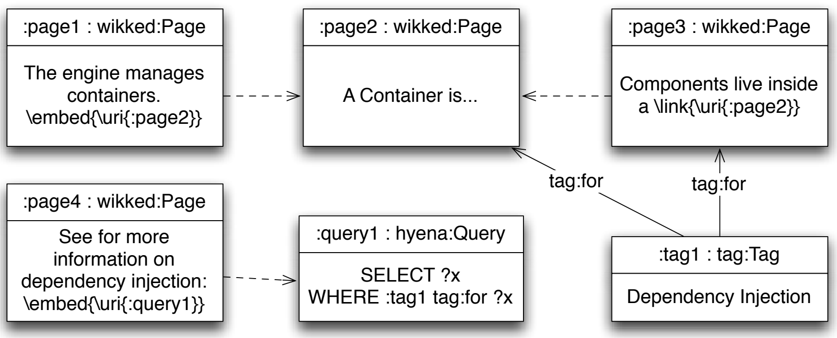


Figure 2: Four wiki pages that use different ways of linking, embedding and tagging: `:page1` embeds `:page2`, `:page3` links to `:page2`. `:page2` and `:page3` are tagged with `:tag1`. `:page4` takes advantage of this fact by embedding the results of a query that returns all nodes marked with `:tag1`. *Legend:* The head of each box contains the ID and the type of the node, the body its content. Dashed lines indicate a URI reference from one node to another.

When `:page2` is selected, its context information asserts that it has the tag `:tag1`. That means that this page is related to all nodes that also have this tag. The user can choose to "zoom out" and display all those nodes which provides her with information that is specific to the context `:tag1`.

2.2 Integrating and exploring the source code

So, for now the RDF database hosts both the JavaDoc and the manual. But what if we want to refer to code locations? Ideally, the code would also live in the RDF database, but in Eclipse, this is currently not practical. Thus, STONEFLY introduces *syntax references*, RDF nodes that refer to code locations and can be embedded or linked to in wiki pages. Similar to running the JavaDoc tool, there is a *synchronization* operation that imports JavaDoc data into the database and makes sure that the code and its representation in RDF are consistent. That means: For each construct that can be JavaDoc-commented, there is one syntax reference in RDF.

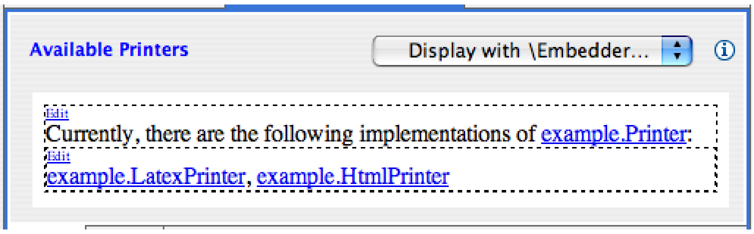
STONEFLY also encodes in RDF relations such as inheritance between the locations. We can thus explore the structure of the code by traversing the RDF graph. Note that this goes beyond what Eclipse offers in that Eclipse views such as "type hierarchy" or "call hierarchy" can only handle a single relation. But as soon as one is looking for locations that are not directly related, it is very likely that more than one relation is involved. Furthermore, the Eclipse views only allow a depth-first search. That is, you can only follow one branch at a time. For many questions that one would ask the code, a breadth-first search is necessary. This kind of search can be handled via SPARQL queries. An example for a "breadth-first question" is: "What methods are invoked by the implementations of interface method `Printer.print()`?". This question also involves *two* relations: inheritance and invocation. It can be answered with the following query. `uri1` identifies the RDF node of the interface method—which can be searched for in the RDF database or jumped to from the source code.

```
SELECT ?x WHERE {
  <uri1> stfl:signature ?s .   <uri1> stfl:isPartOf ?c1 .
  ?c2 stfl:inherits ?c1 .      ?m2 stfl:isPartOf ?c2 .
  ?m2 stfl:signature ?s .      ?m2 stfl:invokes ?x
}
```

One can also impose purely conceptual relations on the code and store them in RDF. For example, we can assign the tag `MainClasses` to `RichText` and `Printer`. This tag thus represents a concern that is a subset of a package and that has not been made explicit in the Java source code. To document the tag `MainClasses`, we can either attach wiki content to the tag or embed a query for the tagged nodes in a wiki page. Many concerns can be marked up this way. For example, "locations that still need to be reviewed" or "all types and methods that have to do with printing". Note that JavaDoc has minimal support for adding conceptual relations via `@see`, but it is only a single relation and cannot be queried. As we have seen before, any kind of query can be stored in an RDF node and embedded in a wiki page. Fig. 3 shows a code-related query: If you want to write about all implementations of interface `Printer`, you query for those classes and embed the result. This way, should more implementations be added, the documentation is always up-to-date.

In order to make the RDF accessible from the source code, STONEFLY taps into the Eclipse infrastructure and uses *markers*, small icons that are attached to the left of a line of text and stay with that line even if it moves. When one clicks on such a marker, a menu appears with which one can go to the RDF node that corresponds to the code location that the marker has been attached to. Markers serve two purposes in STONEFLY: First, if any data has been added to a syntax reference or if other nodes refer to it, it is called an *annotated syntax reference*. A marker at the referenced location is a visual clue that annotated information is available in the RDF. Second, a marker implements a second kind of source code reference: *line references*. The marker tracks the location of the line and makes sure that the link does not break. Navigating from RDF to Java is performed by commands that are available under HYENA. That is, when a syntax reference is selected, the command `gotoJava` jumps to the line or syntactic element that the reference denotes.

```
Currently, there are the following implementations of
\link{\uri{uri1}}: \embed{\uri{uri2}}
```




```
uri2: SELECT ?c WHERE { ?c stfl:inherits <uri1> }
```

Figure 3: The text on top is the content of a wiki page that is displayed as in the screenshot in the middle. It refers to a query whose content is shown on bottom. `uri1` is the URI of the syntax reference to interface `Printer`, `uri2` is the URI of the query node. Note the dashed lines that are only shown on mouse-over. They make it easy to quickly edit embedded content, even if it is deeply nested.

2.3 Visualization and legacy integration

HYENA offers generic graph visualization that can be used by STONEFLY. For example, we can temporarily filter the RDF graph for all classes that have been tagged with `MainClasses` and then visualize their relationships. The resulting graph can either be displayed and navigated (which includes panning and zooming) inside an Eclipse view or it can be exported as a `.dot` file. Dot files are converted via the GraphViz [5] tool suite to SVG, PDF, etc.

STONEFLY makes its RDF data accessible from JavaDoc by placing links inside the JavaDoc entries for annotated syntax references. Clicking on such a link selects the corresponding syntax reference in HYENA (details are mentioned in Sect. 3.5).

3 Technical foundations of Stonefly

As we have mentioned before, STONEFLY has been implemented as an extension for the RDF editor HYENA. HYENA itself is an Eclipse plugin. It manages several RDF databases and allows one to plug in editing support for RDF vocabularies in the form of *vodules* (*vocabulary modules*). A vodule packages everything necessary to edit a given RDF vocabulary and comprises the following elements:

– RDF graph: Some HYENA mechanisms are based on definitions encoded in RDF. By bringing an RDF graph, a vodule can make these definitions. When the vodule is active, its RDF data is merged under read-only access with the other RDF data.

- Embedders: Hyena internally supports the an abstract presentation syntax called Wikked (see below). An embedder translates RDF nodes to this syntax which can then be displayed by translating it again, to either HTML or LaTeX. Thus, embedders provide a read-only visualization of an RDF node. An embedder also specifies the type of the RDF nodes it can translate.
- Inspectors: are very similar to embedders, but instead of presenting an RDF node as Wikked syntax, it displays them using a graphical widget. These widgets normally allow one to edit the RDF data.
- Functions: Whenever a vodule wants to make functionality available to the end user, it implements a *function*. A function's signature defines what data it applies to. That usually means that a parameter is restricted to a certain type of RDF node, but functions can also have other parameters, such as text strings. Hyena displays all available functions in a view and optionally filters them, so that only those appear that can be used in the current context (that is, if the type of the currently selected node does not match a function argument, the function is not shown).
- Internal API: Sometimes other vodules need programmatic functionality related to a vocabulary. One example is the Wikked vodule for wiki pages which includes a parser for wiki syntax.

Stonefly is just another plugin for Hyena, but it only reaches its maximum potential when combined with several other RDF vocabularies. The following subsections describe those vocabularies and the Stonefly vodule itself.

3.1 Vodule "Wikked": Wiki pages

Wikked is a vodule that stores wiki pages [3] in RDF nodes. Its API handles many presentation tasks in Hyena. Wikked has an internal abstract syntax that can be displayed as either HTML or LaTeX. This presentation syntax is returned by an embedder.

3.1.1 Marking up text

Wikked's concrete syntax is a mixture of wiki markup and LaTeX. Traditional wikis have a simple language for quickly marking up text with styles, tables or lists. Here, Wikked follows the *Creole* standard [6] which has a semi-formal, line-based syntax: how text should be displayed is often determined by looking at the beginning of a line: if it starts with an asterisk, the whole line is an entry in a bullet list, if it starts with a pipe, it is a table row, etc. But we want to use an extensible set of commands in a wiki page. So we additionally support a more formal syntax; a subset of the LaTeX (which is inherently not line-based): A command is a backslash and the name of the command. It is followed by zero or more arguments, where each argument is arbitrary text (other—nested—commands, newlines, etc.) surrounded by braces. Fig. 4 shows an example where wiki markup and LaTeX are mixed. During parsing, we initially translate the wiki markup to LaTeX and then build a proper abstract syntax tree. This syntax tree can be exported to either HTML or LaTeX, giving us the option to produce good-looking printed information from wiki content. Note that this hybrid syntax is only targeted at technical users, for the "masses", we are still planning to provide a WYSIWYG editor whose format is translated to and from abstract WIKKED syntax. Even though the hybrid syntax feels a bit awkward at first, we have found that it actually increases usability once one is familiar with its conventions.

```
\embed{\uri{:inspectorDef}}
* Code: \linkdo{\gotoJava{\uri{handle:h1}}}
* Related: \link{\uri{:bookmark1}}
```

An *inspector* is a graphical widget for editing a specific kind of RDF resource. Vodules can provide their own implementations.

- Code: GuiInspector
- Loosely related: Smalltalk Inspectors

Figure 4: For linking and embedding, we need LaTeX syntax. On top, you see how the text is entered, on bottom what it looks like when rendered. We first embed the wiki page stored in node `:inspectorDef`, then we insert a link that, when clicked, brings the user to the Java location that is encoded in node `handle:h1`. Finally, we have a link that will just go to the resource `:bookmark1` that contains a bookmark. Note that the enumeration is written in wiki syntax. We thus have a mixture of LaTeX and wiki markup in this example.

3.1.2 Referencing RDF

References to an RDF node with the URI `u` are written as `\uri{u}` in WIKKED. If such a reference only exists inside the text, we have several problems: using blank node references is not possible, because blank node IDs often change. The referenced node cannot find out about the referencer by examining the RDF data. And renaming the URI of the referenced node breaks the reference. Thus, HYENA manifests the the references in RDF: While the user enters a reference as

\uri{u}, it is internally stored as \uri{i} where i is an index into a container (rdf:Seq). That is, the wiki page contains both the text and indexed edges to all nodes that are referenced from within that text. In Fig. 2, these edges have already been hinted at by dashed lines.

All latex commands internally invoke HYENA functions. If a function returns a text string, the text is displayed in place. Void functions are ignored as far as the wiki page is concerned, but might have a side effect. Three built-in functions are frequently used with RDF references (Fig. 4):

- embed: uses an embedder to translate the referenced node to WIKKED syntax and show it inside the current page. Vodules often bring their own embedders so that the RDF data they support can be correctly displayed inside a WIKKED page. In Fig. 4, the embedder for WIKKED pages is used to inline wiki content.
- linkdo: wraps a function invocation that is only evaluated after the user clicks on a link. In Fig. 4, we use the STONEFLY function gotoJava to create a link that will transport us to a source code location.
- link: renders a node reference as a link. When clicking on the link, the referenced node is selected.

3.2 Vodule: Tagging

Tags are attached to RDF nodes to further describe them; light-weight metadata, if you will. For example, in Fig. 2, :tag1 marks information pertaining to dependency injection—in this case, two wiki pages. Tags are very similar to types, but are more of an attribute of a node than a classification. In HYENA, tags point to the tagged node. HYENA has a graphical view that displays all tags that exist in the currently listed RDF nodes and allows one to filter that list by any one of those tags. By default, a node whose type is tag:Tag is considered a tag. If other kinds of nodes should be recognized as tags by HYENA, one registers them via a (type, predicate) pair. The type specifies the tagging node and the predicate specifies the edge that points to the tagged node.

3.3 Vodule: Embedding query results

Query nodes have two basic ingredients: The query itself, stated in SPARQL and a markup definition. A markup definition states how to translate the query result to WIKKED syntax. To predefined options are to display the result nodes as clickable links that are separated with commas or to embed them. The former can be used to query for a set of code references and embed them inside a text. The latter can be used to create a compound wiki page from smaller pieces.

3.4 Vodule "Fresnel": Declaratively defining editors

RDF can represent different kinds of data, for example knowledge expressed in simple assertions, graphs or record-based data such as contacts or bookmarks. The last kind of data is usually manipulated via form-based editors, where the role models are MS Access and Filemaker. STONEFLY provides a vodule for quickly creating this kind of custom editor. This is done by defining a *lens*, in RDF, with the *Fresnel display vocabulary* [7]. A lens specifies the type of RDF node it applies to, which properties to display and in what order and how to edit property values. As a result, the lens vodule does not provide a single inspector, but a whole family of inspectors, where each lens in the RDF database results in a new family member. Furthermore, the Fresnel vodule also provides lens-based embedders. So any RDF data can be embedded or the results of a query can be presented via a lens. This merging of structured and semi-structured data is very useful for keeping data and its presentation modularized.

3.5 Vodule "Stonefly": Code references

The STONEFLY vodule provides help if one wants to refer to source code. This is obviously a common need when writing technical documentation. STONEFLY introduces one level of indirection here: each code location is manifested as an RDF node. A synchronization operation that is periodically invoked makes sure that the RDF is up-to-date with regard to the code base: New locations result in RDF nodes being created, while nodes that do not refer to an existing location are removed. Now that source code locations have been "imported" into RDF, the usual HYENA features can be brought to bear: The referenced source code can be embedded, it can be rendered as a clickable link that transports one to the code location or one can simply link to the reference RDF node. Going to the reference results in its incoming links being shown. They are displayed as part of a selected node's context information in HYENA and make it easy to find out where in the documentation a given source code location is mentioned. Several functions help with handling the Java side of code references. `gotoJava` jumps to the Java source code denoted by the current node, `sync` performs synchronization and `clearAll` removes all STONEFLY data from the current RDF database.

STONEFLY knows two kinds of source references: syntax references and line references. On one hand, it uses synchronization to parse the source code and to add a syntax reference for each syntactic entity such as types and their members. It refers to them via the Eclipse-assigned *handle ID* which is an identifier string that consists mainly of the name of the project and the signature of the entity. For example, the following is the handle ID for method `example.Printer.print(String)` from the project `stfl-example`.

```
=stfl-example/src<example{Printer.java[Printer˜print˜QString;
```

On the other hand, while a Java editor is active, one can add a line reference to the RDF database that points to the line that the cursor is currently in. To keep constant track of the line, STONEFLY attaches an Eclipse *marker* to it, a small icon that stays on the right side of the text, even when the line moves. When the user clicks on the icon, a menu appears with which one can jump to the RDF node. We call a syntactic reference *annotated* if either another RDF node points to it or it contains additional content such as wiki data. STONEFLY also adds a marker to the source code for each annotated syntactic reference. The marker provides thus a visual clue that there is RDF documentation for the source code location. While non-annotated syntax references do not have a marker, STONEFLY still allows the user to jump from the syntactic entity under the cursor to the corresponding RDF node.

As an external tool, there is a *taglet* that inserts links for annotated references into the JavaDoc-generated pages. Clicking that link selects the corresponding node in HYENA. This works as follows: The link goes to a special port on localhost, where a thread, that has been started by the STONEFLY vodule, listens. For each request, it decodes the URI that has been stored inside the path part of the URL and then selects the node that has that URI.

STONEFLY also records the following relationships between syntactic entities in RDF: `hasParameter` (relationship between a method and a type), `inherits` (relationship between two types), `uses` (field access or method invocation from a method), `hasAnnotation` and `isPartOf` (a relationship between a field or a method and a type). This gives us the possibility to list syntactic elements indirectly, depending on how they relate to other elements. Obviously, this is done via queries, whose results can be embedded inside a wiki page.

4 Stonefly in web mode and future research

The Eclipse frontend for HYENA is only one way of running it. There is also a second frontend that is web-based and implemented using Ajax [8]. We are currently working to bring this frontend up to par, feature-wise, with the Eclipse frontend. By bringing an Eclipse project online, HYENA can function as a collaborative project management tool. Its abilities in this regard could be further bolstered by integrating blogs, forums and mailing lists. This is a very natural extension to what is already available, because the groundwork for the necessary meta-data and editing mechanisms is already present. Moreover, it would be nice to have a mode that is a mix between the (offline) Eclipse frontend and the web frontend: A frontend that can be used offline, but also connect to a server and synchronize with its data. This would support a working style similar to version control systems. Thanks to the universality of RDF, a single synchronization mechanism (such as [9]) can be used to synchronize many kinds of data, at a very fine-grained level.

5 Conclusion

In this paper we have presented STONEFLY, which uses RDF to store wiki pages, meta-data and a syntax abstraction and provides flexible means to browse, query and publish them. It also ensures that source code and RDF data stay in sync. All this serves to make documentation much more versatile.

Acknowledgments

We thank Don Batory and the anonymous reviewers whose comments were instrumental in improving this paper.

References

1. Frank Manola and Eric Miller. RDF primer. `http://www.w3.org/TR/rdf-primer/`. W3C Recommendation.
2. Axel Rauschmayer. Wikifying an rdf editor. 2007. Submitted for publication.
3. Bo Leuf and Ward Cunningham. *The Wiki Way: Quick Collaboration on the Web.* Addison-Wesley, 2001.
4. Eric Prud'hommeaux and Andy Seaborne. SPARQL query language for RDF. `http://www.w3.org/TR/rdf-sparql-query/`, July 2005. W3C Working Draft.
5. AT&T Research. Graphviz—graph visualization software. `http://www.graphviz.org/`.
6. Creole: A common wiki markup language. `http://wikicreole.org/`.
7. Christian Bizer, Emmanuel Pietriga, David Karger, and Ryan Lee. Fresnel: A browser-independent presentation vocabulary for rdf. In *Proc.* 5^{th} *Int. Semantic Web Conf. (ISWC)*, 2006.
8. Jesse James Garrett. Ajax: A new approach to web applications. `http://www.adaptivepath.com/publications/essays/archives/000385.php`, 2005.
9. Giovanni Tummarello and Christian Morbidoni. Rdfsync efficient synchronization of rdf models. `http://semedia.deit.univpm.it/tiki-index.php?page=RdfSync`, 2006.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Accessing RDF Knowledge Bases via LDAP Clients

Sebastian Dietzold
(University of Leipzig, Germany
dietzold@informatik.uni-leipzig.de)

Sören Auer
(University of Pennsylvania, USA
auer@seas.upenn.edu)

Abstract: LDAP based directory services constitute important parts of the IT infrastructure of many organizations and enterprises. They act as a central service for integrating new applications into an IT infrastructure and can be accessed by many different types of clients ranging from content management systems to e-mail clients. In order to make use of widely deployed tools for accessing directory services we present an approach which enables standard directory clients to obtain information from RDF stores (e.g. information about people, organizational structures and addresses). We describe how directory queries can be transformed into queries on an RDF knowledge base and give an overview of our implementation.

Key Words: knowledge management, directory services, LDAP, RDF, SPARQL
Category: H.2.7, H.2.8, H.3.5

1 Introduction

LDAP based directory services are an important part in the IT infrastructure of most organizations and enterprises. They act as a central service for integrating new applications into an IT infrastructure and can be accessed by many different types of clients ranging from content management systems to e-mail clients. The power of LDAP is its broad application support, its usage as an authorization and authentication instance and the flexible, object-oriented data model with a complex schema language, global unique object identifiers and sophisticated syntax matching rules. The most common operations on directory services are simple queries for testing the existence of an attribute of a given directory object, the search for an object based on a single attribute value and the authentication mechanism (bind) on a directory server.

The questions we try to tackle in this paper are: How can directory services benefit from the strengths of semantic web technologies? How can reasoning and rules enhance directory service data? What are the advantages of querying knowledge bases based on the resource description framework (RDF, see [Lassila and Swick1999]) with LDAP clients?

In Section 2 we motivate our work with a usage scenario. We describe in Section 3 how directory queries can be transformed into queries on an RDF knowledge base. We give a short overview of our implementation also with regard to the execution of rules in Section 4 and conclude in Section 5 to future and related work.

2 Usage scenario

We expect that in the near future knowledge bases are the results of collaborative processes. These processes are supported by social, semantic web tool such as OntoWiki [Auer et al.2006] or [Krötzsch et al.2006]. An integral part of such knowledge bases is information about people, projects and organizations. The FOAF project [Brickley and Miller2004] aims to provide an RDF schema for such information. By now, it is heavily used in the blogosphere and by many social networking services.

On the other hand, directory services based on Lightweight Directory Access Protocol [Wahl et al.1997] are widely used within organizations and enterprises. The most common information in LDAP directories is user and contact information forming a shared address book. The primary directory schema for this use case is the inetOrgPerson schema [Smith2000].

While inetOrgPerson focuses on a detailed and granular contact description, FOAF and its extensions describe relations between people, projects and organizations more elaborately.

In order to exploit the client support of LDAP directories with the semantic expressivity of RDF knowledge bases, we aim at integrating RDF knowledge bases with FOAF data into LDAP directories on the basis of the inetOrgPerson schema.

The use of Semantic Web technologies is in particular beneficial when new information can be inferred by processing rules. An example is when work phone information for a person can be inferred from her membership with regard to an organization or company. Such rules can be easily encoded in N3 [Berners-Lee et al.2005] and executed e.g. by the rule processor of cwm[1].

The ultimate goal is to use standard LDAP clients such as e-mail and address book software as well as software relying on LDAP authentication with RDF knowledge bases as backends.

3 Transformation and Mapping process

We analyze how LDAP queries can be translated into corresponding queries based on the query language SPARQL [Prud'hommeaux and Seaborne2006]. The LDAP standard [Wahl et al.1997] defines a syntax for query filters which

[1] `http://www.w3.org/2000/10/swap/doc/cwm.html`

are used by LDAP clients to search for directory objects. This query filter together with some parameters for the search operation (e.g. the starting point of the search in the directory information tree) are the basis for the LDAP to SPARQL transformation process.

This process is divided into two parts.

1. *Query transformation*: The transformation of a given LDAP query into a SPARQL query is done in a straightforward way. The LDAP filter definition allows the compositions of more complex filters with conjunction (`&`), disjunction (`|`) and negation (`!`) operators. Any simple attribute filter consists of an attribute name, a filter type (equality, presence, order, ...) and a filter value. An example for a simple attribute filter is `(surname = "Mewes")`.

   The corresponding SPARQL query uses group graph patterns for conjunction, `UNION` for disjunction and `OPTIONAL`, `FILTER` and `BOUND` for negation. A list of example queries for complex filter compositions is presented in Tab. 1. The namespace `ldap`, which is used in the table, describes vocabulary for the usage of LDAP schema as OWL ontologies [Dietzold2005a]. We use this namespace as the generic representation of all LDAP schema elements in RDF.

LDAP:	`(&(gn="Jason")(sn="Mewes"))`	(conjunction)
SPARQL:	`?Entry ldap:gn "Jason" . ?Entry ldap:sn "Mewes"`	
LDAP:	`(|(cn="Jay")(sn="Mewes"))`	(disjunction)
SPARQL:	`{?Entry ldap:cn "Jay"} UNION {?Entry ldap:sn "Mewes"}`	
LDAP:	`(!(cn="Silent Bob"))`	(negation)
SPARQL:	`OPTIONAL { ?Entry ldap:cn ?Var1} FILTER {!bound(?Var1) || ?Var1 != "Silent Bob"}`	

Table 1: Query transformation: complex filter compositions

   In most cases, the queried RDF model is not an exact representation of a directory information tree (including the object hierarchy as described in [Dietzold2005b]). Instead, the RDF model is based on a semantic web ontology like FOAF so the different schemata have to be mapped as in a second step.

2. *Mapping of RDF schema identifiers to LDAP schema identifiers*: In this step of the transformation process we map the properties and classes, which are used in the LDAP directory, to properties and classes which are used in the

RDF model. For example, we would map `inetOrgPerson` to `foaf:Person`, `organization` to `foaf:organization` and `givenName` to `foaf:givenname`.

Additionally, we need some data transformation functions which are executed on the fly on the data. This is required because some attributes match semantically but have different encoding rules. An example is the mail attribute, which is in FOAF encoded as a resource with an IRI (`mailto:`) and in LDAP as an string attribute. These data transformation functions can be attached to RDF properties to allow a customized handling of these values.

We illustrate the transformation with an example. The following RDF statements are encoded in N3 and form the RDF model we want to query:

```
@prefix : <#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
:sd a foaf:Person;
  foaf:firstName "Sebastian";
  foaf:surname "Dietzold";
  foaf:phone <tel:+493419732366>;
  foaf:mbox <mailto:dietzold@informatik.uni-leipzig.de>.
```

The query for entries with first name Sebastian, last name Dietzold and a saved telephone number will looks as follows:

```
(&(gn="Sebastian")(sn="Dietzold)(telephoneNumber=*))
```

Note that this query filter string is only one part of the LDAP query. Another part is the list of attributes which are to be returned. In this example we want only the telephone numbers.

The resulting SPARQL query with mapped schema elements is

```
@prefix foaf:   <http://xmlns.com/foaf/0.1/> .
SELECT ?Entry ?Var1
WHERE {
  ?Entry foaf:firstName "Sebastian".
  ?Entry foaf:surname "Dietzold".
  ?Entry foaf:phone ?Var1.
}
```

The SPARQL result set contains a list of identifiers with a corresponding telephone number. It has to be transformed to LDAP Data Interchange Format [Good2000] which is a intermediate syntax for the directory server.

4 Implementation

The transformation component is implemented as a backend for the widely used OpenLDAP server[2].

[2] `http://www.openldap.org`

The source code is available as open source software[3]. The implementation works in conjunction with OntoWiki or any web accessible SPARQL endpoint such as Joseki[4].

Fig. 1 shows screenshots of an OntoWiki knowledge base containing information adhering to the FOAF schema. The screenshot on the right hand side shows the evolution address book accessing the OntoWiki knowledge base via LDAP.

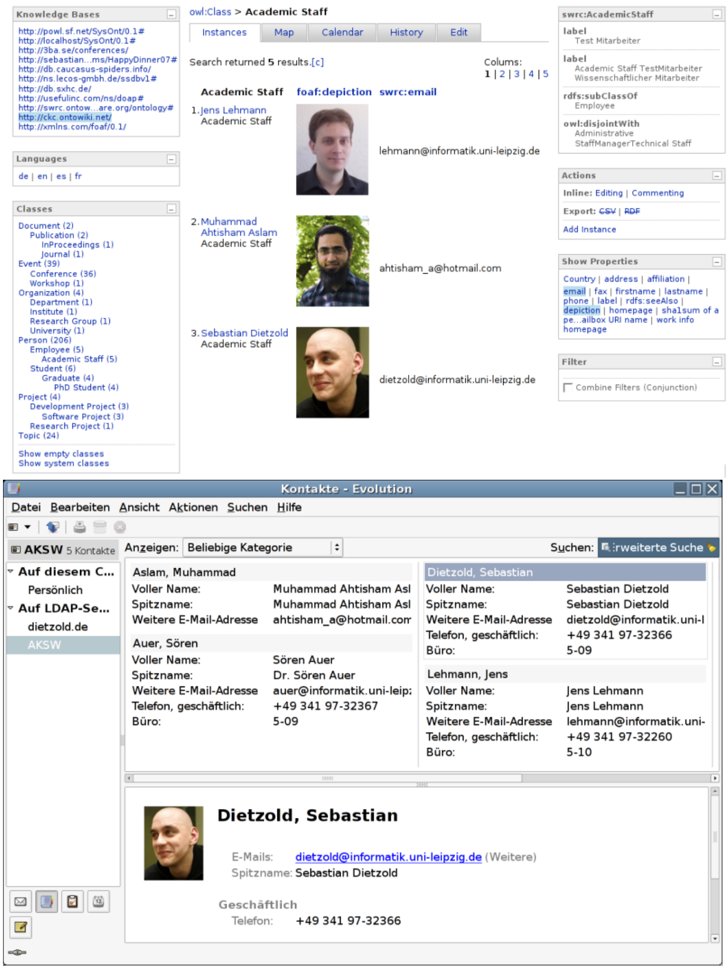


Figure 1: "Same same but different": Visualization of a common RDF model in OntoWiki and over LDAP in the email-reader evolution

[3] `http://aksw.org/Projects/LDAP/Backend`
[4] `http://www.joseki.org`

The example rule described in Section 2 can be expressed in N3 as follows:

