Witold Abramowicz
Robert Tolksdorf
Krzysztof Węcel (Eds.)

LNBIP 57

Business Information Systems Workshops

BIS 2010 International Workshops
Berlin, Germany, May 2010
Revised Papers

Springer

Lecture Notes in Business Information Processing 57

Series Editors

Wil van der Aalst
Eindhoven Technical University, The Netherlands

John Mylopoulos
University of Trento, Italy

Michael Rosemann
Queensland University of Technology, Brisbane, Qld, Australia

Michael J. Shaw
University of Illinois, Urbana-Champaign, IL, USA

Clemens Szyperski
Microsoft Research, Redmond, WA, USA

Witold Abramowicz Robert Tolksdorf
Krzysztof Węcel (Eds.)

Business Information Systems Workshops

BIS 2010 International Workshops
Berlin, Germany, May 3-5, 2010
Revised Papers

Springer

Volume Editors

Witold Abramowicz
Krzysztof Węcel
Poznań University of Economics
Department of Information Systems
Al. Niepodległości 10, 61-875 Poznań, Poland
E-mail: W.Abramowicz@kie.ue.poznan.pl;Krzysztof.Wecel@ue.poznan.pl

Robert Tolksdorf
Freie Universität Berlin
Institut für Informatik
Takustr. 9, 14195 Berlin, Germany
E-mail: tolk@ag-nbi.de

Library of Congress Control Number: 2010934236

ACM Computing Classification (1998): J.1, H.3.5, H.4

ISSN 1865-1348
ISBN-10 3-642-15401-8 Springer Berlin Heidelberg New York
ISBN-13 978-3-642-15401-0 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

springer.com

© Springer-Verlag Berlin Heidelberg 2010
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 06/3180 5 4 3 2 1 0

Preface

Business information systems is a rapidly developing domain. There are many topics that deserve attention but have not yet found a place in canonical research. Workshops give researchers the possibility to share preliminary ideas, first experimental results, or to discuss research hypotheses. Discussions held during presentations strengthen the paper and prepare it for publication. From our experience, workshops are a perfect instrument with which to create a community around very specific research topics, thus offering the opportunity to promote it. Topics that do not find critical feedback at the main International Conference on Business Information Systems (BIS) may experience fruitful discussion when confronted with a well-focused audience.

Over the last few decades, business information systems have been one of the most important factors of the transition toward a knowledge-based economy. At the same time they have been subject to continuous rapid development and innovation driven both by industry and by academia. For the last 12 years these innovations were carefully observed but also shaped by researchers attending BIS yearly.

Recently, apart from the main conference, workshops covering specific topics in the area of business information systems have been organized. Since 2007, BIS has featured several workshops. Their proceedings were published on-line, later in the LNBIP series, and outstanding workshop papers were included in special issues of international journals. This year, four workshops were successfully organized in conjunction with BIS 2010, covering topics of applications and economics of knowledge-based technologies (AKTB), business and IT alignment (BITA), information logistics (ILOG), and legal information systems (LIT). This volume contains papers that were accepted and presented at the BIS 2010 workshops. Additionally it features an invited speech given at the LIT workshop by Tom Gordon.

Workshop papers included in this volume underwent a through review procedure. Each of the submitted papers received between two and three reviews (a total of 172 reviews were received, with an average of 2.5 reviews per paper) by over 89 members of the Program Committees of the workshops. Out of 74 submitted papers and work-in-progress reports, 33 were accepted and presented at the conference (accounting for 48% of all submissions).

On this occasion, we would like to express our thanks to everyone who made the BIS 2010 workshops a success: all Workshop Chairs, members of the Workshop Program Committees, authors of submitted papers, invited speakers and finally all workshop participants. We cordially invite you to visit the BIS website at http://bis.kie.ue.poznan.pl/ and to join us at future BIS conferences.

May 2010

Witold Abramowicz
Robert Tolksdorf
Krzysztof Węcel

BIS 2010 Conference Organization

Program Co-chairs

Witold Abramowicz, Robert Tolksdorf

Organizers

Freie Universität Berlin, Networked Information Systems
Poznań University of Economics, Department of Information Systems
`http://bis.kie.ue.poznan.pl`

Organizing Committee

Birgit Hadel	is-research GmbH, Germany
Ralf Heese	Freie Universität Berlin, Germany
Szymon Łazaruk	Poznań University of Economics, Poland
Monika Starzecka	Poznań University of Economics, Poland
Piotr Stolarski	Poznań University of Economics, Poland
Małgorzata Mochól (Co-chair)	T-Systems Multimedia Solutions GmbH, Germany
Adam Walczak	Poznań University of Economics, Poland
Krzysztof Węcel (Co-chair)	Poznań University of Economics, Poland

BIS 2010 Workshop Organization

AKTB 2010 – Second Workshop on Applications of Knowledge-Based Technologies in Business

Workshop Chairs

Kriksciuniene, Dalia	Vilnius University, Lithuania
Sakalauskas, Virgilijus	Vilnius University, Lithuania

Workshop Organizers

Vilnius University, Lithuania Kaunas Faculty, Department of Informatics, Lithuania

Program Committee

Schoop, Eric	Dresden University of Technology, Germany
Gudas, Saulius	Vilnius University, Lithuania
Krebs, Irene	Potsdam University, Germany
Simutis, Rimvydas	Vilnius University, Lithuania
Meinberg, Uwe	Cottbus Technical University, Germany
Kiss, Ferenc	Foundation for Information Society, Hungary
Xavier, Adilson Elias	COPPE- Federal University of Rio de Janeiro, Brazil
Sakalauskas, Leonidas	Institute of Mathematics and Informatics, Lithuania
Bassa, Lia	Foundation for Information Society, Hungary
Augutis, Juozas	Vytautas Magnus University, Lithuania
Gustas, Remigijus	Karlstad University, Sweden
Lee, Soo-Young	KAIST, Korea
Yin, Hujun	University of Manchester, UK
Jordaan, Jaco	Tshwane University of Technology, South Africa
Zhang, Du	California State University, USA
Corchado, Emilio	University of Burgos, Spain
Shi, Yong	Graduate University of the Chinese Academy of Sciences, China
Tan, Margaret	Nanyang Technological University, Singapore
Mazeika, Arturas	Free University of Bozen-Bolzano, Italy
Bauerle, Peter	IBM Deutschland, Germany
Bergeron, Francois	Universite Laval, Canada

BITA 2010 – First Workshop on Business and IT Alignment

Workshop Chairs

Seigerroth, Ulf	Jönköping University, Sweden
Sandkuhl, Kurt	Jönköping University, Sweden

Workshop Organizers

Jönköping University, Sweden

Program Committee

Butleris, Rimantas	Kaunas University of Technology, Lithuania
Deiters, Wolfgang	Fraunhofer ISST, Germany
Dubois, Eric	CRP Henri Tudor, Luxembourg
Ekstedt, Mathias	KTH, Sweden
Goldkuhl, Göran	Linköping University, Sweden
Gustas, Remigijus	Karlstad University, Sweden
Halpin, Terry	LogicBlox, Australia
Holmberg, Lennart	Metria, Sweden
Hoppenbrouwers, Stijn	Radboud University Nijmegen, The Netherlands
Jørgensen, Håvard	Commitment AS, Norway
Kaczmarek, Tomasz	Poznan University of Economics, Poland
Kirikova, Marite	Riga Technical University, Latvia
Krogstie, John	Norwegian University of Science and Technology, Norway
Lankhorst, Marc	Novay, The Netherlands
Lind, Mikael	Borås University, Sweden
Opdahl, Andreas	University of Bergen, Norway
Peppard, Joe	Cranfield School of Management, UK
Persson, Anne	Skövde University, Sweden
Proper, Erik	Radboud University Nijmegen, The Netherlands
Rosemann, Michael	Queensland University of Technology, Australia
Shilov, Nikolay	SPIIRAS, Russia
Stirna, Janis	KTH, Sweden
Ward, John	Cranfield University, UK

ILOG 2010 – Third Workshop on Information Logistics and Knowledge Supply

Workshop Chairs

Deiters, Wolfgang	Fraunhofer ISST, Dortmund and Berlin, Germany
Mirijamdotter, Anita	Linnæus University, Sweden
Sandkuhl, Kurt	Jönköping University, Sweden

Workshop Organizers

Fraunhofer ISST, Dortmund and Berlin, Germany
Linnæus University, Sweden
Jönköping University, Sweden

Program Committee

Blomqvist, Eva	CNR-ISTC, Italy
Bucher, Tobias	PHOENIX group, Germany
Busse, Susanne	FH Brandenburg, Germany
Deiters, Wolfgang	Fraunhofer ISST, Dortmund and Berlin, Germany
Eriksson, Henrik	Linköping University, Sweden
Haftor, Darek	Linnæus University, Sweden
Han, Yanbo	Chinese Academy of Science, China
Kajko-Mattsson, Mira	Stockholm University, Sweden
Karim, Ramin	Luleå University of Technology, Sweden
Koch, Oliver	Fraunhofer ISST, Germany
Kutsche, Ralf-D.	TU Berlin, Germany
Levashova, Tatiana	St. Petersburg Institute for Informatics and Automation, Russia
Mirijamdotter, Anita	Linnæus University, Sweden
Sandkuhl, Kurt	Jönköping University, Sweden
Seigerroth, Ulf	Jönköping International Business School, Sweden
Smirnov, Alexander	St. Petersburg Institute for Informatics and Automation, Russia
Stirna, Janis	Stockholm University, Sweden
Tarasov, Vladimir	Jönköping University, Sweden
Willem, Jan	Nashuatec/Ricoh Group, The Netherlands

LIT 2010 – Third Workshop on Legal Informatics and Legal Information Technology

Workshop Chairs

Stolarski, Piotr	Poznan University of Economics, Poland
Tomaszewski, Tadeusz	Poznan University of Economics, Poland
Zeleznikow, John	Victoria University, Australia

Workshop Organizers

Poznan University of Economics, Department of Information Systems, Poland

Program Committee

Bellucci, Emilia	Victoria University, Australia
Bourcier, Daniele	CNRS-CERSA, Germany
Brodecki, Zdzisław	University of Gdańsk, Poland
D'Agostini Bueno, Tania Cristina	Presidente da Diretoria Executiva IJURIS, Brazil
Lodder, Arno	Free University of Amsterdam, The Netherlands
Morek, Rafal	University of Warsaw, Poland
Poblet, Marta	Autonomous University of Barcelona, Spain
Schweighofer, Erich	University of Vienna, Austria
Stolarski, Piotr	Poznan University of Economics, Poland
Stranieri, Andrew	University of Ballarat, Australia
Szumlicz, Tadeusz	Szkoła Główna Handlowa w Warszawie, Poland
Tomaszewski, Tadeusz	Poznan University of Economics, Poland
Zeleznikow, John	Victoria University, Australia

Table of Contents

AKTB Workshop

BITA Workshop

ILOG Workshop

LIT Workshop

AKTB 2010 Workshop Chairs' Message

Virgilijus Sakalauskas and Dalia Kriksciuniene

Vilnius University, Lithuania
{virgis,dalia.kriksciuniene}@vukhf.lt

The successful start of the scientific workshop on Applications of Knowledge-Based Technologies in Business (AKTB) last year in Poznan, inspired to call for the second time the researchers and practitioners, specialists and market analysts to share their research expertise and advanced knowledge in modeling innovative solutions for enterprise systems and processes, provide analytic insights and demonstrate experimental research results of implementing computational intelligence methods in various areas of business- related problems.