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix log: <http://www.w3.org/2000/10/swap/log#> .
@forAll :person, :org, :phone.
{ :org a foaf:Organization.
  :person a foaf:Person.
  :org foaf:member :person.
  :org foaf:phone :phone. }
log:implies
{ :person foaf:phone :phone }.
```

Our implementation uses cwm to execute such rules. The inferenced triples are fed back into the RDF store and are subsequently accessible for all clients.

5 Related and future work

SquirrelRDF[5] is an implementation which allows LDAP servers and relational databases to be queried using SPARQL. It is part on the Jena Semantic Web Framework [Carroll et al.2004] and allows the manually mapping of directory attributes to RDF properties. This project implements the opposite direction transforming SPARQL queries into LDAP queries.

Due to the direct mapping of LDAP queries into SPARQL queries, the performance largely depends on the performance of the SPARQL endpoint. In order to make the implementation more scalable it is furthermore desirable to perform the mapping between the LDAP and the RDF schemes directly within the RDF store and not by the query translator.

Acknowledgments

This research was supported in part by the following grants: BMBF (SE2006 #01ISF02B).

References

[Auer et al.2006] Auer, S., Dietzold, S., and Riechert, T. (2006). OntoWiki - A Tool for Social, Semantic Collaboration. In Cruz, I. F., Decker, S., Allemang, D., Preist, C., Schwabe, D., Mika, P., Uschold, M., and Aroyo, L., editors, *The Semantic Web - ISWC 2006, 5th International Semantic Web Conference, ISWC 2006, Athens, GA, USA, November 5-9, 2006, Proceedings*, volume 4273 of *Lecture Notes in Computer Science*, pages 736–749. Springer.

[Berners-Lee et al.2005] Berners-Lee, T., Connolly, D., Prud'hommeaux, E., and Scharf, Y. (2005). Experience with n3 rules. In *W3C Workshop on Rule Languages for Interoperability, 27-28 April 2005, Washington, DC, USA*. W3C.

[Brickley and Miller2004] Brickley, D. and Miller, L. (2004). FOAF Vocabulary Specification. Namespace Document 2 Sept 2004, FOAF Project. http://xmlns.com/foaf/0.1/.

[5] http://jena.sourceforge.net/SquirrelRDF/

[Carroll et al.2004] Carroll, J. J., Dickinson, I., Dollin, C., Reynolds, D., Seaborne, A., and Wilkinson, K. (2004). Jena: implementing the semantic web recommendations. In Feldman, S. I., Uretsky, M., Najork, M., and WillsSemanti, C. E., editors, *Proceedings of the 13th international conference on World Wide Web - Alternate Track Papers & Posters, WWW 2004, New York, NY, USA, May 17-20, 2004*, pages 74–83. ACM Press.

[Dietzold2005a] Dietzold, S. (2005a). Basic vocabulary to use LDAP data in RDF. OWL ontology. `http://purl.org/net/ldap`.

[Dietzold2005b] Dietzold, S. (2005b). Generating RDF Models from LDAP Directories. In Auer, S., Bizer, C., and Miller, L., editors, *Proceedings of the SFSW 05 Workshop on Scripting for the Semantic Web , Hersonissos, Crete, Greece, May 30, 2005*, volume 135 of *CEUR Workshop Proceedings*. CEUR-WS.

[Good2000] Good, G. (2000). The LDAP data interchange format (LDIF) - technical specification. Rfc, The Internet Engineering Task Force (IETF). http://www.ietf.org/rfc/rfc2849.txt.

[Krötzsch et al.2006] Krötzsch, M., Vrandecic, D., and Völkel, M. (2006). Semantic mediawiki. In Cruz, I. F., Decker, S., Allemang, D., Preist, C., Schwabe, D., Mika, P., Uschold, M., and Aroyo, L., editors, *The Semantic Web - ISWC 2006, 5th International Semantic Web Conference, ISWC 2006, Athens, GA, USA, November 5-9, 2006, Proceedings*, volume 4273 of *Lecture Notes in Computer Science*, pages 935–942. Springer.

[Lassila and Swick1999] Lassila, O. and Swick, R. R. (1999). Resource Description Framework (RDF) Model and Syntax Specification. W3c recommendation, World Wide Web Consortium (W3C).

[Prud'hommeaux and Seaborne2006] Prud'hommeaux, E. and Seaborne, A. (2006). SPARQL Query Language for RDF (Working Draft). W3c working draft, World Wide Web Consortium (W3C).

[Smith2000] Smith, M. C. (2000). Definition of the inetOrgPerson LDAP Object Class. RFC 2798, The Internet Engineering Task Force (IETF). http://www.ietf.org/rfc/rfc2798.txt.

[Wahl et al.1997] Wahl, M., Howes, T. A., and Kille, S. (1997). Lightweight Directory Access Protocol (v3). RFC 2251, The Internet Engineering Task Force (IETF). http://www.ietf.org/rfc/rfc2251.txt.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

ConTag: A Semantic Tag Recommendation System

Benjamin Adrian[1,2], Leo Sauermann[2], Thomas Roth-Berghofer[1,2]
[1](Knowledge-Based Systems Group, Department of Computer Science,
University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern)
[2](German Research Center for Artificial Intelligence DFKI GmbH,
Trippstadter Straße 122, 67663 Kaiserslautern Germany,
firstname.lastname@dfki.de)

Abstract: ConTag is an approach to generate semantic tag recommendations for documents based on Semantic Web ontologies and Web 2.0 services. We designed and implemented a process to normalize documents to RDF format, extract document topics using Web 2.0 services and finally match extracted topics to a Semantic Web ontology. Due to ConTag we are able to show that the information provided by Web 2.0 services in combination with a Semantic Web ontology enables the generation of relevant semantic tag recommendations for documents. The main contribution of this work is a semantic tag recommendation process based on a choreography of Web 2.0 services.
Key Words: Ontology, Web 2.0, Semantic Web, Social Software, Tagging
Category: H.1.1, H.3.3

1 Introduction

In this paper we describe ConTag, a recommendation system to tag or annotate documents with concepts of a Semantic Web ontology. In ConTag, Web 2.0 services providing text and term analysis functions such as phrase extraction, dictionaries, thesauri, classifications and term associations are used to extract the information content of a document. This approach shows that the convergence of Web 2.0 and Semantic Web is worthwhile regarding Web 2.0 tagging and Semantic Web ontologies. The information provided by Web 2.0 services combined with a Semantic Web ontology enables us to recommend semantic tags for documents.

In Section 2, we explain the state of the art of tagging in a Semantic Web environment. Section 3 describes the architecture of ConTag, including different possibilies of retrieving relevant similarities between document topics and ontology instances. Section 4 provides concrete implementation details. It illustrates the extraction of document topics based on Web 2.0 services and the recommendation of similar ontology instances as semantic tags. The evaluation in Section 5 confirms the statement that the information provided by Web 2.0 services in combination with a Semantic Web ontology enables the generation of relevant semantic tag recommendations for documents. Finally, Section 6 summarizes the approach and denotes future goals.

2 Related Work

ConTag generates tag recommendations based on an underlying Semantic Web ontology. The recommendations may be used, e.g in a Semantic Desktop application for classifying documents with a personal information model. Tag recommendations are generated by using existing Web 2.0 services. At the moment, we are not aware of any other system performing this task. Therefore we describe the state of the art of tagging in semantic environments.

The haystack project [Quan et al., 2003] was an early approach of Personal Information Management developed with Semantic Web techniques comparable to the Personal Information Model Ontology (PIMO) [Sauermann, 2006]. NEPOMUK - The Social Semantic Desktop[1] is a project using and building on experiences with *gnowsis* and the PIMO language/ontology. Tagging systems such as the bookmarking manager *del.icio.us*[2], the reference manager Connotea [Lund et al., 2005] or the photo sharing service *flickr*[3], enable users to annotate documents with self defined keywords called tags.

The studies [Golder and Huberman, 2005] and [Kipp and Campbell, 2006] point out patterns in tagging systems. Tags are more than just keywords but symbols for personal concepts. They also point out existing semantic difficulties such as managing polysemies and synonyms. In an analysis of tag usage, [Sen et al., 2006] demanded private tags in tagging systems to be used as personal concepts. Bridging the gap beteen tags and ontologies, the approach of [Schmitz, 2006] described the development of ontologies based on tag usages. The general problem of relating tags and ontologies based on social services is called Folksonomy [Wal, 2004]. In order to define tags in Semantic Web ontologies, Richard Newman introduced a first idea of a tagging ontology in [Newman, 2005]. Existing folksonomies are mined for association rules to retrieve semantic relations between tags using co-occurances [Schmitz et al., 2006]. PiggyBank [Huynh et al., 2005], CREAM [Handschuh and Staab, 2003] and Annotea [Kahan and Koivunen, 2001] provide RDF compliant tag or annotation repositories. [Bloehdorn and Hotho, 2004] describes techniques to optimize text classification using semantic information.

As a result of this state of the art analysis, it can be said that by now it is possible to annotate documents with tags, being symbols for personal concepts. These expressions may be stored as semantic relations in a semantic web ontology.

[1] http://nepomuk.semanticdesktop.org
[2] http://del.icio.us
[3] http://www.flickr.com

3 The semantic tag recommendation system ConTag

In order to generate tag recommendations we used concepts formalized in PIMO vocabulary. In PIMO, concepts are separated between the two classes `Thing` (e.g. persons, events, locations, etc.) and `ResourceManifestation` (music files, documents, etc). A relation `occurrence` connects `Things` to `ResourceManifestations`, using the following semantic: *A thing occurs in a document.* Instances in a PIMO ontology are called *things.* Entities occurring in documents, are called *topics.* Expressing relevant similarities between things and topics may assume four different shapes in ConTag:

Equivalence A topic corresponds directly to a thing.

Classification If a topic's class corresponds directly to an ontology class, the topic is recommended as new thing of the ontology class.

Superordination If a topic's class does not correspond to any ontology class, the topic is recommended as new thing of a new ontology class.

Relation If a topic is semantically related to a thing without being equivalent, a suitable relationship between topic and thing should be proposed.

In the actual version of ConTag we focus on realising the similarity case *Equivalence.* Other semantic relations can be found in [Horak, 2006] and are discussed in future work.

Generally, the idea of using things as tags (instead of labels) entails some basic advantages. Things are identified by URIs and labeled by `rdfs:label` or alternative labels `pimo:altLabel`. This design overcomes existing semantic problems such as synonyms, homonyms, acronyms and different spelling, which current tagging systems suffer, by separating the tag's label from its identification. Additionally, things may possess a set of further describing RDF properties providing the capability to better retrieve similarities.

ConTag is based on a Semantic Tag Recommendation Process (see Fig. 1):

1. During the first step, *Normalisation*, the document's content is tranformed to RDF format to gain a fulltext description. We use the Aperture[4] framework to extract data and metadata such as author, creator and creation date.

2. During the second step, *Topic Extraction*, topics are extracted by requesting Web 2.0 services. This results in a *topic map* using SKOS vocabulary (Simple Knowledge Organisation System) [Miles and Brickley, 2005]. In succeeding lookup iterations, each topic entity is enriched by a set of semantic properties, such as definitions and synonyms.

[4] `http://www.aperture.sourceforge.net`

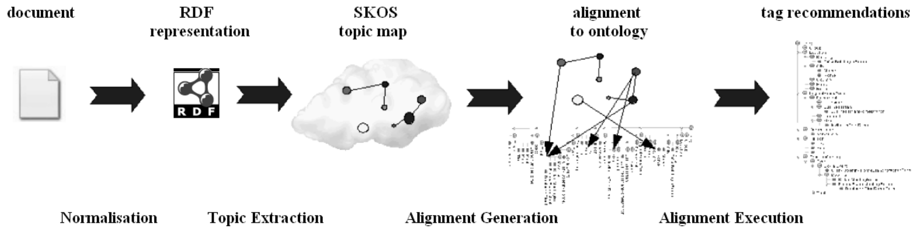


Figure 1: ConTag's Semantic Tag Recommendation Process

3. The *Alignment Generation* is based on document classification methods. For each topic in the topic map, several weighted alignment possibilities are computed to retrieve similar things.

4. The forth step is called *Alignment Execution*. The alignment scheme is visualized as tag recommendations. The user decides whether to accept or reject recommendations. Accepted recommendations are processed to: (1) create new `occurrence` relations in case of *Equivalence*, (2) create new instances in case of Classification, (3) create new classes in case of Superordination, and (4) create new relation types in cases of other semantic relations.

4 Implementation datails

The following sections describe parts of the Semantic Tag Recommendation Process, namely *Topic Extraction* and *Alignment Generation*. We used the RDF store Sesame 2 to manage ontologies in RDFS and topic maps in SKOS.

4.1 Topic Extraction

The topic extraction step is the most valuable step in the Tag Recommendation Process. It results in developing a document specific topic map by executing a Web 2.0 service choreography to extract document entities. The SKOS vocabulary distinguishes topics between instances and classes similar to PIMO language using relations (`broaderInstantive`, `narrowerInstantive`). Each topic possesses a name `prefLabel` and alternative labels `altLabel`. Each topic may be further explained by fulltext definitions written in natural language using `definition`.

The topic extraction step is based on querying Web 2.0 services. The choreography starts with extracting relevant keyphrases of the document. At the moment Web 2.0 services such as Tagthe.net[5], Yahoo's Term Extraction service and Topicalizer[6] are used to extract relevant keyphrases. The results are stored into a document specific topic map.

In a succeeding iteration, for each topic in the topic map, three succeeding lookups request Web 2.0 services to gather for more information:

1. A *definition lookup* queries web dictionaries such as WordNet for existing definitions. These definitions are copied and attached to their grounding topics to be used in the succeeding hypernym extraction and to further provide explanations.

2. A succeeding *hypernym lookup* requests a self written hypernym extraction service called *DefTag*[7] to extract topic classes. These classes are stored as topics and link to instances using `broaderInstantive` and `narrower-Instantive` relations.

3. A third *association lookup* requests services for *word associations* concerning each topic. This lookup considers four different services at the moment: (1+2) Two web services hosted by Ontok Wikipedia provide an access to *Wikipedia Online Encyclopedia*, a collaborative web dictionary system. (3+4) Two web dictionary services (*Moby Thesaurus II*, *WordNet Dictionary*) are requested using the DICT protocol to extract a set of synonyms for a given term.

The topic extraction step results in a document specific topic map written in SKOS. It describes each topic with definitions and word associations. See [Horak, 2006] for more information about the used services.

4.2 Aligning topics to things

The *alignment generation* searches for similarities between topics and things. It results in an alignment scheme which is visualized as a list of tag recommendations. In order to express and weight similarities with confidence ratios, we used an ontology alignment vocabulary[8].

Due to a topological analysis of PIMO ontologies and document topic maps we assume that an ontology contains more entities than a topic map. Additionally, ontologies contain class hierarchies, whether topic maps are rather flat structured. Therefore we focussed on aligning topics to things by applying hierarchical document classification techniques instead of using topolical ontology matching methods. In this paper, we describe a rather simple alignment approach. Other appraches can be found in [Horak, 2006].

[5] `http://tagthe.net`
[6] `http://www.topicalizer.com`
[7] `http://www.dfki.uni-kl.de/~horak/2006/contag`
[8] `http://phaselibs.opendfki.de/wiki/AlignmentOntology`

To retrieve *equivalencies* between topics and things, we compared feature vectors using string matchings. A thing's feature vector is an aggregation of existing describing properties such as `label` or `altLabel`. A topic's feature vector is a list of extracted labels namely `prefLabel` and `altLabel`. We used SPARQL select queries with regular expressions to match both vectors and then computed a string similarity using the dice metric [Rijsbergen, 1979] to gain a confidence ratio. If this confidence ratio exceeds a threshold, an equivalence relation is created.

To retrieve *classification* relations we compared topic classes with thing classes using the upper defined method. If this comparison results in any matching *equivalence*, all concerning topic instances are recommended as new instances of the equivalent ontology class.

Finally, the *Alignment Execution* visualizes the resulting alignment scheme as tag recommendations. Each thing being involved in an alignment relation is concerned to be a tag for the document.

5 Evaluation

In order to provide further evidence for our statement that information provided by Web 2.0 services in combination with a Semantic Web ontology enables the recommendation of relevant semantic tags for documents, we evaluated ConTag by user ratings according to Precision and Recall [Rijsbergen, 1979] ratios. We used an existing ontology with information about projects, employees and partners in PIMO language. Then, we interviewed eight persons, working on subjects being described in the ontology. They got a summary of the ontology content and eleven documents with tag recommendations. The documents were web sites about employees or projects, existing as things in the ontology. The interviewees rated the quality of the given tag recommendations with Precision and Recall ratios (see Fig.2). As a result they rated recommended things of classes Projects, Persons and Organisations with Recall ratios above 80%, in general. These things were based on *Equivalences*. Things of class Location were rated with Recall ratios above 60%. These things were based on *Classifications*. Precision was rated above 70%, in general.

These ratios validate that the information provided by Web 2.0 services in combination with a Semantic Web ontology enables the generation of relevant semantic tag recommendations for documents. Following this result it can be said that the convergence of Web 2.0 and Semantic Web is worthwhile regarding Web 2.0 tagging and Semantic Web ontologies

More detailed evaluation results concerning Precision and Recall progressions in different configuration scenarios and the distribution of tagging recommendations in dynamic and nested class hierarchies are described in [Horak, 2006].

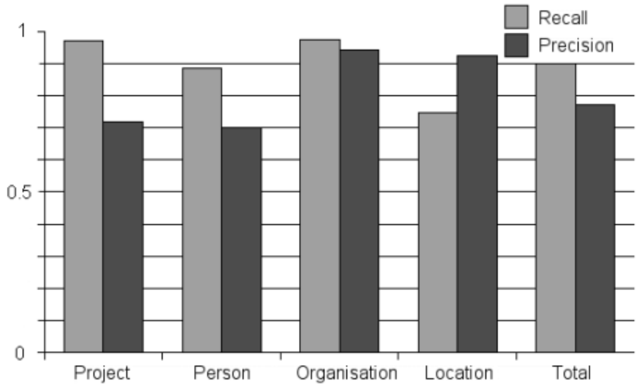


Figure 2: Precision and Recall ratios.

6 Conclusions and Outlook

In this paper we presented *ConTag*, a system to generate semantic tag recommendations for documents based on Semantic Web ontologies. We designed a process to normalize documents to RDF format, extract document topics using Web 2.0 services and finally match extracted topics to instances of a Semantic Web ontology. We use ontologies written in *PIMO language* to formalize instances and classes. Based on our evaluation, we provide evidence that it's possible to create relevant tag recommendations for documents by using Web 2.0 services in combination with a Semantic Web ontology. The implemented system is available under a GPL license for download at the first author's homepage[9].

In future work we want to look for additional similarity metrics to further enhance the alignment creation. The use of additional web services such as Google Glossary and existing tagging services providing accessible APIs is planned. At the moment ConTag concentrates on retrieving equivalent things occuring in documents and the ontology. In future work, we are going to further develop and enhance remaining similarity cases.

Acknowledgments

ConTag was awarded with the IBPM Award at IBPM Kongress 2007 (see `http://www.ibpmkongress.de`) for being the best diploma thesis in the category Document Management. We would like to thank all members of the knowledge management department for all their time, spent on discussions and support. This

[9] `http://www.dfki.uni-kl.de/~horak/2006/contag`

work was supported by the German Federal Ministry of Education, Science, Research and Technology (bmb+f), (Grant 01 IW C01, Project EPOS: Evolving Personal to Organizational Memories) and by the European Commission IST fund (Grant FP6-027705, project Nepomuk).

References

[Bloehdorn and Hotho, 2004] Bloehdorn, S. and Hotho, A. (2004). Boosting for text classification with semantic features. In *Proceedings of the Workshop on Mining for and from the Semantic Web at the 10th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2004)*, pages 70–87. Joint Session with WebKDD workshop. Reprinted in Proceedings of WebKDD 2004, LNCS 3932, Springer.

[Golder and Huberman, 2005] Golder, S. and Huberman, B. A. (2005). The structure of collaborative tagging systems. *Journal of Information Science*, 32(2):198–208.

[Handschuh and Staab, 2003] Handschuh, S. and Staab, S. (2003). Cream - creating metadata for the semantic web. *Computer Networks*, 42:579–598. Elsevier.

[Horak, 2006] Horak, B. (2006). Contag - a tagging system linking the semantic desktop with web 2.0. Diploma thesis, University Kaiserslautern, http://www.dfki.uni-kl.de/˜horak/mypubs/ConTag.pdf.

[Huynh et al., 2005] Huynh, D., Mazzocchi, S., and Karger, D. (2005). Piggy bank: Experience the semantic web inside your web browser. In and Motta, E., Benjamins, V. R., and Musen, M. A., editors, *International Semantic Web Conference*.

[Kahan and Koivunen, 2001] Kahan, J. and Koivunen, M.-R. (2001). Annotea: an open RDF infrastructure for shared web annotations. In *Proceedings of the 10th International World Wide Web Conference*, pages 623–632.

[Kipp and Campbell, 2006] Kipp, M. E. I. and Campbell, D. G. (2006). Patterns and inconsistencies in collaborative tagging systems : An examination of tagging practices. In *Annual General Meeting of the American Society for Information Science and Technology*.

[Lund et al., 2005] Lund, B., Hammond, T., Flack, M., and Hannay, T. (2005). Social Bookmarking Tools (II): A Case Study - Connotea. *D-Lib Magazine*, 11(4).

[Miles and Brickley, 2005] Miles, A. and Brickley, D. (2005). SKOS core vocabulary specification. W3c working draft, World Wide Web Consortium.

[Newman, 2005] Newman, R. (2005). Tag ontology design. blog entry at http://www.holygoat.co.uk/projects/tags.

[Quan et al., 2003] Quan, D., Huynh, D., and Karger, D. R. (2003). Haystack: a platform for authoring end user semantic web applications. In *Second International Semantic Web Conference (ISWC2003), Proceedings*.

[Rijsbergen, 1979] Rijsbergen, C. J. v. (1979). *Information retrieval*. Butterworths, London, 2 edition.

[Sauermann, 2006] Sauermann, L. (2006). Pimo - a pim ontology for the semantic desktop. draft article at http://www.dfki.uni-kl.de/ sauermann/2006/01-pimo-report/pimOntologyLanguageReport.html.

[Schmitz et al., 2006] Schmitz, C., Hotho, A., Jaeschke, R., and Stumme, G. (2006). Mining association rules in folksonomies. In *Proc. IFCS 2006 Conference*, pages 261–270, Ljubljana.

[Schmitz, 2006] Schmitz, P. (2006). Inducing ontology from flickr tags. In *Collaborative Web Tagging Workshop at WWW2006, Edinburgh, Scotland*.

[Sen et al., 2006] Sen, S., Lam, S. K. T., Rashid, A. M., Cosley, D., Frankowsk, D., Osterhouse, J., Harper, F. M., and Riedl, J. (2006). tagging, communities, vocabulary, evolution. In *Proceedings of CSCW 06*. ACM.

[Wal, 2004] Wal, T. V. (2004). Would we create hierarchies in a computing age? blog entry at http://www.vanderwal.net/random/entrysel.php?blog=1598.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

X-KIF – New Knowledge Modeling Language

Michal Ševčenko
(Czech Technical University in Prague
sevcenko@vc.cvut.cz)

Abstract: This paper proposes X-KIF – a new modeling language intended for implementing large, formal knowledge bases. The language originated from the SUO-KIF language, which has been used for implementing a large family of ontologies based on the Suggested Upper Merged Ontology (SUMO). We address some of the problems with the SUO-KIF language, and try to formally define the semantics of some higher-order constructs frequently used in SUMO.
Key Words: KIF, knowledge representation, knowledge modeling language, ontology
Category: I.2.4

1 Introduction

The field of knowledge representation and modeling has long history that went through several stages with different knowledge modeling paradigms. In the 80s and 90s, the primary paradigm was some sort of higher order logic, attempting to maximize expressive power of the language, giving up the completeness of inference. These efforts resulted in languages like KIF [Genesereth 1991] or CycL [CycL].

The last decade brought motion called Semantic Web, which focuses on less expressive, but highly scalable structures, with expressive power not exceeding first-order logic. The primary languages of the Semantic Web are RDF and OWL. The Semantic Web paradigm, however, is not intended for AI-intensive knowledge representation. Semantic Web also promotes strongly decentralistic approach to ontology development.

In the meanwhile, another effort within the IEEE Standard Upper Ontology Working Group resulted in the Suggested Upper Merged Ontology (SUMO) [Niles and Pease 2001]. SUMO defines about a thousand common-sense concepts, accompanied with axioms that partially define their meaning. SUMO uses simplified dialect of the KIF language called SUO-KIF [SUO-KIF]. SUMO concepts have been mapped to the synsets of the WordNet lexical database, making it a valuable computational linguistic tool.

In our recent project "Knowledge support for modeling and simulation" [Ševčenko 2003], we decided to use the alternative SUMO project rather than the Semantic Web languages. We didn't want to completely give up the AI-intensive approach to knowledge representation, and also the more centralistic approach, based on standardized upper ontologies, seemed more appropriate. However, the original SUO-KIF language has couple of problems as well, so we decided to develop a new knowledge modeling language X-KIF, having modern features known from the Semantic Web languages, while retaining the expressivity and philosophy of the SUO-KIF language. The key enhancements with respect to the SUO-KIF language are the following:

1. X-KIF has namespace support. This allows different parties to develop ontologies concurrently without concern of name clashes with different author's concepts.
2. All concepts in X-KIF must be explicitly defined. This prevents errors like "defining" new concept by misspelling its name in an axiom.
3. X-KIF provides support for organizing ontologies into modular sections, making large knowledge bases more manageable.
4. Abstract data model of X-KIF permits for different concrete syntaxes, including XML-based syntax for information exchange, binary data format for scalable, high-performance randomly-accessible representation, and pseudo-text-based format for entering axioms by knowledge engineers.
5. Knowledge base axioms are first-class objects, which can be associated with metadata such as documentation or inference engine hints.
6. Data model is appropriate for wide range of resources, from large, weakly formalized resources like the Wordnet, to frame-based ontologies or highly-formalized ontologies like SUMO.

In Section 2, we describe the overall structure of X-KIF, and particularly its meta-level layer. Section 3 describes the abstract syntax of X-KIF that is used to state formal semantics of concepts. This syntax is similar to the SUO-KIF language syntax. In Section 4 and 5, we address some problems with higher-order semantics of X-KIF stated in [Pease 2003], and propose possible solutions for these problems.

2 X-KIF structure

X-KIF serves for description of knowledge bases, which are collections of *concepts* and *axioms*. Concepts are fixed points to which formal semantic is attached using axioms. Axioms may be simple facts stating e.g. that certain concept is a subclass of another concept, or complex logical formulas stated in first- or higher-order logic.

Concepts denote objects in the *universe of discourse*, i.e. in some interesting part of the world to be modeled. Concepts may denote classes of objects, such as motors, instances of those classes, such as motor in my car, or relations and functions among objects. Note that unlike first-order logic, functions and relations are considered part of the universe of discourse.

The X-KIF data model is split into two layers, see Figure 1. The lower, *axiomatic layer*, is used to describe the actual semantics of concepts in the knowledge base. The semantics is represented by *logical sentences*, specified in Section 3. The axiomatic layer is very similar to the SUO-KIF language. The second, *meta-level layer*, is used to organize large ontologies into modular parts, and to provide means for naming concepts and axioms, so that they can be referenced from logical sentences.

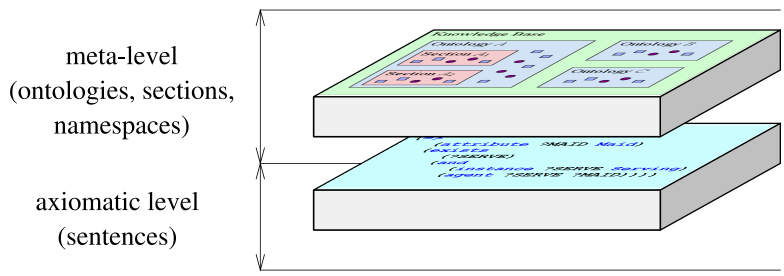


Figure 1: Two layers of X-KIF

The meta-level layer has been added to X-KIF to facilitate management of large knowledge bases. The layer allows to organize concepts and axioms hierarchically into a system of *sections*. Each section may contain concepts and axioms, and may be further divided into subsections to an arbitrary depth. *Ontology* is a special kind of section, containing implementation of some compact, well defined part of the knowledge base, and may be therefore thought of as a separate "package" that is authored and distributed independently. Ontology, however, may refer to concepts and axioms defined in different ontologies. For a conceptual view of knowledge base structure, see Figure 2.

A knowledge base section may be associated with a Uniform Resource Identifier (URI), which defines a *namespace* for all objects defined within that section, including itself. For ontologies, the namespace URI is mandatory. Every knowledge base object, that is, section, concept or axiom, has a local identifier, which must be unique within the object's namespace. The local identifier of an object may be combined with the object's namespace URI to form object's globally unique URI. For instance, if concept Human is contained in section Objects of ontology SUMO, and SUMO has namespace URI `org.sumo.1.0`, the URI of con-

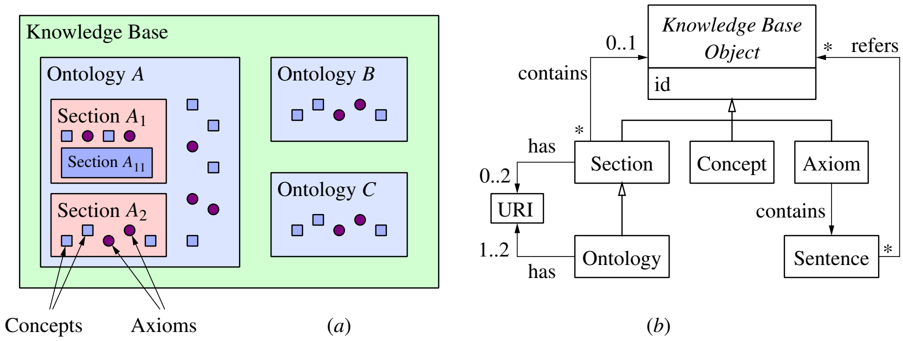


Figure 2: Meta-level layer of X-KIF: (a) conceptual view, (b) UML diagram

cept Human is `org.sumo.1.0#Human`. Note that the ontology SUMO has also its URI, which is different from the URI of its namespace. URI of SUMO may be for example `org.sumo.1.0#sumo`.

The namespace URI of an ontology is intended to refer to specific version of the ontology. If a different version of some ontology is published, it should be given different URI. It is therefore convenient to incorporate some kind of version information or release date directly into the URI of an ontology. A section may also have associated alternative URI called *last version URI*, which is intended to refer to the last known version of that section. Different versions of the same section have therefore identical last version URIs. When knowledge engineer references a concept in external section, he may decide whether it is more appropriate to refer to definition in particular version of that section, and uses its primary URI, or whether the most recent version of the section should be used, and uses the last version URI instead.

All relations depicted in Figure 2b are meta-level relations that are not specified using axioms, and therefore are not "visible" from the axiomatic level of X-KIF. Also note that axioms may refer not only to concepts, but also to sections or other axioms. This allows, for example, to attach metadata such as documentation to concepts, sections and axioms uniformly.

3 Abstract syntax of sentences

In this section, we specify the abstract grammar describing logical sentences comprising axioms. The grammar contains the following terminal symbols: logical connectives $\neg, \vee, \wedge, \Rightarrow, \Leftrightarrow$, quantifiers $\forall$ and $\exists$, logical equality sign $\simeq$, parentheses (and), variables, which we denote with letters x, y, z, possibly with indexes, row variables, typeset in bold, such as $\mathbf{a}, \mathbf{b}$, literal constants, and finally, the knowledge base objects, usually typeset in sans-serif, such as Human.

There are two basic non-terminal symbols. *Terms*, which we denote with letter t, possibly with indexes, and *sentences*, which we denote with Greek letters φ and ψ. Terms are defined inductively as follows.

1. All knowledge base objects are *terms*.
2. *Literal constants* are terms. Literal constants refer to certain well-known objects in the universe of discourse, such as numbers, characters, or strings.
3. *Variables* are terms.
4. Expression $f\,(t_1, \ldots, t_n)$, where f is function symbol, and $t_1, \ldots, t_n$ are terms or row variables, is term.

Sentences are defined inductively as follows.

5. Expression $r\,(t_1, \ldots, t_n)$, where r is either a relation symbol, or a variable, and $t_1, \ldots, t_n$ are terms or row variables, is *sentence*. This particular case of sentence is called *atomic sentence* or *atom*.
6. Expression $r\,(\ldots, \varphi, \ldots)$, where r is either a relation symbol, or a variable, and φ is a sentence, is a sentence. The predicate may contain other terms or row variables as arguments, but may not contain more than one sentence argument. This type of sentence is called *higher-order atom*.
7. A variable is a sentence.
8. Expression $t_1 \simeq t_2$, where t_1 and t_2 are terms, is a sentence called *equation*.
9. If φ, ψ are sentences, then $\neg\varphi, \varphi \vee \psi, \varphi \wedge \psi, \varphi \Rightarrow \psi, \varphi \Leftrightarrow \psi$ are also sentences.
10. If φ is sentence and x is a variable, then $\forall x\varphi, \exists x\varphi$ are also sentences.
11. If φ is sentence, then (φ) is also a sentence. Parentheses may be used to clarify priority of logical connectives and quantifiers.

Note the difference between plain variables and row variables. While plain variables are intended to refer to objects in the universe of discourse, row variables denote *sequences* of objects, which may be used inside expressions 4, 5 and 6. Note that row variables are *not* terms; in particular, they cannot be used in expression 8.

4 Semantics of X-KIF

In this section, we describe the model-based semantics of X-KIF, by extending commonly-used semantics of first-order predicate logic. Formal definition of semantics of first-order predicate calculus uses a notion of interpretation. *Interpretation* $\mathcal{M}$ is a total function from sentences to truth values: $\mathcal{M} : \Lambda \rightarrow \{true, false\}$, where Λ is set of all well-formed sentences of the language. In first-order logic, every interpretation must satisfy the following properties: $\mathcal{M}(\varphi) \neq \mathcal{M}(\neg\varphi)$, $\mathcal{M}(\varphi \wedge \psi)$ is true iff $\mathcal{M}(\varphi)$ is true and $\mathcal{M}(\psi)$ is true, $\mathcal{M}(\varphi \vee \psi)$ is true iff $\mathcal{M}(\varphi)$ is true or $\mathcal{M}(\psi)$ is true, $\mathcal{M}(\varphi \Rightarrow \psi)$ is true iff $\mathcal{M}(\neg\varphi \vee \psi)$ is true, $\mathcal{M}(\varphi \Leftrightarrow \psi)$ is true iff $\mathcal{M}(\varphi \Rightarrow \psi)$ is true and $\mathcal{M}(\psi \Rightarrow \varphi)$ is true. $\mathcal{M}(\forall x : \varphi)$ is true iff $\mathcal{M}(\varphi')$ is true for every sentence obtained by replacing all occurrences of x in φ with some ground term, and $\mathcal{M}(\exists x : \varphi) = \mathcal{M}(\neg\forall x : \neg\varphi)$.

$\mathcal{M}$ is said to be a *model* of sentence φ if $\mathcal{M}(\varphi) = true$, denoted as $\mathcal{M} \models \varphi$. A sentence is said to be *satisfiable* if it has a model. A set of sentences T is satisfiable if there exists an interpretation that is a model for all sentences in T. A sentence ψ is said to be *entailed* by a set of sentences T, written as $T \models \psi$, if each interpretation that is a model for all sentences in T, is also a model for ψ.

We can extend the notion of interpretation to higher-order sentences. We shall call a higher-order interpretation $\mathcal{H}$, which must satisfy at least those properties of $\mathcal{M}$. Additionally, we require that if $\mathcal{H} \models P(\ldots, \varphi, \ldots)$, then $\mathcal{H} \not\models P(\ldots, \neg\varphi, \ldots)$. That is, if certain sentence holds in some context, represented by predicate symbol and its arguments other than the sentence, the negation of the same sentence may not hold in the same context. This requirement holds only for *positive* atoms. It is therefore possible that both $\neg P(\varphi)$ and $\neg P(\neg\varphi)$ have the same model. We also require that all contexts are *closed under entailment*, that is, if $\mathcal{H} \models P(\ldots, \varphi, \ldots)$, and $\varphi \models \psi$, then $\mathcal{H} \models P(\ldots, \psi, \ldots)$. For instance, $P(\varphi) \models P(\varphi \vee \psi)$. The notion of entailment is then extended to higher-order interpretation: a higher-order sentence ψ is entailed by higher-order set of sentences T, if all higher-order interpretations that model all sentences in T also model ψ.

To actually test the entailment, we usually use some inference system. Inference system is a set of syntactic rules that allows to generate new sentences from given set of sentences. If the system allows to generate sentence ψ from a set of sentences T, we denote $T \vdash \psi$. An inference system is said to be *sound* if all sentences generated from T are entailed by T, that is $T \vdash \varphi \Rightarrow T \models \varphi$. An inference system is said to be *complete* if it guarantees to generate every sentence that is entailed, that is, $T \models \varphi \Rightarrow T \vdash \varphi$. In first-order logic, there exists an inference system called *resolution*, which is known to be both sound and complete. We extend the resolution to accommodate higher-order sentences of X-KIF.

First-order resolution contains the following rule: $\{\varphi \vee A, \psi \vee \neg B\} \vdash (\varphi \vee \psi)\,\theta$, where A and B are atoms, and θ is the most general unifier of A and B. We extend the inference system with the following extra rules:

1. $\{\varphi \vee P(\ldots, \alpha, \ldots), \psi \vee \neg P(\ldots, \beta, \ldots)\} \vdash (\varphi \vee \psi)\,\theta$, if $\alpha\theta \vdash \beta\theta$. The sentences $\alpha\theta$ and $\beta\theta$ must be closed sentences (not containing free variables), although the sentences α and β need not be closed. It is therefore possible to resolve literals $\mathsf{believes}(x, \mathsf{teacher}(x))$ and $\neg\mathsf{believes}(\mathsf{John}, \mathsf{teacher}(\mathsf{John}))$, even though the first literal contains a sentence that is not closed (is bound to the outer context).

2. $\varphi \vee P(\ldots, \psi, \ldots) \vdash \varphi \vee \neg P(\ldots, \neg\psi, \ldots)$. This rule specifies that if a clause contains a positive higher-order literal, the literal may be replaced with negative literal, that has the sentence argument negated as well. This is, in fact, an application of resolution rule for general axiom $\neg P(\ldots, \psi, \ldots) \vee \neg P(\ldots, \neg\psi, \ldots)$, which must hold for all predicate symbols, according to our definition of higher-order semantics.

Other operations are done the same way as in first-order resolution. For instance, if a variable is to be unified with a sentence argument, the sentence is treated as term expression, and is simply substituted for the variable.

5 Semantics of other higher-order X-KIF extensions