The thematic areas of the workshop were concentrated to solving complex tasks of contemporary business by applying intelligent and knowledge-based technologies as expressed by the following topics:

- Advanced knowledge-based business information systems;
- Computational intelligence for business (artificial neural networks, fuzzy systems, expert systems);
- Decision support systems in business enterprises, financial institutions and e-management;
- Knowledge-based models of data mining in business;
- Business process and information requirements analysis;
- Information technologies and software developments for business process modeling;
- Agent-based and embedded systems in business applications;
- Information systems in e-business, e-banking and marketing;
- Online trading by using evolution-based methods, neural networks and rule-based systems;
- Advanced computational approaches to portfolio optimization and selection;
- Analysis of financial time series;
- Estimations, modeling, algorithms of application of investment strategies in financial markets;
- Advanced research and case studies of application computational methods in banking, insurance and credit risk evaluation, company rating systems.

The 2nd AKTB 2010 workshop was selected by 43 authors, who submitted 26 articles. Each paper was evaluated in the double-blind review process by at least two independent reviewers of the 28 members of the Program Committee, representing research institutions from 14 countries. The quality of the articles was defined by their conformity to the workshop topics, originality and novelty, methodological background,

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 1–2, 2010.

relevance of the article, and substantiation of the conclusions. The overall acceptance rate for the AKTB workshop was 0.54. There were 12 articles, evaluated as not corresponding to the workshop themes or requirements.

The top- ranked 14 articles were accepted for presenting during the conference. The most mature research articles were assigned to the category of long papers, the biggest group of work-in-progress reports was presented by six authors, and four demo papers expressed the research directions of recently defined problem areas. The intensive scientific discussion during workshop sessions, and numerous questions from the audience served for improvement of the articles, approved for submitting to the post-conference proceedings.

The table below illustrates the statistics of AKTB workshop acceptance rate by number of countries and authors:

Country	*Authors*	*Submitted papers*	*Accepted*	*Acceptance rate*
Austria	2	0.67	0.67	1.00
Bangladesh	1	3.00	0.00	0.00
Brazil	6	1.00	1.00	1.00
Germany	1	0.50	0.50	1.00
Lithuania	23	15.00	8.00	0.53
Norway	3	1.00	1.00	1.00
Poland	1	0.50	0.50	1.00
Romania	3	2.00	1.00	0.50
Sweden	1	1.50	0.50	0.33
Tanzania	1	0.50	0.50	1.00
United Kingdom	1	0.33	0.33	1.00

We would like to express our appreciation to all authors of submitted papers, members of the program committee and reviewers, the Department of Information Systems of the Poznan University of Economics, Freie Universität Berlin and to acknowledge the outstanding efforts of the Organizing Committee of the 13th International conference, BIS2010.

May 2010

Virgilijus Sakalauskas
Dalia Kriksciuniene
Vilnius University, Lithuania

Case-Based System for Innovation Management in Electric Power Enterprises

Tânia C.D. Bueno[1], César R.K. Stradiotto[1], Sonali P.M. Bedin[1], Vinícius O. Mirapalheta[1], Hugo C. Hoeschl[1], Thiago P.S. de Oliveira[1], Luiz A.P. Athayde Filho[2], and Pedro A.S. de Carvalho Júnior[2]

[1] I3G - Instituto de Governo Eletrônico Inteligências e Sistemas, Florianópolis, SC, Brasil
{tania.bueno,cesar.stradiotto,sonali.bedin, vinicius.mirapalheta,hugo.hoeschl,thiago.paulo}@i3g.org.br
http://www.i3g.org.br

[2] CELESC - Centrais Elétricas de Santa Catarina, Florianópolis, SC, Brasil
{luizpaf,pedroascj}@celesc.com.br
http://www.celesc.com.br

Abstract. The CELESC R&D Management Systems is a case-based system for information management and retrieval associated with an ontological process of knowledge representation. The system was built to support R&D project management and was developed inside CELESC Company, to meet requirements of XML information standards of the National Agency for Electric Energy Companies (ANEEL). The database system has projects developed since 2001, inside CELESC. Developed for web environments, the system allows a wide knowledge diffusion of its content and results.

Keywords: Knowledge management; Case-based reasoning, R&D Project Management, Electric Sector, Shared Management.

1 Introduction

In Brazil the Federal Government imposes, by means of the Law, minimal investments in R&D from electric power companies. CELESC is a governmental enterprise responsible for the electric sector, inside the state of Santa Catarina – Brazil. The present project was constructed a system based on a knowledge management framework developed by the Institute i3G [3, 4, 5, 6], and was created an Information Management Model based on the ANEEL procedure manual [1, 2]. For this reason, to explain the developed research work, we divided this job on the following sections: In item 2 it is presented the CELESC R&D Management System. In item 3 it is presented the ontology-based information management model. In item 4 it is presented the construction of cases in the range of the management system. In item 5 it is presented the information retrieval techniques. In item 6 are presented the conclusions and future wok. Finally the acknowledgements and the consulted references are presented.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 3–6, 2010.
© Springer-Verlag Berlin Heidelberg 2010

2 The CELESC Innovation Management System

The system is composed by two environments based on Web: One for management, and the other for searching information on R&D documents. These technological structures will serve as management instruments to consult about the R&D information. With these environments, it will be possible to keep track – daily – about the produced job by all these people involved on production, storage and recovering of strategic information, inside the parameters defined by ANEEL. The search environment is composed by three sub-modules: **Search by similarity**, where the user entry a text in natural language, which is compared with the documents by ontologies; **Search by subject**, where the user looks for texts inside predefined specific domains organized by ontologies; and **Simple search**, where the system searches for an exact string.

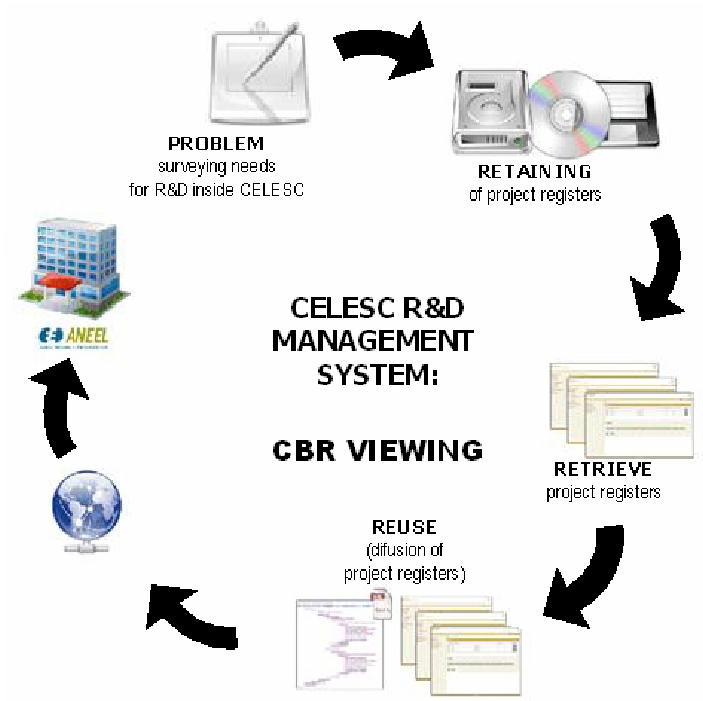


Fig. 1. CBR Viewing for CELESC R&D Management System

3 Ontology Construction and Case Representation

The participation of the ontology structure in the Innovation Management systems occurs in three moments. First, the system extracts information from different previously selected sources. Each of these documents is indexed based on ontologies defined by the specialists and knowledge engineers during the knowledge engineering process: The system will mark the documents with all indicative expressions found in the text, storing them in an organized way in the knowledge base. Thus, it is possible to make a pre-classification of the cases in the base according to what was defined in the knowledge organization helped by ontologies. In a second moment, the ontologies are used in the analysis interface available to the user. The process starts when user types the input text for search. At this point, the indicative expressions defined by the user, which are inside the ontology, are identified. These expressions in the entry case determine the stream of relations that will be used by the system. The cited relationships are the following: Synonyms, Related Terms: Representing terms appearing together in a great amount of texts. A type of: Same as class relation. A part of, also known as the Meronimy relation. These relations are used to compare terms from a user document and the documents inside the knowledge base.

3.1 Indexing Process

The present system was developed using CBR method, which allows the representation of texts in form of cases, through the use of indexes. Each case is structured as a set "indicative expressions" defined by ontologies and all words in the documents from current cases and target cases. Also, it was defined indexes, representing characteristics of the cases that indicate your usefulness in a specific situation. Those indexes have been determined to facilitate the retrieval process, which consists in four attribute-value pairs to index: Date, Source, List of Ontologies and List of Dictionary Words. The system uses the ontologies created in the editor to realize the information retrieval of *Projects, Documents in Library, Innovations and R&D Needs*. The system converts the stored information in a word vector, removing unuseful words to our information retrieval model (stop words), like pronouns, preposition, articles, gender, plural and special characters. After that, the system processes the word stemming. The storage uses an open structure established in language XML, directly related with the database indexed automatically in an intelligent way. The ontologies assist the process of automatic indexation (case definition) and supply the indexes and expressions for the similarity degree evaluation between the query and the case (document) in the database. After the automatic indexation each case is valuated, the cases are classified in accordance with the values obtained with the metric of similarity.

4 Case Constructions and Application of IR Techniques

The documents are represented in form of cases, which consist of the original document text and a set of indexes in form of attribute-value pairs, and a general domain knowledge is included in form of an ontological representation. As a result, the indexed ontologies and relations represent information related to the content of the texts, and the information necessary to identify the tuples (source, number and date) is represented. Documents manually entered into the system follow a sequence of literal operations, which reduces the complexity for representation and allow the transformation of texts into a set of terms and indices. After the reduction in the document size, the system extracts the significant information using the dictionary of ontologies, created for a specific domain. Based on this information, a formal description is generated by the definition of the indices for a specific case. The description is partially compared with all the cases in the base using CBR techniques [7], in order to identify the most equivalent case. To compute the similarity for each case, the Ontology Dictionary is used and it indicates the similarity between the individual terms and each relation between the Indicative Expressions.

5 Conclusions and Future Work

The use of ontologies is responsible for the high quality of results in the information retrieval process. Ontologies are important in the analysis interface, available to the user. Although there are a model to represent knowledge from texts inside CELESC R&D knowledge base, this base is not completed with enough documents to go for experiments. In future works we are going to fill it with R&D related documents to make the tests for the representation with ontologies.

Acknowledgements

Our thanks to Centrais Elétricas de Santa Catarina – CELESC, and to i3G Team: Mr. Thiago Paulo de Oliveira, Angela Iara Zotti and Claudia Bueno, which helped us to finish this job.