In the previous section, we described the semantics of one particular extension of X-KIF, namely the usage of nested sentences. In this section, we describe how we treat other higher-order extensions, namely row variables and usage of variable in the predicate position.

Generally speaking, we consider the axioms using these two extra features as *axiom schemata*, that is, structures that describe sets of axioms. Each such axiom schema can be therefore *instantiated* to generate a set of axioms that no longer contain higher-order features. Although the instantiation process may be thought of a kind of preprocessing step that transforms a particular theory before it is sent to some inference system to construct a proof, we propose to directly modify the inference system itself to accommodate the extensions. Semantically, the inference system extension corresponds to generating axioms from their corresponding schemata on the fly as they are needed during the proof. The proposed extensions are implemented as follows:

1. **Variables in the predicate position.** For sentences using variables for predicate symbol, we simply extend the unification procedure of the resolution. In first order logic, two literals $t_1(\ldots)$ and $\neg t_2(\ldots)$ can be resolved only if t_1 and t_2 are the same predicate symbols. The extension is that t_1

and t_2 can be arbitrary terms that are unifiable. For instance, the axioms $\mathsf{instance(isSibling,SymmetricRelation)}$ and $\mathsf{instance}\,(p,\mathsf{SymmetricRelation}) \wedge p\,(x, y) \Rightarrow p\,(y, x)$, allow to deduce fact $\mathsf{isSibling}\,(x, y) \Rightarrow \mathsf{isSibling}\,(y, x)$. This simple extension does not introduce full higher-order semantics of quantification over subsets, but allows to specify axiom schemata for predicates in a natural way. In fact, the latter axiom is an axiom schema that can be instantiated, using the extended resolution rule, for every symmetric relation, into plain first order axiom. Similar technique can be used for functions.

2. **Row variables.** We use row variables mostly for specifying axiom schemata for concrete relations, which usually have arity bound with some reasonable number. It is therefore feasible to simply replace axiom schema that uses row variable with a small number of axioms, replacing the row variable with a sequence of plain variables. Sometimes, row variables are used in axioms where the length of sequence denoted by row variable cannot be bound. For instance, in SUMO row variables are used for specifying semantics of lists, which may be of arbitrary length. We therefore propose a simple extension of the unification procedure, which should allow to unify terms containing expressions with row variables. For instance, term $f\,(1, \mathbf{a})$ can be unified with term $f\,(x, y, 3)$ with substitution $x \leftarrow 1, \mathbf{a} \leftarrow y, 3$.

6 Conclusion

In this paper, we introduced new knowledge modeling language X-KIF, which combines expressive power of the SUO-KIF language with modern features of contemporary XML-based languages like OWL. The language can represent wide variety of resources like OWL ontologies, SUMO, or the WordNet. We also defined the semantics of some higher-order extensions of the language.

References

[CycL] “The syntax of CycL”; http://www.cyc.com/cycdoc/ref/cycl-syntax.html.

[Genesereth 1991] Genesereth, M.: “Knowledge interchange format”; Proc. Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference, pages 599–600. Morgan Kaufmann, (1991).

[Niles and Pease 2001] Niles, I., Pease, A.: “Towards a standard upper ontology”; Proc. of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001), Ogunquit, Maine, (Oct 2001).

[Pease 2003] Pease, A.: “The sigma ontology development environment”; Working Notes of the IJCAI-2003 Workshop on Ontology and Distributed Systems. Volume 71 of CEUR Workshop Proceeding series, (2003).

[SUO-KIF] “Standard upper ontology knowledge interchange format”; http://sigmakee.cvs.sourceforge.net/*checkout*/sigmakee/sigma/suo-kif.pdf.

[Ševčenko 2003] Ševčenko, M.: “Intelligent user support for modeling and simulation of dynamic systems”; Technical Report DC-PSR-2003-01, Czech Technical University in Prague, (Jan 2003).

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

DKP-OM: A Semantic Based Ontology Merger

Muhammad Fahad, Muhammad Abdul Qadir
Muhammad Wajahat Noshairwan, Nadeem Iftikhar
(Mohammad Ali Jinnah University, Islamabad, Pakistan
{fahad, aqadir, wajahat, nadeem}@jinnah.edu.pk)

Abstract: Accurate mapping and merging of multiple ontologies to produce consistent and coherent merged global ontology is a very important process to enable heterogeneous multi-vendors semantic-based systems to communicate and understand each other. Current systems for ontology mapping and merging are very restricted in term of resolving mismatches or proposing accurate matches with no or minimum human intervention. The suggestions made by these systems do not consider all information available in the semantic knowledge of the ontologies. In this paper we are proposing a system for merging and alignment of OWL-DL ontologies that uses semantics of concepts especially the disjoint knowledge in order to resolve conflicts automatically. We show that in order to provide a complete, consistent and coherent merged ontology, preservation of disjoint knowledge in the definition of ontologies is very helpful. Our system uses heuristics-based approach for mapping concepts of ontologies by analyzing description logic (DL) of concepts and preserves the disjoint knowledge in the merged ontology, too. The concept of validation during the initial stages of ontology mapping not only distinguishes our system from the existing ones' but also reduces the users' dependability for validating the consistency of the generated mappings. We also discuss experimental findings to prove the effectiveness of our approach in merging process.

Keywords: Ontology Mapping and Merging, Disjoint Knowledge Preservation, Ontology Engineering, Knowledge Modelling, Semantic Computing
Categories: H.3, H.4.2, I.2.4

1 Introduction

Ontology is one of the important components of a semantic based information processing systems [Gomez-Perez, 04]. Information processed by such systems is modelled in the form of ontology. To enable these heterogeneous multi-vendors systems communicate and understand each other in order to develop a globally compatible semantic based information system, we need to map and merge the ontologies and generate a unique global ontology [Noy, 01]. Ontology creation is mainly dependant on how the ontologist interprets the domain and models the interpretation, thus source ontologies may model the same knowledge in different ways [Klein 01]. Wiederhold [Wiederhold 94] highlights the ontological mismatches between source ontologies that may arise when ontologists model the same knowledge in different ways. Grosso et al. [Grosso, 98] and Bowers et al. [Bowers, 00] narrate linguistic mismatches when integrating source ontologies that were built using different languages. Klein [Klein 01] combines the work of many researchers and categorizes the mismatches between ontologies as language level mismatches and ontological level mismatches. He also mentioned that ontology mapping and merging systems should be capable enough to find these mismatches and resolve them with minimal or no human intervention to produce consistent and coherent merged ontology. Besides these mismatches, we identified sufficient knowledge omission error [Noshairwan, 07] that would also create problems in onotolgy merging process.

Developing ontology mapping and merging algorithms which can resolve the mismatches with minimum human intervention is a challenge for the developers. Many semi-automatic approaches have been proposed for this purpose. These techniques use many features that can be found in ontologies (labels, structures, instances, semantics) to produce initial mappings between source ontologies for human user [GomezPerez, 04]. Many of these systems produce inaccurate mappings because they assume semantically distinct concepts to be the same as they are using linguistic analysis or instance based matching to find similarities between concepts. Therefore, validation of suggestions and resolution of conflicts is totally dependant on human experts.

Pottinger and Bemstein [Pottinger, 03] formally describes the requirements of a merging process based on similarities. Current ontology merging systems follow these requirements such as element preservation, equality preservation, extraneous item prohibition etc. However, we have observed that one of the important requirements of ontology merging has been overlooked. This requirement is the preservation of Disjoint Knowledge. Qadir [Qadir, 07] raises the significance of disjoint knowledge within the ontology and demonstrates that disjoint knowledge omission could be catastrophic in various situations. Therefore, a system has been proposed to detect disjoint knowledge omission and generate warnings for ontologists. In this paper, we provide an ontology mapping and merging approach, which uses the features of the existing approaches and in addition to these features it also incorporates disjoint knowledge preservation mechanism based on description logic (DL). Analyzing disjoint axioms within the ontologies would raise a new class of conceptualization conflict known as alignment conflict among disjoint relations. These conflicts arise when lexically same concepts within the source ontologies contradict each other with respect to their semantics descriptions of disjointness. When the existing systems merge these ontologies, they ignore the restriction of disjointness and the result would be an incorrect merged ontology. Due to use of more semantics provided in the ontologies to be merged, our approach enhances the accuracy of ontology mapping and merging process and also reduces the human intervention.

Rest of the paper is organized as follows: Section 2 discusses related work. Section 3 presents our framework, “DKP-OM”: the architecture, the working of system and the main mechanisms for finding the mappings between concepts. Section 4 contains our experimental findings and initial report on its effectiveness. Section 5 concludes the paper and shows future directions.

2 Related Work

There are many techniques for ontology mapping and merging based on features that are used for finding similarities between concepts such as instances, labels, structures, etc. Instance based techniques like FCA-Merge [Stumme, 01], and IF-Map [Kalfoglou, 03] consider concepts having the same instance as candidate to be merged. The major drawback is observed when semantically distinct concepts having the common instance are considered to be the same. GLUE [Doan, 04] follows the instance-based and multi-strategy machine learning approach that analyses taxonomic structure for ontology integration. It calculates the probabilities of concept matching to combine results of multistrategies by its subcomponents. QOM [Ehrig, 04] adopts many similarity measures and follows a dynamic programming approach to find correspondences between source ontologies. It uses heuristics by only choosing promising candidate mappings and thus reduces the runtime complexity.

The tools PROMPT [Noy, 00], AnchorPROMPT [Noy, 01] and Chimaera [McGuinness, 00] use the similarity of labels and to a certain extent the structure of ontologies. Prompt follows the cyclic process by finding the initial suggestions based on identical labels, synonyms, and superclasses for top-level classes (structural indications) during ontology merging process. It was designed for specialized terminology for medical ontology; hence it produces less accurate mappings for general ontologies. Moreover when semantically same concepts are modelled with different names, it has no ability to find those matchings. ANCHORPROMPT exploits the graph analysis of ontology where classes are considered as nodes and slots as links between them. It analyses the paths in the subgraph limited by the anchors and determines those classes that are frequently appearing in similar positions on similar paths. Chimaera suggests the user to merge concepts by finding pairs having same labels, same prefix or suffix, substrings and acronym usage. ONION [Mitra, 02] exploits the taxonomic analysis, local definition analysis and formalization of articulation ontologies to merge different ontologies. Articulation ontologies contain concepts and relationships expressed as rules. These rules provide links between source domains while ontology merging process. Cupid [Madhavan, 01] integrates the linguistic and structural approaches to compute normalized correspondence coefficients with the assistance of a precompiled thesaurus. OLA [Euzenat, 03, 04] uses Alignment API that finds correspondences based on terminological, structural and extensional approaches for OWL-Lite ontologies. It finds recursive relationships to find the best match through iteration and searches similarities between similar categories i.e. concepts with concepts, attributes with attributes and so on.

Our sytems follows the hybrid approach and uses linguistic matching and semantic-based formal definition analysis to find correspondences between concepts. It follows the OLA [Euzenat 03, 04] by covering all the possible characteristics of OWL ontologies i.e. terminological, structural and extensional and CUPID [Madhavan, 01] approach in a sense of finding linguistic and structural matching strategies. The use of heuristics that guides system to find mappings and resolves conflicts is similar to the constraints followed by GLUE [Doan, 04]. The use of disjoint knowlegde axioms and validation of initial mappings by considering disjoint relations between concepts distinguishes it from rest of the systems and lessens users' dependability for validating the consistency of the generated mappings.

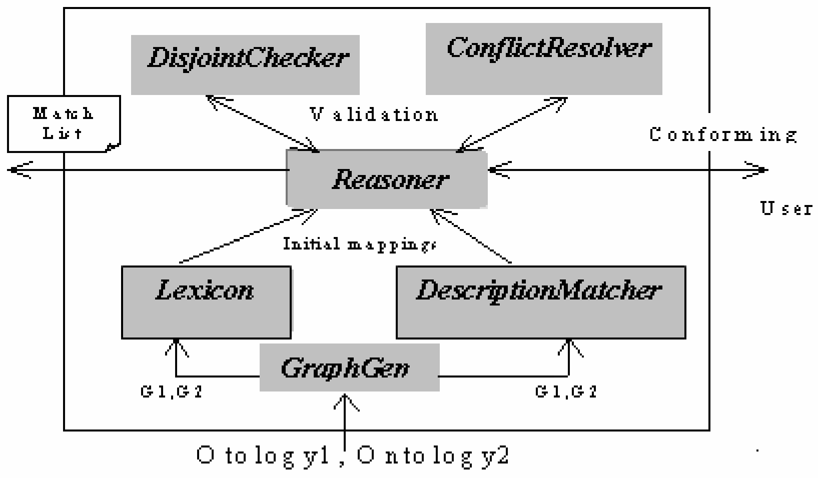


Figure 1: Top-level view of DKP-OM

3 The Design and Working of DKP-OM

Disjoint Knowledge Preserver (DKP) based Ontology Merger (OM) merges ontologies by exploiting semantic knowledge available in the ontologies especially the disjointess of the concepts. Figure 1 shows the top-level view of DKP-OM. In DKP-OM, initial mappings are found by *Lexicon* and *DescriptionMatcher*. *Lexicon* uses string based algorithm and thesaurus lookup to find mappings between concept names. It also does the structural analysis to find correspondences between structural taxonomy between source ontologies. *Lexicon* generally detects (i) concepts having same names, or using same substrings, suffixes, prefixes or acronyms, (ii) concepts having synonyms (same meanings), hyponym less general) or hypernym (more specific) relationships, (iii) concepts having structural similarity with each other i.e. occurrence of concept in the ontologies under consideration, having same parent, siblings, children etc.

DiscriptionMatcher performs the description logic analysis of concepts that reflects semantic similarities by analysing their formal definitions and relations. During analysis it may use the more generic definitions of the concepts in the hirearchy. Generally it detects (i) concepts having same definitions and restrictions or usage of concepts in its definition in intersection or union etc. (ii) concepts having same relationship with other concepts such as parent, child, sibling etc. (iii) overlapping concepts by using heuristics, which increases the chance of correct mappings among them.The probability of correct mappings sent from *DiscriptionMatcher* is greater than the mappings suggested by *Lexicon*. The initial mappings from *Lexicon* and *DiscriptionMatcher* are propagated to the *Reasoner.*

Reasoner does the validation of each mapping found in the previous stage so that the merged ontology stays consistent with reference to the source ontologies. It rejects *lexicon*'s mappings that are linguistically same but semantically different by analysing description logic and then finds one to one correspondence between the concepts of ontologies. Finally it produces consistent, coherent and non-redundant mappings between concepts of source ontologies. The working of *Reasoner* is further explained in the coming sections.

3.1 Reasoning in DKP-OM

Reasoner analyses the concepts to the degree permitted by the formal semantics of the OWL-DL [Gomez-Perez, 04] and validates mappings found by *Lexicon* and *DescriptionMatcher*. It focuses on checking the consistency based on disjoint relationships in merged ontologies. It gets confidence about the accuracy of mappings, which are produced both by *Lexicon* and *DescriptionMatcher*, and detects the conflicting situations caused by wrong mappings. It preserves disjoint knowledge between concepts, and avoids disjoint knowledge omission error. Alignment conflicts among disjoint relations between concepts are detected and suggestions are predicted for conflict resolution.

Reasoner uses *DisjointChecker* to check whether mapped concepts are directly or indirectly (in the hierarchy) disjoint with each other. It rejects the mappings that give rise to conflicting situations. For example consider the ontologies in Figure 2 and assumes that *Lexicon* suggests B and C concepts of ontology 1 are same as the concepts Y and Z of ontology 2, children of B mapped on the children of Y. Assume *Lexicon* also suggests that some child of C is mapped on a child of Y based upon textual similarities. Our disjoint checker would raise an alarm for such mappings because the parents of these two mapped children are disjoint (indirectly) and the mapping would be inconsistent. The conflicts could be resolved either by the conflict resolver or by human expert.

The system has build-in capability to resolve alignment conflicts during ontology merging process with the help of heuristics in the knowledge base as suggested in table 1. *ConflictResolver* checks the knowledge base to get a solution for such conflicts. Some of these heuristics guide the system about the concepts that are candidate for merging but some only predict about their partial mapping by analysing overlapping descriptions. Such overlapping descriptions are useful while analysing mappings came from *Lexicon*. As *Lexicon* only suggest about their linguistic similarity, overlapping heuristics increases the probability and chances of these mapping between concepts to be correct.

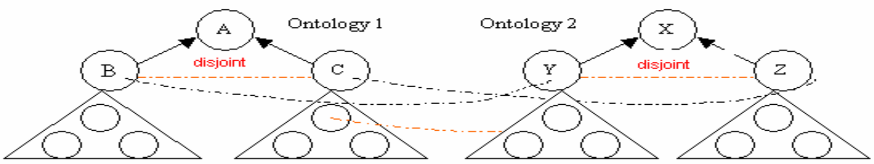


Figure 2: Contradictory Mappings between Concepts of source ontologies

• A concept C_a is made up (union) of n concepts and a concept C_b has same n children, then C_a and Cb are supposed to be candidate for merge.
• If the concept (C_a) is formed by same restriction as restriction in C_b and the type of restriction is some/all value restriction then concepts C_a and C_b are candidate for merge.
• Concept C_a is made up (intersection) of n concepts and a concept C_b has same n children, then the chance of Ca and Cb to be merge increased.
• Concept C_a is union or intersection of some concepts. The concept C_a and C_b are overlapped, if C_b equal to any concept that is used in definition of C_a.
• Concepts C_a and C_b are overlapped; If C_a is sub-concept or super-concept of C_b.
• Concept C_a is based on all value restriction and concept C_b is union of some concepts. The concepts C_a and C_b are overlapped, If the domain concept of restriction is equal or sub concept of C_b or any concept that is used in definition of $C_{b.}$
• Concept C_a is based on maximum or minimum cardinality restriction and concept C_b is union of some concepts. Concepts C_a and C_b are overlapped, If the domain concept of restriction is equal or sub concept of C_b or any concept that is used in definition of C_b.
• Concept (C_b) is formed by same restriction as restriction in C_a and the type of restriction is minimum/maximum cardinality restriction then the concepts C_a and C_b are overlapped.
• Concept (C_b) is formed by same restriction as restriction in C_a and the type of restriction is some/all value or minimum/maximum cardinality restriction then the concepts C_a and C_b are overlapped.

Table 1: Examples of Heuristics

Reasoner suggests the mappings that preserve disjoint knowledge between concepts in merged ontology as they are in source ontologies. While preserving disjoint knowledge between concepts, some conflicts may occur. Resolution criteria for such conflicts should be asked from the user by giving all the possible suggestions. Consider a similar situation with a slight modification to the situation in Figure 2, by removing disjoint knowledge between Y and Z concepts as shown in Figure 3. From the Figure we come to know that B and C are mapped to the Y and Z concepts. Furthermore concept G1 that is a child of Y is mapped onto a child of C. Such conflicts can be resolved automatically by the *Reasoner*. *Reasoner* checks whether concept G or any child of G overlap with concept B or with children of B. If they do not overlap then *Reasoner* accept that mapping to be true and suggest user to map Y to B, and extract subconcepts hierarchy G to be merged with subconcepts hierarchy of C. But if they overlap then it extracts only overlapping concepts of G to be merged with subconcepts of B.

[1] consider G as concept *student* in ontology 2. Due to semantic heterogenity some of the *students* (children of G) are employeed and some are not. In ontolgy 1 *student* is a subchild of C (unemployeed person) and disjoint with B (employees). *DisjointChecker* highlights this situation and *ConflictResolver* suggests employeed *students* to be kept under B and unemployeed *students* under C.

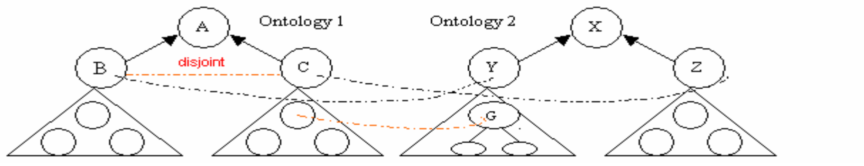


Figure 3: Mappings between Concepts

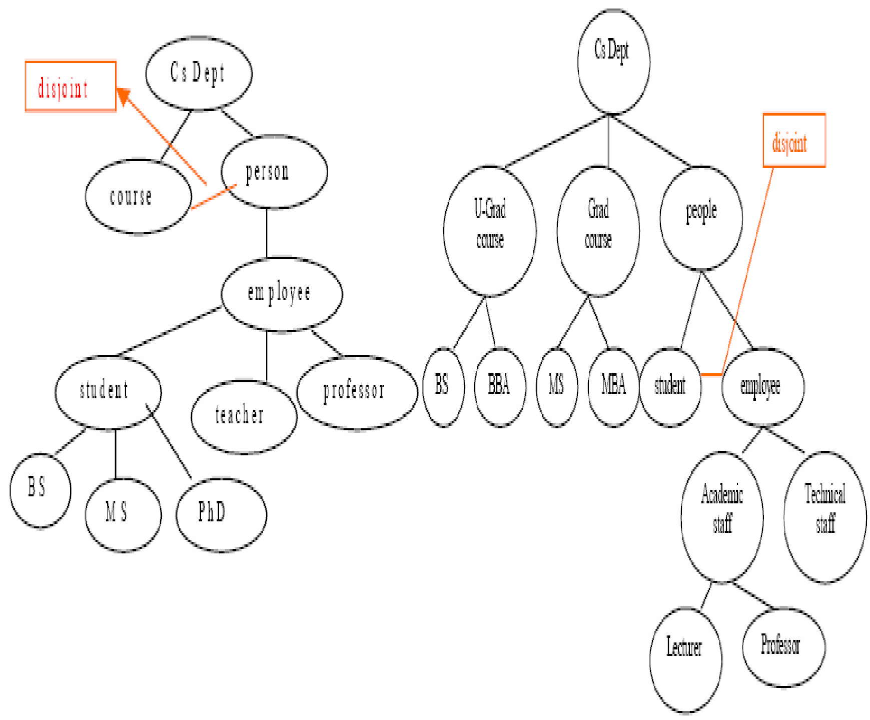
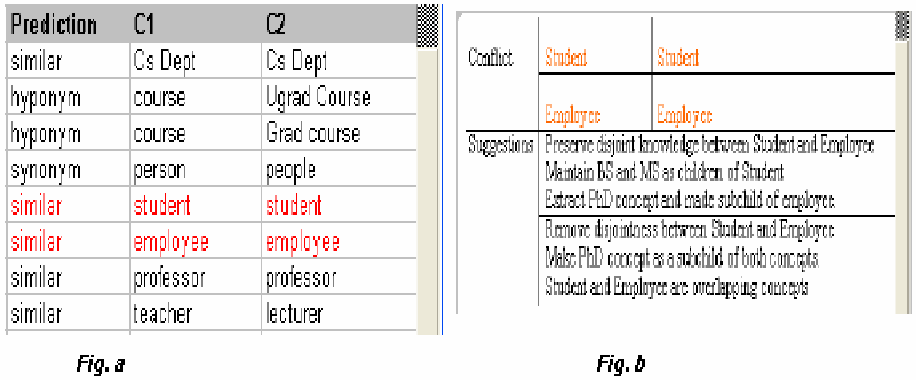


Figure 4: a) University Ontology A b) University Ontology B

4 Evaluation

We have evaluated DKP-OM on several real-world domains to check its matching accuracy and effectiveness. We performed an experiment with DKP-OM by giving two university ontologies having disjoint-axioms. Due to space limitations we only show top-level view of both ontologies in Figure 4 and discuss our high level findings. Suggestions by DKP-OM are represented in Figure 5a. The conflict arises during the merging is shown in Figure 5b. Inaccurate mappings that were rejected are shown in Figure 6. The alignment conflict occurs while mapping the concept *student* of ontology A, with *student* of ontology B. As an instance of only concept *PhD student* can be an *employee* in ontology A but no instance of *student* can be the instance of *employee* as they were made disjoint in ontology B. DKP-OM prompts the user to take the decision for conflict resolution whether to preserve disjoint knowledge as in ontology B or preserve parent child relationship as in ontology A. In the language of the example considered here, one is to preserve disjoint knowledge between *student* and *employee* as in ontology B with children *BS* and *MS*, and extract *PhD* concept and make subconcept of *employee*. Second is to ignore disjointness between *Student* and *Employee* and make *PhD* concept as a child of both concepts by multiple inheritance.

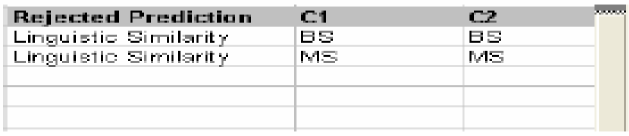


Prediction	C1	C2
similar	Cs Dept	Cs Dept
hyponym	course	Ugrad Course
hyponym	course	Grad course
synonym	person	people
similar	student	student
similar	employee	employee
similar	professor	professor
similar	teacher	lecturer

Conflict	Student	Student
	Employee	Employee
Suggestions	Preserve disjoint knowledge between Student and Employee Maintain BS and MS as children of Student Extract PhD concept and made subchild of employee	
	Remove disjointness between Student and Employee Make PhD concept as a subchild of both concepts Student and Employee are overlapping concepts	

Fig. a Fig. b

Figure 5: Mapping results– a) Mappings between source ontologies b) Suggestions for Alignment Conflict in disjoint relations



Rejected Prediction	C1	C2
Linguistic Similarity	BS	BS
Linguistic Similarity	MS	MS

Figure 6: Rejected Mappings

5 Conclusion and Future Work

This paper presents the DKP-OM system that exploits linguistic matching and semantic-based formal definition analysis to find correspondences between concepts to improve accuracy of mappings. DKP-OM preserves disjoint knowledge in merged ontology, thus avoids disjoint knowledge omission error [Qadir, 07]. It checks the consistency of the merged ontology by taking disjoint knowledge axioms into consideration. The system is capable of finding alignment conflicts among disjoint relations and resolving these conflicts by using set of heuristics or user intervention. Our General-purpose heuristics and description logic analysis make the system work with different contexts, domains and subject ontologies. Domain specific heuristics[1] about disjoint knowledge in source ontologies minimize the search space while finding correspondences between concepts and thus reduce the runtime complexity. However domain specific heuristics can only be applied on well-known ontologies where we do not expect any alignment conflict within the subhierarchies of disjoint concepts.

Our ongoing research direction on this topic is to conduct empirical studies to assess the effectiveness of using disjoint knowledge axioms in ontology merging and to compare our system with other ontology merging systems. At the same time, we will further enhance the system based on some heuristics that lead towards consistent merged ontology with avoidance of higher level of user intervention during ontology merging process. We will further enhance the system for dynamic extraction of mapped code chunks to produce automatic merged ontologies.

References

[Bowers, 00] Bowers, S., Delcambre, L.: "Representing and Transforming model-based information", In First Workshop on the Semantic Web at the Fourth European Conference on Digital Libraries, Lisbon, Portugal(2000), 5-18.

[Doan, 04] Doan, A., Madhaven, J., Domingos, P., Halevy, A.: "Ontology matching: A machine learning approach", Handbook on Ontologies in Information Systems, Springer-Verlag,(2004), 397-416.

[Ehrig, 04] Ehrig, M., Staab, S.: "QOM - quick ontology mapping", In Proc. of the Third International Semantic Web Conference (ISWC2004), Springer, LNCS 3298(2004), 683–696.

[Euzenat, 03] Euzenat, J. and Valtchev, P.: "An integrative proximity measure for ontology alignment", In Proc. of the 1st Intl. Workshop on Semantic Integration(ISWC-2003), Sanibel Island, USA(2003), 33–38.

[Euzenat, 04] Euzenat. J. and Valtchev, P.: "Similarity-based ontology alignment in owl-lite", In Proc. of *16th* ECAI-04, Valencia, Spain(2004), 333–337.

[Gomez-Perez, 04] Gomez-Perez, A., Fernández-López, M., Corcho,O.: "Ontological engineer-

ing: With examples from the Areas of Knowledge Management, E-Commerce and the Semantic Web"; Springer, London(2004)

[Grosso, 98] Grosso, W., E., Gennari, J., H., Fergerson, R., W., Musen, M., A.: "when knowledge models collide (how it happens and what to do)", In Proc. of the 11th Workshop on Knowledge Acquisition, Modeling and Management (KAW '98), Banff Canada (1998), 18-23

[Kalfoglou, 03] Kalfoglou, Y., Schorlemmer, M.: "If-map: an ontology mapping method based on information flow theory", Journal of data semantics, Springer Berlin / Heidelberg, LNCS 2800(2003), 98–127.

[1] For example consider family ontology O_a where Male concept is disjoint with Female concept and ontology O_b where Men concept is disjoint with Women concept. If we get a top level mapping of concept Male into Men and Female into Women, then search space would be reduced by only seeking mapping of subchildren concepts of Male into subchildren concepts of Men only and vice versa rather than in all the taxonomy of concepts.

[Klein, 01] Klein, M.: "Combining and relating ontologies: an analysis of problems and solution", In Proc. of Workshop on Ontologies and Information Sharing (IJCAI-01), Seattle, USA(2001), 53-62.

[Madhavan, 01] Madhavan, J., Bernstein, P., Rahm, E.: "Generic schema matching using Cupid", In Proc. of the 27th Intl. Conference on VLDB, Roma, Italy(2001), 49–58

[McGuinness, 00] McGuinness, D., L., Fikes, Rice, J., Wilder, S.: "An environment for merging and testing large ontologies", In Proc. of the 7th Intl. Conference on Principles of Knowledge Representation and Reasoning, Breckenridge, CO, USA(200), 483–493.

[Mena, 00] Mena, E., Illarramendi, A., Kashyap, V., Sheth, A., P.: "OBSERVER: An approach for query processing in global information systems based on interoperation across preexisting ontologies", International Journal DAPD, 8, 2 (2000), 223-271.

[Mitra, 02] Mitra, P., Wiederhold, G.: "Resolving Terminological Heterogeneity in Ontologies," In Proc. of Workshop on Ontologies and Semantic Interoperability at the 15th ECAI, Lyon, France (2002), 45-50.

[Noshairwan, 07] Noshairwan, W., Qadir M., A., Fahad, M.: "Sufficient Knowledge Omission error and Redundant Disjoint Relation in Ontology", InProc. 5th Atlantic Web Intelligence Conference, Fontainebleau, France (June 25-27, 2007)

[Noy, 00] Noy, N. F., Musen, M. A.: "The PROMPT suite: interactive tools for ontology merging and mapping", IJHCS, Elsevier, 59, 6 (2003), 983–1024.

[Noy, 01] Noy, N. F., Musen, M. A.: "Anchor-PROMPT: Using nonlocal context for semantic matching", In Proc. of Workshop on Ontologies and Information Sharing (IJCAI-01), Seattle, USA(2001), 63–70.

[Pottinger,03] Pottinger, R., Bernstein, P.: "Merging Models Based on Given Correspondences" In Proc. of 29th Intl. Conference on VLDB'03, Berlin, Germany (2003), 862-873.

[Qadir, 07] Qadir, M., A., Noshairwan, W.: "Warnings for Disjoint Knowledge Omission in Ontologies" In Proc. of 2nd ICIW'07, Morne, Mauritius (2007), 45-45

[Stumme, 01] Stumme, G., Mädche, A.: "FCA-merge: bottom-up merging of ontologies", In Proc. 17th IJCAI '01, Seattle, WA, USA (2001), 225–230.

[Wiederhold, 94] Wiederhold, G.: "An algebra for ontology composition", In Proc. of Monterey Workshop on Formal Methods, U.S. Naval Postgraduate School, Monterey CA (1994), 56–61

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Semantic Information in Medical Information Systems – from Data and Information to Knowledge: Facing Information Overload

Andreas Holzinger
(Research Unit HCI4MED, Institute of Medical Informatics, Statistics & Documentation
Medical University Graz, Austria,
andreas.holzinger@meduni-graz.at)

Regina Geierhofer
(Research Unit HCI4MED, Institute of Medical Informatics, Statistics & Documentation
Medical University Graz, Austria,
regina.geierhofer@meduni-graz.at)

Maximilian Errath
(Research Unit HCI4MED, Institute of Medical Informatics, Statistics & Documentation
Medical University Graz, Austria,
maximilian.errath@meduni-graz.at)

Abstract: Although we live in the age of multimedia, most information is still only available in text format. This is especially true in medical informatics. Consequently, research in text mining is an essential area of computer science. With the aid of statistic and linguistic procedures, text mining software attempts to dig out (mine) information from plain text. The aim is to transform data into information. However, for the efficient support of end users, facets of computer science alone are insufficient; the next step consists of making the information both *usable* and *useful.* Consequently, aspects of cognitive psychology must be taken into account in order to transform information into knowledge and finally wisdom. In this paper we describe some experiences we made with text mining in medicine and present a view of current and future challenges in Research & Development in order to match aspects of both Informatics and Psychology. We are aiming to provide *performance support* in clinical decision making. Our end users include professionals in the area of medicine and health care. Our goal is improved presentation of information – in order to enable enhanced cognitive information processing, resulting in easier and more rapid building of knowledge.

Keywords: Semantic Usability, Information Overload, Performance Support, Information Presentation, Medical Documentation
Categories: H.3.1, H.5.2, I.2.7, J.3

1 Introduction and Motivation

Information retrieval is an important part of the daily work of medical professionals and the body of knowledge in medicine is growing enormously. Clinicians must rely on various sources of medical information and increasingly they are faced with the problem of *too much* rather than too little information [Slawson et al., 1994]. The problem of *information overload* is rapidly approaching, however, it was both interesting and motivating for our work, that most of the past research about information overload in health care dealt with information from medical literature search [Sullivan et al., 1999], [Hall and Walton, 2004], and only a few dealt with information overload from clinical information systems [Noone et al., 1998]. Interestingly, people often compare the search in clinical databases to a classical Web search. However, the widespread format of medical documentation is strongly oriented towards individual patients and therefore the number of documents found is measurably smaller than one is accustomed to receiving from, for example, a typical Web search, although they may also comprise several hundred single documents [Holzinger et al., 2007]. Contrary to a Web information search, there is a much higher recall and precision expectancy [Baeza-Yates and Ribeiro-Neto, 2006], [Witten and Frank, 2005] from a search within the medical environment. Furthermore, it is of crucial importance that all documents necessary for the patient's treatment are at the disposal of the medical professionals within the shortest possible time, because often – for example in an emergency – the time for viewing the data is severely limited. A simple data base and/or text search is frequently not a sufficient information system with which to support the doctors and physicians effectively.

It is a crucial factor that a large part of the information necessary to make an appropriate sample is only available in text form – the usual electronic equivalent of medical texts. Support by expert systems is therefore practically impossible without first processing those textual documents. Although, current patient documents also contain codes, among other things for diagnosis or treatments, these only represent a very small part of the actual information available. This has two main causes: the routine coding is often for primarily administrative reasons, for example, accounting or controlling; and mapping information to a standardized system of concepts usually involves a loss of information, which is unacceptable in a field where exactitude is vital. An important step, in order to analyze the texts of, for example, Medical Finding Reports to the fullest extent, is to extract the medical statement, diagnoses and other texts. These facts can be reproduced in a suitable system of concepts and be combined with the (medical) codes originating from administrative and accounting documentation. When this *semantic information extraction* is completed, an appropriate form of *Information Presentation* for the interaction between medical professionals and the system must be developed.

2 Theoretical Background: Text Analysis

The most important means of communication in patient treatment – after the spoken word – is the written word, whereby given that dictation is usual within the medical environment, the intersection between the two is, from the viewpoint of the doctors and physicians, somewhat indistinct. All essential documents of the patient records contain at least a certain portion of data which has been entered in free-text fields. Although free text can be created simply and intuitively, it makes an automatic analysis enormously difficult [Gregory et al., 1995], [Holzinger et al., 2000], [Lovis et al., 2000]. One possible solution for this is the coding and/or the use of keywords for the information contained in the free texts.

In contrast to everyday language, medical language affords a multitude of existing and controlled vocabularies, classifications and nomenclatures.

Here, medicine has a long tradition, partly with schemata which is recognized and accepted world-wide, although current ways of producing text for documents in medicine are not optimal; natural language aims to generate text from a conceptual representation of not only certain facts, but also knowledge [Huske-Kraus, 2003].

Therefore, an analysis of medical texts can not be limited to a simple classical information retrieval by means of the search for words and word components, even when they have syntactically and orthographically correct contents. It is imperative that we take into account the medical connotation of the respective term and, above all, the *context,* which is the information that is being presented at the moment the information need is occurring [Ruan et al., 2000]. A semantic search can be supported optimally only when these measures have been taken.

3 Typical Procedures

The procedures used at present comprise many different methods from medical informatics, including: 1) the use of algorithms from the field of Natural Speech Processing or Natural Language Processing (NLP), 2) Information Retrieval and 3) Machine Learning; in addition, relational data bases and Topic Maps can be used [Bohm et al., 2002].

In a first step, the structured sections of the data, available in the form of semi-structured patient documents, are processed. Here, it is possible to work with the classical methods of Data Mining. An assessment of the data quality and data plausibility is meaningful, because it is not uncommon to find that the data, even if already structured, is incomplete and/or incomprehensible. Our experience shows that at least two factors are of crucial importance: The non-intuitive organization of the user interface and the priority of saving human life over imputing data. For doctors and physicians, the priority naturally lies in medical achievements

When the quality of the data has been examined, checked for feasibility and presented in the required format, then measures are taken to process the non-structured data. In order to extract the information from this data, it is necessary to carry out the usual Text Mining steps. Texts, which come from the daily clinical routine and were not subject to a systematic, linguistic review process, require more procedures than, for example, articles from scientific magazines. Which steps are necessary, and how these must be carried out in detail, depends on the kind and quality of the raw material and the further processing and/or their intended use. It is important to notice that the majority of patient documents were never examined for linguistic weaknesses. Orthographic errors are fairly common during the clinic routine and mistakes, which can obscure the meaning to a certain extent, must be taken into consideration. Physicians in general, combine different languages within a document, in particular Latin, German and English. Purely formally therefore, a multilingual approach is necessary. Abbreviations are used very frequently: Since the medical language is a technical language, it has also its own vocabulary. However, the problem is that this vocabulary varies regionally very strongly. This strongly differing vocabulary is, in addition, regularly abbreviated, particularly in the case of internal documentation.

As long as we are dealing with relatively small sets of data, - for example: many medical studies fall into this category – the automation can be considered unnecessary and the Keywords can be set manually. However, larger data sets make the utilization of tools inevitable and in a hospital, where very large data sets are the general rule, they are necessarily almost essential. Here, software solutions from the area of Text Mining can provide valuable functionality.

4 Usefulness of Medical Information

In [Slawson et al., 1994] the usefulness of medical information is expressed as