References

1. ANEEL (Agência Nacional de Energia Elétrica). Atlas de Energia Elétrica do Brasil, ANEEL, Brasília, Brasil. Disponível em (2002), `http://www.aneel.gov.br/arquivos/PDF/livro_atlas.pdf`
2. Manual do Programa de Pesquisa e Desenvolvimento Tecnológico do Setor de Energia Elétrica. Superintendência de Pesquisa e Desenvolvimento e Eficiência Energética – SPE - Agência Nacional de Energia Elétrica – ANEEL - Educação - Pesquisa e Desenvolvimento, p.16, `http://www.aneel.gov.br/aplicacoes/leitura_arquivo/arquivos/ManualPeD_2008pdf`
3. Bueno, T.C.D., Wangenheim, C.G., Mattos, E.S., Hoeschl, H.C., Barcia, R.M.: JurisConsulto: retrieval in jurisprudential text bases using juridical terminology. In: ICAIL 1999, pp. 147–155 (1999)
4. Bueno, T.C.D., Bedin, S., Cancellier, F., Hoeschl, H.C.: Evaluating the Effectiveness of a Knowledge Representation Based on Ontology in Ontoweb System. In: Larsen, H.L., Pasi, G., Ortiz-Arroyo, D., Andreasen, T., Christiansen, H. (eds.) FQAS 2006. LNCS (LNAI), vol. 4027, pp. 625–633. Springer, Heidelberg (2006)
5. Bueno, T.C.D., Ontojuris: Multi-Language Ontology Edition. Acessado em Maio de (2009), `http://www.i3g.org.br/ontojuris/`
6. Stradiotto, C.R.K., Bueno, T.C.D., Hoeschl, H.C.: Ontology Graphical Editor for Multilingual Document Search System. In: FQAS 2009, pp. 453–464 (2009)
7. Wangenheim, A.V., Wangenheim, C.G.V.: RaciocÍNio Baseado Em Casos. ISBN 85-204-1459-1 Editora Manole. São Paulo, Barueri (2003)

Credit Risk Evaluation Using SVM-Based Classifier

Paulius Danenas and Gintautas Garsva

Department of Informatics, Kaunas Faculty of Humanities,
Vilnius University, Muitines St. 8, LT- 44280 Kaunas, Lithuania
paulius.danenas@vukhf.lt,
gintautas.garsva@vukhf.lt

Abstract. This article presents a method joining Support Vector Machines (SVM), genetic search and multivariate analysis for identification of bankrupt companies. This study proposed to join widely used Altman Z-Score with Support Vector Machines to create a classifier that might be used to evaluate and forecast possible bankrupt companies. A genetic search algorithm is employed for relevant attribute selection to reduce the dimensionality of data.

Keywords: Support Vector Machines, SVM, artificial intelligence, machine learning, credit risk, evaluation, bankruptcy, Altman.

1 Introduction

Credit risk evaluation is a very important subject as it is a difficult task to decide if the loan can be given to the customer or if the credit request has to be rejected. It has become one of the most important topics as a lot of companies are going bankrupt recently and it is very important for the creditor to choose the correct principle or model for evaluation. The rise of artificial intelligence and computational intelligence which simulates human decision making processes by using various computational techniques offered a possibility to create more sophisticated and precise models. Examples of these methods, often referred as natural computing methods, are neural networks (NN), genetic algorithms (GA), evolutionary computing (EC), artificial immune systems (AIS), swarm intelligence (SI) and many others. Support Vector Machines (SVM) is one of these methods which is widely applied as an effective solution to many various pattern recognition, classification, regression and forecasting problems, including financial forecasting and credit risk evaluation.

2 Related Work

Many of the earliest researches were based on discriminant analysis; the most widely known and used was developed in 1968 by Altman [1]. Ohlson used logit approach to construct his model [2].

Neural networks have been used for research in credit evaluation field since they were presented as computational technique. Recent researches include research on evaluation in banks of particular country [3], in particular area [4], as well as

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 7–12, 2010.
© Springer-Verlag Berlin Heidelberg 2010

researches on hybrid methods based on neural networks [5]. Self-organizing maps often referred as Kohonen maps are also widely researched and applied [6, 7].

Lai et. al. [8] applied LS-SVM as standalone method. Other authors combined it with other soft computing techniques in the field of credit risk including ant colony optimization [9], particle swarm optimization [10], rough sets [11, 12], fuzzy logic [13,14]. A metaresearch conducted on previous credit risk evaluation researches based on SVM showed that best results were achieved by using optimization techniques such as genetic algorithms or swarm intelligence and fuzzy logic together with SVM [15].

3 The Method

3.1 Description of Algorithms Used in This Experiment

C-SVC (LibSVM). This is the classical SVM implementation based on the original work of Vapnik; a detailed description of this algorithm can be found in [16,17].

LIBLINEAR. LIBLINEAR is an open source library and a family of linear SVM classifiers for large-scale linear classification which supports logistic regression and linear SVM. In our experiment we apply L2-loss linear support vector machines for classification task. A detailed description can be found in [18].

SMO. SMO is a Sequential Minimal Optimization principle based SVM method, introduced by Platt in 1997 [19].

Genetic algorithm (GA). It is the most widely used evolutionary computing technique which is often applied to solve various optimization problems and find best solutions in the search space by taking best individuals and forming new populations until a particular limit is reached. There are many various sources that can give more information on this technique, i.e. [20].

3.2 Proposed Method

In this section a method based on genetic search, SVM and discriminant analysis is introduced. Genetic algorithm is used for ratio selection; the relevant attributes are selected from various primary and secondary financial ratios.

The algorithm of proposed technique is presented in Figure 1. The main steps are as follows:

1. Results of every instance are calculated using discriminant analysis and are converted to bankruptcy classes;
2. Instances with empty outputs (records, which couldn't be evaluated in Step 1 because of lack of data or division by zero) are eliminated;
3. Data imputation (fill missing values with company average of the attribute with missing data);
4. Companies are divided to disjoint sets whose data will be used for training and testing (for company list $C = C\_train \cup C\_test$, and $|C\_train| > |C\_test|$);

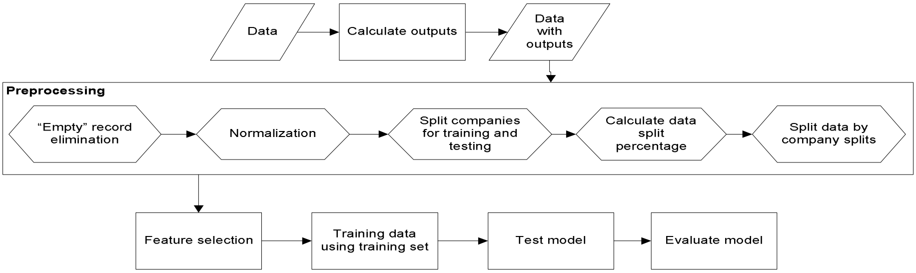


Fig. 1. Algorithm workflow

5. Calculate training and testing data split percentage – it's important if the number of financial entries is different for each company or if some instances were rejected as not evaluated by discriminant analysis;
6. Create disjoint sets as training and testing data by splitting data of selected companies in the sector by a percentage calculated in Step 5 ($C_D = C_D\_train \cup C_D\_test$ and $| C_D\_train| > | C_D\_test|$);
7. A genetic algorithm for feature selection is used to select the most relevant ratios. Evaluation uses correlation-based feature subset selection based on consideration of individual predictive ability of each feature along with the degree of redundancy between them [21, 22].
8. Model is trained using one of classification algorithms, then tested and evaluated.

The results of this algorithm are a model (a decision rule as a list of support vectors and model parameters) and a list of selected attributes which form the model.

3.3 Methods for Evaluation of Instances and Results

Altman's Z-Score was chosen for evaluation and forming of class variable as a popular and widely used evaluation technique. It predicts whether or not a company is likely to enter into bankruptcy within one or two years [1].

The results were evaluated using both accuracy and Type I error that represents the accuracy with which the model didn't identify "bad" debtors. To properly evaluate overall performance, weighted mean was used. Here we evaluated the weighted mean according only to amounts of testing instances; however, if the proportion of training and testing instances varies significantly, it might be useful to evaluate weighted mean with both training and testing instances.

4 The Experiment

4.1 Research Data

The experiments were made by using data from EDGAR database of over 8600 companies from year 1999-2006. It consists of yearly financial records with 79 financial ratios and rates used in financial analysis. This data was split into sectors according to SIC classification.. We used 3:2 split in our experiment (60% percent of companies were selected for training). The sectors and company splits are given in Table 1.

4.2 Experiment Configuration

The experiment was run using LibSVM [16], LIBLINEAR [18], SMO and genetic search implementations in Weka [21] software. The algorithm in Section 3.2 was applied for every sector. The training also was done with the whole dataset for the overall evaluation. The number of all data is different from data in sectors as it also includes entries that weren't labeled as belonging to one of sectors in Table 1 or the number of entries in particular sector was too small for training and evaluating a classifier.

Genetic search used for attribute selection was run using the following parameters: crossover probability equal to 0.6, number of generations equal to 20, mutation probability of 0.033 and population size of 20. SMO and LibSVM (C-SVC) classifiers were run using polynomial kernel with parameters C = 4 for SMO; C = 7 and γ = 7 for C-SVC. LIBLINEAR was run with C = 20. These parameters were chosen in an experimental way.

4.3 Experiment Results

The results of these experiments using data from different sectors as well as all data are presented in Table 1.

Table 1. Experiment results

Sector	Number of companies for training	Number of companies for testing	Training percentage	No of training instances	No of testing instances	SVM (LIBLINEAR)		SVM (LibSVM)		SVM (SMO)	
						Accuracy	Type1 Error	Accuracy	Type1 Error	Accuracy	Type1 Error
Agriculture, Forestry and Fishing	20	13	59,56	134	91	85,71	0,3	75,82	0,561	86,81	0,265
Mining	281	188	59,78	1660	1117	71,17	0,612	77,17	0,408	72,78	0,559
Construction	50	33	60,67	324	210	84,76	0,534	72,38	0,596	88,57	0,349
Manufacturing	1816	1211	60,10	11641	7727	88,92	0,523	90,86	0,336	89,93	0,388
Transportation, Communications, Electric, Gas, Sanitary Services	472	314	59,63	2838	1921	70,85	0,666	81,05	0,294	76,37	0,453
Wholesale Trade	172	115	59,80	1095	736	91,58	0,555	90,08	0,363	92,66	0,473
Retail Trade	243	162	60,20	1570	1038	95,38	0,775	95,09	0,586	94,99	0,706
Finance, Insurance and Real Estate	1112	741	61,18	6413	4069	68,86	0,512	82,13	0,206	74,07	0,389
Services	1027	685	60,22	6241	4122	80,81	0,524	85,30	0,286	84,43	0,382
All data	5199	3466	**63,53**	**36674**	**21054**	**70,45**	**0,667**	**71,67**	**0,635**	**72,46**	**0,613**
Weighted mean						**81,21**	**0,552**	**86,39**	**0,318**	**83,95**	**0,420**

The classifiers achieved good results using both small and large number of instances; the best results were obtained by using LibSVM. Therefore, it must be noted that training LibSVM-based classifier using polynomial kernel might be slow when a large number of instances is used; in these cases SMO or LIBLINEAR might be considered as on option as they achieved similar results. The table also shows that SMO classifier is a good option when the number of instances and attributes is high as it also performs this task in a significantly shorter time than LibSVM based classifier used in this research.

For overall evaluation weighted mean errors of all these three classes were calculated; they are given in Figure 2. These graphs show that LibSVM produced smallest errors in most of cases; however, there were sectors which resulted in extremely different errors (i.e., the sector of Finance) although the performance in other sectors was significantly better.

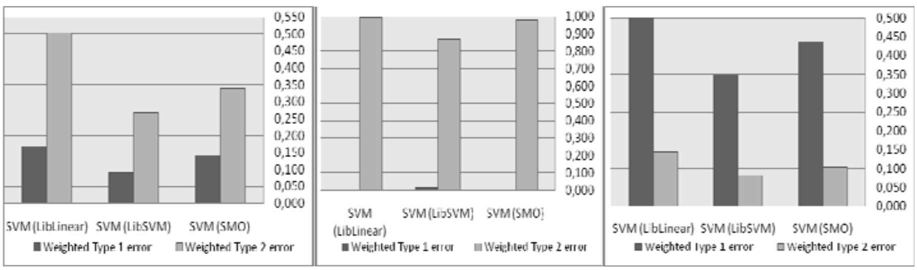


Fig. 2. Weighted mean error values for all classes (bankrupt, average, healthy)

5 Conclusions and Further Research

This article presents a hybrid method for credit risk evaluation that uses evaluation by widely used discriminant methods, such as Altman Z-Score, genetic search for financial ratio selection and model training using SVM based classifiers. It might be applied in cases when there are no label attributes as class evaluation is done by using discriminant analysis methods. The experiment results show that this method gives promising results and might be used as a basis for further research. Its performance might be improved by using various techniques for parameter selection and thus optimizing the classifier.

References

1. Altman, E.: Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The Journal of Finance 23(4), 589–609 (1968)
2. Ohlson, J.A.: Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research 18(1), 109 (1980)
3. Wei, G., Meiyan, C., Jianfeng, Z.: Study on Chinese Banks of Credit Risk Evaluation Models of Real-Estate Based on the BP-Neural Network Model. In: Proceedings of the 2009 WRI World Congress on Computer Science and Information Engineering, vol. 2, pp. 288–292 (2009)
4. Jian, H.: Personal Credit Rating Using Artificial Intelligence Technology for the National Student Loans. In: 4th International Conference on Computer Science & Education, pp. 103–106 (2009)

5. Huang, F.: A Particle Swarm Optimized Fuzzy Neural Network for Credit Risk Evaluation. In: Proceedings of the 2008 Second International Conference on Genetic and Evolutionary Computing, pp. 153–157 (2008)
6. Merkevicius, E., Garsva, G., Simutis, R.: Neuro-discriminate Model for the Forecasting of Changes of Companies Financial Standings on the Basis of Self-organizing Maps. In: Shi, Y., van Albada, G.D., Dongarra, J., Sloot, P.M.A. (eds.) ICCS 2007. LNCS, vol. 4488, pp. 439–446. Springer, Heidelberg (2007)
7. Merkevicius, E., Garsva, G., Girdzijauskas, S.: A hybrid SOM-Altman model for bankruptcy prediction. In: Alexandrov, V.N., van Albada, G.D., Sloot, P.M.A., Dongarra, J. (eds.) ICCS 2006. LNCS, vol. 3994, pp. 364–371. Springer, Heidelberg (2006)
8. Lai, K.K., Yu, L., Wang, S., Zhou, L.: Credit Risk Evaluation with Least Square Support Vector Machine. In: Wang, G.-Y., Peters, J.F., Skowron, A., Yao, Y. (eds.) RSKT 2006. LNCS (LNAI), vol. 4062, pp. 490–495. Springer, Heidelberg (2006)
9. Zhou, J., Bai, T.: Credit Risk Assessment Using Rough Set Theory and GA-Based SVM. In: Proceedings of the 2008 The 3rd International Conference on Grid and Pervasive Computing-Workshops, pp. 320–325. IEEE Computer Society, Los Alamitos (2008)
10. Jiang, M.J., Yuan, X.Y.: Construction and Application of PSO-SVM Model for Personal Credit Scoring. In: Shi, Y., van Albada, G.D., Dongarra, J., Sloot, P.M.A. (eds.) ICCS 2007. LNCS, vol. 4490, pp. 158–161. Springer, Heidelberg (2007)
11. Wang, B., Liu, Y., Hao, Y., Liu, S.: Defaults Assessment of Mortgage Loan with Rough Set and SVM. In: Proceedings of the 2007 International Conference on Computational Intelligence and Security, pp. 981–985 (2007)
12. Zhou, J., Tian, J.: Credit risk assessment based on rough set theory and fuzzy support vector machine. In: Advances in Intelligent Systems Research, ISKE-2007 Proceedings (2007)
13. Hao, P., Lin, M., Tsai, L.: A New Support Vector Machine with Fuzzy Hyper-Plane and Its Application to Evaluate Credit Risk. In: Eighth International Conference on Intelligent Systems Design and Applications, vol. 3, pp. 83–88 (2008)
14. Chong, W., Yingjian, G., Dong, W.: Study on Capital Risk Assessment Model of Real Estate Enterprises Based on Support Vector Machines and Fuzzy Integral. In: Control and Decision Conference, pp. 2317–2320 (2008)
15. Danenas, P., Garsva, G.: Support Vector Machines and their Application In Credit Risk Evaluation Process. Transformations in Business & Economics 8(3(18)), 46–58 (2009)
16. Chang, C., Lin, C.: LIBSVM: a Library for Support Vector Machines (2001)
17. Vapnik, V.: The Nature of Statistical Learning Theory. Springer, Heidelberg (2000)
18. Fan, R., Chang, K., Hsieh, C., Wang, X., Lin, C.: LIBLINEAR: A library for large linear classification. The Journal of Machine Learning Research 9, 1871–1874 (2008)
19. Platt, J.: Sequential minimal optimization: A fast algorithm for training support vector machines. In: Advances in Kernel Methods-Support Vector Learning, vol. 208, pp. 98–112 (1999)
20. Haupt, S.E., Haupt, R.L.: Practical genetic algorithms. Wiley-IEEE (2004)
21. Weka 3: Data Mining Software in Java, `http://www.cs.waikato.ac.nz/ml/weka/`
22. Hall, M.A.: Correlation-based Feature Subset Selection for Machine Learning. Hamilton, New Zealand (1998)

Logistic Analysis of Price Bubble and Current Situation in USA Real Estate

Stasys Girdzijauskas[1] and Marius Dubnikovas[2]

[1] Vilniaus universitetas, Muitinės g. 8, LT-44280 Kaunas, Lithuania
stasys.girdzijauskas@vukhf.lt
[2] "Elektronine komercijos baze", A.Mickevičiaus g. 29,
LT-44245, Kaunas, Lithuania
marius@ekb.lt

Abstract. Several recent years have been marked by challenges to financial markets. It is especially applicable to real estate markets. The article analyses the development of prices in real estate market in the USA during last 35 years concentrating on the changes in the last decade. Employing the logistic theory of capital management, the authors identify the problematic stages of crises and to highlight the stages of their formation and problems occurred. The article emphasizes the problem of market capacity and sequential growth. The data is processed by using the system 'Loglet Lab2'. Modeling reveals the main issues in market saturation.

Keywords: Logistic, real estate, USA, price bubble, saturation.

1 Introduction

Capital markets are likely to float. Considering the data of the past, one can easily distinguish some obvious stages of a sharp growth and fall which are described as the formation and burst of economic bubble. The period of 2007-2009 are characterised by intense price floats and falls in almost each asset class which brought investors to the biggest damages since The Great Depression. Price bubbles were common to that period and the dynamics of USA real estate was under harsh discussion.

The article analyses price changes in USA real estate during last 35 years concentrating on the functions of limited growth or logistic functions which describe the process of capital accumulation (growth). Having identified the bubble during the last crisis, the authors of the article attempt to reveal the limit point at which the saturation was reached and the point a sharp growth of the market started at. The authors also identify the point which the market of USA real estate stands at because of its still huge importance on other parallel markets and consumer expectations at the range of global scale.

Logistic growth is common not only to the capital but also to any other populations the growth speed of which is proportional to their size.

Overall, logistic models are widely applied in biological system research. These models have been hardly applicable to the analysis of economic phenomena until

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 13–18, 2010.
© Springer-Verlag Berlin Heidelberg 2010

now. Thus only single attempts of economic system research have been known (Omar Campos Ferreira, 1998, Ronald Shone, 2001, John Sterman, 2000). The drawback of these models is the lack of growth function expressed by complex percentage. In Lithuania, such researches started in 2002 (S. Girdzijauskas, 2002).

The authors also try to determine the point when USA household market was at the bursting point and if the marginal saturation of the capital which potentially impelled the situations of the sudden fall and the credit trap effect which enforced the negative consequences was reached.

The aim of the research: is to determine the bubble formation in USA real estate and its relation to the limit of market saturation.

The object of the research: the prices of USA real estate during last 35 years.

The method of the research: is analytical, additional the methods of comparative and mathematical analysis as well as econometric calculations based on the possibilities of information technologies.

2 Logistic Modeling

In contemporary economics and precisely in investment science (apart of rare exceptions) economic growth is considered as unlimited, whereas in reality each growth is doomed to stop sooner or later. This may be observed in nature as well. In the science of biology, the models analysing the growth of populations were invented more than a hundred years ago. Cyclical development of regions and countries also proves the limitation of economic growth. The Logistic Theory of Capital Management invented in recent years by Stasys Girdzijauskas, the professor of Vilnius University, estimates the fact that economic development has its own limits as well.

While analysing the capital growth, it is usually supposed that there is a specific financial capacity (habitat) of limited size which can be occupied by that capital. The capital usually does not occupy the whole capacity (habitat) in investment, only a part of it. This part of the capital is called the coverage. The rest part of the habitat (capacity) is left for the capital growth and is called the resource of growth.

As the coverage of investments increases, the area for growth (resource of growth) decreases because the capacity of investment capital is limited. Thus, investment capacity limits the growth of investment coverage. Sometimes investment capacity may be equal to the capacity of entire economics (Sterman, 1998).

Logistic model of growth estimates the change of capital coverage and detects a very important change of development as growth resource diminishes. When investment coverage occupies almost all the capacity, the symptoms of economics saturation occur which may lead to the formation of economic bubbles.

The created Logistic Theory of Capital Management reveals the possibilities to avoid the saturation phenomenon in economics or to soften its negative effects. The capacity of the capital must be enlarged. This can be done in several ways: the extensive method – the expansion of investment markets using marketing and other means (globalization process); the intensive method – the renewal and expansion of the markets using the achievements of new technologies and science (innovation process). The second method is, of course, much more perspective.

Generally, the value of either present or future capital is calculated in order to analyse various financial issues regarding payments or money value on the current time period or to model the capital, investments or any other money flows. Such calculations are generally based on the so-called complex interest formula (Bodie Z. et al., 2001):

Usually the growth of the capital (population, product) is calculated according to the formula (1), but this can be done till the growth of the capital is not restricted by external factors (Merkevičius E., Garsva G., Girdzijauskas S., 2006). The capital cannot grow at the same speed for a long time, particularly, if the system is entirely or partially closed. The capital growing in such a system exhausts limited resources in that environment, becomes a competitor for itself and makes its own growth slower. The system is saturated then.

Then the growth of the capital is expressed by logistic function of the growth (S. Girdzijauskas, 2002):

$$K = \frac{K_m \cdot K_0 \cdot r^t}{K_m + K_0\left(r^t - 1\right)} \text{.,} \tag{1}$$

Where K_0 – initial (existing) value of the capital, r – the coefficient of the accumulation speed, t – time expressed by the same units as the time estimated in the norm of growth percentage (generally are the integer recalculation periods of percentage norm).

It is important to note that if the maximum product value Km increases and approximates to infinity , i.e. if the equation (2) has the limit calculated, as it is true to form, the equation (2) would become an ordinary rule of complex percentage (1). Thus, the formula (1) of complex percentage is a single case of the equation (2) of logistic accumulation when the maximum capital value Km is endlessly large.

In economic calculation, not the capital growth, i.e. its future value, but its present (current) value is calculated. Thus, the logistic function of present value is employed (S. Girdzijauskas, 2002):

$$K_0 = \frac{K_m \cdot K}{K + (K_m - K) \cdot r^t} \tag{2}$$

where K_0 – present capital value, K – capital value on the current time period, t, r – the coefficient of accumulation speed with interest rate i), t – accumulation duration expressed by time units which are fixed in interest rate. The expression form is the formula of logistic discount.

Differentiating the function (2) of logistic accumulation in respect of the coefficient of growth speed r, the dependence of logistic capital growth speed on the size of the accumulation coefficient r may be observed:

$$\frac{dK}{dr} = \frac{K_0 \cdot K_m \cdot r^{t-1} \cdot t \cdot \left(K_m - K_0\right)}{\left(K_m + K_0 \cdot \left(r^t - 1\right)\right)^2} \tag{3}$$

The logistic model apparently shows when economics undergoes globalization pressure. This pressure starts after reaching the top of the growth speed (also called marginal growth in economics) in the diagram and starting to go down, i.e. economics growth slowing down and the crisis coming. Only fast progress in science and technology, which opens the possibility to start new stage of economics growth and install new clearer resource-economical technologies may solve the problem.