```
Usefulness = (Relevance x Validity) / Work
```

Wherein Relevance = 0 when the document is irrelevant to the current question, and the Validity = 0 when the document itself is invalid. Work is the work required to retrieve and interpret the information. Although Slawson's work (and many others in this area) refer to *medical literature*, in our experience this also holds for *medical documentation.* The information retrieval **unit of work** is known in Human–Computer Interaction as **time to perform task,** which is clearly measurable [Stary and Peschl, 1998]. This includes the four stages of user interaction, query definition, database retrieval, and presentation. Appropriate user interfaces can influence this process and thereby reduce the work and increase the usefulness. Consequently, it is important during the design of such interfaces to learn about the end users, who will be using it; different types of end users require different types of interfaces and the time to perform task is dependent on various metrics, including level of skill, goals, computer literacy, frequency of use and familiarity with the domain (expert/novices) [Johnson et al., 2005], [Holzinger, 2005]. The primary goal of performance support ist to provide adaptive user interfaces, which assist end users to accomplish their work accurately. They must be specifically adapted to the end users individual perception of work and methods. By addressing the end users' needs, the quality of their work, as well as their individual performance, can be improved. The adaptation of behavior of interactive applications to individual work processes can enable the achievement of this and other goals. According to [Stary and Stoiber, 2003], user interfaces can be considered as *process portals* reflecting the individual views on the work in which the end users are involved. Thereby, it is essential to focus on individual end user performance, rather than on purely functional specification [Stary and Stoiber, 2003].

5 Usability of Medical Information

In order to make the information received accessible, one last step is crucial: the usable *presentation of the information* in order that the end users are able to transform the perceived information into knowledge. Here, we must emphasize that there are some deviations in the definitions of knowledge and here, we follow the psychological definition, which deviates from the general definition in the area of knowledge management [Schneider, 2002]. Knowledge in that respect is the result of the interpretation of information [Holzinger, 2002]. Not until the information has been psychologically perceived, assimilated and cognitively processed and accepted in the memory – through integration into existing knowledge – can the information become knowledge.

While end users are confronted with a flood of data, the designers of the user interface must learn to differentiate between information required, which must consequently be integrated into the chain of procedure, and the (currently) irrelevant data, in order to support the cognitive development of knowledge. It is easy to acknowledge that an information presentation, which has been adapted to the medical profession, can contribute considerably towards performance support for the end users. This research into the inclusion of human factors in the development of such systems is of vital importance [Kohn et al., 1999].

One serious issue concerned with textual information is the problem of *Information Overload* [Eppler and Mengis, 2004], when end users are confronted with too much irrelevant information, this is similar to *Memory Overload*, [Koppel et al., 2005], [Nielsen, 2005]. The flood of information is actually rising constantly; more and more information becomes available within our medical information systems, while the cognitive capabilities of humans during the same period of time fails to increase to the same extent [Lewis et al., 2006]. Information overload reduces the quality of work and is a major cause of difficulties, which medical professionals experience during diagnosing.

In cognition research, one of the principles used to explain this phenomenon is referred to as *Cognitive Load* [Sweller, 1988]. This principle, originally based on the limitations of the human short-term memory [Miller, 1956], is probably the most prominent problem in human information processing.

Therefore, the minimization of the *cognitive load,* which we believe to be an important factor in the maximization of the end users cognitive performance, and facilitating the assessment of the most substantial information are the most prominent goals and remain the principle of good and accepted information systems. The form of the information presentation depends considerably on the working habits, the working environment and on the existing IT System arrangements. In order to make a holistic perspective possible, the end users workflow must be carefully examined in detail and, above all, the experiences and experimental results from *real life* must flow into the informatics development at systemic level.

6 Tools for the analysis of medical texts

A distinctive feature for tools designed for the semantic analysis of medical texts is their original range of primary application. It is interesting that some these tools were first used mainly in accounting and for administrative documentation. Others originate from the area of document research. Typical representatives of the first group would be "Semfinder" or "ID DIACOS", a typical representative of the second group "Morphosaurus" [Marko et al., 2005]. The knowledge base of tools, which originates from the routine, is based on years of experience with medical documentation conforming to clinical convention. Local language characteristics need to be allowed for, just as the regional legislation and naturally the rules of the respective financing system must be taken into consideration. Through this close connection with the hospital business, the tools are particularly well-engineered and therefore particularly suitable for these areas.

The second group of tools was originally used to retrieve scientific texts. Previous to publication, these texts were reread, peer-reviewed, edited etc., of course, the correctness of the contents were also implicitly controlled.

As a result, the raw material has a substantially higher quality. Since scientific publications are, per se, intended for a wider public, they exhibit a far smaller regional specificity than the texts from the clinical routine and patient treatment. Naturally, texts retrieved by the second group offer far better conditions for key-wording etc., than texts from routine medical work. A further differentiating factor is the width of the approach. Here the implementations, which are intentionally limited to an extract of the possible documents and/or contents, are divergent to those with a more comprehensive approach. For example, on the basis of sentence fragments, "Semfinder" supports a diagnostic coding – primarily interactive but also in batch when desired – and only includes other aspects as for example: treatments, medicines and patient occupation, to the extent that this is relevant to the diagnostic coding.

Other programs focus on some special document types, for example for key-wording biomedical texts or annotating pictorial material. However, some products – such as the "ID DIACOS" linked to "ID Tasmed" – exclude no document and/or no aspect of medical documentation. The goal is to make the patient document accessible in its entirety.

The different computer-linguistic procedures, which are used for the analysis of natural speech, can also be drawn on for the differentiation of documents. These are tools that can access a specific corpus (e.g. the MorphoSaurus mentioned above) and furthermore, they can work with semantic nets and/or ontologies (e.g. ID DIACOS). However, our experience showed that a rule-based approach supplied quite good results – particularly when the area is relatively small.

7 Conclusion and Future Outlook

In the course of the further development of the organization-spanning, interlaced patient document (keyword: e-Health), it is to be anticipated that the semantic development of the medical documentation will increase in significance. The synergy with the Semantic Web group will probably continue to accelerate the development. Similar to the World Wide Web, the "World Wide Health Net" would offer the possibility of a more purposeful search. In addition to this improvement, the use of information in expert systems would actually be possible to a wider extent than ever before. A possible application here would be, for example, a "Ranking" of medical documents, according to their relevance with regard to a definite question. The steps described above – to convert a semi-structured patient document into a machine processable database – are an important advance on this path. However, a further component for the successful implementation of any health net is the expert knowledge, as to the relevance of different combinations of observations and findings for the respective medical problem. To date, no generally accepted specified set of rules to specify the characteristics of relevant findings exists.

Even for wide-spread diseases, several different sets of rules as to the correct procedure often exist – in medicine; these are also referred to as clinical paths. It is essential to acquire the existing expert knowledge for each area considered. However, expert knowledge is not always simple to verbalize, and often not available – at least consciously. Everyone has, at some time, explained their method by phases such as: "it just looks right" or "that way feels right".

In future, the conscious knowledge of the respective relevance of medical facts combined with the articulation of unconscious knowledge which is manifested in the actual observation pattern of the medical findings will become increasingly more important. From these and other questions, which arose during our work, it is obvious that there is a need for interdisciplinary cooperation between computer scientists, psychologists and educationalists.

References

[Baeza-Yates and Ribeiro-Neto, 2006] Baeza-Yates, R. and Ribeiro-Neto, B.: "Modern information retrieval", ACM Press, New York, (2006).

[Bohm et al., 2002] Bohm, K., Heyer, G., Quasthoff, U. and Wolff, C.: "Topic map generation using text mining", Journal of Universal Computer Science, 8, (2002), 623-633.

[Eppler and Mengis, 2004] Eppler, M. J. and Mengis, J.: "The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines", Information Society, 20, (2004), 325-344.

[Gregory et al., 1995] Gregory, J., Mattison, J. E. and Linde, C.: "Naming Notes - Transitions from Free-Text to Structured Entry", Methods of Information in Medicine, 34, (1995), 57-67.

[Hall and Walton, 2004] Hall, A. and Walton, G.: "Information overload within the health care system: a literature review", Health Information and Libraries Journal, 21, (2004), 102-108.

[Holzinger, 2002] Holzinger, A.: "Multimedia Basics, Volume 2: Learning. Cognitive Fundamentals of multimedial Information Systems", Laxmi, New Delhi, (2002). Available also in German: www.basiswissen-multimedia.at

[Holzinger, 2005] Holzinger, A.: "Usability Engineering for Software Developers", Communications of the ACM, 48, (2005), 71-74.

[Holzinger et al., 2007] Holzinger, A., Geierhofer, R. and Errath, M.: "Semantische Informationsextraktion in medizinischen Informationssystemen", Informatik Spektrum, 30, (2007), 69-78.

[Holzinger et al., 2000] Holzinger, A., Kainz, A., Gell, G., Brunold, M. and Maurer, H.: "Interactive Computer Assisted Formulation of Retrieval Requests for a Medical Information System using an Intelligent Tutoring System", ED-MEDIA 2000, Montreal, (2000), 431-436.

[Huske-Kraus, 2003] Huske-Kraus, D.: "Text generation in clinical medicine - a review", Methods of Information in Medicine, 42, (2003), 51-60.

[Johnson et al., 2005] Johnson, C. M., Johnson, T. R. and Zhang, J. J.: "A user-centered framework for redesigning health care interfaces", Journal of Biomedical Informatics, 38, (2005), 75-87.

[Kohn et al., 1999] Kohn, L. T., Corrigan, J. M. and Donaldson, M. S.: "To Err Is Human: Building a Safer Health System ", The Institute of Medicine, Washington (DC), (1999).

[Koppel et al., 2005] Koppel, R., Metlay, J. P., Cohen, A., Abaluck, B., Localio, A. R., Kimmel, S. E. and Strom, B. L.: "Role of computerized physician order entry systems in facilitating medication errors", Jama-Journal of the American Medical Association, 293, (2005), 1197-1203.

[Lewis et al., 2006] Lewis, R. L., Vasishth, S. and Van Dyke, J. A.: "Computational principles of working memory in sentence comprehension", Trends in Cognitive Sciences, 10, (2006), 447-454.

[Lovis et al., 2000] Lovis, C., Baud, R. H. and Planche, P.: "Power of expression in the electronic patient record: structured data or narrative text?" International Journal of Medical Informatics, 58, (2000), 101-110.

[Marko et al., 2005] Marko, K., Schulz, S. and Hahn, U.: "MorphoSaurus - Design and evaluation of an interlingua-based, cross-language document retrieval engine for the medical domain", Methods of Information in Medicine, 44, (2005), 537-545.

[Miller, 1956] Miller, G. A.: "The magical number seven, plus or minus two: Some limits of our capacity for processing information", Psychological Review, 63, (1956), 81-97.

[Nielsen, 2005] Nielsen, J.: "Medical Usability: How to Kill Patients Through Bad Design In: Jakob Nielsen's Alertbox, April 11", (2005),

[Noone et al., 1998] Noone, J., Warren, J. and Brittain, M.: "Information overload: opportunities and challenges for the GP's desktop", Medinfo, 9, (1998), 1287-1291.

[Ruan et al., 2000] Ruan, W., Burkle, T. and Dudeck, J.: "A dictionary server for supplying context sensitive medical knowledge", Journal of the American Medical Informatics Association, (2000), 719-723.

[Schneider, 2002] Schneider, U.: "The knowledge-attention-gap: Do we underestimate the problem of information overload in knowledge management?" Journal of Universal Computer Science, 8, (2002), 482-490.

[Slawson et al., 1994] Slawson, D. C., Shaughnessy, A. F. and Bennett, J. H.: "Becoming a Medical Information Master - Feeling Good About Not Knowing Everything", Journal of Family Practice, 38, (1994), 505-513.

[Stary and Peschl, 1998] Stary, C. and Peschl, M. F.: "Representation Still Matters - Cognitive Engineering and Task-Based User Interface Development", Behavior and Information Technology, 17, (1998), 338-360.

[Stary and Stoiber, 2003] Stary, C. and Stoiber, S.: "Model-based electronic performance support", Interactive Systems: Design, Specification, and Verification, Springer-Verlag Berlin, Berlin, 2003, 258-272.

[Sullivan et al., 1999] Sullivan, F., Gardner, M. and Van Rijsbergen, K.: "An information retrieval service to support clinical decision-making at the point of care", British Journal of General Practice, 49, (1999), 1003-1007.

[Sweller, 1988] Sweller, J.: "Cognitive load during problem solving: Effects on learning", Cognitive Science, 12, (1988), 257-285.

[Witten and Frank, 2005] Witten, I. H. and Frank, E.: "Data Mining: Practical Machine Learning Tools and Techniques. Second Edition", Morgan Kaufmann, Amsterdam et al., (2005).

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Creating an Annotated Set of Medical Reports to Evaluate Information Retrieval Techniques

Regina Geierhofer
(Research Unit HCI4MED, Institute of Medical Informatics, Statistics & Documentation
Medical University Graz, Austria
regina.geierhofer@meduni-graz.at)

Andreas Holzinger
(Research Unit HCI4MED, Institute of Medical Informatics, Statistics & Documentation
Medical University Graz, Austria
andreas.holzinger@meduni-graz.at)

Abstract: During the evaluation of Information Retrieval Algorithms, measurements are often made against a Gold Standard. Since the construction of such a Gold Standard requires considerable resources, a high validity and a high degree of reusability would be an advantage. From a scientific point of view, a high comparability between different Gold Standards would be advantageous, in order to enable the comparison of the results of different measurements. Unfortunately, the validity of such a standard can be negatively affected by many different factors and the reusability and comparability is often limited. In this article, we enumerate some problems, which can negatively affect the result. We show, by the creation of an annotated set of pathology reports, in which form these problems can emerge and how we have tried to minimize their influence.

Keywords: Information Retrieval, Text Mining, Semantic Usability, Medical Documentation
Categories: H.3.1, H.3.3, I.2.7, J.3

1 Introduction and Motivation for Research

The *Recall*, *Precision* and *Fall-out measurements* have proved useful in the evaluation of the tools and algorithms used for information retrieval [Holzinger et al., 2007]. Essentially, they correspond to the yardsticks used by statistics: Sensitivity, positive predicative value and 1-specificity. One common method of accomplishing this is the measurement against a so-called Gold Standard [Darmoni et al., 2003]. Thus, this Gold Standard is a *truth* within the context of this evaluation.

In medicine, *truth* is classically determined – when not clearly establishable through measurements – by the consensus of several specialists. This procedure is also applied, in a modified form, by the production of an annotated set of findings. Several selected persons annotate the selected documents. Should, thereby, deviations arise between the specialists' decisions; there are several possibilities for solving these and arriving at a consensus.

Usually, there is a further person, who makes the final decision, however there are also other methods, which can be used, such as the Delphi-method [Graham et al., 2003], average values [Michaelis et al., 1990], or the majority decision [Zhang et al., 2006].

The annotation procedure appears relatively simple and in publications there is usually only a short reference to the creation of the reference set, although a faulty reference set has the same influence on an evaluation as a badly calibrated measuring instrument on a physical measurement. In our study we investigated the possible sources of error and problems arising with the creation of a Gold Standard, and show how we can avoid these.

2 Problems identified

2.1 Problem: Badly or insufficiently specified questions

All scientific work must begin with a hypothesis or the statement of a problem. When measuring the quality of the Information Retrieval, the measurement is naturally not the actual question but only a gauge of how far the actual goal – the (re-) procurement of the relevant information – was reached. The design of the evaluation should, therefore, be determined by the intended goal of the information retrieval. The selection of the annotation strategy, the catalogs and the range of the set to be annotated must be dependent on this. Since the designs differ in detail, according to the setting, it is usually not possible to use the same reference set for different questions without revision.

The central aim of our project is to reduce the *information overflow* in electronic health records. Consequently, we are developing algorithms, which annotate reports with regard to the following characteristics for refining the pathologists' search:

a) Is it possible to conclude the presence of a neoplasm from this pathological report and/or
b) Does the report contain a statement, which indicates an inflammation?

During our project, we were faced in particular with the problem of whether to classify as relevant a pathological report, which contains a clear reference to an existing neoplasm but which either, makes no statement about a neoplasm or at least fails to mention one explicitly.

Examples can be found in the area of tests (e.g.: HER-2-Test) carried out by existing neoplasm or by the explicit declaration of a connected neoplasm (e.g.: investigation of normal lymph nodes with known Mamma Carcinoma). With regard to a ranking of the relevance of a report, this question would be answered with a clear "Yes"; with regard to a collective structuring for further processing (e.g.: administrative statistics or Quality Management) with "No". Naturally, in this case it is possible to increase the number of regarded attributes but the expenditure rises, lineally at the least, with each new question. We are of the opinion that a difference must be made between *certain* diagnoses (found/not found) and *uncertain* diagnoses. Since the organs involved represent also an important refinement criterion for pathologists' report retrieval we also decided to realize an annotation of these organs. As a secondary goal, we examined our data set building algorithms for validity, whereby we accepted the possible necessity of a re-annotation and/or manual control.

2.2 Problem: Unsuitable Catalogue for Annotation

In medicine, there are many schemas, with which it is possible to code medical issues and entities [Strang et al., 2002]. The Unified Medical Language System (UMLS) currently encompasses over 140 terminologies [Chen et al., 2007]. Nevertheless it is also thoroughly possible, and also often meaningful, to create an individual annotation system, particularly by adapting one of internationally usage. Therefore, choosing the correct one from this abundance of possibilities is not straightforward. The fitting system must at the least:

a) Be exact enough to be able to act as a clear reference for the question
b) Be simple enough for practically usage
c) Be unique with regard to the facts to be annotated
d) Be complete, in order to be able to annotate all relevant facts

The following example from our project clarifies the form in which the above questions can appear in practice. For the primary question, a binary annotation of our reference set would actually have been sufficient.

However, with regard to the secondary question, it was necessary to refine the statements in particular those referring to neoplasm. We decided, on the basis of neoplasm, for the topology and/or morphology axis of the ICD-O-3. However we thereby limited the neoplasm, which we could annotate, to those which were oncologically interesting. With regard to the granularity, we effected a compromise. We only required from our annotators the indication of an ICD-O group, not however the selection of a specific code. However, by an appropriate configuration of the user interface, we succeeded in receiving, the accurate ICD-O Morphology code as annotation. In order to make a clear allocation to a group possible, we had to suspend the hierarchical structure of the ICD-O, as they are specified, for example, in the area of the Hodgkin Lymphomas and Non-Hodgkin Lymphomas. Also, we did not explicitly require the indication of a four digit organ code but were satisfied with the three digit code. Both of these were a result of the circumstance that accurate codes cannot also be clearly derived retrospectively from the pathological text and increased granularity meant greater expenditure. With regard to the coding rules, we oriented ourselves to the instructions for coding available over the DIMDI (Deutsches Institut für Medizinische Dokumentation und Information), however, in some cases we deliberately deviated from it, especially if the coding rules recommend accumulative codes like *C26[.8] digestive system – several parts overlapped*. In order to minimize the sources of error, we also used the topology axis of the classification for the annotation with regard to the inflammations.

In order to annotate the diagnostic certainty, we selected the appropriate codes from the SNOMED CT [Schulz et al., 2006]. However, practical preliminary tests soon showed that the values are not clearly interpretable with regard to the degree of certainty expressed (see table 1).

In order to make a reproducible annotation of all documents possible, and to be able to differentiate between diagnoses found and not found respectively, it was necessary for a conclusion of neoplasm as well as for inflammation, to insert two further values in the axis diagnostic certainty - "not found" and "not mentioned".

2.3 Problem: Unsuitable Selection of Annotators

Humans differ from each other and their sensory impressions influence both their behavior and their perception, in particular when these are embodied in their long-term memory [Holzinger, 2002]. This statement appears trivial; however it is often not taken into account during annotation.

Most studies, insofar as they indicate the context of the annotators at all, try to avoid this problem of different mental models by selecting annotators with comparable context. What is nearly always missing is the indication as to why this group of annotations was chosen. Usually the answer might be trivial, for example: “because they were the best that were obtained”, however, an exact definition of “the best” is not always possible.

To take an example of an actual case, which arose during the annotation of a pathological report at our Institute:

`“Stomach mucosa of antrum type without substantial pathological changes. No appreciable inflammation. No helicobacter pylori provable”.`

The report was to be annotated, with regard to the presence of an inflammation. The following annotators with divergent prior knowledge annotate differently:

a) A medical computer scientist with sound, predominantly theoretical, knowledge of coding in medicine and a basic knowledge of medicine: Without substantial pathological change → an insignificant change exists; no appreciable inflammation → there is an inflammation, but it is not significant → inflammation available – certain diagnosis.

b) A pathologist

Without substantial pathological change → thus only insignificant pathological change, thus practically no inflammation; no appreciable inflammation → no inflammation is still supported; no inflammation – certain diagnosis.

c) Accounting personnel

no considerable inflammation → there is a diagnosis with suspicion of tentative diagnosis of inflammation, also a test was made, this caused expenditure and the suspicion was not explicitly excluded → inflammation existent – uncertain diagnosis.

The meaning of levels of conclusiveness or degrees of certainty in a diagnosis offers another example:

We took the 7 codes for the indication of diagnostic certainty provided by SNOMED, and ask 14 different people; (physicians, computer scientists, biometricians). to sort the codes in descending order. We received 11 different lin-ups (lu).

Definite (1)			Consistent with (2)			Equivocal (3)			Inprobable (4)	
Probable (5)			Possible (6)			Uncertain (7)				
lu1	**lu2**	**lu3**	**lu4**	**lu5**	**lu6**	**lu7**	**lu8**	**lu9**	**lu10**	**lu11**
1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	3	3	3	3	4
3	3	3	4	4	6	2	2	2	4	2
4	4	5	3	3	5	4	4	4	2	3
5	6	4	5	6	4	5	6	7	6	5
6	5	7	7	5	3	7	5	5	5	7
7	7	6	6	7	7	6	7	6	7	6

Table 1: Diagnosis Certainty Look-ups

Consequently, it makes sense to select end users with a similar background to annotate, since one would otherwise have to expect a degeneration of the inter-rater reliability. However, the fact that another group, with different background can also annotate just as reliably and congruently and nevertheless come to different results with "flexible" formulations, remains unheeded. The specification of strict annotation rules only appears to solve this problem, since these regulations were provided, in principle, by people who are shaped and influenced by their own background. We have tried to moderate this dilemma by allowing different groups (computer scientists; biometricians; medical students and physicians) to work together on defining the relatively strict annotation rules.

2.4 Problem: Quality control of the Gold Standard

A Gold Standard stands for *the truth*, in evaluations, however it must not necessarily be correct from a specialist point of view. Continuing from our example of the 2+1 annotators; if the two primary annotators come to the same, nevertheless wrong result, the report is not submitted to the third annotator and is therefore adopted unchecked. An equally wrong result can arise when a characteristic attribute occurs very frequently or many findings are attribute-free and therefore high probability exists that inattentiveness can cause this annotation code to be falsely selected. Since Gold standards can also be incorrect, for other reasons, it is recommended to submit it to a quality control.

The measurement of the inter-rater reliability provides a first measure for the quality of the annotation, in order to principally determine the measure of agreement.

2.4.1 Problem: Relativity of Inter-rater Reliability Measure according to the Cohen's Kappa example

The Inter-rater reliability states the measure of homogeneity of categorisation, between two or more Raters, in excess of the incidental congruency. Therefore, the larger the coincidental agreement – the smaller, the Kappa value [Thompson and Walter, 1988].

In order to determine these, there must be a valid model, which permits estimation of the size of the incidental conformity. If this rate of incidental conformity is not known, the usual measure of Inter-rater reliability cannot be used, even when there are approaches for other methods [Hripcsak and Rothschild, 2005].

The size of the inter-rater reliability gives some indication of how high the reliability of the rater is, but does not indicate how large the actual match is or how correct the annotation is. One reason why it does not represent a measure of the actual identicalness is that also statistic methods have restrictions. The readily used Kappa coefficient [Cohen, 1968], for example, is encumbered with a number of problems, which make a comparison between the Kappa values of the individual studies difficult. On the one hand, the value of Kappa depends on the distribution of the non-identical values [Feinstein and Cicchetti, 1990], on the other hand, also it also depends on the prevalence of the characteristic. Example of the dependence on the prevalence of the characteristic: Pathology reports are to be annotated with regard to the tumor dignity of possible existing neoplasms. When the annotations are distributed as in the following table we receive a Kappa of 0.800:

dignity	*benign*	*in situ*	*malign*	*metastasis*	*without neoplasm*	*total result*
benign	14	1	1	1	1	18
in situ	1	2	1	1	1	6
malign	1	1	31		1	34
metastasis	1	1		1		2
without neoplasm	1	1	1		61	64
total result	18	6	34	3	64	125

Table 2: Example Kappa 0.800

However, when most pathology reports contain no neoplasm, which means that the incidental congruency has a clear maximum in one place, the Kappa value in the example below is drastically reduced to 0.574. Although the rater agreed just as often, and the number of non-identically annotated reports, to be examined by the 3rd annotator remained the same, the result was a fundamentally different reliability.

In our case, we made direct comparisons of equality and made certain that each deviating error was accomplished either due to single usage of the Delphi-method or, in case that failed, an annotation by a third specialist lead to clarification. To our third specialist it was also recommended to secure his judgment, where possible, by references and afterwards replied to the questions of the other annotators. This occasionally led to a change of the annotation and, by indicating references, to a higher level of comprehension. Unfortunately however, it also led to a rise in the expenditure of time.

dignity	*benign*	*in situ*	*malign*	*metastasis*	*without neoplasm*	*total result*
benign	2	1	1	1	1	6
in situ	1	2	1	1	1	6
malign	1	1	2		1	5
metastasis	1	1		2		3
without neoplasm	1	1	1		101	104
total result	6	6	5	4	104	125

Table 3: Example Kappa 0.574

2.4.2 Problem: Examination of conforming false annotations

When the above mentioned primary annotators came to the same incorrect result, then the annotation was adopted in the standard. One possibility of reducing the probability of this error would be to increase the number of primary annotators. However, this only lowers the rate of coincidental errors, *not the systematic errors.* Another method would be to submit ALL reports to the third specialist for examination. This can reveal both systematic and coincidental errors. Both methods together are responsible for the increase in personnel expenditure for the annotation, examination and co-ordination.

The systematic errors can be decreased by exact annotation rules. A disadvantage of this method is, however, the necessity of determining rules in advance, which must be medically well-founded, and which essentially cover the range to be annotated. Unfortunately, it is not easy to find such rules in medical science.

When it is possible to develop such rules, at least for a sub-division of the field, then it becomes possible to design a computer program that makes a parallel annotation possible and thus to effect a further annotation, which is free of the results of inattentiveness, fatigue and variable interpretation. However, this program must be designed by people other than the developers of the program to be tested, in order to prevent the algorithm being tested from reproducing the annotations, rather than using the fundamental hypotheses und techniques. Unfortunately, this method also clearly increases the expenditure of time and personnel. We tested this approach, and had good results. Due to our good inter-rater reliability, we found non systematic errors, probably caused by fatigue or inattentiveness.

3 Conclusion and Future Outlook

We observed various difficulties and factors of influence on the validity and reusability of a Gold Standard. It appeared necessary to indicate at least the benchmark data of the factors of influence mentioned above, in order to enable the reader to compare the results of the evaluation with others. This does not by any means claim to provide a complete list of all influencing factors; such as the necessary size of the random sample or how the sample is chosen. Both these factors of influence are of crucial importance for the extrapolation of the results of the total. With the increasing size of data collections, they already represent a problem, due to the necessarily increasing sample size or the unknown total volume and composition (e.g.: WWW). Generally, we discovered that the effort to achieve higher quality is always connected to higher expenditure and thus to higher costs. It therefore appeared meaningful, to reuse *such reference sets* in the future, including spanning groups of researchers, in order to lower the costs per evaluation for a well documented Gold Standard to an extent, thereby permitting the evaluation of ever larger amounts of data. However, much further research is necessary within this area.

References

[Chen et al., 2007] Chen, Y., Perl, Y., Geller, J. and Cimino, J. J.: "Analysis of a study of the users, uses, and future agenda of the UMLS", Journal of the American Medical Informatics Association, 14, (2007), 221-231.

[Cohen, 1968] Cohen, J.: "Weighted Kappa - Nominal Scale Agreement with Provision for Scaled Disagreement or Partial Credit", Psychological Bulletin, 70, (1968), 213-220.

[Darmoni et al., 2003] Darmoni, S. J., Amsallem, E., Haugh, M., Lukacs, B., Leroux, V., Thirion, B., Weber, J. and Boissel, J. P.: "Level of evidence as a future gold standard for the content quality of health resources on the Internet - A preliminary study", Methods of Information in Medicine, 42, (2003), 220-225.

[Feinstein and Cicchetti, 1990] Feinstein, A. R. and Cicchetti, D. V.: "High Agreement but Low Kappa .1. the Problems of 2 Paradoxes", Journal of Clinical Epidemiology, 43, (1990), 543-549.

[Graham et al., 2003] Graham, B., Regehr, G. and Wright, J. G.: "Delphi as a method to establish consensus for diagnostic criteria", Journal of Clinical Epidemiology, 56, (2003), 1150-1156.

[Holzinger, 2002] Holzinger, A.: "Multimedia Basics, Volume 2: Learning. Cognitive Fundamentals of multimedial Information Systems", Laxmi, New Delhi, (2002).

[Holzinger et al., 2007] Holzinger, A., Geierhofer, R. and Errath, M.: "Semantische Informationsextraktion in medizinischen Informationssystemen", Informatik Spektrum, 30, (2007), 69-78.

[Hripcsak and Rothschild, 2005] Hripcsak, G. and Rothschild, A. S.: "Agreement, the F-measure, and reliability in information retrieval", Journal Of The American Medical Informatics Association, 12, (2005), 296-298.

[Michaelis et al., 1990] Michaelis, J., Wellek, S. and Willems, J. L.: "Reference-Standards For Software-Evaluation", Methods Of Information In Medicine, 29, (1990), 289-297.

[Schulz et al., 2006] Schulz, S., Hanser, S., Hahn, U. and Rogers, J.: "The semantics of procedures and diseases in SNOMED (R) CT", Methods of Information in Medicine, 45, (2006), 354-358.

[Strang et al., 2002] Strang, N., Cucherat, M. and Boissel, J. P.: "Which coding system for therapeutic information in evidence-based medicine", Computer Methods and Programs in Biomedicine, 68, (2002), 73-85.

[Thompson and Walter, 1988] Thompson, W. D. and Walter, S. D.: "Kappa And The Concept Of Independent Errors", Journal Of Clinical Epidemiology, 41, (1988), 969-970.

[Zhang et al., 2006] Zhang, J., Hsee, C. K. and Xiao, Z. X.: "The majority rule in individual decision making", Organizational Behavior and Human Decision Processes, 99, (2006), 102-111.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

A Semantic Web Framework for the Life Sciences Based on Foundational Ontologies and Metadata Standards

Matthias Samwald
(Section on Medical Expert and Knowledge-Based Systems, Medical University of Vienna
Austria
matthias.samwald@meduniwien.ac.at)

Klaus-Peter Adlassnig
(Section on Medical Expert and Knowledge-Based Systems, Medical University of Vienna
Austria
klaus-peter.adlassnig@meduniwien.ac.at)

Abstract: This article describes an ontology framework called *bio-zen,* which can be used for the representation of information from biomedical research on the Semantic Web. The ontology framework adheres to the OWL DL format and is based on existing foundational ontologies and metadata standards like DOLCE, SKOS and Dublin Core. It is optimised for the usage in distributed environments like the Internet. Novel ontological design patterns in bio-zen allow the unification of good ontological consistency with a flexible, clean and intuitive structure. A unique feature of the bio-zen ontology is that it allows the seamless integration of mathematical descriptions and simulation parameters into qualitative information, making a quick transition from plain data to model simulations possible.
A growing number of extension packages are available for the ontology, including concepts from taxonomies such as the Gene Ontology, Medical Subject Headings or the NCBI Taxonomy.

Keywords: Semantic Web, Life Sciences, ontology, simulation, data integration
Categories: J.3, H.3.0

1 Introduction

The development of the *bio-zen ontology framework* is an attempt to represent data and information from research in all facets of the *life sciences* on the Semantic Web. The goal of this project is the unification of information that is now scattered over a multitude of different data structures, exchange formats and databases. Through the use of Semantic Web technologies, the decentralised and barrier-free development and exchange of experimental data, hypotheses and biological models becomes possible. A unique feature of the *bio-zen* ontology is that it allows for a seamless integration of mathematical descriptions and simulation parameters into qualitative information, enabling a quick transition from plain data to model simulations and back.

The development of this ontology addresses several pressing needs that are not fulfilled by current ontologies for the life sciences. For example, most of the currently available Semantic Web ontologies for information exchange in the life sciences (e.g.

BioPAX [biopax][Luciano, 05], MGED ontology [Whetzel, 06]) are based on foundational ontologies. This has several disadvantages:

First, it slows down the development of the ontologies. Without a basic ontological base to build on, every project has to re-invent basic relations and classes (at least implicitly). This is a time-consuming task, especially in projects where many participants are developing the ontology in collaboration and each participant possibly implies different and incompatible ontological foundations.

Second, the absence of a sound ontological basis can result in poor design choices. For instance, a feature that can be observed in many domain-specific Semantic Web ontologies is the use of many different properties where the same information could also be conveyed with very few generic properties like ‘part of’ or ‘participant in’ or ‘attribute of’. One drastic example is the MGED ontology that contains over 100 properties that could be reduced to just a few properties without losing expressivity [Soldatova, 04]. Such redundancies do not only make it very complicated to understand, use and maintain ontologies; they also complicate the construction of queries and the interoperability between ontologies.

2 Methods

The ontology was designed with the Stanford Protégé ontology editor and its OWL plugin [Protégé].

The basic structure of the ontology was created by mapping classes from the BioPAX ontology to the classes of the *Descriptive Ontology for Linguistic and Cognitive Engineering* (DOLCE [Gangemi, 02][dolce]), a foundational ontology available in OWL DL format. It soon became apparent that the BioPAX ontology is mostly focused on an abstract, conceptual representation, as opposed to a direct ontological description of biological reality. This led to some major reinterpretations of the classes imported from BioPAX.

Where possible without significant loss of expressivity, properties and classes from BioPAX were replaced by more generic properties and classes from the DOLCE ontology. Besides the mapping of BioPAX classes and properties, many new structures were added to make the ontology more expressive.

After most of the ontology development was done, a considerable number of classes and properties from DOLCE that were deemed not useful or too complicated for the scope of the ontology were removed. Among the things that were removed are most of DOLCE’s advanced modules, all constructs dealing with ‘quality spaces’ and all inverse properties.

Properties defined in the RDF version of the Dublin Core metadata standard [dublincore] were used to replace some properties from the BioPAX ontology, mostly for the description of database entry provenance and bibliographic information. The core and extended ontologies of the *Simple Knowledge Organisation System* (SKOS [skos]) were added to *bio-zen* for the description of concepts and taxonomies. Because of their special importance, all concepts from the OBO evidence ontology [evidence] were also included in the core *bio-zen* ontology in the form of SKOS concepts.

A basic design requirement for the *bio-zen* ontology is conformance with the OWL DL standard. While the DOLCE ontology is already valid OWL DL, the SKOS and Dublin Core are not: SKOS uses a mixture of constructs from OWL and generic RDF; the official version of Dublin Core is pure RDF Schema. Some minor modifications were made to these ontologies to make them conform to OWL DL. All of the ontologies were merged into a single OWL file to avoid problems with ontology imports when a client is disconnected from the internet.

3 Basic principles and design patterns

The ontology is built upon existing foundational ontologies and metadata standards (DOLCE, SKOS, Dublin Core). Through the use of established foundational ontologies and metadata standards, *bio-zen* is rooted in a sound ontological framework, easing the interoperability with ontologies from other domains.

Statements made in *bio-zen* are direct ontological descriptions of biological reality and not of some abstraction of biological reality. This is in contrast to many other bioinformatics projects based on RDF, which are focused on the description of such abstractions, e.g. the organisation of database records. On the contrary, *bio-zen* is focused on the description of *spatio-temporal particulars* (concrete biological things existing in a certain time and space).

Users of the ontology only need to make OWL individuals to represent information. The definition of new classes is not necessary, which helps to keep the class structure clean and simple. It also helps to delineate ontology developers, i.e. people that are educated about the development of ontologies and that make use of ontology editors like Protégé [Protégé], from end-users, i.e. people that are not educated about ontology development and that make use of specialised software or internet portals. Furthermore, this design avoids many other grave problems that arise with OWL ontologies that have an overly complex class structure [Samwald, 06]. If the user of the ontology intends to make statements about some general observed phenomena (e.g. 'Drosophila has two wings'), these general principles are *exemplified* with a certain spatio-temporal-particular that acts as a canonical reference.

A characteristic feature of the ontology is that it integrates two different approaches of information representation in a common framework: 'realist' ontological descriptions and 'conceptual' taxonomies and concept hierarchies. The difference between the two approaches lies in the levels of abstraction: whilst the former focuses on describing reality itself, the latter focuses on describing the conceptualisations humans have made about reality. Unifying both approaches in one common framework makes it possible to combine the specific advantages of each approach in the best way possible. The consistency of the realist approach is complemented with the flexibility of the conceptualist approach. Making a clear distinction between both approaches reduces the susceptibility to inconsistencies.

'Realist' ontological descriptions in *bio-zen* are based on DOLCE and *spatio-temporal particulars* and employ a rich set of classes and properties. 'Conceptual' descriptions are based on SKOS, and mainly use only one class (*'skos:Concept'*) and only a few properties to describe the relations between concepts (*'broader'*, *'narrower'*, *'related'* etc.). The two forms of descriptions are almost fully disconnected. The only property that can be used to connect the two is called *described-by*. It can be used to annotate and identify a *spatio-temporal-particular* with one or more concepts (Fig. 1). This design pattern and the *described-by* property are innovations of the *bio-zen* ontology.

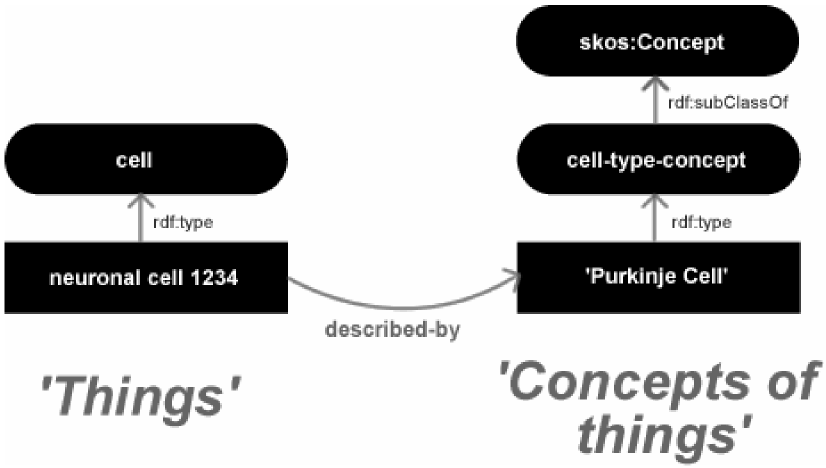


Figure 1: The two ‚worlds' in the bio-zen framework – the world of real things located in a certain space and time and the world of abstract concepts about things. Both worlds can only be connected through the 'described-by' property – otherwise, they are completely separated.