3 The Situation in USA Real Estate

The United States of America (USA) are usually considered as the strongest market of real estate with the system of investors managing it. It in turn creates the most powerful base of consumers in the world which often determines all processes and targets in global economics. A consumer is an economic moving force which stimulates the demand development and offers new solutions. USA market is distinct at a global scale because a consumer usually spends his/her future income. Sequential growth of real estate during last 35 years enabled the enforcement of consumer base in the USA as well as its influence at a global scale. The diagram below shows the average price change of one semi-detached house in the United States from the beginning of 1975 till the beginning of 2010.

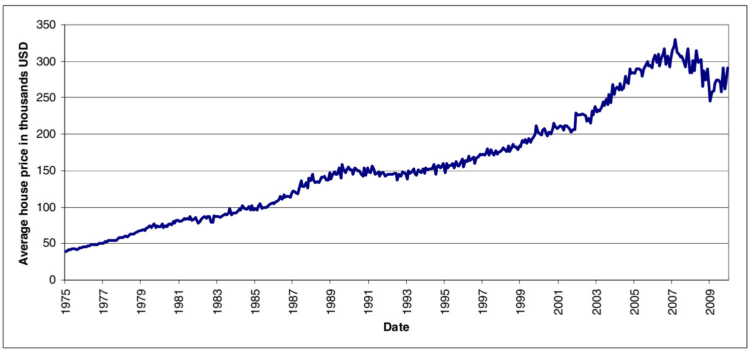


Fig. 1. One family house average price 1975 – 2010. Data source: Bloomberg.

The diagram clearly illustrates the sequential growth of prices which got more intense in the last decade and ended in a sharp fall. It obviously shows the existence of global financial crisis in the period of 2007-2009. The latter is often described as the largest fall since The Great Depression.

Such dynamics of real estate and negative changes that are not common to the last 35 years are obvious stimulates a thorough analysis at various aspects.

Having performed the calculations of the saturation, it gets clear that during the last decade, the saturation limit of one USA habitation was reached at the point of 300 thousand USD. Despite of sequential development of the market, the price growth in the market of real estate surpassed the market growth and the saturation limit was even exceeded.

In the further performance of the research, inflation effects were attempted to eliminate calculating the prices in the market of USA real estate. During the last 35 years, the inflation of the USA dollar has floated from 1,5 percent inflation to 1,5 percent deflation.

Inflation effects are especially relevant in the estimation of a 35-year-lasting market of USA real estate. Having estimated the changes of inflation, the change of USA habitation price seems much more sequential.

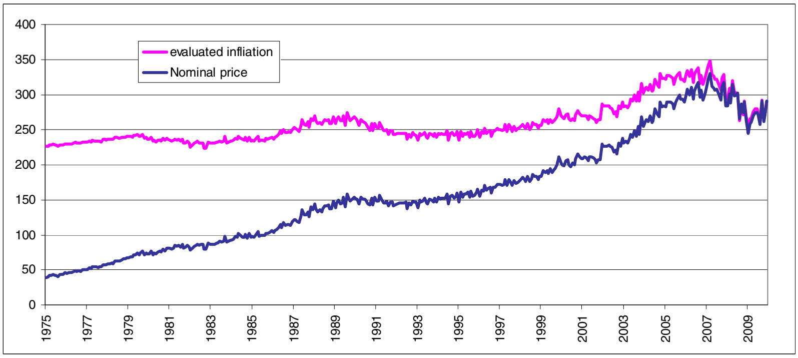


Fig. 2. One family house average price (inflation evaluated) 1975 – 2010. Data source: Bloomberg.

Having estimated habitation prices and converted all the 35-year-old values into current prices, a very sharp rise and a sudden fall of prices reveals. The data got with converted prices were processed into the system 'Loglet Lab2' and revaluated again.

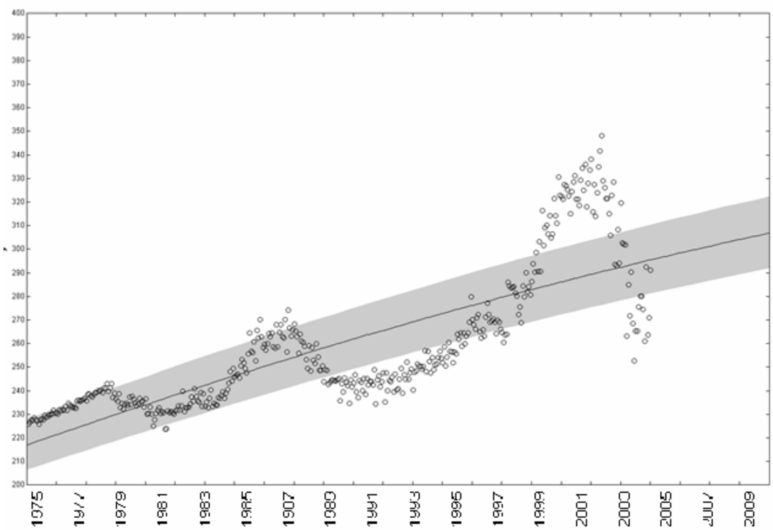


Fig. 3. One family house average price and market saturation line (inflation evaluated) 1975 – 2010. Data source: Bloomberg and Loglet LAB2.

The result was positively surprising. After the estimation of inflation and the present price of real estate, the saturation curve revealed that the market of USA real estate is likely to grow especially sequentially. Whereas economic cycles and economic bubble bursts in other economic niches make habitation prices floating in sinusoid and seldom surpassing the saturation curve. This confirms the logistic theory of capital management which states that the price bubble starts to blow and then bursts returning at the saturation curve eventually when the market limit is exceeded.

The analysis also reveals that during the last economic overheat a gap break from the saturation curve was the largest during the last 35 years.

The further fall below the saturation curve allows thinking that the market is getting closer to certain stabilization at the present. The logistic analysis reveals all the peculiarities of USA real estate in the past.

4 Conclusions

1. Having analysed the 35-year period, the logistic capital analysis confirms the formation of the price bubble in the market of USA real estate during the last decade. The burst of the economic bubble was induced by an extremely large gap break between habitation price and the saturation curve in the market.
2. It is notable that the market of USA real estate seems to grow very sequentially during the last 35 years. This is illustrated by almost straight curve which was calculated using the system and analysing the data of prices and inflation. The growth of the capacity of USA habitation market is also very sequential.
3. In order to avoid mistakes in the future, it is essential to develop the market or derive the capital outside the market limits and thus to avoid the saturation of marginal capital. As the market of real estate develops extremely sequentially, the entrance of redundant capital must be limited by restrictive means stimulating its export to other economic segments or real estate markets.
4. At the moment, the market of USA real estate is at a very low level in comparison with the past price floatation and the gap break from the saturation curve, which may allow to predict the price stabilization.

References

1. Bodie, Z., Kane, A., Marcus Alan, J.: Essentials of Investmens, p. 982. McGraw-Hill, New York (2001)
2. Bodie, Z., Merton Robert, C.: Finance, p. 592 (2000)
3. Edvards, C.H., Penney, D.E.: Elementary Differential Equations with Applications, p. 632 (1985)
4. Girdzijauskas, S.: Logistic capital management theory; determined methods, Vilnius, p. 232 (2006)
5. Girdzijauskas, S., Boguslauskas, V.: The Possibilities for the Application of the Logistic. Engineering economics 1(41), 7–15 (2005)
6. Varian, H.R.: Intermediate Microeconomics; A Modern Approach, 6th edn., p. 688. W.W Norton & Company, Inc., New York (2003)
7. Sterman, J.D.: Business Dynamics: Systems Thinking and Modeling for a Complex World, 950 (2000)

Using Multi-criteria Decision Making to Choose Process Representation Format for a Process Repository

Mturi Elias and Khurram Shahzad

Department of Computer and Systems Science (DSV),
Royal Institute of Technology (KTH) / Stockholm University, Sweden
{mturi,mks}@dsv.su.se

Abstract. Reuse of business process models is the act of designing business processes by using existing process models. Reuse of business process models has been considered as a way to reduce the cost of modeling business processes from scratch. In order to support reuse of process models, a critical mass of process models is required which justifies the effort of maintaining a process model repository. While there are several process modeling languages, no single language is widely accepted. So in order to make process models usable, the stored process models must be presented in a language understandable by users. The purpose of this work is to apply multi-criteria decision making to choose process representation format for a process model repository.

1 Introduction

In recent years organizations started adopting Business Process Management (BPM) for unifying business and IT activities, and coordinating the actions and behaviors of people and systems around the common context of business processes. Business process modeling plays a key role in the adoption of BPM. However, modeling of business processes from scratch is a complex, time consuming and error prone task [1, 2]. A natural alternative to that is reuse of existing process models [1, 3].

The first step towards supporting reuse of business processes is to store process models and make them available for reuse. Therefore, a repository is necessary for storing and managing process knowledge [4]. On the other hand, there are several process modeling languages like YAWL, BPMN and EPC etc that are used for modeling processes [6]. However, there is no single process modeling language that is widely accepted.

There are several formats in which a process model may be represented in the repository [4, 8]. Selection of the process format cannot be done literally. In order to increase the likelihood of process model reuse, it is important to involve potential users i.e. process and enterprise modeling experts (stakeholders) in selecting the best alternative. Therefore the purpose of this study is to apply a stakeholder centered multi-criteria decision making approach to select the format in winch a business process should be represented in the repository.

The paper is organized as follows: the methodology used is briefed in section 2. Alternatives in which a business process may be represented in the repository are detailed in section 3. Relationships between criteria and alternatives are explained in

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 19–24, 2010.
© Springer-Verlag Berlin Heidelberg 2010

section 4. In section 5, we employee multi-criteria approach for choosing an alternative. Finally the conclusion is discussed in section 6.

2 Process Representation Selection Methodology

In this section we present the methodology we have used for deciding, how business processes should be provided or represented in the repository using a multi-criteria decision making method. The methodology consists of four steps (shown in figure 1): Defining alternatives, identification of selection criteria, assigning weights to the criteria, and selection of alternative.

The identification of alternatives aimed at eliciting the possible format in which a business process can be represented in the process model repository. The alternatives are elicited from the literature and also through surveying existing repositories [4, 5].

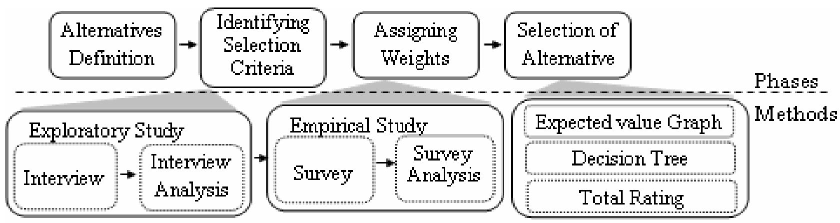


Fig. 1. The Proposed Methodology

Exploratory Study: In order to identify criteria for selecting the alternative, an exploratory study was carried out. In this step, comments, opinions, suggestions and ideas on process model repositories were elicited from both researchers and practitioners. The study was designed to collect as many comments, opinions, ideas and suggestions as possible encouraging the respondents not to restrict their responses. Therefore, the exploratory study consisted of an open-ended questionnaire which was used in an oral interview. The responses were collected, analyzed and reformulated into criteria for measuring the alternatives. A small number (five) of experts were involved as the purpose was to identify the criteria. The identified criteria include *understanding, modeling effort, reusability, bridging the gap* and *granularity*. Due to space limitation the details of these criteria are not presented.