This design pattern has the major advantage that it allows users to put *spatio-temporal-particulars* into a hierarchy without resorting to complicated OWL class hierarchies. Furthermore, it allows users to extend the hierarchy by simply making new instances of *skos:Concept* without needing to define new OWL classes. This keeps the distinction between ontology developers and end-users intact. It gives end-users the ability to create new categorisations while avoiding the potential pitfalls associated with the definition of new OWL classes [Rector, 04].

The concept annotations can be used to express similarity between different *spatio-temporal-particulars* and to 'glue' graphs from different sources (and possibly different ontologies) together (Fig. 2).

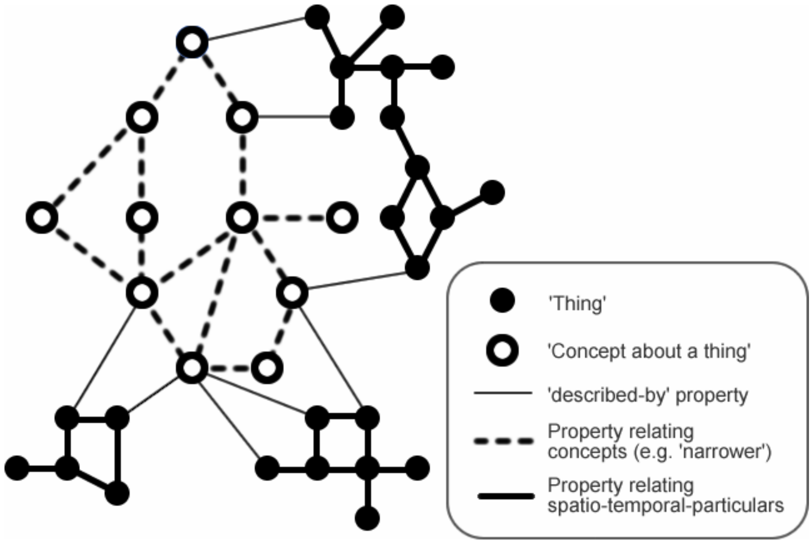


Figure 2: Concepts can act as a kind of ‚glue' between different models. Similarities among different OWL individuals can be inferred from similar concept annotations.

Unnecessary over-specialisation of properties is avoided where possible. Generic properties like 'has part', 'has constituent', 'broader concept than', 'caused by' etc. are the backbone of the whole ontology. The direction of the properties is used consistently, in that they are pointing from the 'larger, containing' thing to the 'smaller, contained' thing (e.g. parts, qualities, participants, features). This allows for the creation of simple query algorithms that can easily capture all parts and details of a resource through iteration.

All of the classes and properties of the ontology are labelled with the *rdfs:label* property. Many other ontologies (e.g. BioPAX) solely rely on the URIs of their resources as human-readable descriptors. Since URIs are essentially not meant to contain meaningful information, this is a grave shortcoming that hampers the development of user-friendly application interfaces.

The ontology allows the seamless integration of mathematical descriptions into existing information. Mathematical formulas can be used to describe the correlation of different qualities (e.g. concentrations of molecules) over a certain timeframe. The mathematical formulas can be expressed in MathML. Physical measurements are strictly represented according to the International System of Units (e.g. kilogram, second, Kelvin) and floating point numbers. This gives *bio-zen* the power to act as a modelling language similar to the popular *Systems Biology Markup Language* (SBML [Hucka, 03]).

Contrary to other ontologies in the field, molecular interactions are modelled as stochastic processes involving populations of molecules, not as single events that involve single molecules. This approach is much closer to biological reality in most occasions, and avoids some grave consistency problems associated with the other approach.

The ontology has basic support for fuzzy logic - like constructs. A property called *'realness'* can be used to assign a measure of uncertainty to every *spatio-temporal-particular* in the ontology. The value of this property should be a floating point number between 0 and 1. This is a metric for how certain it is that the described entity really exists in nature. If a user applies a realness-value of 1 to an entity this essentially means that he or she is sure that the entity exists in nature. Lower values mean that he or she is less certain. A value of 0 implies that there is no known evidence that the entity exists in nature.

Another property called *'interestingness'* can be used to make subjective ratings about how interesting a given resource is. This might seem unusual but is a very important feature, as not all proven and true facts are relevant for scientific discourse. The realisation of such a system for fuzziness values and ratings based on simple datatype properties does produce only a small triple overhead and can be queried with good performance, as compared to other design patterns that make use of RDF reification. As most relations between biological entities in *bio-zen* are reified to start with, the use of RDF reification is not necessary.

4 Ontology specifications and status

The ontology is designed to conform to the OWL DL standard, which guarantees computability and eases the development of tools that work with the ontology. Currently, the core ontology defines around 130 classes, 40 datatype properties and 60 object properties. This includes entities that have been taken over from DOLCE, SKOS, Dublin Core and the evidence code ontology.

There are several extensions available in the form of OWL files. These extensions represent concepts from other biomedical ontologies and taxonomies as concepts based on the SKOS ontology. All of the current extensions are derived from taxonomies that are part of the *Open Biomedical Ontologies* repository [OBO]. Currently, the following extension packages are available:

The *Gene Ontology* [Ashburner, 00] extension (defines 20.000 concepts), the *Medical Subject Headings* [MeSH] extension (23.000 concepts), the *NCBI Taxonomy* [Wheeler, 00] extension (340.000 concepts), the *celltype ontology* [Bard, 2005] extension (800 concepts), the *sequence ontology* [Eilbeck, 05] extension (1000 concepts) and the *INOH Molecule role ontology* [inoh] extension (7.200 concepts).

With all extension packages taken together, *bio-zen* is among the largest structurally coherent ontologies currently available in OWL. Further extension packages will be added in the future.

The ontology, all extensions and a manual can be downloaded from http://neuroscientific.net/index.php?id=download

5 Discussion

The experience of mapping BioPAX to DOLCE showed that such a mapping process can reveal undiscovered problems in the original ontology. Ontology developers in the field of the life sciences should therefore be encouraged to conduct such mappings to test the consistency of their ontologies. It also became apparent that the real value of foundational ontologies lies in the definition of a small set of the most basic concepts (e.g. ‘occurent’, ‘continuant’, ‘process’, ‘part of’). The value of a foundational ontology can be severely decreased by unnecessary details and specialisations.

The mapping process also necessitated minor changes to some of the foundational ontologies in order to make them valid OWL DL. There seem to be no widely accepted agreements on whether such minor modifications to existing ontologies are acceptable or not and how they should be handled. Such an agreement will become necessary when ontology use and therefore also ontology *reuse* on the web will increase.

A project related to *bio-zen* that is worth mentioning is the *Ontology of Biomedical Investigation* [OBI]. The development of OBI is based on the *Basic Formal Ontology* [BFO]. Compared to the BFO, DOLCE has the advantage that it already has an established formalisation in OWL, while the formalisation of the BFO in OWL is still under development. As many of the concepts of the BFO have a counterpart in DOLCE, interoperability between both ontologies should be easily achievable when necessary.

6 Conclusions

The *bio-zen* ontology is among the first functional Semantic Web ontologies for molecular biology that are based on widely accepted foundational ontologies. It is also the first metadata format that attempts to unite taxonomies, ontologies, qualitative data and mathematical modelling in a coherent data structure. It introduces new design patterns, e.g. strategies of avoiding overuse of OWL classes and the ability to easily bind together incompatible ontologies through concept annotations. These design patterns might prove useful for future ontology developments in many different knowledge domains, not only the life sciences.

References

[Ashburner, 00] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet (2000), 25(1):25-29.

[Bard, 2005] J. Bard, S.Y. Rhee, M. Ashburner. An ontology for cell types. Genome Biol 2005; 6(2):R21.

[BFO] http://ontology.buffalo.edu/bfo/

[biopax] http://www.biopax.org/

[dolce] http://www.loa-cnr.it/DOLCE.html

[dublincore] http://dublincore.org/documents/dces/

[Eilbeck, 05] K. Eilbeck, S.E. Lewis, C.J. Mungall, M. Yandell, L. Stein, R. Durbin et al. The Sequence Ontology: a tool for the unification of genome annotations. Genome Biol (2005) 6(5):R44.

[evidence] http://obo.sourceforge.net/cgi-bin/detail.cgi?evidence_code

[Gangemi, 02] A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, L. Schneider, Sweetening ontologies with DOLCE, in: Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02), Springer-Verlag, London, UK, 2002, pp. 166-181

[Hucka, 03] M. Hucka, A. Finney, H.M. Sauro, H. Bolouri, J.C. Doyle, H. Kitano et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics 2003; 19(4):524-531.

[inoh] http://www.inoh.org/download.html

[Luciano, 05] J.S. Luciano. PAX of mind for pathway researchers. Drug Discov Today (2005), 10(13):937-942.

[MeSH] http://www.nlm.nih.gov/mesh/

[OBI] http://obi.sourceforge.net/

[OBO] http://obo.sourceforge.net

[Protégé] http://protege.stanford.edu/

[Rector, 04] A. Rector, N. Drummond, M. Horridge, J. Rogers, H. Knublauch, R. Stevens et al. OWL Pizzas: Practical Experience of Teaching OWL-DL: Common Errors & Common Patterns. 3257 ed. 2004.

[Samwald, 06] M. Samwald, Classes Versus Individuals: Fundamental Design Issues for Ontologies on the Biomedical Semantic Web, in: Proceedings of the STC2006, AKA, Berlin 2006, pp. 335-340

[skos] http://www.w3.org/2004/02/skos/

[Soldatova, 04] L.N. Soldatova, R.D. King. Are the current ontologies in biology good ontologies? Nat Biotechnol (2005), 23(9):1095-1098.

[Wheeler, 00] D.L. Wheeler, C. Chappey, A.E. Lash, D.D. Leipe, T.L. Madden, G.D. Schuler et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res (2000), 28(1):10-14.

[Whetzel, 06] P.L Whetzel, H. Parkinson, H.C. Causton, L. Fan, J Fostel, G. Fragoso et al., The MGED Ontology: a resource for semantics-based description of microarray experiments. Bioinformatics (2006), 22(7):866-873.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Concepts for an Intelligent Information Portal in Pharmaceutical Research

Alex Kohn, François Bry
(Inst. Inf., University of Munich, Germany)

Stefan Klostermann, Alexander Manta
(Roche Diagnostics GmbH, Penzberg, Germany)

Abstract: Once upon a time scientists were experts in their field. They knew not only the "hot questions" but also the scientists involved and the various approaches investigated. More important, they were well informed of novel research results. Gone are these favorable times! Hot issues and active research teams emerge with high pace and being informed within days or even hours might be essential for success. Furthermore, no one can any longer keep an eye on the research publications, patents, and other information that might be relevant for one's research. As a consequence, scientists often feel - and in fact they sometimes are - rather unaware of areas that are of prime importance for their research. High diversity, considerable amounts of information and extremely fast communication are key characteristics of today's research - especially in medical biology. An automatic tracking of technical and scientific information is a way to cope with these aspects of today's research. Such a system is made possible by emerging techniques such as "Semantic Web". This article describes the corner stones of such an "Intelligent Information Portal" currently developed at Roche Diagnostics GmbH for scientists in Pharmaceutical Research. The article stresses the salient aspects of the envisioned system and its focus on personalization/adaptation.
Key Words: adaptation, data integration, knowledge management, life sciences
Category: I.2.0, I.2.1, H.3.0, H.5.0

1 Introduction

Since the beginning of the information age, the data output has steadily increased. Hot research topics emerge with such a high pace that being informed fast might be eminent for success. The mere number of 1700 new publications appearing daily on Medline[1] only (in 2006), shows how dramatic the information overload is that a scientist has to deal with nowadays. Extrapolating to other domains like patent or sequence databases worsens the problem. Obviously, nobody can continuously keep track of all the information being published that might be relevant for one's research. The quantities are too high, changes too fast and data too heterogeneous. The situation observed in the public domain has analogies in many corporate intranet environments, as it is the case for instance at Roche Diagnostics GmbH.

The present problems of information retrieval in a Pharma Research department will be outlined by a scenario. Let's assume that a scientist has identified

[1] After the key Medline indicators: http://www.nlm.nih.gov/bsd/bsd_key.html

an oncology target (ReceptorX) expressed in human breast cancer cells, which is a target for therapeutic proteins (e.g. antibodies). Inhibiting/activating the receptor results in a modification of the signaling pathway. The consequence is apoptosis of breast cancer cells expressing ReceptorX. Prior to clinical studies in humans, model organisms are used to test the biological effects of the potential drug. The identification of suitable models requires a homology search of ReceptorX in different species. A toxicological model is only suitable if the homolog is recognized by the therapeutic antibody in question. Answering the scientist's query, which could be phrased as "homologs ReceptorX", involves several steps:

Fundamentals: The starting point for the query is an in house gene database. Here, fundamental information like gene structure, synonyms, some homologs, literature, etc. is available.

Homologs: The homologs returned by the gene database will usually not suffice and therefore a homology search against other sequence databases (species) is necessary. For that task a couple of web-based sequence analysis tools (e.g. BLAST[2]) are available to the scientist.

Experts: Sequence analysis tools have many parameters and false settings can lead to poor results. Assuring quality requires finding out who is an expert in this area. Furthermore, the expert's contact details need to be acquired by looking into the company's telephone book.

Candidates: Having found all homologs, the scientist would like to know if any of them is currently part of a project. Hence, a query against corporate databases, tracking all projects and lab experiments in protein research has to be performed.

Reflecting on this scenario, several issues arise: S/he must know which resources to use, s/he must know where to find them and s/he can't query all four resources at once. These observations can be generalized, such that the characteristics of the information landscape we face, get apparent. We have many *heterogeneous resources* like databases, applications, websites, web portals (SharePoint, LiveLink, etc.) and file shares. The resources generally have a *low linkage*, hence *low visibility*: users are often not aware of the existence of relevant information resources. However, one of the biggest shortcomings is the *lack of meta information* on these resources. All that makes information retrieval quite difficult [Mühlbacher 08].

Traditional intranet search engines and data integration approaches fail to cope with these issues, hence they fail to fulfill the information needs of a scientist. Intranet search engines perform poor [Xue 03] due to several reasons: the linkage structure is hierarchical thus the ranking-assumption of popularity equaling relevance can't be applied, access barriers (security) are common and

[2] Basic Local Alignment Search Tool: http://www.ncbi.nlm.nih.gov/BLAST/

the "deep intranet" is usually not indexed. In data integration, a traditional approach is data warehousing. Here several databases are integrated by replication into a central store. Besides its merits, this approach has also drawbacks: data is mirrored and thus out of date or can't be mirrored at all, IT costs are high, etc.

How would a perfect information landscape look like? It would consist of a system S which amongst others would have the following features: it has a user-centric interface, it can be easily used for search & navigation, it knows all the resources and their semantics, it has domain knowledge, it learns from its users, it is context-aware, it knows the users' information needs and last but not least it has more knowledge about itself than the user. S is a vision. Although not feasible in the near future, we could at least begin to move into this direction. A first step towards S will be the development of an Intelligent Information Portal (IIP), whose corner stones are described in the following sections.

2 Intelligent Information Portal

In light of S we identified five key cornerstones of an IIP: Resource Integration, Semantics, Ontology Management, Querying and Adaptation.

The resources (databases, applications, websites, web portals, file shares, etc.) need to be integrated and made accessible via a standardized interface. Many concepts for integrating resources exist. An example is the Kleisli [Davidson 97] query system developed at the University of Pennsylvania and marketed today by several companies (see [Section 3]). Another example is the concept of Service-Oriented Architecture (SOA). The paradigm of SOA is to distribute the system logic over several independent services rather than building a monolith system.

Semantic technologies enable the computer systems to understand and reason on data. The semantic web standards RDF (Resource Description Framework) and OWL (Web Ontology Language) provide a framework for adding meta information to both, the resource structure and the data itself. In the public domain many ontologies are already available, especially in biology which is a driving force. Alone on the Open Biomedical Ontology Foundry[3] website a huge collection of publicly available ontologies can be found. Amongst others, the well known GeneOntology[4] is available. To illustrate the doors semantic technologies can open, consider e.g. two databases DB_1 and DB_2 both storing taxonomical data. In DB_1 data is stored in a field called "species" while in DB_2 the field is called "organisms". By applying an ontology mapping to each database schema, the computer will infer that both fields in fact refer to the same concept. Besides the annotation of data and its structure, one could also provide meta information to the resource itself, i.e.: a description of what kind of data is available,

[3] http://obofoundry.org/
[4] http://www.geneontology.org/

where a resource is located, how a resource can be accessed, etc. Given this meta information and the mapping, the resource becomes a self-described modular data resource. Hence, by wrapping semantics around the existing systems, previously unconnected resources become related and a cross-domain query becomes possible. It has to be noticed, that a high-scale usage of ontologies requires an Ontology Management System, which has to address several curation tasks: it must define methods for storage, versioning, up-to-dateness, mediation, etc.

The next two subsections will describe adaptation and an advanced keyword query approach in more detail as these are the key components the user interacts with.

2.1 Adaptive Personalization

Provided that integration succeeds, a huge amount of information will be available at the researcher's desk. Obviously the risk of information overload remains, leading to poor precision and recall when doing inquiries. A promising technique to mitigate these issues is adaptive behavior. An *adaptive system* is a system which adapts its communication patterns to the current actor. *Recommender systems* are a specific implementation of this technique. They guide the user in a personalized way through the complex information landscape with its large number of options [Burke 02]. *Personalization* in this context means to adapt the communication pattern to the user's characteristics [Baldoni 05]. An essential part of adaptive systems are therefore *user profiles* which store user preferences in attribute-value pairs. The data stored in a user profile can contain amongst others [Baldoni 05]: device information, preferred settings, goal, current task, information needs, required information depth, time constraints, previously regarded information, previously gained knowledge and much more. The maintenance of a user profile, i.e. the acquisition of data, the update and inconsistency checking is accomplished by a *user model*.

A detailed description of recommendation techniques is given in [Burke 02] who distinguishes between five basic approaches, namely Collaborative-based, Content-based, Demographic-based, Utility-based and Knowledge-based recommendation. Relying on these basic approaches, several hybrid-based systems have been proposed. In order to give an idea of how recommender systems work, the collaborative filtering (CF) and demographic-based approaches are described briefly.

User or item-based CF is the most mature technique. It is used today by many applications especially in e-commerce (e.g. Amazon or E-Bay). The basic idea of user-based CF is to recommend previously unknown items to a user based on the items preferences in his neighborhood. Let U be the set of users, I the set of items and r_u a rating vector with $u \in U$, mapping items to a value of unity if it is considered relevant by u and zero otherwise. Given users

$u_1, u_2, u_3 \in U$, items $A, B, C \in I$, the ratings $r_{u_1} = \{(A, 1), (B, 1)\}$, $r_{u_2} = \{(C, 1)\}$ and $r_{u_3} = \{(A, 1), (B, 1), (C, 1)\}$. A significant correlation between u_1 and u_3 can be detected since both have rated item A and B positive. Thus, u_3 is in the neighborhood of u_1 and the unknown item C can be recommended to u_1. The idea of item-based CF is very similar to user-based CF. In item-based CF the perspective is opposite, meaning that highly correlated items are found according to the preferred items. Commonly used techniques in CF are Pearson correlation, vector similarity, clustering, Bayesian networks, etc.

Demographic recommenders classify users into classes based on their personal attributes. Therefore, it is eminent that users provide explicit personal information about their preferences. The information for categorization can be gathered with surveys or by the usage of machine learning algorithms which analyze user profiles. Given the demographic data, a system can identify demographically similar users to extrapolate for instance from their ratings.

Personalization relies on user profiles so that privacy issues arise. The following policies describe options of how to abate them. Most importantly, the works council and the users have to be elucidated about the stored data. Keeping the profiles transparent is also crucial. Users should have the possibility to view their profiles and eventually delete data that they don't want to be stored. A third approach is anonymization. This could be achieved by applying personalization on the group level instead of the individual. Here, roles, tasks, projects, etc. are pre-defined and the user can select between them in a multiple-choice manner. Depending on the groups a user has subscribed to, the information portal is adjusted. Given that a set of users belong to a common group, their actions will contribute to changes of the group profile. Thus in group recommenders, the system tries to fulfill the needs of all group members by maximizing the average member-satisfaction. While this approach guarantees anonymity, it has several drawbacks: (a) the initial choice of the proper groups is difficult, (b) group members drifting away with their interests will gain a poor personalization and (c) personalization can't be as accurate as applied on an individual level.

The pilot will apply only group-based recommendation as default. However, people will have the freedom to decide if they want an individual personalization activated or not. In both cases, the personalization will be transparent. Hence, inference rules will be shown, profile data is viewable and in case of individual personalization also erasable.

2.2 Advanced Keyword Query

A search in IIP will be keyword-based with an easy to use structural and/or semantic prefix extension, such that scientists are able to specify what they are interested in. On the one hand, the more detailed a query language is, the more accurate the delivered answers are. On the other hand, formal query languages have to be learned and are therefore not readily used. Simple keyword-based query interfaces (e.g. Google or Yahoo) have without a doubt a reason for their success: little is required to enter a few keywords. However, keyword search, how simple and therefore appealing it might be, is unspecific. Looking e.g. for ReceptorX, one cannot distinguish between articles that accidentally mention the concept somewhere, and those that have the concept in their *main title*, *section titles*, etc. Thus, refining keyword query with some structural elements while keeping the appealing simplicity of keyword querying has been proposed by [Cohen 03].

Structural information can be added to a query by providing the user with a small number of structural concepts. This can be done with a simple textual query interface, e.g. "mt:ReceptorX" expressing that ReceptorX should occur in the *main title* or "t:ReceptorX" meaning that ReceptorX should occur in a *title* at any depth. Dependencies between components can be expressed as well, e.g. "s:ReceptorX > a:John Q. Public" expressing that a *section* contains ReceptorX authored by John Q. Public is sought for.

Accordingly, we propose semantic prefixes to be added to keyword-based query. Instead of using keyword prefixes expressing structure, it is sufficient to select a few keyword prefixes expressing semantics. For example, "homologs: ReceptorX" telling to search for homologs of ReceptorX. Because the term homologs is a concept of the MeSH[5] thesaurus, the system infers that in fact both, orthologs and paralogs might be important to the user.

The central issue in using structure and/or semantics to refine keyword search, is the choice of the relevant concepts. Too many make the approach inherently complicated, similar to a full-fledged structure-oriented query language like XPath or XQuery. Too little or the wrong choice of prefixes turns the approach useless. And here adaptation comes into play. Tracking queries of an individual or a group of users and hence their interests, enables the suggestion of exactly the relevant prefix refinements. Given two scientists A, B working in the same pharmacology department on toxicological models. Hence, their demographic classes are correlated. Let's assume that in the search history of B the entry "homologs:ReceptorX" exists, i.e. s/he searched once using the prefix extension *homologs*. As the system has semantic annotations, it knows that ReceptorX is an instance of the *Protein* concept. Now let's assume scientist A, having an empty search history, is using the traditional keyword search to query ReceptorZ. Again, the semantics provide a way to detect ReceptorZ as an instance of the *Protein* concept. Since the adaptive system knows that A and B are correlated, it searches for items unknown to user A. Both scientists have searched for Proteins but only B has refined the query with the prefix semantics. Thus the system suggests A to search for "homologs:ReceptorZ". In conclusion,

[5] Medical Subject Headings: http://www.nlm.nih.gov/mesh/

a recommender system could dynamically select and suggest structure and/or semantic keyword prefixes to a scientist.

3 Related work

Data integration techniques have always been of great interest for industry and research. Interesting integration approaches in the biological domain are e.g. Kleisli [Davidson 97], Tambis [Baker 98] and BioMediator [Donelson 04].

Kleisli is based on the Collection Programming Language (CPL). CPL uses sets, bags, lists and records to describe any data. In addition, it offers functions for manipulating data. Even though this approach is very powerful, it has the handicap of not offering a mediated schema over the data sources, i.e. semantics are missing. Therefore the task of choosing the appropriate data source remains at the user.

Tambis is the first semantic approach in the area of biology for data integration. It is based on the Tambis ontology (TaO) which models in parts the molecular biology domain. The TaO concepts are used to model the database sources as well as to express source independent declarative queries. Tambis uses Kleisli's CPL for accessing the data sources. Drawbacks are the limited number of databases which can be queried and the static TaO. The fundamental Tambis ontology can't be customized, making it difficult for users with different schemata to use the same system [Donelson 04].

The BioMediator approach uses an annotated mediated schema to model data sources and their relationships. A source knowledge base contains the mediated schema (which describes entities, attributes and relationships of interest to a particular group of researchers), a list of all data sources and the mapping rules. In contrast to Tambis, the mediated schema can be customized by editing with the Protégé[6] Ontology Editor. Schema adaptation requires modeling knowledge thus remaining an expert task.

The described approaches differ in degree of user guidance. While Kleisli requires the user's knowledge to decide which database to choose for querying, BioMediator makes the choice itself, i.e. the system knows more about itself and its data reservoirs. Therefore the user is unburdened in decision making. We propose to further reduce the discrepancy between the knowledge a system has about itself and a user needs to know about a system. This could be achieved by the combination of resource integration, semantic technologies, adaptive systems and an advanced query engine. The problems depicted in the introductory scenario might be solved, thus improving dramatically a scientist's information gain. If the adaptive system knows the scientists information needs it can tailor

[6] http://protege.stanford.edu/

navigation pathways to their specific requests and help by suggesting relevant extensions for keyword-based query. Therefore precision & recall of search results can be improved. In contrast to systems in the public domain, our approach addresses a closed domain, namely the corporate intranet. Here, we have the advantage, that we could use a priori knowledge about a user's roles, tasks, educational background, current department, involved projects, etc.

4 Summary and Outlook

Information overload has become a severe problem in the public domain and in companies. Traditional search and integration approaches perform poor in answering a scientist's queries. New techniques such as semantic technology offer means to apply meta information to data and resources, thus enabling computers to reason on the data. A user-centric interface to the information is still missing even though semantics have been added. Adaptation can close the gap between a user's interface and the underlying data reservoirs by customizing the communication patterns.

The extension of traditional keyword-based query with structural and/or semantic prefixes offers a simple interface for building more complex queries. The proposed combination of prefix extensions with adaptation could emerge as a useful concept for improving information access. As this idea is still in its infancy, it is the task of further research to exploit its full potential.

References

[Baker 98] P. Baker et al.: "Tambis: Transparent access to multiple bioinformatics information sources"; Proc. 6th ISMB, AAAI Press, Menlo Park, 25–34, 1998.

[Baldoni 05] M. Baldoni et al.: "Personalization for the semantic web"; Proc. Summer School Reasoning Web, LNCS, 3564, 173–212, 2005.

[Burke 02] R. Burke: "Hybrid recommender systems: Survey and experiments"; User Modeling and User-Adapted Interaction, 12(4), 331–370, 2002.

[Cohen 03] S. Cohen et al.: "Xsearch: A semantic search engine for xml"; VLDB, 2003.

[Davidson 97] S. Davidson et al.: "Biokleisli: A digital library for biomedical researchers"; Int. J. on Digital Libraries, 1(1):36–53, 1997.

[Donelson 04] L. Donelson et al.: "The biomediator system as a data integration tool to answer diverse biologic queries"; Medinfo, 72, 768–772, 2004.

[Mühlbacher 08] S. Mühlbacher: University of Regensburg / Roche Diagnostics GmbH Penzberg, Dissertation, 2008, forthcoming

[Xue 03] G. Xue et al.: "Implicit link analysis for small web search"; Proc. 26th ACM SIGIR, 56-63, 2003.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Clinical Ontologies Interfacing the Real World

Stefan Schulz
(Department of Medical Informatics, University Medical Center Freiburg, Germany
stschulz@uni-freiburg.de)

Holger Stenzhorn
(Department of Medical Informatics, University Medical Center Freiburg, Germany
and
Institute for Formal Ontology and Medical Information Science, Saarbrücken, Germany
holger.stenzhorn@uniklinik-freiburg.de)

Martin Boeker
(Department of Medical Informatics, University Medical Center Freiburg, Germany
martin.boeker@uniklinik-freiburg.de)

Rüdiger Klar
(Department of Medical Informatics, University Medical Center Freiburg, Germany
klar@mi.ukl.uni-freiburg.de)

Barry Smith
(Department of Philosophy, University of Buffalo, State University of New York, USA
and
Institute for Formal Ontology and Medical Information Science, Saarbrücken, Germany
phismith@buffalo.edu)

Abstract: The desideratum of semantic interoperability has been intensively discussed in medical informatics circles in recent years. Originally it was assumed by many that this issue could be addressed simply by insisting on the application of shared clinical terminologies. More recently however the use of the term 'ontology' has been steadily growing. We here address the issue of the degree to which the use of ontologies represents any real advance on the road to semantic interoperability.

Keywords: Clinical Ontologies, Knowledge Representation
Categories: I.2.4, SD J.3

1 Introduction

The desideratum of semantic interoperability has been intensively discussed in medical informatics circles over the past decade [Rossi-Mori 98, Ingenerf 01, Garde, 07]. Consider for example the evolution of the Unified Medical Language System

(UMLS)[1], of the Open Biomedical Ontologies (OBO)[2], of the HL7 Common Document Architecture (CDA) [Dolin 06], of openEHR [Kalra 05], or of SNOMED CT[3].

Originally the issue of semantic interoperability was supposed to be addressed mainly by applying clinical terminologies. More recently, however, we have seen steady growth in usage of the term ‘ontology’. The issue addressed here is: does this constitute any real advance or advantage? There is indeed good reason to cast at least some doubt on the claims made on ontology’s behalf: Too many recent publications, calls for research proposals and project descriptions have embodied what are in our view (and have sometimes already proven themselves to be) insupportable expectations. It is thus understandable that some have been tempted to see in ontology just one more new and flashy buzzword. To begin, we must first ask where the proper difference lies between terminology on one side and ontology on the other. Since neither one of those two terms has been unambiguously defined so far, we will adhere in the context of this paper to the following definitions:

- Terminology: A set of terms representing the system of concepts of a particular subject field. [ISO 00]
- Ontology: The study of what there is. Formal ontologies are theories that attempt to give precise formulations of the types of entities in reality, of their properties and of the relations between them. [Quine 1948]

2 Delimiting The Concept of Ontology

2.1 Contrasting Ontologies with Terminologies

We summarize our position on clinical ontologies and terminologies as follows:

Terminologies are term centered. They relate the senses or meanings of linguistic entities. Classes of (quasi-)synonymous terms are commonly referred to as ‘concepts’. In many terminology systems (often called thesauri or semantic lexicons), concepts are furthermore related by informal semantic relations which are often closely related to natural language predicates[4]. In medical informatics, this language-centered view characterizes the UMLS legacy. In spite of its well-known shortcomings, the UMLS can be seen as a robust and highly useful platform for the retrieval of terms belonging to of heterogeneous, context-dependent, informal terminology systems.

Ontologies are intended to describe a portion of reality that exists independently of human language. Their constituent nodes are (entity) *types*, not concepts. Types (often also referred to as ‘categories’, ‘kinds’ or ‘universals’) are well suited to hierarchically order the particular entities (patients, lesions, surgical procedures) which exist on the side of reality. The existence of certain types and the basic structure of ontological principles are subject to major philosophical disputes. However, at any given stage in the development of science, there is a consensus core of scientific understanding of reality, and it is this (on our view) which should serve as starting point

---

[1] http://www.nlm.nih.gov/research/umls
[2] http://obofoundry.org
[3] http://www.snomed.org/snomedct/documents/january_2007_release.pdf
[4] In the context of this paper we do not demand the existence of inter-concept relations as a necessary criterion for terminologies.

in developing science-based ontologies. Examples of statements belonging to this consensus core are: that humans are vertebrates, that cells contain membranes, that adenosine diphosphate is phosphorylated in mitochondria, that the retina contains photosensors. However, types of the sort with which we are concerned in medicine are elucidated not only by such observation-based descriptions of nature, but also often result from a prescriptive, definitory process: appendectomy (type) is *defined* as the "*surgical removal of an appendix*", and hepatitis (type) is *defined* as an "*inflammation of liver tissue*". (Such axioms may be ignored but cannot be falsified.) Types in an ontology apply to classes of entities in the real world (also called individuals or instances) since classes indeed include collections of the entities instantiating a given type. The main construction tenet for ontologies is the taxonomic principle: a type *S* is a subtype of a type *T* if and only if all instances of *S* are also instances of *T*.

2.2 Ontology Constraints

Based on understanding ontologies as representations of types of entities and of relations between them, we can rule out some common misconceptions that often obscure the sharp difference between ontologies, terminologies and other artifacts such as representations of contextual knowledge and information models for data acquisition.

As a fundamental principle, *all* properties of a given type in any ontology are true for *all* instances of this type. Thus all instances of appendectomy are performed on some instance of appendix; all instances of water molecules contain oxygen and hydrogen. This restricts the ability of an ontology to express seemingly obvious assertions such as "*hands have thumbs*" or "*aspirin alleviates headache*", because there are hands without thumbs and not all aspirin tablets are used to alleviate any headache.

A further restriction is that probabilistic assertions, which are of tremendous importance for everyday clinical reasoning, cannot be expressed in an ontology in a simple way. For example, if a prevalence of 1% is ascribed to lung cancer then this is not a property inhering in any instance of lung cancer. It is rather a factual statement about some given population with respect to the occurrence of this disease. However, the common consensus in science is in many areas based on probabilistic theories which describe results in terms of probabilistic states, processes and events. So is the assessment of risks (of signs, symptoms, and therapies for specific diseases) commonplace in medical practice. (e.g., arterial hypertension is considered a risk for stroke). Unsatisfactorily, the related entity types and relations cannot be straightforwardly represented in formal ontologies following the principles described above. A possible solution is to introduce probabilistic dispositions [Jansen, 07] into the ontology, i.e., dispositions to do something (under certain circumstances) with a certain probability. Such dispositions are related to events by the relation of *realization*. They are special kinds of dependent entities, in that they need not be realized in order to exist. E.g., "risk for stroke" could be represented in such a way.

These fundamental constraints are corollaries of the fact that all assertions of relations between types in ontologies should be of the basic form of universal statements: "*for all instances of type T there is some...*". We could, of course, consider types and instances as two different ranges for our quantifiers. Then, however, we would have to accept some higher-order logic, which would cause problems for machine reasoning since such logics are known not to be computable in all circumstances using current algorithms. By contrast, languages from the family of Description Logics [Baader et al. 03] are computable and therefore frequently used in the ontology development.

2.3 Epistemological Classification Criteria

Classes are the basic building blocks for clinical classification systems as the ICD [ICD 07], which, for the time being, provide the most significant support of semantic interoperation of clinical data. It has repeatedly been observed that medical classification systems (even claiming to classify entities in reality) are distinguished from ontologies because of their use of "*un-ontological*" classification criteria (i.e., such as to represent the knowledge independent reality of the entities) but rather "*epistemological*" (i.e., to represent the knowledge one has about these entities) [Bodenreider et al. 04]. Thus the current ICD makes a classificatory distinction between cases of tuberculosis diagnosed by bacterial culture and those diagnosed by histology. But, a particular disease is not different in nature only because of a different diagnostic method.

Epistemological issues are nevertheless crucial for medical documentation. Diagnostic statements tend to be error-prone with vital decisions often based upon brittle evidence. Necessary or desirable information may simply be missing. So a place must exist to encode the information one actually has in the practically available form. An ontology is, however, not the right place for this. Classes such as "*unspecified tumor stage*" or "*infection of unknown origin*" do not stand for more specific subclasses. They just manifest lack of adequate knowledge, mixing up "*what is*" with "*what we know*". Such knowledge is important in the clinical scenario but requires additional means to represent contextual knowledge in encoding specific clinical instances.

3 Clinical Ontologies in Practice

All of the questions addressed above arise, to different degrees, in cases where formal domain ontologies are expected – in the framework of clinical research projects but also in routine documentation – to improve data acquisition, standardization, interoperation, as well as data analysis. We report on experiences within the projects ACGT (Advancing Clinico-Genomic Trials on Cancer)[5] and @neurist (Integrated Biomedical Informatics for the Management of Cerebral Aneurysms)[6], in which customized ontologies are required and are being currently developed.

Both ACGT and @neurIST aim at setting up integrated information technology infrastructures by implementing common software platforms to improve disease management through a more efficient processing and presentation of knowledge and data. ACGT focuses on nephroblastoma and breast cancer, basing its work on a master ontology for cancer supporting the facility to create clinical report forms automatically to support clinical trial research in cancer genomics. @neurIST is concerned with acquiring and estimating the risk of intracranial aneurysms and subarachnoid hemorrhage based on multimodal data. The goal of the developed platforms is the integration of data from various sources and disciplines within the projects (e.g., clinical studies, genomic research and patient management). These data are highly

[5] http://www.eu-acgt.org
[6] http://www.aneurist.org

fragmented and heterogeneous in regard to format, scale and their particular content reflecting the projects' specific sub-domains.

Consequently, it is a big challenge to design ontologies that acknowledge this broad scope and are capable of integrating available data. Semantic interoperability here means that all data collected for each individual patient, for each experiment, or of each literature abstract considered relevant for the domain, should point to nodes in a domain ontology. One issue that particularly complicates this task is the multitude of entity types and the different scales of spatial and temporal granularity (i.e., medical, biomolecular or epidemiological entities from single cell division to human life). Further, the ontologies have to integrate various levels of description in the available data (e.g., literature, clinical databases, imaging databases and terminologies).

4 Interfaces of Ontologies

Lessons learned from ACGT and @neurIST have shown that the shortcomings and problems described in section 2 can be alleviated by clearly defining the interfaces between the ontologies and other artifacts in the semantic interoperation environment.

4.1 The Interface between Clinical Ontologies and Terminologies

Ontologies, in a strict sense, are domain descriptions that are independent of human language, so they need not incorporate any lexical or term information at all. The fact that – for practical reasons – they commonly employ human-readable names is not a contradiction to this claim. These names may, but do not have to, coincide with actual domain terms. They do so because ontologies need to be maintained by humans, and are often used by humans in expressing their results without any intervention of a machine. Terms and descriptions in ontologies should be precise, unambiguous and self-explanatory (which is often not the case with typically used domain terms).

The interface between domain ontologies on the one hand and domain term lists on the other hand is characterized by a many-to-many relationship: Several terms may be connected to one ontology class due to the phenomena of synonymy and cross-language translation, and polysemy has to be accounted for to link an ambiguous term to more than one ontology class. For instance, the natural language terms "*mamma carcinoma*", "*breast cancer*" and "*Brustkrebs*" are linked to the same node in the ACGT ontology, whereas "*ulcer*" points to two different ontology nodes, *viz.* first, the process of ulceration and then, the pathological structure (the *result* of this process).

4.2 The Interface between Clinical Ontologies and Upper Ontologies