Empirical Study: In this step, the defined criteria were assigned weights through an empirical study. A total of 25 participants (16 researchers and 9 practitioners) were involved in the study. Participants of the study were researchers and practitioners who participated in the 2nd IFIP Working Conference on The Practice of Enterprise Modeling, (PoEM´09) Stockholm, Sweden [11].The weights were computed and assigned to criteria based on the analysis of the data collected from the study.

The data collected from the empirical study were analyzed to compute the weights of the defined criteria. The participants who agreed or strongly agreed on each criterion were counted. The percentages of participants were computed and adjusted over a scale of 1 to 100. The steps taken for calculating the weight of each criterion (shown in table 1) are as follows:

1. Let P_{Ci} be the number of participants (in percentage) who agree or strongly agree on criteria *Ci*.
2. The sum of number of participants who agree or strongly agree for all criteria denoted as $SUM_{PCi} = \sum P_{Ci}$ where i = 1 to n, n is the number of criteria.
3. The weight of criteria *Ci* denoted as $W_{Ci} = Ci / SUM_{PCi} * 100$,

Table 1. Criteria, sub-criteria and their weights

Criteria	Sub-Criteria	W_{Ci}
Understanding		22.2
	Graphical Form	80
	Textual Form	20
Modeling Effort		24.4
	Modeling Cost	68
	Modeling Complexity	32
Reusability		21.1
	Within Language	76
	Cross language	24
Bridging the Gap		13.3
	Communication Gap	92
	Designing Gap	8
Granularity		18.9

3 Possible Process Representation Alternatives

In this section the possible alternatives in which a business process can be represented in the repository are discussed.

***Alternative 1 (Alt 1)*. Use one modeling notation to store and retrieve process models (existing notation).** In the first alternatives, the repository supports one modeling notation (like BPMN, YAWL, EPC, etc) in storing and retrieval of process models.

***Alternative 2 (Alt 2)*. Use of a common modeling notation to store and retrieve process models.** In the second alternative, process models are stored in one common modeling notation language. The common notation is developed based the common features of existing modeling languages [8], leaving out the specificities of process modeling language.

***Alternative 3 (Alt 3)*. Use multiple modeling notations to store and retrieve process models.** In third alternative, the repository supports multiple modeling notations on storing and retrieving process models.

4 Relationship between Criteria and Alternatives

In this section, the relationship between the identified alternatives and the defined criteria is provided. For each alternative, we describe how it relates to every defined criterion (see table 2). The table given below (table 3) formally presents, the relationship between the alternatives based on the criteria.

Table 2. Relationship between criteria and alternatives

Criteria	Alternative 1	Alternative 2	Alternative 3
Understanding	Understanding will be less if users don't understand the supported modelling notation	The understanding will be increased to users of all modelling notations	The understanding will be increased to users of all modelling notations
Modelling Effort	The effort for adopting on different notation will be very high.	Less modelling effort is required as models are available to user's choice	Less modelling effort is required as models are available to user's choice
Reusability	Will be less reusable	Increased reusability	Increased reusability
Bridging the Gap	The gap will not be bridged.	Bridges the gap	Bridges the gap
Granularity	If the supported modelling notation allows hierarchical model structuring, different levels of details will be provided, otherwise this will be hampered.	Difficulty to provide different levels of granularity with different modelling notations.	Difficulty to provide different levels of granularity with different modelling notations.

Table 3. Relationship between criteria and alternatives

Criteria	Sub-Criteria	A1
Understanding		
	Graphical Form	A1=A2=A3
	Textual Form	A1=A2=A3
Modelling Effort		
	Modelling Cost	A3>A2>A1
	Modelling Complexity	A3>A2>A1
Reusability		
	Within Language	A1>A2=A3
	Cross language	A3=A2>A1
Bridging the Gap		
	Communication Gap	A3=A2>A1
	Designing Gap	A3>A2>A1
Granularity		A1>A2=A3

5 Selection of Alternative

In this section, we presented the developed multi-criteria model for the repository case. For analyzing the alternatives we employ expected value graph and total ranking to select an alternative. A criteria tree is a hierarchical structure that allows decision maker to model complex decision problems involving multiple objectives [7, 9], which in our case are called criteria. Figure 2 shows the developed multi-criteria models (Criteria Hierarchy Trees) implemented in DecideIT [12]– a decision making tool developed at our department [10].

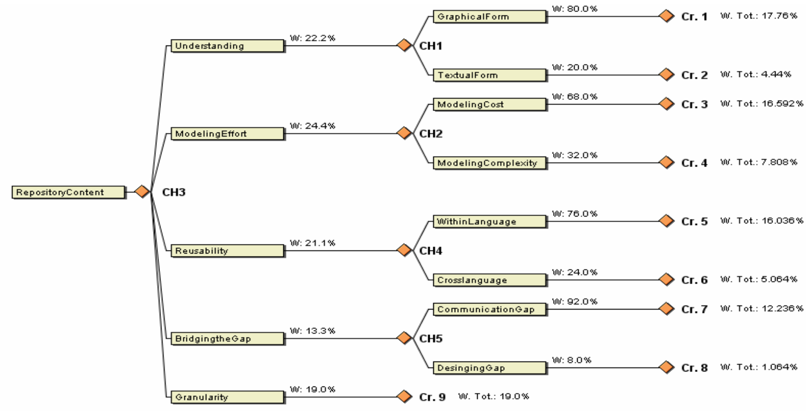


Fig. 2. Multi-Criteria Model

For analysis of the alternatives we use a pairwise comparison. Pairwise comparison shows the possible differences between two alternatives at a time. Some results of comparison are shows in figure 3 (a, b). The figure shows that alternative 3 is preferred over alternative 2. Similarly, alternative 2 is preferred over alternative 1.

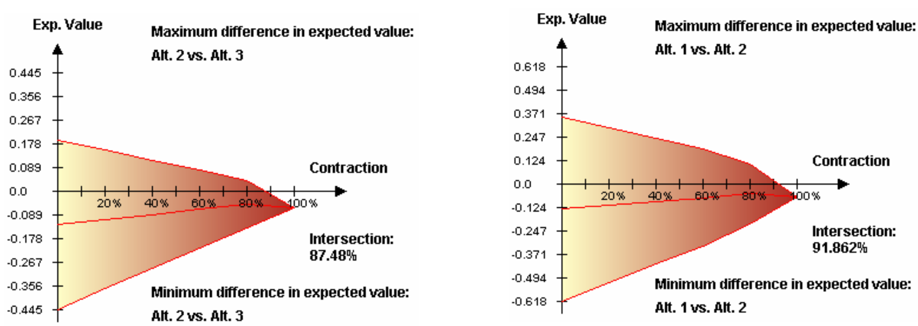


Fig. 3 (a, b). Expected value graph for alternatives 1, 2 and 3

The total ranking shows in figure 4 shows the overall relation between the three alternatives. From the brief analysis, it is concluded that alternative 3 is superior to both alternative 2 and 1.

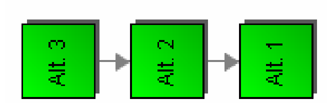


Fig. 4. The total ranking becomes

6 Conclusion

In this paper some results from our ongoing work on process repository are presented. We applied a multi-criteria decision making approach to select the best alternative in which a business process may be represented. For that, we defined the alternatives based on the literature, identified the criteria through the exploratory study, assigned weights to the criteria through the empirical study and finally used expected value graph and total ranking as methods to select the alternative.

From the study, we have found that alternative 3 is better than other alternatives i.e. use multiple modeling notations to store and retrieve process. Therefore, the repository should allow storing process models in multiple modeling languages and this mean the stored process models are available for use in different languages. One of the limitations of the study is that the selection have not considered to technical challenges in implementing the repository. Also, an in-depth analysis of the selection of the alternative is presented. Future researches aims at implementing the repository and evaluate the intended purpose.

References

[1] Markovic, I., Pereira, A.C.: Towards a Formal Framework for Reuse in Business Process Modeling. In: ter Hofstede, A.H.M., Benatallah, B., Paik, H.-Y. (eds.) BPM Workshops 2007. LNCS, vol. 4928, pp. 484–495. Springer, Heidelberg (2008)

[2] Hornung, T., Koschmider, A., Oberweis, A.: A Recommender System for Business Process Models. In: Proceedings of the 17th Annual Workshop on Information Technology and Systems (WITS'07), Montreal, Canada, pp. 127–132 (2007)

[3] Rodrigues, J.A., Souza, J.M., Zimbrao, G., Xexeo, G., Neves, E., Pinheiro, W.A.: A P2P Approach for Business Process Modelling and Reuse. In: Eder, J., Dustdar, S. (eds.) BPM Workshops 2006. LNCS, vol. 4103, pp. 297–307. Springer, Heidelberg (2006)

[4] Yan, Z., Dijkman, R., Grefen, P.: Business Process Model Repositories - Framework and Survey. Beta Working Papers, vol. 292. Eindhoven University of Technology (2009)

[5] Shahzad, K., Andersson, B., Bergholtz, M., Edirisuriya, A., Illayperuma, T., Jayaweera, P., Johannesson, P.: Elicitation of Requirements for a Business process Model Repository. In: Ardagna, C., et al. (eds.) BPM 2008 Workshops. LNBIP, vol. 17, pp. 42–53. Springer, Heidelberg (2009)

[6] List, B., Korherr, B.: An Evaluation of Conceptual Business Process Modelling Languages. In: Proceedings of the 21st SAC'06, Dijon, France, pp. 1532–1539 (2006)

[7] Danielson, M., Ekenberg, L., Gronlund, A., Larsson, A.: Public Decisino Support – using a DSS to increase Democratic Transparency. International Journal of Public Information Systems 1(1), 3–25 (2005)

[8] Shahzad, K., Elias, M.: Towards Cross Language Process Model Reuse – A Language Independent Representation of Process Models. In: Proceedings of the 2nd IFIP WG8.1 PoEM'09, Stockholm, Sweden. LNBIP, vol. 39, pp. 176–190. Springer, Heidelberg (2009)

[9] Cleman, R.T., Reily, T.: Making Hard Decisions with Decision Tools. Duxbury Thomson Learning, USA (2001)

[10] `http://pvu.dsv.su.se/Decide/Index` (last accessed March 14, 2010)

[11] `http://poem.dsv.su.se` (last accessed March 14, 2010)

[12] Danielson, M., Ekenberg, L., Johansson, J., Larsson, A.: The DecideIT Decision tool. In: Bernard, J.-M., Seidenfeld, T., Zaffalon, M. (eds.) Proceedings of ISIPTA'03, pp. 204–217. Carleton Scientific (2003)

On Valuing Advanced Technologies in Greenfield Projects

Vilde Hushovd Godø, Darijus Strasunskas, and Asgeir Tomasgard

Dept. of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway
{vilde.h.godo,darijuss}@gmail.com, tomasgar@iot.ntnu.no

Abstract. Adoption of advanced technology, improved access to real-time data, integrated people and organizations and changed work processes enable better and faster decision-making. However, the challenge is how to measure them, which becomes even more complex in new business areas that are not influenced by earlier experience. In this paper we propose a pragmatic decision analytic valuation method for monetary evaluation of integrated operations.

Keywords: business value, integrated operations, Greenfield, ICT value.

1 Introduction

Integrated operations (IO) in the petroleum industry involve adoption of advanced information and communication technology (ICT). A key characteristic of IO is improved access to real-time data through sensors, aggregation and visualization techniques, which in relation to multi-disciplinary teamwork enables more proactive work practices and increases the quality of decisions. However, in order to stimulate adoption of IO, identification and communication of IO value is required [6].