In our application contexts the project-specific ontologies are rooted in upper ontologies. According to the Standard Upper Ontology Working Group[7] upper ontologies provide generic categories or types suited to address a broad range of domain areas at a high level in a way which can support integration of the underlying data. These upper-level types provide a highly general structure which in turn helps the consistent representation of the entity types in the associated domains. Following this, the use of

[7] http://suo.ieee.org

an upper ontology is intended to improve the development of the actual domain ontologies by providing a consistent and sound top-level framework. Whereas the ACGT ontology uses the Basic Formal Ontology (BFO) [Grenon et al. 04] as its upper level, the @neurIST ontology employs the Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) [Gangemi et al. 02]. For a comparison of the two upper level ontologies see [Mascardi et al. 07].

4.3 The Interface between Clinical Ontologies and Non-Ontological Knowledge

As we said before, the domain representations generated by large research projects must extend what can be expressed by ontologies. This extension is here referred to as "*non-ontological knowledge*". Typical examples are assertions such as "*A treats B*" and "*C is a risk for D*". This knowledge is still at the terminological level (rather than at the level of instance data), but it is not knowledge which should properly be included in an ontology because it does not express what holds universally of the given types. There are, in principle, two different ways to represent such knowledge:

- The first follows the representational scheme of the UMLS Metathesaurus. Non-ontological knowledge is represented in a thesaurus-like terminology, often based on *concept–relation–concept* triplets. Thus it does not support logic-based reasoning about *classes*, as in logic-based representations, but is anyhow available for informal reasoning about *concepts*, e.g. related concept search by graph transversal.
- The second solution has been applied in the @neurIST ontology and consists in a parallel system of non-ontological, reified classes (e.g. the class "*suspected risk factor for aneurism rupture*") thus inserting epistemological categories into the taxonomy. These categories are irrelevant for the correct ontological entity description but needed for the specific retrieval requirements of users in the project. From the representation perspective, however, this difference is ignored. So "*hypertension*" is a subclass of the above class just as it is a subtype of "*cardiovascular disorder*".

4.4 The Interface between Clinical Ontologies and Information Models (and the World)

Information models are templates for the acquisition of clinical data which enable semantic interoperability in the scope of the given information model but not between different information models. They can be based, e.g., on openEHR archetypes [Beale et al. 01, Kalra et al. 05], and are built in such a way as to involve reference to ontologies but they are not by themselves ontologies. In an information model we encode what we know about concrete instances in a certain situation and under certain circumstances. Besides offering a template for the facts to be reported, the models may further include the conditions of measurement, the certainty of an assertion, or other contextual factors. This is why information models are necessary and the simple instantiation of ontology classes is not sufficient in clinical documentation. Ontologies provide the types for the particular instances to be recorded in an information model. This relation has recently received increased attention in the context of openEHR Archetypes, HL7 Version 3 Clinical Document Architecture, and SNOMED CT [Rector et al. 06], and has been further discussed by including experiences from large scale implementation attempts such as the UK Connecting for Health project.

4.5 The Interface between Clinical Ontologies and the Ontology Engineer

The actual interface between clinical ontologies and their developers is an ontology editing environment that ideally offers supports ontology development and maintenance by a graphical user interface and releases the ontology engineer from the need to access and edit the actual source code, e.g. the Web Ontology Language (OWL)[8]. Most editors allow users to further describe ontology nodes both with textual information and logical definitions. The latter can be used by terminological reasoners to enable automated checking of the structural and (to some extent) semantic correctness of the ontology. Both ACGT and @neurIST use open-source software, *viz.* Protégé[9] (currently the most widely-used ontology editor) together with the reasoner Pellet[10].

4.6 The Interface between Clinical Ontologies and the Application Builder

Application builders need a way to programmatically access the content and structure of an ontology in order to create software systems that refer to this ontology, as in the ACGT and @neurIST projects. Therefore generic application programming interfaces (API) have been developed that can be used by application builders for example to take a given entity type from the ontology and link it with a multilingual terminology system. Another example is the development of an easy-to-use retrieval interface for the ontology content, since it turned out that ontology editors such as Protégé are too complex and therefore less suited for application builders.

5 Conclusion

Ontologies are important informatics resources for large multicentric clinical research projects because they foster semantic interoperability. They offer a stable, language-independent vocabulary that helps standardize and explain the meaning of domain terms. However, ontologies are often mixed up with terminologies, thesauri, and representations of contingent or probabilistic domain knowledge, as well as with database-centered information models which serve recording of instance data. This often creates exaggerated expectations on the part of the users of an ontology. We argue in the above that in order to minimize such expectations we should clearly delimit the scope of ontologies from that of other knowledge and information resources. To this end we defend the introduction of clearly defined interfaces between ontologies and other supporting artifacts. Only after their success being proven in clinical research projects, formal ontologies can be expected to be seriously used to enable semantic interoperability in the clinical routine, as well.

Acknowledgements

We thank our colleagues and ontology co-developers: Susanne Hanser (@neurIST) and Mathias Brochhausen and Cristian Cocos (ACGT). Work on this paper has been carried out cooperatively within the @neurIST and the ACGT integrated projects funded by the European Commission (IST-027703, IST-026996).

---

[8] http://www.w3.org/TR/owl-features
[9] http://protege.stanford.edu
[10] http://pellet.owldl.com

References

[Baader et al. 03] Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P.: "The Description Logic Handbook: Theory, Implementation and Applications"; Cambridge University Press, Cambridge, United Kingdom (2003)

[Beale et al. 01] Beale, T., Goodchild, A., Heard, S.: "EHR Design Principles"; London, United Kingdom (2001)

[Bodenreider et al. 04] Bodenreider, O., Smith, B., Burgun, A.: "The Ontology-Epistemology Divide: A Case Study in Medical Terminology, Proc. FOIS-2006, Torino, Italy (2004)

[Dolin et al. 06] Dolin, R., Alschuler, L., Boyer, S., Beebe, C. Behlen, F., Biron, P., Shabo Shvo, A.: "HL7 Clinical Document Architecture, Release 2"; J Am Med Inform Assoc, 13, 1 (2006) 30-9.

[Gangemi et al. 02] Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., Schneider, L.: "Sweetening Ontologies with DOLCE"; Proc. EKAW-2002, Siguenza, Spain (2002)

[Garde et al. 07] Garde, S., Knaup, P., Hovenga, E., Herd, S.: "Towards Semantic Interoperability for Electronic Health Records"; Method Inf Med, 46, 3 (2007) 332-343

[Grenon et al. 04] Grenon, P., Smith, B., Goldberg, L.: "Biodynamic Ontology: Applying BFO in the Biomedical Domain"; in Pisanelli, D. (ed.): "Ontologies in Medicine", IOS Press, Amsterdam, Netherlands (2004).

[Ingenerf et al., 01] Ingenerf, J., Reiner, J., Seik, B.: "Standardized Terminological Services Enabling Semantic Interoperability between Distributed and Heterogeneous Systems"; Int J Med Inform, 64, 2-3 (2001) 223-40.

[ISO 00] International Organization for Standardization: "ISO 1087-1: Terminology work – Vocabulary – Part 1: Theory and applications", Geneva, Switzerland (2000)

[Jansen 07] Jansen, L. On Ascribing Dispositions. In Bruno Gnassounou and Max Kistler, editors, Dispositions and Causal Powers, pages 161–177. Ashgate, Aldershot, 2007.

[Kalra et al. 05] Kalra, D., Beale, T., Heard, S.: "The openEHR Foundation"; Stud Health Technol Inform, 115 (2005) 153-73

[Mascardi et al. 07] Mascardi, V., Cordì, V., Rosso, P.: "A Comparison of Upper Ontologies"; Technical Report DISI-TR-06-2, Genova, Italy (2007)

[Quine 1948] On What There Is. Review of Metaphysics. (1948)

[Rector et al. 99] Rector, A., Zanstra, P., Solomon, W., Rogers, J., Baud, R.: "Reconciling Users Needs and Formal Requirements: Issues in Developing Re-Usable Ontology for Medicine"; IEEE Transactions on Information Technol in BioMedicine, 2, 4 (1999) 229-242

[Rector et al. 06] Rector, A., Qamar, R., Marley, T.: "Binding Ontologies and Coding Systems to Electronic Health Records and Messages"; Proc. KR-MED-2006, Baltimore, USA (2006)

[Rossi-Mori et al. 98] Rossi-Mori, A., Consorti, F.: "Exploiting the Terminological Approach from CEN/TC251 and GALEN to Support Semantic Interoperability of Healthcare Record Systems"; Int J Med Inform, 48, 1-3 (1998) 111-124

[Uschold and King 96] Uschold, M., King, M.: "Ontologies: Principles, Methods, and Applications"; Knowledge Eng. Rev., 11, 2, (1996) 93-155

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Utilizing Text Mining Techniques to Analyze Medical Diagnoses

Felix Mödritscher
(Institute for Information Systems and Computer Media,
Graz University of Technology, fmoedrit@iicm.edu)

Roland Tatzl
(IT and IT Marketing, Campus 02,
University of Applied Sciences Graz, roland.tatzl@campus02.at)

Regina Geierhofer
(Research Unit HCI4MED, Institute of Medical Informatics, Statistics & Documentation,
Medical University Graz, Austria, regina.geierhofer@meduni-graz.at)

Andreas Holzinger
(Research Unit HCI4MED, Institute of Medical Informatics, Statistics & Documentation,
Medical University Graz, Austria, andreas.holzinger@meduni-graz.at)

Abstract: Due to the increasing amount of medical patient data collected in hospitals, technology-based methods are of increasing interest for processing and analyzing such materials. Therefore, computer supported techniques have to be evaluated by means of their efficiency for this application area. In this paper, we introduce an approach for analyzing expert comments on magnetic resonance images (MRI) diagnoses by applying a text mining method in order to scan them for regional correlations. Consequently, we propose a calculation of significant co-occurrences of diseases and defined regions of the human body, in order to identify possible risks for health, and we present a special tool, which we have implemented in order to test this approach.

Keywords: Information Retrieval, Text Mining, Performance, Medical Documentation

Categories: H.3.1, H.3.3, I.2.7, I.7, J.3

1 Introduction and Motivation

The application of sophisticated medical information systems amasses large amounts of medical documents, which must be reviewed, observed and analyzed by human experts [Holzinger et al., 2007]. All essential documents of the patient records contain at least a certain portion of data which has been entered in free-text fields. Although text can be *created* simple, the support of automatic analysis is extremely difficult [Gregory et al., 1995], [Holzinger et al., 2000], [Lovis et al., 2000]. In order to support the end users during their daily work, both technological performance and cognitive performance must be considered [Holzinger, 2002] and the integration of usability methods on systemic level is essential [Holzinger, 2005]. Against this background, the Institute for Medical Informatics, Statistics and Documentation (IMI) at the Medical University Graz (www.meduni-graz.at/imi) has been carrying out a variety of projects to analyze and process medical documents by applying computer-based techniques and to present the information human centered. One of these approaches is described in this paper, which is structured as follows: Section 2 outlines some theoretical background of text mining in the field of medical informatics as well as some related projects in this area. In section 3 we present our approach and the web-based tool. Finally, we discuss the utility of our methodology on the basis of specific results, experiences, problematic aspects and opportunities. Finally, the paper is concluded by providing an outlook on further research activities.

2 Text Mining for Medical Documents

Contrary to structured information, textual information is characterized by its inherently *unstructured and fuzzy nature,* being language and domain dependent, as well as consisting of sentences and sub-sentences. Basically, text mining approaches apply statistical or pattern based algorithms in order to extract significant key-word associations or to mine for prototypical documents (e.g. for parts-of-speech tagging or term extraction). Text mining is considered a sub-specialty of Knowledge Discovery from Data (KDD) and has been headed strongly in the direction of natural language processing (NLP) during the last decade.

In accordance with [Granitzer, 06], the following stages of the text mining process can be identified: (1) Pre-processing is necessary to prepare texts, e.g. by removing layout information, or to improve the text quality by utilizing methods like stemming. (2) Information extraction comprises the stage which transforms unstructured text entities into structured elements such as database entries. (3) By applying statistical methods, features of an information space can be extracted from the structured text. Hereby, methods such as frequency analyses (e.g. of words), collocations or co-occurrences are utilized. (4) As a result, each text object is described with several features, which, for instance, spans an *n-dimensional vector* space with *n* equals to the *number of features for a text object.* This feature space can be utilized for further operations, e.g. comparing texts, visualizing relations in the text corpus, calculating similarities or rankings, etc.

In the last years, mining in medical digital libraries has come up with new findings and hypotheses. [Srinivasan, 04] reports on the development of text mining methods on the basis of medical subject headings (MeSH). These algorithms generate hypotheses by identifying potentially interesting terms related to the specific input. Additionally, new protein associations have been found by clustering learning and vector classification techniques [Fu, 03]. Another approach utilizes a rule-based parser and co-occurrence for extracting and combining relations [Leroy, 03]. Further, a new way to use thalidomide has been discovered by mapping phrases to concepts of the Unified Medical Language System [Weeber, 01]. Finally, co-occurrences are useful to build up gene networks [Jenssen, 01] and to discover gene interactions [Stephens, 01]. Derived from these experiences, three kinds of application areas for text mining can be outlined in the field of medical documentation:

Firstly, such methods are applied in order to build up an infrastructure or models for biomedicine, i.e. by finding patterns or relations in texts and generating a feature space. Inspecting the projects on the basis of the well-known text mining framework GATE (http://gate.ac.uk/projects.html), MultiFlora or myGRID can be identified as examples for this approach. Secondly, text mining is used to observe and retrieve documents with innovative ideas in the scope of a restricted domain (cf. projects like BioRAT or InESBi). Thirdly, text mining techniques are utilized to extract information or features out of a medical text corpus, which is generated as product of clinical documentation for further operations such as information retrieval. The MedDictate software comprises one solution in this scope. Although this last category of applications areas overlaps partially with the first one, there are only a few reports about the mining of medical diagnoses. In addition to these experiences, we want to report on our approach towards aiming at the detection of diseases in MRT diagnoses. In the following two sections, this project and its outcome are described in detail.

3 The Solution Approach

The success of text mining methods in medical research was the origin of our idea to apply statistical techniques in order to find hints for possible locations of diseases in MRT diagnoses. Therefore, we aimed at the topological proximity between anatomic structures and pathologic expressions and implemented a tool which calculates the significant co-occurrences of anatomic and pathologic terms within the diagnoses.

3.1 Basic Algorithm and Methodology

This calculation of significant co-occurrences is based on the Poisson distribution. In accordance with [Heyer, 06], the original formula can be simplified for two different ranges of the input parameters (see also figure 1). Hereby, *a* stands for the number of sentences containing term *A*, *b* for the number of sentences containing term *B*, *n* for the number of all sentences and *k* for the number of sentences containing both terms.

Be	$\lambda = \frac{a \cdot b}{n}$	then:	$sig(A,B) = \frac{-\log\left(1 - e^{-\lambda} \sum_{i=0}^{k-1} \frac{1}{i!} \cdot \lambda^i\right)}{\log n}$
If	$\frac{(k+1)}{\lambda} > 2.5$	then:	$sig(A,B) \approx \frac{\lambda - k \cdot \log \lambda + \log k!}{\log n}$
If	$k > 10$	then:	$sig(A,B) \approx \frac{k \cdot (\log k - \log \lambda - 1)}{\log n}$

Figure 1: The three formulas to calculate significant co-occurrences of two terms

Due to performance reasons, we decided to implement the calculation of the significant co-occurrences independently, instead of re-using existing text mining modules. The different ways to calculate the co-occurrence allow the usage of the fastest algorithm for each diagnose, as the simplified formulas (the last two in the figure) require less time and processing power. However, we also had to consider pre-processing steps of the MRI diagnoses in order to complete calculations on the initial text corpus within a reasonable period of time.

3.2 Text Corpus and Pre-Processing

Accompanying measures during the evaluation of an information extraction tool revealed the need for additional assistance in finding topological relations between anatomic structures including regional indicators and diseases such as tumors. Given these requirements, there was a demand for domain specific databases for anatomic structures and pathologic expressions.

At the starting point of this project, we used a text corpus of about 6.000 diagnoses, which comprises comments of medical experts on magnetic resonance imaging (MRI) material. These findings derived from different radiologists, are completely written in capital letters and full of medical terms. Thus, we also faced the problems of synonyms, medical dialects and abbreviations. Further, the diagnoses are spread over a period of 17 years. As a consequence, time-dependent changes of terminology might be possible. These textual diagnoses were made anonymous and imported into a database. Considering the systemic performance and the sentence-based statistical calculation, the texts were also split up into sentences. Pre-processing of the free text is so realized that the occurrences of expression pairs are counted and stored. Consequently, the calculation of the co-occurrences can be executed by means of one of the three formulas which are shown in figure 1. Performance issues demanded a full-text indexing of the diagnoses as well as a reduction of the anatomic terms. Unfortunately, two or three character words are found in many other words, so they have to be excluded from the reference database. Additionally, we were in need of anatomic and pathologic terms for our approach. A corpus of approximately 6800 anatomic structures was generated from an anatomic dictionary [Dauber, 05]. This dictionary offers a rough allocation of anatomic structures to anatomic regions. More precision in finding such structures can be reached by using synonyms, the gathering of which is also time consuming. Efforts have been made to start with a synonym enhancement for the anatomic data at IMI, which has been used for the calculation. On the other hand, a Pathology database has been set up manually due to a lack of accessible resources. These corpuses represent the domain specific database sources for the statistic calculation and can be maintained in special application modules.

3.3 The Web Application

A basic user access control system is used for logging purposes. Maintenance operations must be executed as administrator. User actions include registration, editing of the registration information, login, logout and observing the login history. The application itself offers modules for the following functionality:

- Diagnoses can be listed and filtered according to two terms.
- Anatomic terms can also be filtered.
- The location for each anatomic term is indicated.
- The synonyms module shows all available synonyms for each anatomic term. These synonyms cover small parts of terms, but will increase in future.
- The menu option “ADD PATHOLOGY” provides a dialogue to add a term.
- Significant co-occurrents are listed in module COOCCURRENTS. Additionally new calculations can be started.
- A maintenance module provides splitting of the diagnoses as well as calculation of the occurrence of the single terms.

3.4 Core Functions and Advantages

For performance purposes, the administrator can initiate a pre-calculation of the occurrence of each single term. Thus the number of sentences and anatomic terms for the statistic calculation can be reduced. The splitting of the diagnoses is the second method of improving performance during the calculation. Additionally the split sentences are reduced by excluding sentences with a character length < 15. Thereby abbreviation sentences are most likely eliminated.

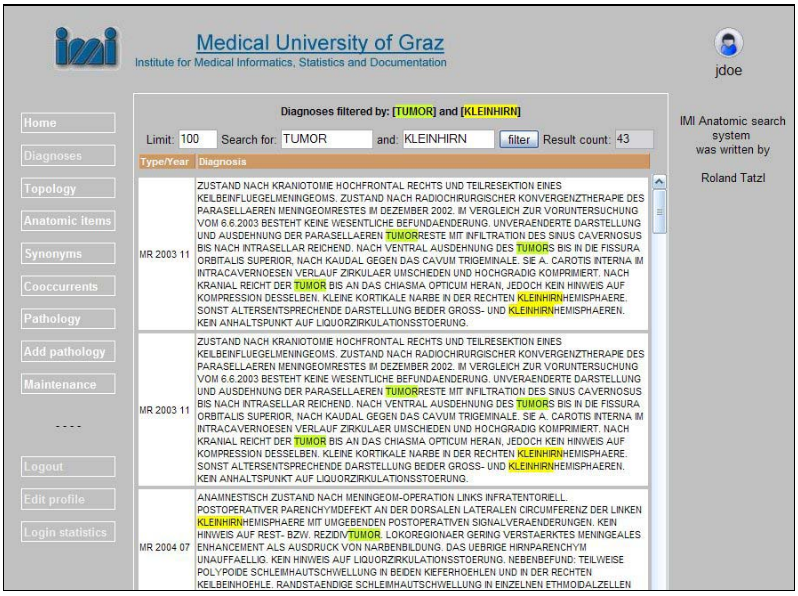


Figure 2: Filtering diagnoses according by the terms "TUMOR" and "KLEINHIRN"

Search for topological relations is based on MRI diagnoses in a heterogeneous context (e.g. cranial and spinal MRI diagnoses), as visualized in figure 2. The table of diagnoses can be filtered manually and retrieved in a list. Simple IR-techniques, such as query term highlighting, are used to visualize the results. These functions support medical experts on comparing the results for the calculated pairs of terms.

Anatomic and pathologic terms can be edited in separated dialogues. Because of a lack of a Pathology reference corpus, authoring functionality has been implemented for the application in order to manually add, modify or remove expressions. The processing of the medical free text is based on the formulas described in subsection 3.1. In order to improve the overall performance during calculation, the query considers only sentences which contain any of the anatomic or pathologic terms and only terms which were previously identified in the diagnose corpus.

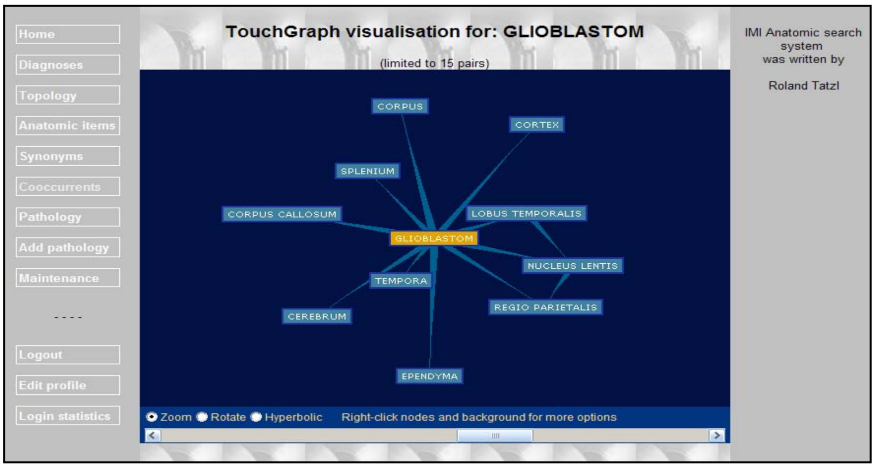


Figure 3: Tough graph visualization for the pathological term "GLIOBLASTOMA"

The resulting pair-list is shown in a table in descending order of significance for each expression. For further investigations, the most promising pairs can be used to filter the diagnoses in order to evaluate the results. For visualization purposes, a *Touchgraph* applet was implemented which enables the medical expert to estimate the proximity at a glance (see figure 3). In addition to the basic relations between the pathologic expression and the anatomic structures, the interconnection significance among the anatomic structures is calculated, to show their potential proximity within the corpus of diagnoses. This visualization clearly shows the topologic relationship.

4 Results and Experiences

The results of the calculation must be analyzed in their special context. Each pair of terms implicates specific anatomic-pathologic issues, and, in addition, the meaning of the co-occurrence must be interpreted individually.

4.1 Results for an Exemplary Scenario

The results are discussed on the basis of the malign brain cancer glioblastoma, which generally occurs in the group of middle-aged men and is located most likely in the corpus callosum and the temporal lobe. According to [Poeck, 87], glioblastoma does not occur in the cerebellum, a fact that is not contradicted in the result set at least. An analysis in a linguistic database (http://wortschatz.uni-leipzig.de) showed a very low occurrence in common language sources, such as newspaper articles and books, and emphasizes the importance of a well maintained domain specific database. The result set for the calculation (see figure 5) showed 10 pairs of terms for glioblastoma. We emphasise, that a highly significant co-occurrence does not prove the affliction of an anatomic structure with the paired disease. Also, a co-occurrence suggests a relation for further investigation. The most significant result suggests the co-occurrence with TEMPORA and SPLENIUM, whereby both locations are well known and located in the most significant decile of the result set. The second group of results does not make sense, because the associated anatomic terms CORPUS or EPENDYMA (thin epithelial membrane lining the ventricular system of the brain and the spinal cord canal) are too common and do not implicate worthwhile location information. The third group showed a combination of terms which are less well known and might be an interesting hint for further investigation.

4.2 Discussion of Experiences, Problematic Aspects and Opportunities

The resulting set of the calculations are listed in descending order according to the significance. The overall statistical evaluation says that the upper decile contains the most significant co-occurrences with a high probability. Some of the less significant results include interesting combinations of terms which may point out valuable new conclusions. Finally, syntax highlighting enables the observer to find the terms of the query quite easily in the result set of diagnoses. Despite such experiences, we also identified the following problematic aspects for applying statistical text mining techniques on diagnoses: (1) The group of results which are not useful is an evidence of the weakness of the anatomic reference corpus. Expressions with a too widespread a meaning should not be considered in order to reduce the wrong results. (2) The amount of diagnoses must be increased. Examples from professional common language research show that about 5 million sentences or more are required to validate our approach. (3) The diagnoses are specific for responsible radiologists. Therefore, they are not thoroughly comparable, as different experts tend to use other terminologies. (4) Words with only few characters (like “OS” or “COR”) are not suitable for searching purposes. Thus, they have to be eliminated. (5) Acronyms and abbreviations (dotted) cause difficulties when splitting the diagnoses into sentences.

Nevertheless, we find that our methodology for analyzing medical diagnoses comprises a promising approach: The analyzing of medical text corpora is fully computer driven and fully automated. The overall calculations require from a few minutes up to some hours, depending of the computational hardware and the amount of data. Thus, this method can be applied on text corpora at any time, e.g. to use other, more accurate anatomic and pathologic expressions for old diagnoses. Secondly, our tool is of interest for clinical professionals in order to support them at their daily tasks, for instance during pre-analyzing diagnoses. We also implemented some functions to support general medical experts in their daily work in order to reduce their cognitive load (see subsection 3.3). Finally, this kind of text mining algorithm could be also valuable for other application areas.

5 Conclusion and Future Work

We emphasize the importance of computer-based methods in medical documentation and the automatization of clinical processes, including analyzing diagnoses. In this context, we developed a methodology for text mining in medical text corpora and implemented a tool to evaluate our idea. The outcome of the calculations showed valuable results although based on a relatively low number of sentences. Observing all the diagnoses, generated in a hospital daily, will definitely improve the diagnostic value. However, we still have no proof that the anatomic structure is affected by the related disease, however, our experiences encouraged us to carry out further research efforts on these co-occurrences. Mining in large amounts of textual medical information can reveal new patterns for various questions. There will be a continued need for new mining assistants solving problems which are not even known today. We identified a huge benefit for the administration in identifying trends and developments in time to come to the appropriate decisions. The appropriate information presentation to the end users is a central future challenge, in order to keep their cognitive load in an optimum level, providing cognitive performance support.

References

[Dauber, 05] Dauber, W., Feneis' Bild-Lexikon der Anatomie, 9th ed, Stuttgart: Thieme, 2005.

[Fu, 03] Fu, T., Mostafa, J., Seki, K., Protein association discovery in biomedical literature, Proc. ACM/IEEE-CS Joint Conference on Digital Libraries (2003), 113-115.

[Granitzer, 06] Granitzer, M., KnowMiner: Konzeption und Entwicklung eines generischen Wissenserschliessungsframeworks, Dissertation, TU Graz, 2006.

[Gregory et al., 1995] Gregory, J., Mattison, J. E. and Linde, C.: "Naming Notes - Transitions from Free-Text to Structured Entry", Methods of Information in Medicine, 34, (1995), 57-67.

[Holzinger et al., 2007] Holzinger, A., Geierhofer, R. and Errath, M.: Semantische Informationsextraktion in med. Informationssystemen, Informatik Spektrum, 30, (2007), 69-78.

[Holzinger et al., 2000] Holzinger, A., Kainz, A., Gell, G., Brunold, M. and Maurer, H.: Interactive Computer Assisted Formulation of Retrieval Requests for a Medical Information System using an Intelligent Tutoring System, ED-MEDIA 2000, Montreal, (2000), 431-436.

[Holzinger, 2002] Holzinger, A.: Multimedia Basics, Volume 2: Learning. Cognitive Fundamentals of multimedial Information Systems, Laxmi, New Delhi, (2002), available also in German: www.basiswissen-multimedia.at

[Holzinger, 2005] Holzinger, A.: Usability Engineering for Software Developers, Communications of the ACM, 48, (2005), 71-74.

[Jenssen, 01] Jenssen, T.K., Laegrid, A., Komorowski, J., Hovig, E., A literature network of human genes for highthroughput analysis of gene expression, Genetics, 28, 1, (2001), 21-28.

[Leroy, 03] Leroy, G., Chen, H., Martinez, J.D., Eggers, S., Flasey, R.R., Kislin, K.L., Huang, Z., Li, J., Xu, J., McDonald, D.M., Ng, G., Genescene: Biomedical text and data mining, Proc. ACM/IEEE-CS joint conference on Digital Libraries, (2003), 116-118.

[Lovis et al., 2000] Lovis, C., Baud, R. H. and Planche, P.: Power of expression in the patient record: structured data or narrative text? Int. Journal of Med. Informatics, 58, (2000), 101-110.

[Poeck, 87] Poeck, K., Neurologie, 7th edition, Berlin: Springer, 1987.

[Srinivasan, 04] Srinivasan, P., Text Mining: Generating hypotheses from MEDLINE, American Society for Information Science and Technology, 55, 5, (2004), 396-413.

[Stephens, 01] Stephens, M., Palakal, M., Mukhopadhyay, S., Raje, R., Mostafa, J., Detecting gene relations from Medline abstracts, Pacific Symp. on Biocomputing, (2001), 483-496.

[Weeber, 01] Weeber, M., Klein, H., Berg, L., Vos, R. Using concepts in literature-based discovery: Simulating Swanson's Raynaud-Fish Oil and Migraine-Magnesium discoveries, Journal of the American Society for Information Science, 52, 7, (2001), 548-557.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Connecting User Interfaces and User Impairments for Semantically Optimized Information Flow in Hospital Information Systems

Shuaib Karim
(Institute of Software Technology & Interactive Systems, TU Wien, Austria
Department of Computer Science, Quaid-i-Azam University, Islamabad, Pakistan
skarim@ifs.tuwien.ac.at)

A Min Tjoa
(Institute of Software Technology & Interactive Systems, TU Wien, Austria
amin@ifs.tuwien.ac.at)

Abstract: The investigation of semantic relationship between the functional entities in a hospital information system, and their formal representation is of high value for automation purposes. The information flow in a health service setup revolves around some key entities such as various functional departments, information items generated from these departments, hospital staff and the patients. User's interactions within the system presume the device and the user profiles to be integral components which correspond to user interface characteristics and user impairments respectively. There is a meaningful relationship between the semantics of the information items, the roles of hospital staff, the functional departments, the user interface characteristics and the user impairments or abilities. Semantic web technology is very helpful in order to formally describe the semantics of these entities in ontologies and then creating rules to formally connect these ensuring a controlled and optimized information flow in the overall system. A test case, by connecting user impairments and user interface characteristics, is presented to show the validity and efficacy of our approach.

Keywords: Knowledge personalization and customization, Semantic Web, Connecting Ontology, Medical Informatics, User Interfaces

Categories: H.3, H.4, H.5.2, H.5.4, I.2.1

1 Introduction

Different information domains in a *Hospital Information System (HIS)* must interact with each other for providing meaningful results to user queries [Holzinger et al.2007a]. The results are transformed into appropriate *User Interfaces* (*UI*) according to user profile, especially the *Impairments*, and device profile for gaining optimal interaction with the system. Here, the *Impairments* signify the user's physical or cognitive limitations such as colour blindness, low visual acuity, problem in hearing, memory problems, specific learning difficulties, problems with motor functions, and problems affecting user's mobility [Holzinger et al.2007b]. The effects of *Impairments* on *UI* become valuable in varied interaction settings. In clinical environments the user to *UI* relationship is *m:n*. Often, the critical patient conditions and time constraints do not permit to shift the interaction settings from one mode to another for gaining optimal access to personalized interfaces. In that case, automatic interface adaptation according to user's needs and abilities is desirable.

The usage of *Style Sheets* for providing personalized *UI* is very beneficial in above interaction settings. Cascading Style Sheet [Bos et al.1998] is a language which separates the presentation from the content by allowing authors and users to associate style information such as fonts, spacing, and aural cues with structured documents. Still, the explicit declaration of style sheet according to the designer's mental model, which bridges the gap between the *UI* design and the user's *Impairments* information, is a manual activity. It can partly be automated by connecting the knowledge about *Impairments* and the *UI* characteristics. For this purpose, the *Semantic Web* technology is exploited which is based upon knowledge representation formalism of *Description Logics*. It is used on one hand to formally describe the semantics about the *Impairments* and *UI* in *OWL-DL*, and on the other hand the rules to connect these domains are described in *OWL Rule language*. The usage of *OWL-DL* ensures that a multitude of open source *DL Reasoners* can be used. The rules are manageable by the domain experts thus taking away a significant amount of overhead related to manual coding of style sheets. The execution of these rules across two models results in another model consisting of *UI suggestions* according to user *Impairments*. In the light of these *suggestions* the style sheets can be generated and / or adapted on the fly. On similar lines, other functional entities in an *HIS* may also be automated.

This paper is in continuation of our initial concept to connect *Impairments* and *UI* [Karim and Tjoa2006]. That resulted in a prototype implementation for an *Accessibility Framework* [Karim et al.2007] while specifying a number of ontology design patterns for our system. Here, we describe the realization of the *Perception Effect Pattern* from those, which shows at an abstract level the relationships between *Impairments*, *device profile* and the *information representation* on the interface.

In the rest of the paper the related work is given in Section 2. Section 3 describes the method, by introducing *Accessibility Framework* in context of our prototype *SemanticLIFE* [Ahmed et al.2004], and a brief description about the *Impairments* and *UI* ontologies. This is followed by their connection mechanism. The results are explained in Section 4, followed by conclusions and future work in Section 5.

2 Related Work

The importance and urgent need for new ways to integrate information in the medical domain, as highlighted by [Chen et al.2004] and other similar discussions, lead to the foundation of *Semantic Web Health Care and Life Sciences Interest Group* [Hclsig2004]. The resolution of semantic heterogeneity over the web is the key to several information integration issues[1]. The concept of connecting ontologies using *Semantic Web* technology is still under investigation to connect heterogeneous domains, especially for improving accessibility that might also be very useful for applications integration in various domains. In [Bouquet et al.2004] the *GALEN* and *Tambis* biomedical ontologies are first aligned with *UMLS* [Umls2007] and then mapped to each other by bridging rules using *Context-OWL*. Our work is different because of connecting different conceptual entities from heterogeneous ontologies.

[1] `http://www.ontologymatching.org/ (5th July 2007)`

Connecting diverse ontologies in an electronic medical record application using *OWL* [Sean et al.2004] and *RDQL rules* [Seaborne2004] is described in [Sheth et al.2006]. The concepts of medical terminology are linked with concepts in drug ontology. Semantic annotations are applied in *XML* files allowing usage of Web technologies such as *XSL* and *XPATH*, and the rules interpretation by *RDQL*. Data integration can be formally described using *Horn clause* rules and *F-Logic* rules [Angele and Gesmann2006] with some benefits of expressivity using *F-Logic*. However, the Semantic Web open world reasoning does not fit very well with *F-Logic* which is frame-based and influenced by object-oriented paradigm [Tetlow et al.2006] thus involving risks of incompatibility and undecidability.

The usage of rules for inferring triples from OWL models representing heterogeneous data sources has been a recent achievement by Oracle [Oracle2007]. Their scalability and performance claims with billions of *RDF* triples is very encouraging for the adoption of *Semantic Web* technology by the industry, especially in *HIS* where the data accumulated over the years is huge. The major benefit due to representation in form of triples of both data and schema, is the possibility to infer and execute queries which were not initially envisaged. Integration of heterogeneous data sources for associating genotype to phenotype information using *RDF* is described in [Sahoo et al.2007]. Rules are used to make associations based upon *isA* and *partOf* relationships because the *Gene Ontology* consists of only these relationships. In our case, the relationships are between classes, and also between individuals. Another very relevant work is introduced by [Obitko2007] about the translation of ontologies in Multi-Agent Systems in the manufacturing domain. The rules are transported via messages and are interpreted in respective agents. In our opinion, when the rules are executed in sequence then the inferred triples are added to the model which are not necessarily being transported or referred. This may not be a requirement in specific manufacturing application but certainly it is an issue if one has to benefit from the open world reasoning provided by ontologies in *DL*.

3 Method

The method operates under the *SemanticLIFE* framework. The setting consists of a personal information management system, *Accessibility Framework*, ontologies of *Impairments* and *UI* characteristics, and *rules* for making ontological connections.

3.1 Introduction to SemanticLIFE and Accessibility Framework

SemanticLIFE is our personal information management system for managing associations between the user's lifetime information items such as user's emails, browsed web pages, documents under process, processes running on user's computer. It has an additional feature to plug-in the *Google Desktop* as another data feed. In other words, it is storing the user's lifetime electronic activities, not merely the documents. It is a *Java*-based open source framework built into the *Eclipse* plug-in environment. *Semantic Web* technology is used to transform and store the information items' metadata into *RDF* triples in accordance with our core ontology consisting of items as classes and their metadata as properties. The queries are either launched by sending *SPARQL* query strings [Prud'hommeaux and Seaborne2007] or programmatically using *Ontology API* [Jena2007]. More sophisticated queries are realized by enhancing our ontology to *RDFS* and applying rules via *Jena Rule Engine.*

In order to provide accessibility in a systematic and generic way, the *Accessibility Framework* [Karim et al.2007] is incorporated within our system which is based upon *Connecting Ontologies* of the contextual components such as those related with user *Impairments*, *device* in use, *representation* on the interface, and the *task* being executed, and some related services. The interface to the *Accessibility Framework* is via *Info-Viz Bridge Service* and the *Accessibility* Service which operates on top of *Connecting Ontologies* formed as a result of rule execution across different contextual ontologies mentioned earlier. As a result of rule execution new triples are generated which are stored separately for reuse by the *Accessibility Service*.

Ontology definition for satisfactorily covering the corresponding domain is not a short-term but a long-term and intensive activity that requires a clear application focus and consensus of the community [Gruber1995]. It should be noticed that *Impairment* and *UI* ontologies do not yet cover the domains of discourse exhaustively.

3.2 User Impairments Ontology

A schematic overview of the ontology is shown in Figure 1. Sample competency questions which were prepared at the start of this activity are mentioned below:

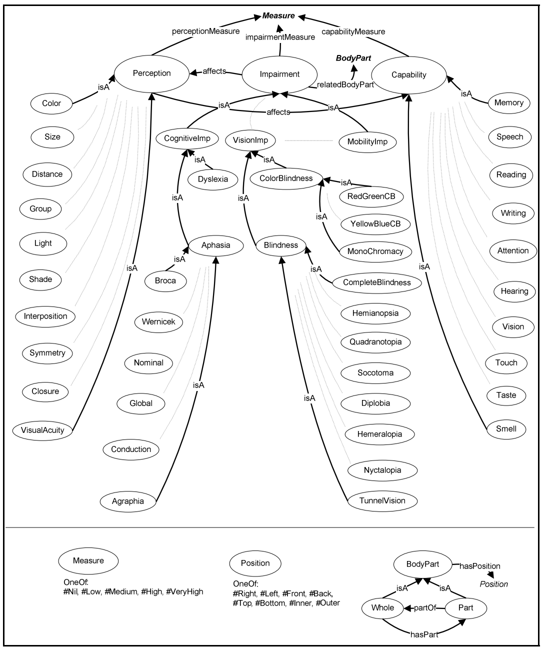


Figure 1: A Schematic Overview of User Impairments Ontology

- What is / are the related body parts?
- What is the impaired side?
- What is the impairment severity (on a predefined scale)?

- What is the perception cue which is affected, and up to what degree?
- What is its effect on another impairment w.r.t. affected perception?

These lead to a conceptual schema (Figure 1), and an *OWL-DL ontology*[2]. It was possible to execute queries such as to find all the "*left sided impairments*", "*right sided impairments*", "*any sided impairments*" and "*both sided impairments*".

- $LeftSidedImapirment \subseteq Impairment$
  $LeftSidedImpairment \equiv \exists relatedBodyPart(\ni hasPosition.\{Left\})$
- $RightSidedImapirment \subseteq Impairment$
  $RightSidedImpairment \equiv \exists relatedBodyPart(\ni hasPosition.\{Right\})$
- $AnySidedImapirment \subseteq Impairment$
  $AnySidedImpairment \equiv \exists relatedBodyPart\begin{pmatrix} \ni hasPosition.\{Left\} \bigcup \\ \ni hasPosition.\{Right\} \end{pmatrix}$

We may also get *AnySidedImpairment* as a union of already derived classes, e.g.,

- $AnySidedImpairment \equiv LeftSidedImpairment \bigcup RightSidedImpairment$

3.3 User Interface Ontology

On similar lines the *UI* ontology[3] was developed which could answer the following:

- Find the part-whole relationship of *UI* components
- Find the attributes of a component and their values (according to predefined usability scale for a normal user in normal conditions)
- Given specific attribute name, find the related *UI* components

It was possible to formulate and execute queries such as "*good usability components*", "*fair usability components*" and "*fair user control components*".

- $GoodUsabilityComponent \subseteq UiComponent$
  $GoodUsabilityComponent \equiv UiComponent \bigcap (\ni hasUsability.\{Good\})$
- $FairUsabilityComponent \subseteq UiComponent$
  $FairUsabilityComponent \equiv UiComponent \bigcap (\ni hasUsability.\{Fair\})$
- $FairUserControlComponent \subseteq UiComponent$
  $FairUserControlComponent \equiv UiComponent \bigcap (\ni userControl.\{Fair\})$

3.4 Connecting Ontology between Impairments and UI

The basic concept of *Connecting Ontology* between *Impairments* and *UI* is explained in [Karim and Tjoa2006]. Some of the competency questions are as follows:

- Avoidance of confusing colors for particular type of user's color blindness
- Font adjustments according to user's visual acuity
- Information presentation on the better part of the screen for a user suffering from Hemianopsia (absence of vision in half of visual field)

[2] http://www.ifs.tuwien.ac.at/~skarim/imp-v2.owl (5th July 2007)
[3] http://www.ifs.tuwien.ac.at/~skarim/ui.owl (5th July 2007)

These queries require the connection of *Impairments* and *UI* ontologies, for finding the *UI* components which are suitable for the prevalent user *Impairments*.

3.5 Rules for making connections

The connections are made by rules via *Jena Inference API* [Jena2007]. For example:

- *Low perception implies suggesting high usability components*
  (?x rdf:type imp:VisualAcuity) (?x imp:perceptionMeasure imp:Low) (?y rdf:type ui:UiComponent) (?y ui:hasLegibility ui:Good) → (?x eg:suggests ?y)
- *High perception implies suggesting fair usability components*
  (?x rdf:type imp:VisualAcuity) (?x imp:perceptionMeasure imp:High) (?y rdf:type ui:UiComponent) (?y ui:hasLegibility ui:Fair) → (?x eg:suggests ?y)
- *High rheumatism implies suggesting easily operatable components*
  (?x rdf:type imp:Rheumatism) (?x imp:impairmentMeasure imp:High) (?y rdf:type ui:UiComponent) (?y ui:userControl ui:Good) → (?x eg:suggests ?y)

4 Results and Discussion

Upon execution of rules the *suggestions* are generated (Figure 2) which can be stored for reuse until there are some further changes in the participating ontologies.

```
<rdf:RDF xmlns:co="http://www.ifs.tuwien.ac.at/ontologies/co#" ...>
  <rdf:Description rdf:about="http://www.ifs.tuwien.ac.at/ontologies/imp#VisualAcuity_High">
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#TextSize_09"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#TextSize_26"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#TextStyle_Italic"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#TextStyle_BoldItalic"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#TextSize_08"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#TextSize_24"/>
  </rdf:Description>
  <rdf:Description rdf:about="http://www.ifs.tuwien.ac.at/ontologies/imp#Rheumatism_High">
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#LabelledButton"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#ComboBox"/>
  </rdf:Description>
  <rdf:Description rdf:about="http://www.ifs.tuwien.ac.at/ontologies/imp#VisualAcuity_Low">
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#TextSize_18"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#TextSize_10"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#TextFont_TimesNewRoman"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#TextStyle_Bold"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#TextSize_22"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#TextSize_11"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#TextSize_20"/>
  </rdf:Description>
<rdf:Description rdf:about="http://www.ifs.tuwien.ac.at/ontologies/imp#ColorBlindness_RG">
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#Color_Grey"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#Color_Orange"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#Color_Cyan"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#Color_Blue"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#Color_Yellow"/>
  </rdf:Description>
  <rdf:Description rdf:about="http://www.ifs.tuwien.ac.at/ontologies/imp#ColorBlindness_YB">
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#Color_Grey"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#Color_Orange"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#Color_Cyan"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#Color_Green"/>
    <co:suggests rdf:resource="http://www.ifs.tuwien.ac.at/ontologies/ui#Color_Red"/>
  </rdf:Description>
</rdf:RDF>
```

Figure 2. Suggestions for Specific Impairments Generated by Applying Rules

These *suggestions* are processed for adapting the *CSS*. For example, there is a range of recommended text sizes for each type of visual acuity. The discrete values in each range are mapped to text size attributes for header tag “*h*” in a *CSS*. The highest suggested *TextSize* is mapped to "*h1*", the next *TextSize* to "*h2*" and so on. For the time being, this mapping is done programmatically using *CSSOM* style sheet parser [van Kesteren ed.2007]. A next step to be performed is the connection between *Impairments* ontology and style sheet elements and attributes. The result set in the form of *RDF* triples is an ontology in itself which is sharable and process able by *Semantic Web* tools for any useful purpose. Also, the *Impairments* and *UI* ontologies, and the method of *Connecting Ontologies* could be usefully exploited towards the automation of usability engineering [Holzinger2005] in general.

5 Conclusions and Future Work

Connecting heterogeneous information spaces has always been problematic and tedious. *Semantic Web* technology provides us the formalism and the tools for making improvements towards its automation. Using a test case of *Impairments* and *UI*, which are altogether heterogeneous domains, it is shown how these can be connected by using rules. Besides utility of this work for *UI*, the technique also can be replicated to semantically automate the integration of other sub-domains in an *HIS*.

Next, we plan to semi-automatically manage the rules with the help of end user's scenarios automation. Also, the automatic population of participating ontologies would be implemented by incorporating task and usability measurement ontologies.

Acknowledgements

This work is supported jointly by the Higher Education Commission of Pakistan, and the ASEAN-EU University Network Initiative.

References

[Ahmed et al.2004] Ahmed, M., Hanh, H. H., Karim, S., Khusro, S., Lanzenberger, M., Latif, K., Elka, M., Mustofa, K., Tinh, N. H., Rauber, A., Schatten, A., Tho, N. M., and Tjoa, A. M. (2004). SemanticLIFE - A Framework for Managing Information of A Human Lifetime. In Proc. of the 6th Int. Conf. on Information Integration and Web-based Applications and Services.

[Angele and Gesmann2006] Angele, J. and Gesmann, M. (2006). Data Integration using Semantic Technology: A use case. In Thomas Eiter and Enrico Franconi and Ralph Hodgson and Susie Stephens, editor, RuleML, pages 58–66. IEEE Computer Society.

[Bos et al.1998] Bos, B., Lie, H. W., Lilley, C., and Jacobs, I. (1998). Cascading Style Sheets, level 2 CSS2 Specification, W3C Recommendation. W3C Recommendation, W3C.

[Bouquet et al.2004] Bouquet, P., Giunchiglia, F., van Harmelen, F., Serafini, L., and Stuckenschmidt, H. (2004). Contextualizing ontologies. Journal of Web Semantics, 1(4):325–343. Web Semantics: Science, Services and Agents on the World Wide Web.

[Chen et al.2004] Chen, H., Colaert, D., and Roo, J. D. (2004). Towards Adaptable Clinical Pathway Using Semantic Web technology. In W3C Workshop Semantic Web for Life Science. Position paper for W3C workshop Semantic Web for Life Science.

[Gruber1995] Gruber, T. (1995). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. Human-Computer Studies, 43(5-6):907–928.

[Hclsig2004] Health Care and Life Sciences Interest Group. http://www.w3.org/2001/sw/hcls/

[Holzinger et al.2007a] Holzinger, A., Geierhofer, R., and Errath, M. (2007a). Semantische Informationsextraktion in medizinischen Informationssystemen. Informatik Spektrum, 30(2):69–78.

[Holzinger et al.2007b] Holzinger, A., Searle, G., and Nischelwitzer, A. (2007b). On some Aspects of Improving Mobile Applications for the Elderly. In Coping with Diversity in Universal Access, Research and Development Methods in Universal Access. Vol. 4554 of LNCS, pages 923–932.

[Holzinger2005] Holzinger, A. (2005). Usability Engineering for Software Developers. Communications of the ACM, 48(1):71–74.

[Jena2007] Jena Inference API, http://jena.sourceforge.net/inference/ (accessed 5th July'07)

[Karim and Tjoa2006] Karim, S. and Tjoa, A. M. (2006). Towards the Use of Ontologies for Improving the User Interaction for People With Special Needs. In Proceedings of ICCHP, volume 4061) of LNCS, pages 77–84, Linz, Austria. Springer - Berlin / Heidelberg.

[Karim et al.2007] Karim, S., Latif, K., and Tjoa, A. M. (2007). Providing Universal Accessibility using Connecting Ontologies: A Holistic Approach. To appear In Constantine Stephanidis, editor, Universal Access to Applications and Services, Vol. 7 of LNCS 4556.

[Obitko2007] Obitko, M. (2007). Translations between Ontologies in Multi-Agent Systems. PhD thesis, Czech Technical University in Prague, Czech.

[Oracle2007] Oracle (2007). Semantic Data Integration for the Enterprise. http://www.oracle.com/technology/tech/semantic_technologies/index.html

[Prud'hommeaux and Seaborne2007] Prud'hommeaux, E. and Seaborne, A. (2007). SPARQL Query Language for RDF. W3C Candidate Recommendation, W3C.

[Sahoo et al.2007] Sahoo, S. S., Bodenreider, O., Zeng, K., and Sheth, A. (2007). An Experiment in Integrating Large Biomedical Knowledge Resources With RDF: Application to Associating Genotype and Phenotype Information. In Proceedings of WWW 2007 Workshop Health Care and Life Sciences Data Integration for the Semantic Web. ACM Press.

[Seaborne2004] Seaborne, A. (2004). RDQL - A Query Language for RDF. W3C Member Submission, W3C. http://www.w3.org/Submission/2004/SUBMRDQL-20040109/.

[Sean et al.2004] Sean, B., van, H. F., Jim, H., Ian, H., McGuinness, D. L., F., P.-S. P., and Andrea, S. L. (2004). OWL Web Ontology Language Reference. Recommendation, W3C. http://www.w3.org/TR/owl-ref/.

[Sheth et al.2006] Sheth, A., Agrawal, S., Lathem, J., Oldham, N., Wingate, H., Yadav, P., and Gallagher, K. (2006). Active Semantic Electronic Medical Record. In Cruz, I., Decker, S., Allemang, D., Preist, C., Schwabe, D., Mika, P., Uschold, M., and Aroyo, L., editors, The Semantic Web - ISWC 2006, volume 4273 of LNCS, pages 913–926. Springer.

[Tetlow et al.2006] Tetlow, P., Wallace, E., Oberle, D., and Knublauch, H. (2006). A Semantic Web Primer for Object-Oriented Software Developers. W3C Note, W3C. http://www.w3.org/TR/2006/NOTE-sw-oosd-primer-20060309/.

[Umls2007] Unified Medical Language System. http://umlsinfo.nlm.nih.gov/ (5th July, 2007)

[van Kesteren ed.2007] van Kesteren ed., A. (2007). Cascading Style Sheets Object Model (CSSOM). Draft, W3C.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Facilitating Knowledge Management in Distributed Healthcare Systems

Bo Hu, David Dupplaw, Madalina Croitoru,
Srinandan Dasmahapatra, Paul Lewis
(University of Southampton, UK
{bh, dpd, mc3, sd, phl}@ecs.soton.ac.uk)

Abstract: We developed a prototype to support data integration and decision making in Breast Cancer domain, wherein the knowledge management plays a critical role. While the decision making needs to rely on different medical expertise, our system experiments with a system ontology and then a process calculus based language. The integration of data gathered from different modalities, represented in heterogeneous canonical formats, and broken down at diverse granularity levels is, therefore, facilitated by interaction models combining implementation flexibility and logic rigidity.
Key Words: knowledge management, ontology, process calculus
Category: I.2.4, I.2.m, H.m, D.2.11

1 Introduction

This paper presents an interaction model based integration of heterogeneous data in the healthcare domain. Our work builds upon the Miakt system, a breast cancer screening and diagnosis support system. However, unlike conventional approaches to heterogeneous healthcare data sharing we do not only look at standardising knowledge but also the sharing task. By sharing knowledge through interactions we indirectly share data. We believe that this approach (based on the Lightweight Coordination Calculus, LCC) is well suited for our problem domain. Indeed, Section 2 will show that healthcare requires dealing with groups of people who have seldom had their systems engineered to do tasks together. We will demonstrate that looking at interaction models as opposed to trying to integrate knowledge in the traditional manner benefits healthcare knowledge management and complements the existing work in Miakt.

Knowledge Management (KM) in healthcare poses a series of interesting and challenging issues. These issues arise from both the heterogenous, distributed nature of the domain and the ethical and security concerns that make the data even harder to access. Another characteristic that distinguishes KM in healthcare applications is that data sharing is done strictly following medical guidelines. While the former challenges could be partially overcome by conventional knowledge integration approaches, the latter requires more sophisticated reconciliation mechanisms. In this context we build upon the Miakt work that successfully addressed knowledge integration for breast cancer diagnosis, and take this work further by looking at interaction model based integration. This follows naturally from the integration guidelines, allowing explicit representation of medical guidelines and protocols. Sharing with interaction models also enables a dynamic way of knowledge combination.

2 Motivation

KM in healthcare presents itself as both an opportunity and a challenge. On the one hand, healthcare data is diverse in format, massive in size, and inconsistent in quality. It provides a good testbed upon which semantic-rich KM methodology can be evaluated and verified. On the other hand, it is a challenge in the sense that with the flux of data from heterogeneous resources and apparatus, many assumptions enjoyed by conventional KM is no longer applicable. Such assumptions include a centralised data repository and a globally accepted data model. The situation is exacerbated when full access to the entire domain knowledge is replaced with fragmented views that are limited by different privileges granted to the users, different usage of the data, and different hardware capacity.

The trend calling upon a non-conventional KM in healthcare is evident in the breast cancer multi-disciplinary meetings (MDMs) which will be used as the exemplar application in this paper. In an MDM, various medical experts come together to make a diagnosis of the patient's disease based on various data types that have been gathered and interpreted. This data may include images from a variety of imaging modalities, such as X-ray mammograms, Magnetic Resonance Imaging (MRI) scans, ultrasound scans and histopathological slides cut from biopsies taken at the suspect area. The data will also include information about the patient obtained from different sources, such as previous examinations and outcomes, medication allergy records, and family history of specific diseases. During the current MDMs in UK hospitals, medical practitioners arrive to the scene with hard-copy versions of their data relevant to the case. They will, in a collaborative manner, present their views and decide together on a course of action for that patient. Much of the data that is brought to the meeting exists only in hard-copies and available only to the concerned experts. Therefore, once the meeting is closed, access to the pre- and post-meeting materials by other medical staff becomes challenging. Digitising such materials is feasible but only partially solves the problem. The natural second step after digitisation is to address issues raised when such information is not readily available in a centralised data repository. This transposes the problem to one of distributed knowledge management.

An important consideration with respect to the new type of KM is the distributed nature of not only the data but also the users accessing the data. Data about a particular patient might be held by different departments within one hospital, from different hospitals and/or even from hospitals locating in different countries. Data requests might come from MDM members from theirs office or home, auditing committee, and patients themselves all with different access privilege. Apart from widely spreading in geographic regions and a diverse landscape of users, the heterogeneity of medical data is also demonstrated in the different levels of granularity, different nomenclatures used in sub clinical domains, and different standards reinforced by industrial manufacturers. In such an environment, knowledge which is a prime capital can only be drawn from distributed and heterogeneous data sources. Users, therefore, need to locate the correct data providers, retrieve the most appropriate parts of the exposed data and glue together all the bits and pieces of information to make sensible conclusions. In the meantime, we need to observe the data integrity and obey the data privacy and ethnic regulations. These constraints which are specific to medical domains suggest that exploitation of the data should not be directly laid upon the data itself but rather through dedicated knowledge services. We emphasis on the "knowledgable" aspect of these services due to the fact that they provide add-on values on top of the merely exposure of the encapsulated data.

3 The Miakt project and Miakt system

The aforementioned data heterogeneity issue is partially addressed in the Miakt project which does so by employing the latest development in Semantic Web (SW) [Berners-Lee et al., 2001] methodologies and technology. Miakt focuses on the breast cancer domain. Breast cancer is the most common cancer for women in the UK [Cancer Research UK, 2007]. Diagnosis of breast cancer normally involves multi-disciplinary meetings with experts from different medical backgrounds. A typical scenario of breast cancer assessment process starts with a report from routine X-ray mammographic examinations or self-reported abnormal symptoms followed by an X-ray mammographic examination. When a definitely benign diagnosis cannot be made, the breast MRI is normally treated as an expensive complementary method to the breast X-ray imaging so as to increase the diagnostic confidence. As the last resort, invasive methods are operated and tissue histopathology is acquired as the ground truth. Miakt aimed to accommodates the needs and challenges in diagnosing and treating breast cancer patients based on the knowledge drawn from all the different modalities.

The design philosophy of Miakt system is solidified with the Miakt architecture (as shown in Figure 1) developed primarily to allow the integration of various knowledge-based tools, that are published as services, into a knowledge management system [Dupplaw et al., 2004]. Exposing services instead of data has the advantage of allowing partners within the Miakt project to retain the integrity of their data while have their services and fragmented knowledge merged into a single KM system.

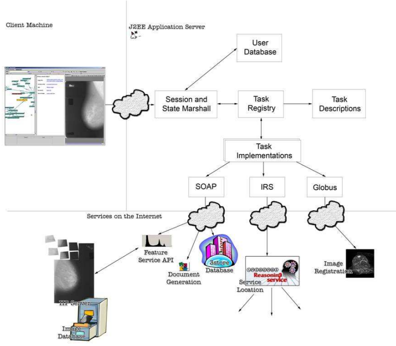


Figure 1: Miakt services

"Semantic flavours" are added into Miakt system through a domain ontology and a system ontology. Different knowledge services are projected upon a common conceptualisation, the Miakt domain ontology (Breast Cancer Imaging Ontology, Bcio for short). The existence of Bcio, however, should not be misinterpreted as the support in favour of a global model. It is not a surprise that human experts have in-depth knowledge on one aspect of the domain and only limited knowledge of the others. Bcio acts more in a manner of a reference point. It provides handles for the relevant information concerning a particular case via different aspects and different grain-sizes. This allows an expert to only focus on the fragments that he/she is really familiar with and makes available his/her evidence, conclusions and the basis of the judgement to the users of other modules [Hu et al., 2007]. However, knowledge captured with the ontology can only achieve its modularly accessible vision if backed up by a mechanism with which integration with other data fragments is facilitated.

As an effort towards defining this mechanism, we developed the Miakt System Ontology (MiSO) that organises the different services and regulates data transfers around the system. Users can invoke the available distributed services described based on their exposed functionality. MiSO also commits to satisfy the requirements on a general task invocation framework that allows remotely developed and appropriately packaged modules to be called upon, should they satisfy the knowledge handling requirements for use. Miakt utilises MiSO to facilitate the exchange of information among different services. Analogous to other communication protocols, MiSO specifies the format of messages passing from one service to another, the message initiators and recipients, the mechanism to parse and understand the contents of the messages, and the structure of the replies. The capability of a remote service is also defined based on MiSO. MiSO relies on a predefined, fixed workflow to regulate message flows. This workflow indicating how the various services should work together is, however, beyond a system ontology. In Miakt, this workflow is specified using a third-party software with descriptors drawn from MiSO and is defined in a deterministic manner scripted based on the working procedure of a UK MDM.

4 From MiSO to Interaction Model

Miakt prototype provides us an ideal platform to investigate the impacts and implications of applying semantic-rich technologies to KM with respect to a distributed heathcare system. Part of the challenges we faced in Miakt was how to exploit domain knowledge in steering the workflow of a particular participant. MiSO provides a feasible solution but at a cost of sacrificing system flexibility and extensibility. As a process ontology, MiSO only regulates *what* can be exchanged by various systems or services. It, however, does not formalise *how* the data should be transferred and *how* a mutual understanding can be established that underpin such a communication. Such weakness makes MiSO less favourite in healthcare domain wherein the major concern is not only at what can be said but also how things should be said and how these are understood by others. The necessity of the latter is evident in the prevalence of national and hospital-local medical guidelines and protocols. We, therefore, need a means to capture both the static aspect and the dynamic aspect of the breast cancer domain.

A solution to address this inefficiency is a formalisation of the interaction that only explicitly specify how services communicate while grant more freedom to the services themselves to decide how the interaction is materialised. One of the exemplar techniques facilitating declarative interaction specification is LCC [Robertson, 2004]. LCC is a process calculus for specifying coordination among multiple participants. It does so by clearly stating what role an individual plays in a messaging process, what messages should be sent and are expected to receive, and what constraints should be satisfied before a message can be initiated. LCC, therefore, is capable of capturing the dynamic characteristics of a domain. We use a few example to explain how LCC interaction models can contribute to data integration in Miakt. We would like to empahsis that the data integration is built upon an awareness of the flow of information within the system, reflecting protocols and guidelines that are driven by legal and ethical concerns given the sensitive nature of medical information. For instance, in Figure 2, we define how a domain expert could join a particular MDM event and how she could retrieve patient records from those holding the data and merge these "foreign" patient records with her local copies. In an interaction model specified using LCC, we use predicate $a()$ to specific the role that an individual is playing, $\Rightarrow$ and $\Leftarrow$ for the direction of message flow, and $\leftarrow$ for constraints. LCC interaction models are interpreted in terms of a Horn clause logic program.

In Figure 2, domain expert E's participation in an MDM starts with an invitation initiated from the meeting coordinator which is denoted as MDMC

$$
\begin{array}{l}
a(\text{expert}, E) :: \\
\quad \mathit{invitation}(E, X) \Leftarrow a(\text{mdmc}, C) \ \textbf{then} \\
\quad \mathit{accept}(E) \Rightarrow a(\text{mdmc}, C) \leftarrow \mathit{satisfies}(X) \ \textbf{then} \\
\quad \ldots \\
\quad \mathit{request}(\text{Patient}, Y, M) \Rightarrow a(\text{mdmc}, C) \leftarrow \mathit{certificate}(Y) \wedge \mathit{trans\_method}(M) \ \textbf{then} \\
\quad \mathit{receive}(\text{Patient}) \Leftarrow a(\text{mdmc}, C) \ \textbf{then} \\
\quad \left(\begin{array}{l}
\mathit{get\_patient\_id}(\text{Patient}, \mathit{ID}) \ \textbf{then} \\
\mathit{retrieve\_local\_record}(\mathit{ID}, \text{Patient}_{\text{local}}) \ \textbf{then} \\
\left(\begin{array}{l}
\mathit{align}(A_l, \text{Patient}, \text{Patient}_{\text{local}}) \leftarrow \mathit{find\_local\_aligner}(A_l) \\
\textbf{or} \\
\mathit{align}(\text{Patient}, \text{Patient}_{\text{local}}) \Rightarrow a(\text{aligner}, A_r) \leftarrow \mathit{find\_remote\_aligner}(A_r)
\end{array} \right)
\end{array} \right) \\
\quad \ldots \\
\quad \left(\begin{array}{l}
\mathit{classify}(\text{Patient}) \leftarrow \neg\, \mathit{missing\_info}(\text{Patient}) \\
\textbf{or} \\
\left(\begin{array}{r}
\mathit{ask\_ehr}(\text{Patient}, M) \Rightarrow a(\text{datahandler}, H) \leftarrow \mathit{missing\_info}(\text{Patient}) \wedge \\
\mathit{found\_new\_handler}(H)
\end{array} \right) \ \textbf{then} \\
\ldots
\end{array} \right) \\
\quad \ldots
\end{array}
$$

Figure 2: Domain experts in an MDM interaction model

and represented using a role introducing predicate, $a(\text{mdmc}, C)$. This invitation specifies that domain experts in an MDM should satisfy a list of restrictions given as X. In an interaction model, this is expressed as a message from the MDM coordinator (represented as an outbound double arrow leading from the coordinator to the expert). An individual is given the full responsibility to decide whether she is capable enough to take the role of a domain expert in a particular MDM instance. An acceptance will be sent off if she is confident in meeting all the requirements raised by the coordinator in X. The source of confidence might come from her education and working experience, her knowledge about this particular patient, and/or her availability during the time this MDM event is to be held. Exactly how the constraints are satisfied and how E's confidence is interpreted are left in the hand of E herself or a software agent acting on behalf of E. For instance, a crawling tool such as semantic squirrel[1] might set off to gather all the information from E's electronic diary, her personal webpage, email, publications, and her resume. This can be done with or without the supervision of a human and the results could be a stand-alone measure or one criteria as a part of a comprehensive measure covering all the aspects of E.

Upon joining an MDM, E's concern can be boiled down to two separated but closely related tasks: data acquisition and data integration. Expert E first sends a request to download the patient's record from local and remote data repositories. Together with the request, she also submits certificates Y for receiving the data and her preferred methods M for data transfer.

Each expert only has access to a small fragment of the patient data. How an MDM team would glue the information together and build up diagnostic decision therefrom then relies on to what extend they overlay their knowledge, together with their general expertise of the field and their experiences, onto the

[1] http://semantic-squirrel.org

body of a particular patient. Data sharing within an MDM team should not be assumed to be on an equal basis. It might be necessary to present the conclusion instead of the raw data to experts from different background. This is under both security and clarity considerations. Certain patient information is sensible and should not be disclosed to those who are not responsible for interpreting the data. Figure 3 illustrates fragments of LCC interaction model that retrieves data based on the request submitted by an arbitrary domain expert. It is evident that whether or not a particular expert is qualified to receive the requested data is subject to data-specific justification using $\mathit{is\_qualified}(E, \text{Patient})$. Meanwhile, this interaction model also emphasises on the customisation of data transfer methods. We use $\mathit{trans\_method}(E, M)$ to state that the data transfer task is specific to a particular expert.

$$
\begin{array}{l}
a(\text{datahandler}, H) :: \\
\quad \mathit{ask\_ehr}(\text{Patient}, M) \Leftarrow a(\text{expert}, E) \ \textbf{then} \\
\quad \mathit{register\_clearance}(E, \text{Patient}) \leftarrow \mathit{is\_qualified}(E, \text{Patient}) \ \textbf{then} \\
\quad \mathit{inform}(\text{Patient}) \Rightarrow a(\text{expert}, E) \ \textbf{then} \\
\quad \mathit{get}(P, F) \Rightarrow a(\text{DataMart}, D) \leftarrow \mathit{registered}(D) \wedge \mathit{contains}(D, P) \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad \wedge \mathit{matches}(P, \text{Patient}) \wedge \mathit{trans\_method}(E, F) \\
\quad \ldots
\end{array}
$$

Figure 3: Interaction Model for a datahandler

Comparing and contrasting localised patient data against that provided by MDMC could be the first step towards establishing a common ground for data integration. Most likely, however, patient records gathered by MDMC are not in a ready-to-use format for E. The received data is, therefore, subject to the alignment against that kept locally by E. If an alignment has already been established and can be reused in the current task, E invokes the local *aligner* to integrate remote patient data with the local records. If, on the other hand, information in the received patient records is beyond the coverage of existing alignments, E needs to locate a dedicated aligning service and submit both the remote and local patient records for aligning. Fragments of the *aligner* interaction model is shown in Figure 4. It is evident that we do not assume a conceptualisation which is globally accepted by all the participants. The existence of a domain ontology as a common reference point is not mandatory but an advantage to incorporate domain knowledge. For instance, one can align against and translate a patient record into existing standards in medical domains, such as HL7, DICOM, etc or a purpose-built application ontology, e.g. the Bcio developed for Miakt.

5 Conclusions

In this paper we reviewed our experience in designing and developing a KM system that provides integration for application domains with distributed data

$$
\begin{array}{l}
a(\text{aligner}, RA) :: \\
\quad \mathit{align}(\text{Patient}_1, \text{Patient}_2) \Leftarrow a(\mathit{expert}, \_) \textbf{ then} \\
\quad \left(\begin{array}{l}
\left(\begin{array}{l} \mathit{global\_align}(\text{Patient}_1, O) \textbf{ then} \\ \mathit{global\_align}(\text{Patient}_2, O) \end{array} \right) \leftarrow \mathit{exist\_domain\_ontology}(O) \\
\textbf{or} \\
\mathit{local\_align}(\text{Patient}_1, \text{Patient}_2) \leftarrow \neg \mathit{exist\_domain\_ontology}(O)
\end{array} \right) \\
\quad \ldots
\end{array}
$$

Figure 4: Interaction Model for data *aligner*

sources. Our approach is tuned in particular to handle heterogeneous data in medical domains wherein data integrity is given a strong emphasis due to the privacy and ethic concerns. Such requirements are met by concealing the data with services and composing the services based on individual applications.

It is our contention that similar data-oriented healthcare systems could be significantly enhanced with emerging semantic-rich technologies. In order to explore such potentials, we experimented with a system ontology to regulate what can be said by a service encapsulating the date and took one step further to enhance this static conceptualisation with a process calculus, the LCC process calculus language. Process calculus presents communications with interaction models declaratively representing the interaction/coordination procedure while leaving plenty room for implementation specificity. It can faithfully reflect medical protocols and guidelines by way of specifying the exact workflow that an event or a task should proceed. The merit of an LCC enabled KM system is in the mixture of process calculus and Horn clauses pair of which can provide a close resemblance of logic programming and the flexibility of implementation.

Acknowledgements

This work was partially supported by the Miakt project funded by UK EPSRC under Grant number GR/R85150/01.

References

[Berners-Lee et al., 2001] Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The Semantic Web. *Scientific American*, pages 28–37.

[Cancer Research UK, 2007] Cancer Research UK (2007). Breast Cancer at a glance. Retrieved in March 2007 from http://info.cancerresearchuk.org/cancerandresearch/cancers/breast/.

[Dupplaw et al., 2004] Dupplaw, D., Dasmahapatra, S., Hu, B., Lewis, P., and Shadbolt, N. (2004). Multimedia distributed knowledge management in miakt. In Handshuh, S. and Declerck, T., editors, *Proceedings of Knowledge Markup and Semantic Annotation, 3rd International Semantic Web Conference*, pages 81–90.

[Hu et al., 2007] Hu, B., Dasmahapatra, S., Dupplaw, D., Lewis, P., and Shadbolt, N. (2007). Reflections on a medical ontology. *International Jour. of Human-Computer Studies*. to appear.

[Robertson, 2004] Robertson, D. (2004). Multi-agent coordination as distributed logic programming. In *ICLP*, pages 416–430.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

StemNet: An Evolving Service for Knowledge Networking in the Life Sciences

Udo Hahn Joachim Wermter Michael Poprat
Jena University Language and Information Engineering (JULIE) Lab, Germany
(http://www.julielab.de)

David S. DeLuca Rainer Blasczyk Peter A. Horn
Department of Transfusion Medicine, Hannover Medical School, Germany
(http://www.mh-hannover.de/217.html)

Asad Bajwa
Clarity AG, Germany (http://www.clarity-ag.net)

Web: http://www.stemnet.de
Contact: {hahn,wermter}@coling-uni-jena.de

Abstract: Up until now, crucial life science information resources, whether bibliographic or factual databases, are isolated from each other. Moreover, semantic metadata intended to structure their contents is supplied in a manual form only. In the StemNet project we aim at developing a framework for semantic interoperability for these resources. This will facilitate the extraction of relevant information from textual sources and the generation of semantic metadata in a fully automatic manner. In this way, (from a computational perspective) unstructured life science documents are linked to structured biological fact databases, in particular to the identifiers of genes, proteins, etc. Thus, life scientists will be able to seamlessly access information from a homogeneous platform, despite the fact that the original information was unlinked and scattered over the whole variety of heterogeneous life science information resources and, therefore, almost inaccessible for integrated systematic search by academic, clinical, or industrial users.
Key Words: text mining, semantic retrieval, biomedical ontologies and databases
Category: I.2.7, I.2.1, H.3.1, H.3.3, J.3

1 State of the Art in Accessing Life Sciences Information

The life sciences, i.e., medicine, biology, chemistry and pharmacology, experience a dramatic growth of the amount of available data. This can be observed, e.g., in the area of genomic and proteomic research in which we witness an exponential growth of available sequence databases. Another source of evidence for this trend is the ever-increasing number of life science publications, i.e., scientific journal articles, patent reports as well as the growing proportion of free-text comments in biomedical databases.

At this point, the sheer volume of biomedical literature makes it almost impossible for biologists, clinical researchers and medical professionals to retrieve all relevant information on a specific topic and to keep up with current research. For example, in the world's largest bibliographic database for the life sciences, PubMed[1], the current number of entries (as of March 2007) already amounts to over 16 million entries, with up to 4,000 new ones being added each day. As a result, in recent years PubMed has attained a truly (and still growing) global impact as the most widely used and queried bibliographic database in the life sciences. This is exemplified by the number of query requests which is steadily rising (see Figure 1).[2]

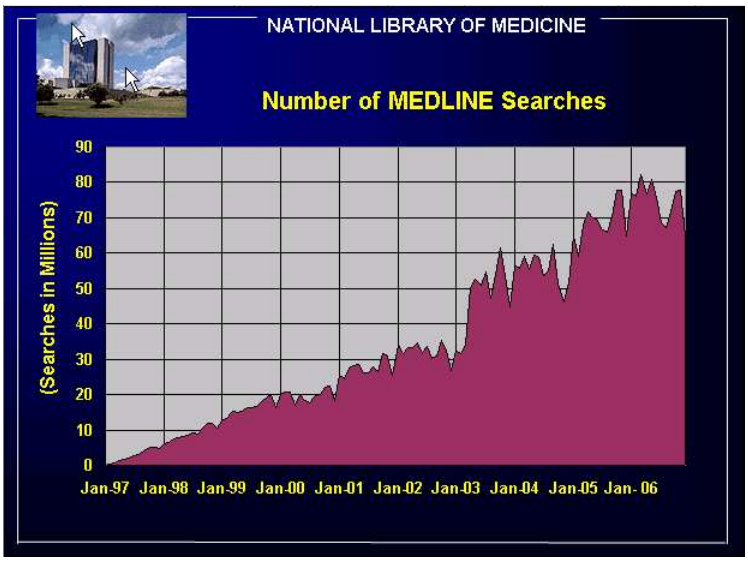


Figure 1: Development of PubMed Search Statistics

Large-scale and fully-interlinked semantic access to this vast amount of knowledge, however, is hampered mainly for two reasons (as illustrated in Figure 2):

- Current retrieval methods are insufficient since they are not geared toward getting deeply at the semantics of the biomedical text. Apart from bibliographic meta-information such as author names and publication years, the user interface of PubMed basically supports keyword-based (Google-like) querying. Due to the terminological and semantic complexity of the life sciences domain, the retrieval results for such queries are typically incomplete and suboptimal [1]. More semantically focused searches (e.g., for certain proteins or biological processes in which such proteins are involved) are only marginally supported, e.g., by manually assigned document-level indices as

[1] http://www.ncbi.nlm.nih.gov/entrez/
[2] http://www.nlm.nih.gov/bsd/medline_growth.html

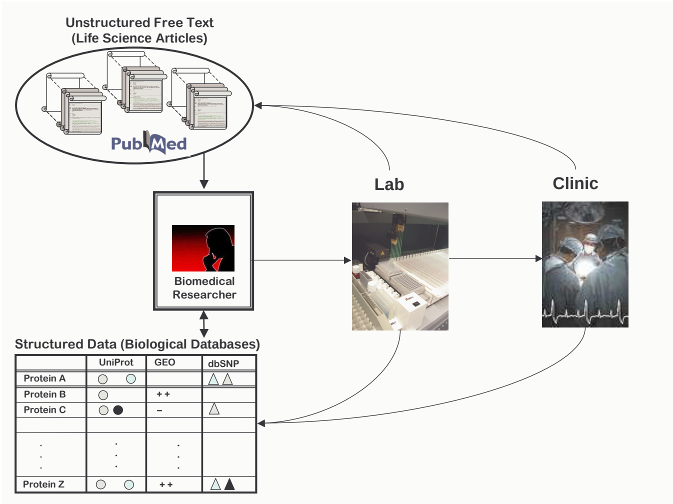


Figure 2: Current Practice: Hampered Access to Life Sciences Information

provided by the biomedical MeSH thesaurus[3] which, however, has proven to be rather inconsistent and incomplete. Furthermore, PubMed needs to service *all* aspects of the global life science community and thus needs to retain a high degree of semantic generality. This, however, stands in contrast to very specific search topics often formulated by researchers and clinical users.

– Apart from being locked in unstructured free text, a substantial amount of biomedical knowledge is housed in structured biological databases. These databases focus on specialized biomedical data, such as (species-specific) sequence information for defined genes and proteins, gene expression information in certain tissues, etc. Unfortunately, the knowledge found both in unstructured text and in structured databases is not linked. Thus, if a biomedical researcher finds information on a certain protein in a scientific article, linking this information to the respective database entry for this protein in a specialized database is usually not supported. An additional knowledge management problem immediately occurs since protein identifiers differ from one database to the other in an unpredictable way. Similarly, links from free-text fields of a database (which contain manually supplied annotations of the data in verbal, i.e., unstructured form) to relevant publications are not supplied on a larger scale. In any case, such linkings or mappings from free-text sources to unique biomedical database entries are hampered by the enormous degrees of ambiguity of biomedical terms and names [2].

[3] `http://www.nlm.nih.gov/mesh`

2 Goals of the StemNet Project

The goals of the StemNet[4] project respond to these two shortcomings. On the one hand, we aim at providing truly *semantic* access to the vast amount of knowledge found in the unstructured free texts of the PubMed bibliographic database. On the other hand, we aim at *linking* this knowledge encoded in free texts to respective knowledge stored in structured biomedical databases. In StemNet we plan to improve the semantic interoperability of currently disconnected information in the life sciences.

The biomedical subdomains the StemNet project focuses on are Hematopoietic Stem Cell Transplantation (HSCT) and Immunology. Both lie at the center of the fast-growing and crucial interface between genomic/proteomic research, on the one hand, and medical/clinical application, on the other hand. HSCT is used for a variety of malignant and nonmalignant disorders to replace a defective host marrow or immune system with a normal donor marrow and immune system. In many cases, the clinical treatment of patients with leukemia and other malignant hematological tumors is only successful, if a HSCT with a genetically different (allogeneic) donor is carried out, thus triggering the therapeutic effect of tumor cell elimination, known as the graft versus leukemia (GVL) effect.

The high risk of HSCT is due to the complex genetic differences of both HLA-genes and non-HLA genes between stem cell recipients and donors [3], which can only be controlled for through a complex and interactive analysis of numerous parameters. Since the GVHD and GVL effect are closely interrelated, the severity of GVHD is inversely related to the risk of relapse and strategies aiming at reducing GVHD may increase relapse rates. Currently, new strategies are being developed to separate these two effects in order to decrease the incidence and severity of GVHD without increasing the risk of relapse.

3 Resources for StemNet's Knowledge Network

Using PubMed as a starting point, the following knowledge resources are essential for the StemNet knowledge network under construction (see Figure 3):

- **OBO – Open Biomedical Ontologies.** OBO[5] is an umbrella organization for ontologies and shared terminologies for use across all biological and biomedical domains. In particular, the *Gene Ontology* (GO) [4] provides a community-wide accepted semantic framework to describe and annotate biomedical knowledge found both in unstructured free text and in biomedical databases. In order to grant semantic access to the scientific literature kept in PubMed, it is essential to annotate textual data with OBO/GO-based

[4] `http://www.stemnet.de`
[5] `http://obo.sourceforge.net`

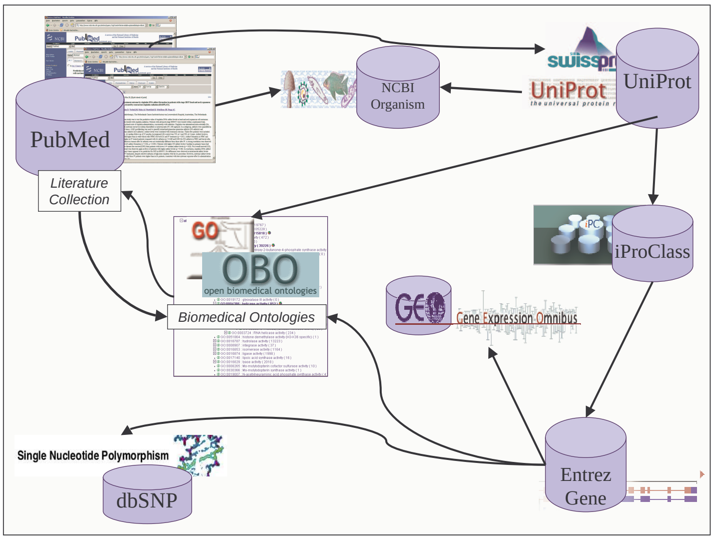


Figure 3: StemNet Knowledge Resources Network

biomedical terms (in particular, the molecular function and location of genes and proteins, as well as the biological processes they are involved in). Once a significant sample of free text is manually annotated using this vocabulary, such an annotated corpus can be exploited to automatically train entity and relation taggers in a supervised way [5]. After successful training, these text analysis engines will perform large-scale semantic annotation of textual data in a fully automatic way.

- **UniProt and iProClass.** The Universal Protein Resource (UniProt) provides the life-science community with a single, centralized, authoritative resource for protein sequences and functional information [6]. Each protein entry is associated with its respective organism (e.g., human, mouse, bacteria, viruses, etc.) and provides a link to the NCBI taxonomy organism database.[6] The iProClass mapping database[7] links UniProt to over 90 biological databases.

- **Entrez Gene** is provided by the U.S. National Center for Biotechnology Information (NCBI)[8] to organize information about genes, and serves as a major node in the nexus of genomic map, sequence, expression, protein structure, function, and homology data. This database serves as a hub of information for databases both within and external to NCBI.

[6] `http://130.14.29.110/Taxonomy`
[7] `http://pir.georgetown.edu/iproclass`
[8] `http://www.ncbi.nlm.nih.gov/`

- **dbSNP – Single Nucleotide Polymorphism.** In collaboration with the National Human Genome Research Institute, the NCBI has also set up the dbSNP database to serve as a central repository for both single base nucleotide substitutions and short deletion and insertion polymorphisms, which are key to genetics research in associating sequence variations with heritable phenotypes (particularly diseases).
- **GEO – Gene Expression Omnibus.** The *Gene Expression Omnibus* (GEO) is a public repository that archives and freely distributes high throughput gene expression data submitted by the scientific community. GEO currently stores some billion individual gene expression measurements, derived from over 100 organisms, addressing a wide range of biological issues.

Up until now, these resources remain, by and large, unconnected. One major goal of the StemNet project will be to interlink the underlying terminological resources used to describe the biological data in the databases and thus develop a conceptual foundation for interoperability based on a carefully designed formal ontology infrastructure (the rationale and progress of this work is described in [7]). Once this link has been fully established, these resources will be integrated into the StemNet system as its conceptual backbone.

4 Semantic Knowledge Networking and Semantic Access

Using state-of-the-art text mining technology [8, 9], we automatically annotate a sample of the PubMed textual data with terms from the OBO/GO ontologies. Adding this semantic metadata to documents empowers and further facilitates *semantic* retrieval of biomedical knowledge [10, 11] beyond the traditional keyword-based search [1]. In this respect, we also annotate the molecular functions of genes and proteins, the key players of biological processes at the molecular level. In preliminary experiments, we achieved an F-score[9] of about 90% in automatically annotating the immunologically relevant *cytokine function* of proteins and about 80% F-score in annotating mentions of *cytokine receptor functions*. Similar results were also obtained for *variation events* (i.e., polymorphisms), *organisms*, *immune cells* and *antigens*.[10] This evaluation data, still in a very early stage of the project, already compares with the performance level that has been reported at BioCreaTive, the latest major BioNLP software competition [12].

After having identified a protein name in a text, its entry in the UniProt database must be located. This is a challenging task because protein names

[9] The F-score is a standard evaluation metric which balances between precision and recall measurements; cf. [1].

[10] Actually, on a semantically more fine-grained level, the StemNet annotations already cover over 60 different semantic categories.

are highly ambiguous on several layers of meaning [2]. Annotating the respective organism and linking it to NCBI taxonomy organism database aids in this disambiguation task. The UniProt identifier opens up the door to several other StemNet-relevant biological databases. Through the iProClass mapping database, a knowledge link to Entrez Gene is established, and from there, additional links to GEO and dbSNP can be constructed. Moreover, both the UniProt and the Entrez Gene database entries for genes and proteins also contain (curated) GO annotations. We then come full circle, as these descriptive items, in turn, may serve as additional semantic metadata for the original PubMed text and thus facilitate semantic access and retrieval for the user.

In this way, the StemNet Knowledge Server semantically links the disparate biomedical knowledge resources and thus provides biomedical researchers with an integrated view of relevant information. In particular, the user accesses information from the homogeneous StemNet server (see Figure 4). This contrasts with the original search paradigm (see Figure 2), where bibliographic and fact databases are strictly isolated from each other and thus each must be searched separately in a complicated, expensive and error-prone manner. The results of these searches largely depend on the ingenuity, experience and time investment of the searcher, who has to battle with different query languages and large amounts of a priori knowledge related to the relation structure and other content issues specific to *each* of the databases involved.

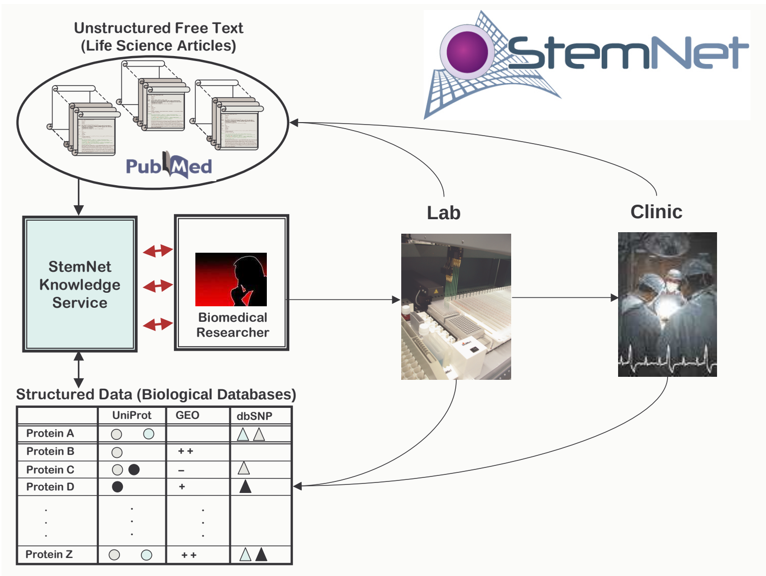


Figure 4: StemNet Scenario for Enhanced User Access to Life Sciences Information

5 Conclusions and Outlook

The StemNet Knowledge Server links disparate biomedical knowledge resources on a semantic layer and thus enables biomedical users to access and search for relevant information in an integrated manner. Starting from the vast amount of life science documents in the PubMed literature database, it provides the user with a semantic view on these documents in terms of annotations (semantic metadata added to the texts). The annotated documents are interlinked with external knowledge resources, such as biomedical ontologies and databases.

While the focus of the StemNet project is on the clinically relevant biomedical subdomain of Hematopoietic Stem Cell Transplantation (HSCT), the underlying methodologies which provide for semantic interoperability are designed and implemented to be easily extensible to other subdomains of the life sciences and, possibly, even translate to other science and technology domains, as well.

Acknowledgement

The StemNet project is funded by the German Ministry of Education and Research (BMBF) via its *e-Science* initiative (funding code: 01DS001A to 1C).

References

1. Hersh WR. Information Retrieval. A Health and Biomedical Perspective. Springer, 2nd ed., 2002.
2. Hirschman L, Colosimo M, Morgan A, Yeh A. Overview of BioCreAtIvE task 1B: Normalized gene lists. BMC Bioinformatics. 6 (Suppl 1: S11) 2005.
3. Horn PA, Elsner HA, Blasczyk R. Tissue typing for hematopoietic cell transplantation: HLA-DQB1 typing should be included. Pediatric Transplantation 10(6) 2006:753–754.
4. Gene Ontology Consortium. The Gene Ontology (GO) project in 2006. Nucleic Acids Research. 34(1) 2006:322–326
5. Feldman R, Sanger J. The Text Mining Handbook. Advanced Approaches in Analyzing Unstructured Data. Cambridge University Press, 2007.
6. Bairoch A et al. The Universal Protein Resource (UniProt). Nucleic Acids Research. 33(1) 2005:154–159.
7. Schulz S, Beisswanger E, Hahn U, Wermter J, Kumar A, Stenzhorn H. From Genia to BioTop: Towards a top-level ontology for biology. In: Formal Ontologies in Information Systems. Proceedings of the FOIS 2006 Conference, pp.103-114. 2006.
8. Hahn U, Wermter J. Levels of Natural Language Processing for Text Mining. In: S. Ananiadou and J. McNaught (Eds.), Text Mining for Biology and Biomedicine, pp.13–41. Artech House Publishers. 2006.
9. Buyko E, Wermter J, Poprat M, Hahn U. Automatically adapting an NLP core engine to the biology domain. Proceedings of the ISMB 2006 "Joint Linking Literature, Information and Knowledge for Biology and the 9th Bio-Ontologies Meeting". 2006.
10. Ferrucci D, Lally A. Building an example application with the Unstructured Information Management Architecture. IBM Systems Journal 43(3) 2004:455–475.
11. Carmel D, Maarek YS, Mandelbrod M, Mass Y, Soffer A. Searching XML documents via XML fragments. Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2003), pp.151–158. 2003
12. Yeh A, Morgan A, Colosimo M, Hirschman L. BioCreAtIvE task 1A: Gene mention finding evaluation. BMC Bioinformatics. 6 (Suppl 1: S2) 2005.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

On the Integration of a Conventional Content Management System into a Semantic System