The contribution of ICT to organizational performance has been examined extensively by many researchers in operations and information systems research during the past decades. Typically, the business value of ICT is decomposed into three main categories that should be evaluated [7]: support of the strategic and operational goals (value is created indirectly); contribution to positive or reduction of negative cash flows (value is created directly); and reduction of technological and organizational risks. However, ICT investments alone do not suffice to create value, other assets such as organizational structure [5] and work processes needs to be changed correspondingly. The main lesson learnt from the related work is that technology shall not be analysed separately from human affairs [8]. Despite of a large number of evaluation frameworks, there is still a quest for improvement of them [9] and assessment of ICT business value remains a challenging task [5].

However, there are approaches that neglect qualitative evaluation and focus on technical component analysis of technology. For instance, vom Brocke et al. [4] introduce an ontology-based approach to represent the relationships among business processes and IT infrastructure components. They use scenarios to simulate technology usage and assess needed capacity (costs) and resulting benefits. The proposed

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 25–28, 2010.
© Springer-Verlag Berlin Heidelberg 2010

approach is entirely based on technical characteristics of process and technology. Bratvold et al. [3] formally analyze value of improved information in isolation from other influencing factors.

A special case is *greenfield* projects that lack any constraints imposed by prior work [11], which is typically the case for a new field development in the petroleum industry. Greenfield projects also imply more uncertainty, which increases the value of information and control gained by ICT. However, in most businesses it is not entirely correct to say that a greenfield project is completely unaffected by prior work. Many petroleum companies have already adopted some IO components, which will affect their future IO implementation strategies for new projects, or should at least be considered when evaluating future IO components.

There are no methods explicitly aiming at assessment of technology implementation in greenfield projects. Partially, the techno-business assessment (TBA) framework [12] is designated to analyze and evaluate new mobile services that are enabled by advanced semantic technologies. The analysis starts from qualitative analysis and move toward quantitative cash flow based valuation. There exist methods designed for valuation of IO (cf. [1, 2]), yet they do not target greenfield projects specifically. In [1] the benefit logic is defined as result of technical analysis, i.e. IO solutions are related to improvements. In [2] the method is tailored for collaborative working environments.

2 Valuation Framework

Operations in the petroleum industry are characterized being decision-centric. The decision-analytic view is based on explicitly representing decisions, and is very different from qualitative approaches. The proposed framework for the earlier described context is depicted in figure 1. It is built upon our valuation framework [10] and adapts techniques to the context of IO valuation in relation to greenfield projects. The framework consists of three phases, which will be described shortly, and aims at quantifying the value added by applying IO in new field developments.

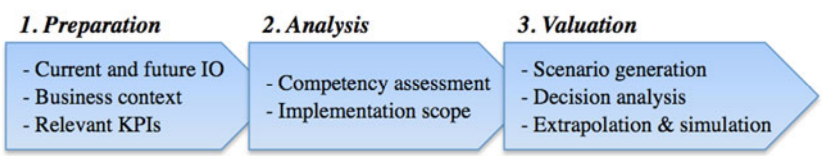


Fig. 1. The proposed valuation framework

1. *Preparation.* The preparation phase consists of collecting background information regarding the company and greenfield project, which is later on used in refining the valuation scope. During this phase we obtain an overview of *current and future IO* components. Future possibilities are the most significant for greenfield projects as few restrictions apply to the implementation scope and all opportunities should be explored. *Business context* includes business (field) characteristics, organisational goals, and analysis of how current and planned IO implementations align with the

company's strategic goals. Further on, a stakeholder analysis is used to assess commercial and local implications of current, planned and possible future IO implementations, e.g. how internal and external relations might change (new team compositions, supplier relationships, etc.) and which regulations and restrictions may apply. This is further on used to refine the valuation scope. Relevant *key performance indicators* (KPIs) are collected and linked to IO components to illustrate how value is created. They are also used later on in the valuation phase to tie results back to the company's own performance measurement system.

2. *Analysis.* The analysis phase is organised as a workshop involving key stakeholders from the host company. The workshop is commenced by a short presentation and discussion of current findings to ensure all participants agree on a common basis, and is followed by a qualitative *competency assessment*. A survey charts respondents' perceptions of former (pre IO), current and future competency levels within all four IO dimensions (people, process, technology and organisation). Results from the survey are then presented to the participants and are used to identify competencies obtained due to IO, and further competencies required for the greenfield project. Further on, we use this information to refine *implementation scope* by discussing how the various IO components can help the company obtain the required competencies, and identify critical success factors for further IO implementation.

3. *Valuation.* The first part of the valuation phase is, similarly to the analysis phase, organised as a workshop, preferably with the same participants. In this session we ask participants to *generate scenarios* of two types: first reflecting their previous experiences with current IO implementations, and second reflecting their expectations towards new IO implementations. Experience scenarios are examples of specific situations in which IO influenced operational decisions. Participants are also asked to assess how the scenario would play out without IO, and with possible future IO components. Expectation scenarios are the participants' anticipation about impact of possible future IO components/combinations of components on operational decisions. Finally, participants are asked to tie scenarios back to the previously identified KPIs. The next step of the phase is to formalise values of the identified scenarios through *decision analysis* using decision trees or alternatively influence diagrams in order to capture value of information, control and flexibility. Again, this is performed on both types of scenarios. The experience scenarios are used to formalise key decisions and identify values created by current IO implementations. The expectation scenarios are used to estimate expected values of implementation strategies. Finally, we *extrapolate* the results based on assessments of how often scenarios could play out during the project's life and use *simulation* to cope with uncertainties related to the intervals.

3 Conclusions and Future Work

In this short paper we have reported our research in progress on a potential framework to value advanced technology implementation in greenfield projects. Greenfield is characterised by higher level of uncertainties and in combination with advanced and novel technology makes the valuation task a huge challenge for many companies. The proposed framework starts from analysis of foundations, such as organisational goals, feasible technological solutions and business characteristics. Then we proceed to qualitative analysis of competencies that constitute a prerequisite for successful development. This provides a good information background for scenario generation and exploring potential impact on future operational decisions. Decisions are formalised using decision trees and valuated. Currently, we are in the process of a pilot case study to test and validate the proposed framework. An initial positive feedback indicates the framework being feasible.

Acknowledgment. This work is supported by the IO Center (the Center for Integrated Operations in the Petroleum Industry, http://www.ipt.ntnu.no/iocenter/).

References

1. Adefulu, Y.: Enhancing Asset Performance: How can the benefits of collaborative work environments be effectively measured? Paper SPE 128713 Presented at the 2010 SPE Intelligent Energy Conference, Ultrecht, the Netherlands (2010)
2. van den Berg, F., Perrons, R.K., Moore, I., Schut, G.: Business Value from Intelligent Fields. Paper SPE 128245 Presented at the 2010 SPE Intelligent Energy Conference, Ultrecht, the Netherlands (2010)
3. Bratvold, R.B., Bickel, J.E., Lohne, H.P.: Value of information in the oil and gas industry: past, present, and future. Paper SPE 110378 RMS presented at the 2007 SPE Annual Technical Conference and Exhibition, Anaheim, CA, USA (2007)
4. vom Brocke, J., Braccini, A.M., Sonnenberg, C., Ender, E.: A process oriented assessment of the IT infrastructure value: a proposal of an ontology based approach. In: Abramowicz, W. (ed.) BIS 2009. LNBIP, vol. 21, pp. 13–21. Springer, Heidelberg (2009)
5. Brynjolfsson, E., Hitt, L.M., Yang, S.: Intangible Assets: Computers and Organizational Capital. Brookings Papers on Economic Activity 1, 138–199 (2002)
6. Govia, D., Carpenter, K.: Valuation Models for Intelligent Strategies. Paper SPE 112237 Presented at the 2008 SPE Intelligent Energy Conference, Amsterdam, the Netherlands (2008)
7. Lech, P.: Proposal of a Compact IT Value Assessment Method. Electronic Journal Information Systems Evaluation 10(1), 73–82 (2007), http://www.ejise.com
8. Orlikowski, W.J.: Sociomaterial practices: exploring technology at work. Organization Studies 28(9), 1435–1448 (2007)
9. Schubert, P., William, S.P.: An extended framework for comparing expectations and realized benefits of enterprise systems implementations. In: AMCIS 2009, Proceedings, Paper 751 (2009)
10. Strasunskas, D., Tomasgard, A.: A Method to Assess Value of Integrated Operations. In: Proceedings of AMCIS 2010 (2010), http://aisel.aisnet.org/amcis2010/
11. Wikipedia, Greenfield project (2010), http://en.wikipedia.org/wiki/Greenfield_pro-ject (accessed April 23, 2010)
12. Zoric, J., Strasunskas, D.: Techno-business assessment of services and service platforms: Quantitative, scenario-based analysis. In: Proceedings of ICT-MobileSummit 2008, Stockholm, Sweden (2008)

Self-adapting Intelligent Business Processes Execution Analysis

Dalia Kriksciuniene and Sandra Strigunaite

Department of Informatics, Vilnius University, Muitines 8, 44280 Kaunas, Lithuania
{dalia.kriksciuniene,sandra.strigunaite}@vukhf.lt

Abstract. The main goal of the article is to explore methods of solving problem of the gaps, occurring during execution of enterprise process workflows due to lack of information for making unambiguous decisions in various stages of business process scenarios. The suggested method integrates application of computational intelligence methods for processing historical information of business processes and extracting scenarios used for their performance, evaluation of process risk and generating fuzzy rules for solving possible process gaps and leading to successful execution of the process workflow. The process risk and fuzzy rules modules register data of process modifications and solutions which were applied while executing business processes.

Keywords: Intelligent business process, BI, business process risk, process execution rules.

Introduction

The urgent need of the enterprises to constantly increase business process performance invokes research in the area of Business process intelligence (BPI) offering wide spectrum of methods: starting from application of business intelligence (BI) and data mining (DM) techniques to searching of fundamentally new methods, leading to solution of high level intelligent tasks [1]. The variety of approaches, including analysis of the gaps occurring during execution of process workflows, application of intelligent analysis methods for mining process log data, prediction of exceptions in the process scenario specification, identification of process patterns, leading to process workflow design from the identified sub-processes.

The problem which is addressed in this article deals with modelling self-adapting business process intelligence system, which could enable feature of evaluating and selecting best possible process scenario for its execution, and reduce risk of incomplete performing of business processes. The suggested approach is based on analysis of performance of process scenarios by applying fuzzy intelligence methods.

The essence of the research methods, their advantages and drawbacks in executing business processes and ability to avoid possible gaps is analysed in the first part of the article. The second part discusses the structure and functioning of the suggested system model.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 29–32, 2010.
© Springer-Verlag Berlin Heidelberg 2010

1 Intelligent Business Process Analysis Solutions

Definition of business processes in the system implementation phase bears high responsibility from the designers, but inevitably numerous exceptions occur during practical application of the system for performing business processes. The research works suggest solutions of various types. According to [4] the requirements for the enterprise systems stress organizational measures, which state, that the decision maker should be informed about all changes in business processes and guaranteed of information supply from all systems across the organization. In [4] the business process is viewed as collection of agents, enabled to interact and negotiate when they have interdependencies. The system solution is presented as agent-based system, which emphasizes the need of richness of negotiation ability among the differently informed agents, which could ensure service fulfilment in alignment to all defined requirements. Such kind of system is able to provide in-depth analysis of separate process details and fragments, but brings highly unpredictable outcome of business process performance, as it does not analyse or optimize entire workflows of the executed processes.