Erich Gams, Daniel Mitterdorfer and Christoph Wieser
(Salzburg Research Forschungsgesellschaft m.b.H.
{egams|dmitter|cwieser}@salzburgresearch.at)

Abstract: Content management systems (CMS) have been established to ease hosting and organizing of large amounts of multimedia content. However, existing systems often have to struggle with heterogeneity and distribution of the content and meta-data over several repositories; Furthermore, these systems often lack user support by offering full-text search in text documents only or support browsing based on predefined-defined static taxonomies not adapting to new content. In this paper we address these issues by facilitating Semantic technologies and introduce a semantic CMS reference architecture. Our approach offers the integration of different CMS services on a semantic level based on a modular architecture. The applicability will be shown by migrating an existing CMS platform containing e-learning content to our semantically enhanced architecture on service and data level.
Key Words: Integration, Legacy System, Semantic Content Management System
Category: Semantic Web Applications

1 Introduction

Current Content management systems (CMS) facilitate the organization, control, and publication of documents and other content. Multimedia content can be enhanced and structured with meta-data to provide full text search or search based on meta-data. However, re-finding multimedia assets in a repository is challenging systems of multimedia repositories. Within the Salzburg NewMediaLab a project addresses these issues in the domain of e-learning content management systems. Currently, a company uses a traditional content management system, which stores several thousand learning material resources in form of documents, videos or animations. The learning material is contained in several distributed repositories. Users are able to access these repositories through different separated portals. In order to ensure high quality of the learning material, the editorial staff assorts as well as categorizes the content according to a predefined taxonomy. Based on this taxonomy users are able to search for and browse the available learning material. During its operation the system revealed some weaknesses of traditional content management systems. In this paper we describe how we intend to address these problems by facilitating Semantic technology to build up a semantically enhanced content management system. We show our approach of a modular architecture of a semantic CMS and of how to migrate from an existing CMS to a semantic system.

2 Open Issues of the Existing CMS

In the last few years new requirements arose and it became apparent that the existing CMS has different shortcomings. As these issues demonstrate, the current search service does not fully address the needs of its users.

- **Single Point of Access:** Users have to consult different applications for searching over different types and formats of content in different repositories.
- **Taxonomy Categorization:** Once defined modifying the taxonomy is difficult. Moreover, it is generally not possible to use more than one category.
- **Browsing:** First, users have to know the taxonomy to search for a specific learning material. Next, a taxonomy provides only a one-dimensional classification. Generally, a third problem is that the editorial staff and the users might have different viewpoints of a domain, known as Semantic gap.
- **Full-text Search:** Currently, only exact matches with guessed search terms are returned as search results.

3 How Semantic Technologies Address the Open Issues

The main idea of Semantic Web [1] is that content and its meta-data can be also interpreted by machines. The approach is to construct a knowledge model over the domain knowledge of the users by assigning learning material to categories or concepts in a background ontology. This approach enhances the user experience significantly by addressing the shortcomings described above:

- **Single Point of Access:** Ontologies can be used to integrate content that adheres to different meta-data standards. This enables users to search over different file types in different repositories without any hassle.
- **Taxonomy Categorization:** A model represented by an ontology is much more flexible. Learning material can be assigned to multiple categories and more associations can be expressed by comparison to a hierarchical model.
- **Browsing:** The semantic gap between the editorial staff and the users can be reduced by allowing personalized views on the data material. As the system evolves, it adapts more and more to each user's experience.
- **Full-text Search:** Using thesauri such as WordNet [1], we are able to detect concepts and can broaden or narrow a search based on concepts or search for synonyms. Misspelled search terms can be handled gracefully using fuzzy string matching.

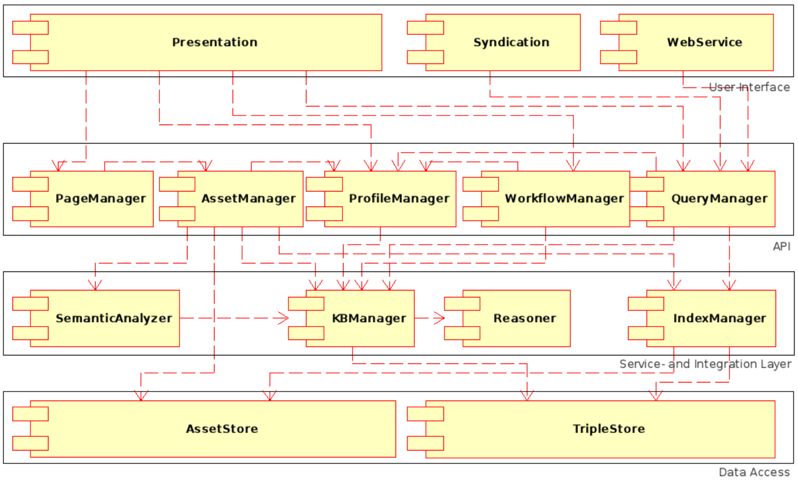


Figure 1: System Architecture.

Data Access Layer. In this architectural layer all logic to store and retrieve entities from any data sources is encapsulated. The *AssetStore* is responsible for persisting asset files. Although many file formats support embedded meta-data this meta information is not taken into consideration for search as is. As already outlined, the rationale is that the meta-data can not be used for efficient search in this form. Therefore, when saving an asset, its meta-data is extracted and stored separately in the triple store. In the *TripleStore* contains the entire system knowledge as a model like meta-data about assets and also information about users, subjects, and suchlike. Each information is represented as an RDF triple [2]). Storing all information as RDF triples allows for efficient search and reasoning. The triple store can either be realized using a database or a file as persistent storage. Technologies such as D2RQ allow to integrate legacy databases.

The **Service and Integration Layer.** consists of components having direct access to the data access layer. Basic data access functionalities are meant to be integrated to libraries for the application programming layer. *SemanticAnalyzer.* The SemanticAnalyzer analyses media data and in particular plain text for a semantic analysis and yields meta-data. *KBManager and Reasoner.* KB-Manager stands for Knowledge Base Manager and is the heart of the system and manages triples of the model are describing. Another feature of the KB-Manager is the abstraction of the reasoner. *IndexManager.* The system uses an index that can be used for full-text queries. However, the most important duty of the IndexManager is to add to and remove entries from the index.

[1] http://wordnet.princeton.edu/, last visited 21st May, 2007

Application Programming Interface. The *AssetManager* is responsible for saving, accessing, and deleting assets. In the context of a semantic system the AssetManager handles also the meta-data of assets and can trigger a semantic analysis and an update of the search index. A specialized AssetManager is the **P**ageManager. As already mentioned above the **P**rofileManager manages access rights to assets and meta-data. Besides access rights the ProfileManager also manages general user data like the e-mail address. The *WorkflowManager* abstracts over triples representing workflows like how user rights are granted. The *QueryManager* is supposed to hide the complexity of queries by deciding whether a query should be answered by a full-text query on the data or by a search on the meta-data.

User Interface. The user interface separates data from presentation. While the *P*resentation component offers the main interface for the user, the *S*yndication and the *W*ebService components offer interfaces for automated access to the systems via aggregators (such as RSS clients) or via agents. The most complex part of these components is the *Presentation* component. The other ones only provide a subset of the functionality of the presentation component. As the architecture evolves, we may add additional functionality to these components.

4 Summary and Conclusion

We discussed requirements and open issues of the content organization functionality of a CMS. We identified four main issues according to categorization and search in a content repository based on the analysis of an operative e-Learning CMS. The system is not able to dynamically adapt taxonomies to new content or provide a personalized view of the content. As a consequence of our analysis we discussed how Semantic technologies may address the open issues identified. We set up a modular reference architecture of a semantically enhanced CMS and discuss an approach of how to migrate from a CMS to a semantic system. Our conclusions are as follows: (1) SOA is well suited for a migration from a monolithic architecture of a legacy CMS to a modular architecture of a semantic CMS. (2) A novelty of our approach is the clear separation of content in the AssetStore and meta-data in the TripleStore. Hence, all meta-data is modeled in a (modular) ontology, which enables a unified interface for reasoning. (3) Interfaces between highly interacting components, e.g. KBManager and Reasoner, have to be tightly coupled to ensure optimal performance.

References

1. T. Berners-Lee et al. The semantic web. *Scientific American*, May 2001.
2. G. Klyne and J. Carroll. Resource description framework (RDF): Concepts and abstract syntax. Technical report, W3C.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Enabling Privacy in Social Communities

Josef Noll, Mohammad M. R. Chowdhury, György Kálmán
University Graduate Center - UniK, Kjeller, Norway,
[josef,mohammad,gyorgy]@unik.no

Abstract: Ubiquitous computing and the pervasive Internet have enabled service access in every situation. However, adaptation to the user needs is purely handled, and service specific security implementations are only found for specific services. This paper presents an approach to combine the I-centric and service centric world based on a semantic description of user relations enabling service access. A prototype using over-the-air key distribution demonstrates the capabilities of the suggested approach.
Key Words: Semantics, privacy, social communities, key distribution, mobile, identity
Category: F.3.2, K.4.1, H.4.3.

1 Introduction

Current developments in service delivery have the focus on Mobile Service Delivery and Semantic Service Delivery. Current reporting from The World Factbook states two to three times as many mobile users as Internet users in UK, France and China [World factbook 2006]. Taking into account that mobile users are always available as compared to an average PC usage of just above 2 h/day[1] shows the importance of mobile service access [Ball State 2006]. This paper explains the principles of the I-centric [Arbanowski et. al. 2004] and service centric world in sect. 2. It then introduces in sect. 3 an identity architecture. Based on the social relations of a user, it will then in sect. 4 provide a concept for role-based service access and a prototypical implementation.

2 I-centric service provision in a social community

The key challenge in personalised service access is the handling of user preferences, context, devices, and connectivity. Personalisation should be supported by *learning* profiles handling the preferences of the user, the *presence* (where is the user, what is he doing), and the social/community characteristic of a user [Noll 2006]. The mobile phone has a central place in this picture, as it supports seamless authentication, out-of-band key distribution, presence and location information. Semantics are introduced to describe user preferences and relations and to characterise the social context of the user as indicated for a

[1] 137.3 minutes/day for male users and 134.2 minutes/day for female users

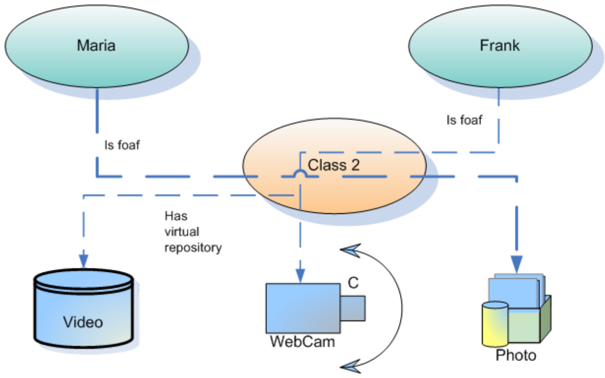


Figure 1: Virtual data and service repository

school scenario. Paul and Anna are members of class two of Sogn school, and their parents, here: Frank and Maria, are linked to the class 2 through a social graph. This paper uses social relations to enable content and service access. We use the semantic description of social relationships to define service/document access rights as presented in fig. 1. Through the relation (here: mother of a child in class 2) Maria gets access to the photos taken by Frank, who is father of a child in class 2. Our approach is to use the mobile phone for authentication services, here based on short messages (SMS) to distribute keys.

3 Identity based service access

In the virtual world identity handling has to take into account the dynamic service requests and privacy requirements of a user. Roccas introduced this in 2002 through the term of *social identity complexity*, defining a new theoretical construct that refers to an individual's subjective representation of the interrelationships among his or her multiple group identities [Roccas and Brewer 2002]. Identity is mainly verified through an authentication mechanism. The Internet was built without such an identity layer. In the current Web2.0 discussion Identity2.0 is introduced to interconnect people, information and software. Identity mechanisms as suggested by e.g. Microsoft, Sxip and Liberty Alliance are tailored towards remote services. In this paper we focus on methods of using different identification mechanisms for the variety of remote and proximity services, thus providing an Identity management for the I-centric and service centric world.

The proposed integrated identity mechanism consists of certificates, keys and preferences stored in a personal device and in the network. These identities are categorized in three groups of identity, personal identity (PID), corporate identity (CID) and social identity (SID) based on the roles exercised by a person in real life [Chowdhury and Noll 2007]. Users are authenticated by identity providers using keys. These keys were distributed only among the members of the group using the mobile environment. They then access the social contents using the proposed role-based ontology with differential access rights. The generic architecture is illustrated in fig. 2. Our service scenario builds on the relation

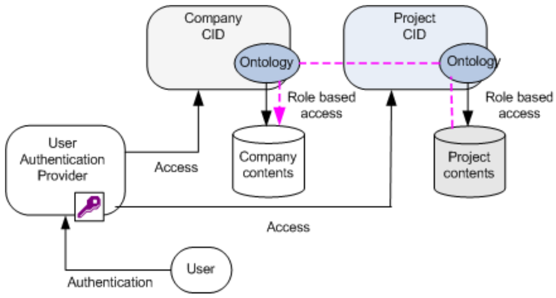


Figure 2: Generic architecture on seamless authentication and role-based service access for a corporate scenario.

between the members of a social community, and establishes access rights to contents and services. The example uses social relations (SID) of Maria, and corporate (here: school) relations (CID) of Paul. Service realisations based on social identities will use the semantic description of relationships, preferences and context information. Service access requiring PID information is subject to user involvement, as outlined in the next section.

4 Prototypical implementation of key distribution

As a key of identity management in our proposal, a mobile based key exchange demonstrator was built [Noll 2006]. The key generation and distribution was modified to support requirements of social communities. The authentication system transmits the authentication keys through the mobile phone system to the mobile terminal. The terminal can either access services based on that key or perform a user identification In our scenario the user wants to get access to remote content. The access request is sent (1) and the access control system of the data service sends (2) a message to the Service Centre. This entity acts sends down (if needed) the required application (3) and a binary key (4). The key is stored in the integrated SmartMX card of the phone and can be transmitted over the NFC interface (5) to use the remote content. Our implementation uses Nokia 3320 mobile phones and keys distributed through Telenor's Innovation lab

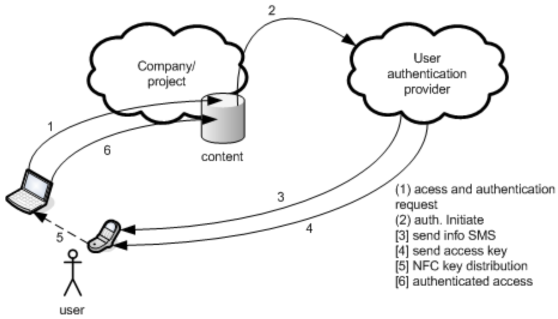


Figure 3: Prototype of key handling for content access control

PATS[2] .

5 Conclusions

The pervasive Internet has enabled service access in every situation. However, adaptation to the user needs is purely handled. The paper introduces a semantic description of user preferences and social relations describing the identity of the user. Sensitive information is suggested to be stored in the SIM card of the user's personal device, while preferences and social relations are stored in the network. Authentication is the key issue for role-based document and service access. A SMS-based key distribution demonstrates how access or group keys can be distributed through the mobile network and used for contactless authentication.

References

[Arbanowski et. al. 2004] S. Arbanowski, P. Ballon, K. David, O. Droegehorn, H. Eertink, W. Kellerer, H. van Kranenburg, K. Raatikainen, and R. Popescu-Zeletin, "I-centric Communications: Personalization, Ambient Awareness, and Adaptability for Future Mobile Services", IEEE Comm. Magazine, Sep 2004, pp 63-69

[Ball State 2006] Ball State study finds computer usage trails only television viewing, News center from Ball State University, `http://www.bsu.edu/news/article/0,1370,-1019-45461,00.html`, [accessed 17.12.2006, 20:41h]

[Chowdhury and Noll 2007] M. M. R. Chowdhury, J. Noll, "Distributed Identity for Secure Service Interaction", The Third International Conference on Wireless and Mobile Communications, ICWMC07, March 4-9, 2007-Gaudeloupe, French Caribbean.

[Noll 2006] J. Noll, "Services and applications in future wireless networks", Telektronikk 3/4.2006, pp 61-71

[Roccas and Brewer 2002] S. Roccas, M. B. Brewer, "Social Identity Complexity", Personality and Social Psychology Review, 2002, Vol. 6, No. 2, 88-106

[World factbook 2006] The World factbook 2006, `https://www.cia.gov/cia/publications/factbook/geos/gm.html`, [accessed 17.12.2006, 20:23h]

[2] `http://www.pats.no`

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Business Process Knowledge Integration – A Semantic Based Approach

Janez Hrastnik
(University of Madeira, Portugal
hrastnik@uma.pt)

Jorge Cardoso
(University of Madeira, Portugal
jcardoso@uma.pt)

Abstract: Knowledge necessary for the creation of business process models is distributed, consists of different types, and expresses different levels of abstraction. Its acquisition and collection into a common knowledge base, which implies integration into a single model, is the goal of the approach we are proposing. In this paper a framework for the integration of business process knowledge is proposed. It is shown how semantic technologies can contribute to the integration of different models, which represent different aspects of an organization, in order to create a more expressive model of business process knowledge.

Key Words: business process management, interoperability, semantics, ontology
Categories: I.2.4, I.2.6, K.4.3

1 Introduction

The knowledge required for the creation of process models is distributed among people, documentation, and systems and must be collected at several levels within and outside of the organization. This knowledge is heterogeneous, isolated, is expressed at different levels of abstraction and scopes, and is often contradictory. The critical connections or dependencies are not always obvious or defined. On the other hand, process models, even if they are feasible, are mostly designed for only one stakeholder and are therefore described in an inappropriate description language, and have inappropriate scope and form in order to be used as a proper means of communication with different stakeholders. Process models that cannot take into account process stakeholders' perspectives can hardly be expected to be easily communicated and properly applied.

For the successful management of process knowledge (*knowledge about the organizational processes*) it is necessary to treat it in a broader context [Hrastnik, 2007]. One possible solution for the generation of an integrated and more expressive business process knowledge (BPK) model that aggregates stakeholder's perspectives is to rely on semantic technologies [Hepp, 2005]. Semantic technologies can extend the capabilities and value of information [Noy, 2001]. Models with richer semantic relationships and strict rules offer a more powerful and flexible basis for knowledge integration, analysis and communication.

In this paper we propose an extendable framework for BPK integration, which takes advantage of semantic technology to enable the mapping and integration of different models into a common knowledge base with the help of an BPK ontology. Those independent models are parts of BPK and represent aspects, dimensions, or abstraction levels of an organization. Once integrated in a new model, they provide a better basis for improved business analysis, the creation of more suitable process models and improved process knowledge communication.

2 Motivation

Current enterprise and business modeling approaches, tools, and modeling languages which in specific contexts and different focuses address business process modeling can only partialy fulfill important requirements regarding integration and communication of BPK. Established process modeling languages (e.g. BPMN, BPEL, UML) do not cover all the important concepts or knowledge elements [Fadel, 2005] which critically contribute to the expressiveness of process models. Due to the fact that these modeling languages were designed for specific purposes and with specific objectives, they only can support a limited number of perspectives [Curtis, 1992]. Extensions to modeling languages have, however, been proposed (e.g. process goals and performance measures [Korherr, 2007]). The enterprise modeling approaches which usually include the process aspect do not offer methods for its integration with the other aspects of the organization to be modeled. The Zachman Framework [Zachman, 1987]) offers classification and proposes modeling languages for different layers only as examples. Therefore the “bridges” between different layers cannot easily be established. Also, popular enterprise modeling tools (e.g. Aris Process Platform[1] or ProVision Modeling Suite[2]) do not integrate the various models or provide only very loose connections between them. Some approaches which offer comprehensive integration (e.g. MEMO [Frank, 2002]) introduce new modeling languages for perspectives and aspects of the organization to be modeled. Solutions applying semantic technologies have already been proposed (e.g. SUPER Project[3]), but they don’t provide any directions regarding knowledge basis instantiation.

3 Integration Approach

To enable a systematic approach, the BPK framework was proposed [Hrastnik, 2004]. It considers BPK as a super set of conventional business process models and also includes knowledge about the motivation behind processes, reasons for their existence, knowledge about constraints, the required resources for their execution, as well as its interfaces, process environment, capability and performance. The systematic approach includes the following management steps: acquisition, synthesis, and communication. In this paper, as a part of the synthesis step, we propose BPK integration.

In our solution we propose a pragmatic procedure for achieving better process models that do not require additional effort or changes to workflow within the organization. The goal of building better business process models and knowledge can

[1] http://www.ids-scheer.com
[2] http://www.proformacorp.com/
[3] http://www.ip-super.org/

be achieved iteratively by enriching the process models designed with information captured in other knowledge sources already available to the organization. Different organizational roles, either at the strategic or operational level, design and model various aspect of the organization, regarding the important aspects of their work (e.g. strategy, people, resources, data). In order not to require additional work or learning new tools or modeling languages, we use those specific models in the form in which they usually already available (e.g. from standard business software file formats). The specific models describe an aspect or layer of the organization or business (e.g. goals tree, value chains, organizational chart) at different levels of abstraction or describe a specific view of it. They often include knowledge which represents a relevant part of or constraints for the BPK model.

In support of the creation of a common BPK model, all applicable knowledge captured in specific models is mapped to a common description language and integrated into a single model. How different knowledge elements from specific models relate to each other and to the knowledge elements in the knowledge base is defined by a BPK meta model, which in our case is an ontology. Several business process and enterprise ontologies (e.g. [Jenz, 2003], [Uschold, 1998] [Fox, 1997]) have been already proposed and can be adopted for the framework application.

In the integration process (for the knowledge base) relevant knowledge elements of the specific models are merged with models already available in the knowledge base (e.g. process goals with process models, process models with their corresponding measurement categories and indicators). It is important to consider that knowledge elements of different models often overlap. Therefore, additional merging rules, which are not part of the ontology, have to be defined. Depending on the usage of the knowledge base, the extraction and backward mapping into a specific model structure format is often needed. Because of the transformation from a semantically richer model to a more specific format, special consideration is required to perform this step.

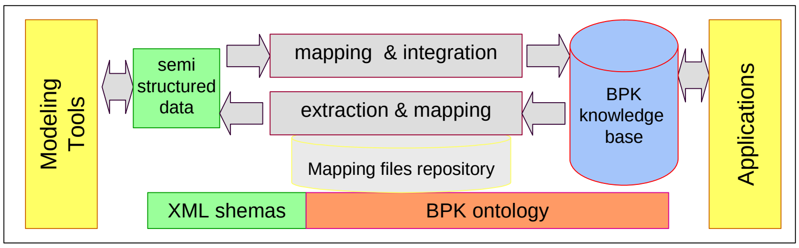


Figure 1: System Architecture

4 Overall System Architecture

The system architecture for the implementation of the integration framework comprises of the following basic parts (Figure 1): (a) business process knowledge ontology, (b) XSLT mapping files repository, (c) mapping and integration component, (d) extraction and mapping component.

The prototype tool under development implements the described architectural design. The BPK ontology was designed in the RDF/OWL format. Presently the prototype provides basic integration functionality and supports the mapping of specific models for processes (in XPDL), organizational and role charts, business and process goals, measurement categories and indicators representative of the process functional, organizational, and strategic perspective.

5 Conclusion and Future Work

In this paper we propose an adaptive framework for BPK integration, which enables the mapping and integration of different organization aspect models into a common knowledge base. The framework is extensible without computer programming in several ways (e.g. adding of new particular models, adopting or exchanging of ontology). Our future work includes completing the framework prototype, focusing especially on defining the merging rules to solve the problem of overlapping particular models.

References

[Curtis, 1992] Curtis B., Kellner M.I., Over J. Process modeling. Communications of the ACM, ACM Press 1992;35(9):75-90.

[Fadel, 2005] Fadel K.J., Tanniru M. A Knowledge-Centric Framework for Process Redesign. ACM SIGMIS CPR. Atlanta, Georgia, USA: ACM Press; 2005.

[Fox, 1997] Fox M.S., Grüninger M. On Ontologies and Enterprise Modelling. ICEIMT 97: Springer- Verlag; 1997.

[Frank, 2002] Frank U. Multi-perspective Enterprise Modeling (MEMO) - Conceptual Framework and Modeling Languages. HICSS'02. Honolulu: IEEE USA; 2002.

[Hepp, 2005] Hepp M., Leymann F., Domingue J., Wahler A., Fensel D. Semantic Business Process Management: A Vision Towards Using Semantic Web Services for Business Process Management. ICEBE; 2005 October 18-20; Beijing, China; 2005.

[Hrastnik, 2007] Hrastnik J., Cardoso J., Kappe F. The Business Process Knowledge Framework. ICEIS. Funchal, Madeira; 2007.

[Hrastnik, 2004] Hrastnik J., Rollett H., Strohmaier M. Heterogenes Wissen über Prozesse als Grundlage für die Geschäftsprozessverbesserung, Prozesswissen als Erfolgsfaktor, Deutscher Universitätsverlag; 2004.

[Jenz, 2003] Jenz D.E. Business Process Ontologies: Speeding up Business Process Implementation: Jenz & Partner GmbH; 2003.

[Korherr, 2007] Korherr B., List B. Extending the EPC and the BPMN with business process goals and performance measures. ICEIS. Funchal, Madeira; 2007.

[Noy, 2001] Noy N.F., McGuinness D.L. Ontology Development 101: A Guide to Creating Your First Ontology. In: Stanford University; 2001.

[Uschold, 1998] Uschold M., King M., Moralee S., Zorgios Y. The Enterprise Ontology. The Knowledge Engineering Review 1998.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Enabling Collaborative Knowledge Modelling by Hiding Inherent Complexity from Non-Experts

Daniel Doegl
(uma information technology Gmbh, Austria
daniel.doegl@uma.at)

Claire Jessel
(uma information technology Gmbh, Austria
claire.jessel@uma.at)

Abstract: Semantic technologies which visualise complete ontologies in network form are manifold; yet such semantic networks remain for the most part exclusively within the grasp of domain experts. In this paper, we suggest that semantic networks representing foreign domains can indeed present many benefits to the common end-user, as long as certain semantic objects from the domain are extracted and only these core objects are displayed to the user. This hypothesis we claim to test using our own self-developed knowledge modelling tool. Indeed, not all representations of foreign domains must overstrain the end-user. The visualisation must be object-centred on a semantic level, such that ontology objects which carry a meaning are represented. In this way, minimal explanation is needed to make it accessible to any end-user, whether domain expert or domain layman, and this increased understanding on behalf of the user turns the exercise of knowledge modelling into a collaborative activity.

Keywords: Knowledge Modelling, Semantic Network, Ontology, Visualisation, User
Categories: H.1.0, H.5.1, H.5.2

1 Introduction

Ontologies are large and complex hierarchies, containing a huge number of entities and relations linking these entities, as well as domain knowledge in the form of rules and axioms. Such large quantities of data overstrain any end-user, who is not familiar with the ontology domain itself, and render the ontology unintelligible and inaccessible.

One problem with which the user is confronted is orientating himself in this large ontology consisting of numerous relations. Hence, semantic technologies aiming to visualise ontologies were developed. These represent ontologies as semantic networks, with the entities as network nodes and the relations between these entities as network edges. Such a visualisation helps the user find his way in this semantic network, enabling him to view the various clusters and traverse the nodes following the relation paths, and hence facilitates overall comprehension of the ontology. In Figure 1 such a semantic network is displayed. It contains three clusters, each with its own cluster centre. Within each cluster all elements are semantically and formally connected. Depending on the user's current main interests, the network can be re-clustered so as to change the cluster centres.

While certain visualisation techniques do succeed in making the ontology slightly less overwhelming to the user, such as the CO-ODE project led by the University of Manchester (http://www.co-ode.org/), the inherent complexity of the ontology remains unchanged whatever the visualisation technique: the number of concepts and relations stays very large, rendering the semantic network rather unreadable to a non-expert. Hence, methods of reducing this complexity must be found, as this is a condition upon which semantic networks can become the object of a social and collaborative exercise [Rector and Seidenberg 2007]. One possibility we suggest consists in extracting certain "core objects" from the ontology domain, i.e. relevant objects which carry a meaning to the non-expert and only display these. Such an object-centred visualisation clearly reduces the complexity level as it fails to represent large parts of the ontology. However it must be ensured that the visualisation remain coherent and understandable by the end user despite the omissions; otherwise the underlying ontology appears inconsistent to the end-user.

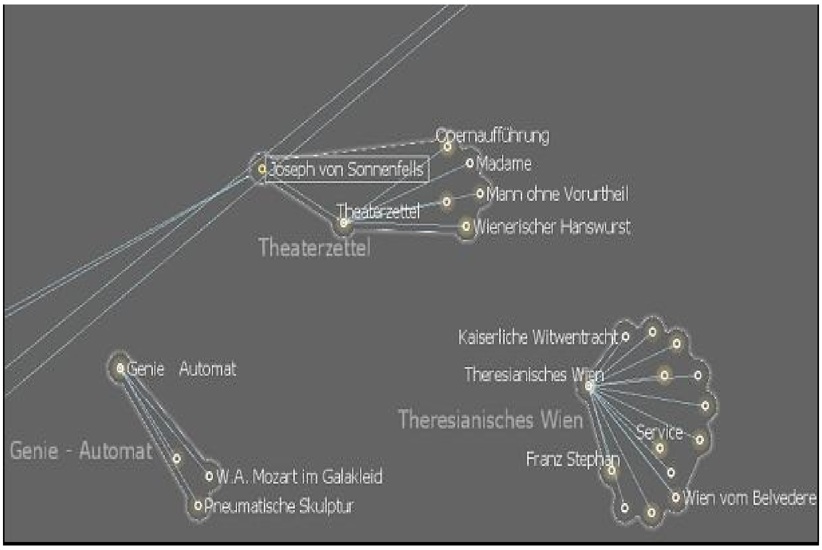


Figure 1: Network for Topic "Mozart"

2 Problematic Issues in Reducing Complexity

Simply ignoring large parts of the ontology in the visualisation is bound to raise many issues, with respect to selection criteria, user comprehension, network readability, consistency and soundness etc. The following sub-sections discuss a few of these issues which must be considered before dismissing part of the ontology in a visualisation.

2.1 Selection of Entities

In the type of reduced visualisation we suggest, only a small set of the entities from the underlying ontology is represented, the selection of which is crucial to the visualisation's quality and the user's understanding. Depending on the end user's needs and interests, certain entities are more relevant and carry more significance than others. We refer to these as "core objects". Our assumption is that if all such semantic objects are extracted from the ontology domain and displayed to the user, the complexity of the visualisation is largely diminished, without affecting the user's comprehension. The visualisation becomes object-centred on a semantic level.

The question remains: how to identify such "core objects"? The ontology consists of a very large range of concepts and relations, all referred to as "base set". The aim is to extract a subset of the ontology which fits the current interests of the user. By symbolising the elements so as to ensure their unique unambiguous interpretation, elements irrelevant to the user needs can be hidden from the visualisation to be accessed only via a drill-down process, rendering the remaining visualisation more readable and less complex without it losing any soundness or consistency. The elements in the final subset can be referred to as 'core objects'.

2.2 Selection of Relations

In an ontology, entities may be linked by several relations, varying in their type and number. Each relation carries a particular meaning and contributes therefore to the overall coherence of the ontology, despite their potentially overwhelming number. Yet, if only a restricted set of the ontology entities is being represented as a semantic network, clearly the set of relations represented must also be restricted. A straightforward restriction is to ignore all relations which do not involve at least one entity from the selected set of entities. Problems arise with relations which connect two elements, one of which is not in the selected set of entities, as the relation's contribution to overall comprehension must be evaluated first before a decision concerning its potential selection can be taken. Also, it may be the case that not all relations linking two entities in the set are necessary; ensuing restriction criteria could include relation type or relation weight. Nevertheless, the resulting graph must remain a connected graph.

To proceed with the method suggested in part 2.1, the 'core relations' are taken to be those which are used in the final subset.

2.3 Importance of Context

By following this hypothesis according to which only a restricted set of entities and relations must be represented so as to reduce inherent complexity and thereby increase user comprehension, the importance of context remains unconsidered. Context can be generally defined as an environment containing intrinsic entities and relations and unique domain knowledge, thereby providing a meaning for any particular entity, event or relation. Hence, for any form of human comprehension, context clearly plays an important role.

To further the method suggested in both parts 2.1 and 2.2, user needs and interests are always relative to a particular context. Hence context is a type of filter which is applied on the semantic network. Each concept in the ontology is associated with one or several contexts (eg. “economics”, “politics”, “social”); the user selects the context in which he is modelling knowledge, and the appropriate concepts are automatically selected. These concepts form the base of the subset selection described in parts 2.1 and 2.2. Indeed, in different contexts, the final set of elements displayed to the user will clearly be different, as each intersection result set is unique. If for instance the topic is “Mozart” and the context is “Health”, then all concepts such as “requiem”, “instruments”, “talent” are not going to be ‘core elements’ as they are not inherent elements to the ‘Health’ context. Establishing and defining a context is therefore primordial when identifying which concepts are “core objects” and which are not.

By definition, an ontology is a hierarchical collection of entities representing a knowledge domain. This domain consists of many sub-domains, each of which contains a number of entities and relations. The user will rarely be interested in the whole ontology domain; rather he will be interested in one particular sub-domain of this ontology. Hence, restricting the displayed nodes and edges to the elements in the concerned sub-domain, must not infringe on the user’s understanding of the visualisation, as the other sub-domains may not contribute in any significant way to comprehension. This implies that despite hiding parts of the ontology in the visualisation, the view must remain consistent.

3 Conclusions

The level of complexity in ontologies must be decreased, so as to become accessible to the end-user. Network representations of ontologies partly fulfill this goal, in that the navigation possibilities offered aid the user by giving him visual anchor points. However, inherent complexity remains unchanged. Our hypothesis discussed in this paper is that by extracting certain elements of the ontology and representing exclusively these, complexity can be reduced without any negative effect on coherence or soundness. Moreover this reduction of the complexity encourages the user to contribute knowledge to this small semantic network, in the form of new relations and concepts, hence enlarging the shared domain knowledge of the ontology. The idea may sound simple, yet many problems pose themselves, as to how to best select these elements. These problems were raised in the paper, as entry points for further research, rather than as solution proposals. If a technique were found to reduce this aforementioned complexity, all semantic networks as well as ontologies would be rendered accessible to any end-user with only a minimal need for explanations. Semantic networks could become an every day tool for knowledge modelling accessible to a very wide range of users, who would then become involved in the modelling process, thereby turning it into a collaborative and social exercise.

References

[Rector and Seidenberg 2007] Rector A., Seidenberg J.: “A Methodology for Asynchronous Multi-User Editing of Semantic Web Ontologies”; to appear in Proc. KCAP 2007, ACM Press, New York, USA.

Proceedings of I-MEDIA '07 and I-SEMANTICS '07
Graz, Austria, September 5-7, 2007

Smart Media Archive – An Approach to Open Archives to Communities

Dietmar Glachs
(Salzburg Research Forschungsgesellschaft m.b.H, Austria
dietmar.glachs@salzburgresearch.at)

Daniel Mitterdorfer
(Salzburg Research Forschungsgesellschaft m.b.H, Austria
daniel.mitterdorfer@salzburgresearch.at)

Abstract: Nowadays, professional Media Asset Management systems are a standard tool for many organizations to manage their media assets. The classification of media must be sophisticated enough to remain stable throughout several decades. However, even the most sophisticated classification scheme covers only particular aspects of the archive domain. Furthermore, the public understanding of the domain may evolve over time. In our approach, the system is opened to communities which classify the media themselves building their own folksonomy. The folksonomy is analyzed using semantic technologies and then compared to the organization's classification. Using this approach, the organization obtains two key benefits. Firstly, knowledge can be adapted as the community notices different aspects of the organization's media. Secondly, the organization can gain more insight about their community and can use this experience as a direct input in the product innovation process and therefore increase market orientation as well as customer retention.

Keywords: Archive Systems, Customer Retention
Categories: H3.3, H3.7

1 Motivation

Popular community portals like Flickr or YouTube show that there is high potential in communicating with communities and learning from them. Moreover, many organizations maintain archives containing multimedia assets and other information assets of interest for their communities. However, by primarily supporting the professional discourse of the given archive domain, the public cannot easily benefit from the provided information due to the semantic gap between professional users and the public thinking of the archive's domain [Trant 2006]. By establishing a community view on the archive, organizations can share their knowledge with their communities, attracting them to participate and even get in contact with the organization itself: Bring the archive to the communities.

In addition, communities can often provide knowledge (e.g. stories, photos …) that is otherwise not available to the organization, but adds significant value to the archive. For example, the archive of a German automotive manufacturer has grown over several decades and contains a lot of material covering several aspects, such as their racing activities or the company's history. Due to the company's activities, different communities have emerged. As the archive is only internal, the experience of the communities evolved almost independently from the experience gained by the company itself. Consequently, both parties have different viewpoints on the domain.

For detecting previously not visible aspects in their material, the company intends to invite the communities to support the archive's evolution. Newly detected aspects to the archive should then support the company's activities: they can launch marketing campaigns or PR activities based on new insights about their community, increase customer retention or thoughtfully adapt the company's view on the archive.

In addition to "informal" knowledge, communities can also contribute "formal" knowledge by annotating content with semantic metadata which would otherwise be very time-consuming and expensive for an organization: Feed the communities' knowledge back to the organization and to the archive.

However, organizational archives have a particular purpose such as supporting marketing or public relation activities on the one hand and storage and retrieval of organizational assets on the other hand. Retrieval of media assets is crucial for the proper function of the archive. Therefore, professional archivists build and maintain an organizational view on the archive which must not be affected by the community activities: Retain the internal archive consistency.

1.1 Related Approaches

Recent efforts to open archives to their community were made in different contexts. As an example, the Steve[1] project allows for social tagging of art images. The community mainly benefits from the tagging process with improved search results. Other efforts such as the current IST project QVIZ[2] target the utilization of social tools like Wiki systems to access the information in a user-adapted context.

The ongoing effort in these two and other related projects shows that both approaches – Wikis and Social Tagging – are well-understood, proven technologies to let the public participate in open archives. In this paper we describe how we intend to apply these technologies in order to support organizations in (re)discovering unrevealed knowledge with the help of their communities

2 Target System – Vision

The proposed target system has to fulfill the above-mentioned requirements in retaining the initial archive system functionality and by extending a given system with community functionality. Figure 1 depicts the "core media archive" usually consisting of the media assets ① connected with a formal knowledge model ②.

To plug in new functionality we add a new layer to the archive system. Beside the organizational knowledge, the collected community knowledge ③ is established and used when annotating media assets. Simple personal and free tagging of media assets is used to build a folksonomy. The combination of a media asset with community aspects is referred to as Smart Media Asset ④. Furthermore, the community knowledge collected is analyzed and compared to the organizational knowledge. Identified semantic gaps will then be reported to the organization ⑤.

[1] http://www.steve.museum (last visited 5th April 2007)
[2] http://www.qviz.eu (last visited 5th April 2007)

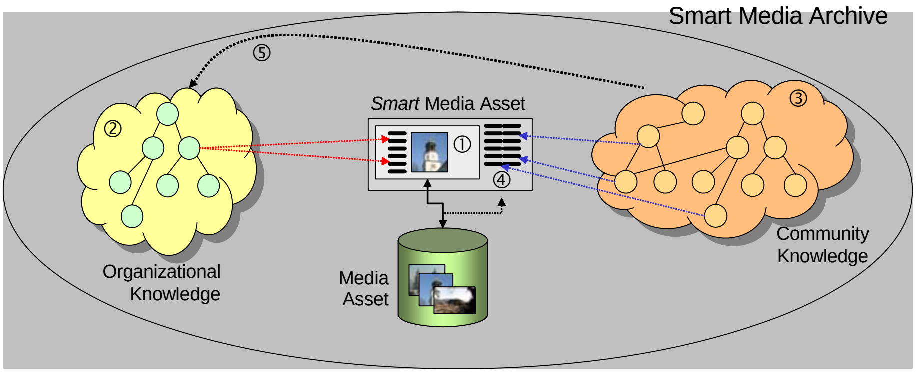


Figure 1: Smart Media Archive Overview

Each media asset gains additional knowledge through its community extension. This community added knowledge is then used to support searching and to identify gaps between the organizational view and the community view(s). Search results are more specific to a given user group because their annotations are considered. Personalized annotations reflect different associations of different communities.

2.1 Smart Media Asset

In order to enrich existing media assets with new functionality, we propose to expand the media assets with new means of classification. In addition to the organizational driven categorization of media assets, we intend to establish a community driven ontology for interlinking with media assets based user annotations.

2.2 Smart Media Asset Wiki

Our approach is to open the archive to the public allowing users to comment, annotate and add stories about media assets. In order to not interfere with the traditional part of the smart media archive, each media asset must also be accessible to the public by allowing editing the media asset's community extension. We follow the approach described in [Schaffert 2006] which allows the creation of community authored knowledge models by facilitating semantic Wikis thus provide a Semantic Media Asset Wiki to the communities.

One important community contribution is the collective knowledge building by populating an ontology covering the archive domain. This resulting ontology may or may not highlight aspects of the archive which are different from the organizational view. It is therefore required to provide means for detecting gaps between the distinct views on the archive.

2.3 Knowledge Feedback Recommender

[Mika 2005] states, "community-based ontology extraction has a great potential in extracting ontologies that more closely match the conceptualization of a particular community". By analyzing the annotations of the community we aim to detect new insights which may not necessarily conform to the organizational view of the archive domain. This gap is made accessible to the organization's archivist via recommendations who can then decide to either adapt the organization's view or not.

However, the recommender technology is subject to further research as a folksonomy does not directly permit applying semantic technologies such as reasoning. Since reasoning is only possible over a formal knowledge model (e.g. facts and rules of an ontology) we need to first transform the terms of the folksonomy into valid ontological statements. An especially promising approach is the integration of Wordnet and DOLCE, as described in [Gangemi et.al. 2003].

3 Conclusions

Combining knowledge of communities with organizational knowledge provides benefits for both parties. Our solution is based on three paradigms:

1. Bring the archive to the communities
2. Feed the communities' knowledge back to the organization and to the archive
3. Retain the internal archive consistency

We showed how to extend an existing media archive system with community functionality whilst preserving the internal archive functionality. Using this approach, organizations can reveal knowledge captured in existing media archives based on community contributions. Semantic technologies enable to incorporate new knowledge from the community contributions. This knowledge can be applied in many different ways, such as adapting the internal view on the archive or by supporting marketing and PR activities with information tailored to target audiences.

Acknowledgements

The reported work has been carried out at Salzburg NewMediaLab, a competence centre in the K_{ind} action line funded by the Austrian Federal Ministry of Economics an Labour (BMWA), the state of Salzburg and by the project partner "mediamid digital services GmbH".

References

[Gangemi et.al. 2003] Gangemi A., Guarino N., Masolo C. and Oltramari A.: "Sweetening WordNet with DOLCE"; AI Magazine 24, 3 (2003)

[Mika 2005] Mika P.: "Ontologies are us: A unified model of social networks and semantics"; Proc. ISWC 2005, Springer (2005), 522-536.

[Schaffert 2006] Schaffert S., "Semantic Social Software, Semantic Systems - From Visions to Applications"; Proc. Semantics 2006, OCG (2006), 99-112.

[Trant 2006] Trant J., "Social Classification and Folksonomy in Art Museums"; 17th SIG/CR Classification Research Workshop (2006)

Author Index