The possibilities of the managers to perform business processes without gaps, misalignments and exceptions are analysed in [2]. Here the suggested method is based on analysis of process execution logs, as a source of learning from the past information. The problem is formulated as necessity to provide service processes with consistent and predictable quality in order to create BPI solution based on mining process execution logs. Therefore the architecture of the system includes three main components, which provide the functions of extracting data from process logs, application of data mining techniques and providing graphical interface for analytical activities of business users. According to [2] the method enables to bridge the gap between the conceptual process as it was initially designed and the actual process implementation. Research in the area of BPI- oriented system architecture includes not only intelligent mining approach or agent-based system applications. In [3] the system architecture integrates intelligent and systems components for management of computational grid and history database of system performance.

By summarizing the scientific research, we can conclude, that the information of process execution history is insufficient if it contains only logs about process steps and reports about the exceptional situations. In order to successfully lead the process to completion it is necessary to select adequate scenario from the process database, which could be used for current process performance as best conforming to its requirements.

2 Self-adapting Business Intelligence Process Model

The analysis of the research leads to the insight that the past information should bring not only predictions of gaps and misalignments, but also to capture decisions and experience of the managers, who solved the unexpected situations and executed the entire scenario of the process, and to enable to extract rules of process performance.

In order to conduct intelligent management of business, the model is designed, which could amend the database of scenarios which were performed during execution of business processes by taking into account the unexpected gaps. The processes with widest variety of modifications of scenarios (as compared to the processes designed according to initial specification) are marked with the highest risk. They are recommended by the system for further analysis of emerging business rules, consequently leading to the possibility to self-adapting of the processes, achieved by applying fuzzy analysis for selecting the best fitting scenario.

Self-adapting business intelligence process model architecture consists of four compounds: data collection tools; analytical tools; risk evaluation tool based on application of fuzzy logic; and intelligent process management tool. In Fig.1 the structure and interrelationships of the compound parts are presented.

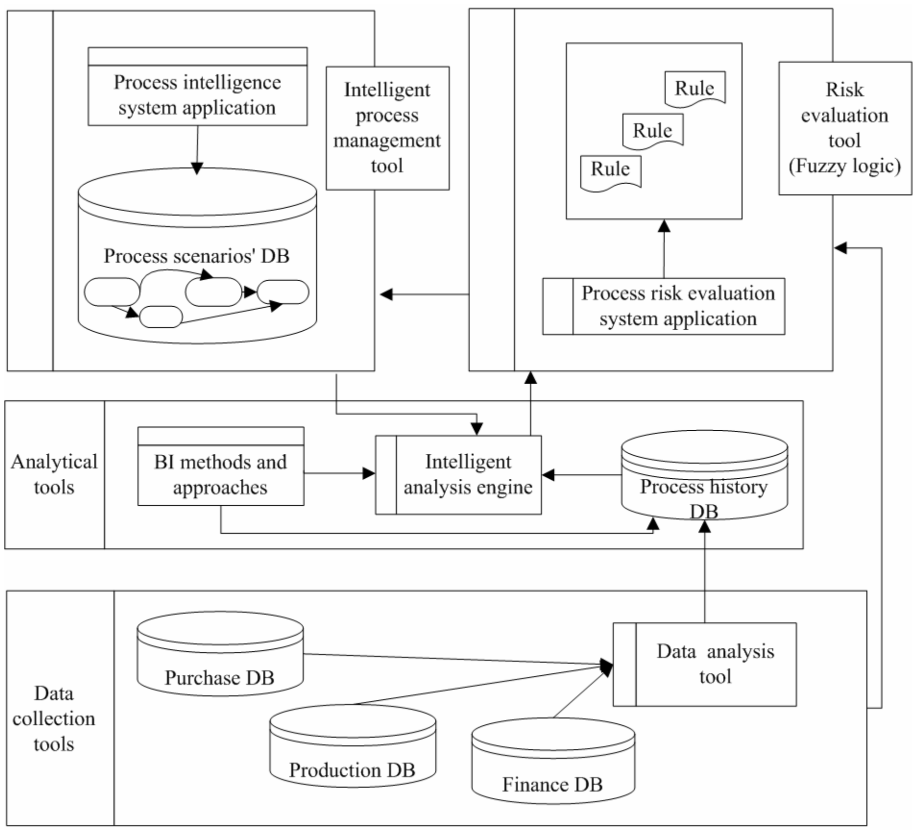


Fig. 1. Self adapting Business intelligence model architecture

The basic layer of the model is data collection tools compound. It includes historical databases of business transactions (such as purchase, production, CRM or finance databases) and the data analysis tool, which explores business process execution experience by registering the historical workflows. The module of analytical tools deals with application of BI methods and approaches for mining process history database and extracting scenarios of the process execution together with the underlying business rules, which reflect variety of the actual workflows of the specified processes. This module interacts with risk evaluation and intelligent process management tools. Risk evaluation tool has to define process risk and consequently it triggers various functions of intelligent engine with different complexity in order to analyse historical rules and refining process scenarios for enabling self-adapting of the business processes.

The model of self- adapting business intelligent process enhances its own performance not only due to capturing processes data but also due to complementing of analytical methods. The history database stores information about the business case and also analytical information about the computational intelligence approaches applied to business processes, which were executed during the case. As risk management tool, rules database and analytical capabilities are constantly updated, the effect of self–adaptability of the process scenarios is achieved.

3 Conclusions and Further Research

The main research methods and approaches of analysis of possible gaps in process execution include application of agent- based information management, mining process logs, analysis of process history database. The methods concentrate on the advancement of the analytical capabilities of the process instances, but do not solve the problem of refinement of entire workflow for each particular enterprise process.

In this article the self-adapting business intelligence process model is suggested, which implements the approach of capturing and reuse of fuzzy rules generated from actual process scenarios. The structure of the model is based on integration of the following functional parts: data collection tools, analytical tools of intelligent interpretation of process history, risk evaluation tool with fuzzy logic component, and intelligent process management tool, which derives refined scenarios and enables self-adapting of business processes. The further research work aims to in-depth analysis of performance of each structural component of the self-adapting business intelligence process model and the refined experimental verification of the suggested model.

References

1. Workshop on Business Process Intelligence (BPI), `http://is.tm.tue.nl/bpi06/`
2. Grigori, D., Casati, F., Castellanos, M., Dayal, U., Sayal, M., Shan, M.-C.: Business Process Intelligence. Computers in Industry 53, 321–343 (2004)
3. Dongarra, J., Eijkhout, V.: Self-adapting numerical software for next generation applications. Lapack working note 157, ICL-UT-02-07, Innovative Computing Laboratory, University of Tennessee, Knoxville TN 37996, USA (2002)
4. Jennings, N.R., Faratin, P., Johnson, M.J., O'Brien, P., Wiegand, M.E.: Using Intelligent Agents to Manage Business Processes. In: 1st Int. Conf. on The Practical Application of Intelligent Agents and Multi-Agent Technology, London, UK, pp. 345–360. Morgan Kaufmann, San Francisco (1999)

MDA Compatible Knowledge Based IS Development Process

Audrius Lopata[1,2] and Martas Ambraziunas[3]

[1] Kaunas University of Technology, Information Systems Department, Studentu St. 50, Kaunas, Lithuania
[2] Vilnius University, Kaunas Faculty of Humanities, Muitines St. 8, Kaunas, Lithuania
[3] Valdoware Inc.
Audrius.Lopata@ktu.lt, Martas.Ambraziunas@gmail.com

Abstract. Enhancement of MDA process with Knowledge Base Subsystem is aimed to reduce risk of project failures caused by inconsistent user requirements caused by insufficient problem domain knowledge. The enhancement of IS development environment with Enterprise Knowledge Base is discussed in this article. The major parts of Knowledge Base Subsystem are Enterprise Meta-Model, Enterprise Model and transformation algorithms.

Keywords: Enterprise Knowledge Based Information System Engineering, Model Driven Architecture, Enterprise Model, Enterprise Meta- Model.

1 Introduction

The majority of IT project failures (about 68% [2]) are caused by inconsistent user requirements and insufficient problem domain analysis. Although new methods of information systems engineering (ISE) are being researched and developed, they are empirical in nature: the project models repository of CASE system is composed on the basis of enterprise problem domain. The problem domain knowledge acquisition process relies heavily on the analyst and user; therefore it is not clear whether the knowledge of the problem domain is adequate. The knowledge stored in repository of CASE tool is not verified through formalized criteria, thus it is necessary to use advanced data capture techniques that ensure iterative knowledge acquisition process during which missing or incorrect data elements are obtained and fixed.

OMG provides Model Driven Architecture (MDA [4]) approach to information systems engineering where MDA focuses on functional requirements and system architecture not on technical details only. Model Driven Architecture allows long-term flexibility of implementation, integration, maintenance, testing and simulation. It means that enterprise modeling and user requirements engineering stages of information system engineering life cycle are not covered enough by MDA yet. There is lack of formalized problem domain knowledge management and user requirements acquisition techniques for composition and verification of MDA models. In order to solve this problem enhancement of MDA approach by the best practices [5], [6], [7]

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 33–38, 2010.
© Springer-Verlag Berlin Heidelberg 2010

of Knowledge Base IS engineering can be used. The proposed enhancement will intellectualize MDA models composition process by improving their consistency and decreasing the influence of the empirical information in composition process. Knowledge Base Subsystem will ensure MDA models verification against formal criteria defined by Control Theory [8]. It will reduce risk of project failures caused by inconsistent user requirements and insufficient problem domain knowledge verification.

2 Related Works

After a complete analysis of enterprise modeling standards including CEN CEN EN 40003 (CIMOSA) [6], UEML [7] is performed the following usage shortcomings of mentioned standards to particular Information System's Engineering steps (user requirements acquisition, analysis and specification) are defined:

- The content of enterprise models, presently used in CASE systems, is not verified according to formalized criteria. Although enterprise models are created on the basis of various notations, their composition is not verified by CASE systems with respect to defined enterprise domain characteristics such as relations among particular processes, functions, actors, flows and etc.
- Traditionally problem domain modeling and user requirements specification processes are performed by the analyst and user, who use their experience and knowledge. The result is empirical, because user requirements specification is not verified according to formalized criteria.

In order to solve these problems a particular Enterprise Meta-Model [9] has been developed in Kaunas University of Technology, Department of Information Systems. Its internal structure is based on Control Theory and best practices of the above mentioned enterprise modeling standards.

Recently, several authors [9], [10] proposed EMM based Use Case [10] and Class [16] models generation methods in order to decrease empirical nature of IS Engineering process already. These methods are able to cover static aspects of particular MDA steps, though mentioned EMM is not MDA compatible enough to fully support MDA process yet.

3 Knowledge Base MDA Principles

Most of MDA related techniques [1] are based on empirically collected problem domain knowledge thus negative influences verification of user requirements specification against actual customer needs. This problem can be solved by implementing Knowledge Base Subsystem (with EMM developed on basis of Control theory [8]) to particular MDA technique.

The Enterprise Knowledge Based Subsystem consists of three parts: the Enterprise Meta–Model (EMM), the Enterprise Model (EM) and model transformation algorithms. EM is a representation of the structure, activities, processes, information, resources, people, behavior, goals, and constraints of a business, government, or other enterprises [15]. EMM is formalized method based specification that regulates the formation order of the EM. Model generation algorithms handle transformation among MDA models (CIM, PIM) and EM.

3.1 Application of Knowledge Base Engineering Principles to MDA

According to survey [1], leading MDA based ISE methodologies need improvement in the following areas: requirements engineering, CIM construction, system models verification against problem domain processes. These problems can be solved by enhancing the MDA approach with Knowledge Base Subsystem. This subsystem is able to handle MDA models (CIM, PIM) verification process against EMM.

EM construction requires formal CIM structure. Although the existing numerous techniques [3] describe CIM construction procedures, most of them are not formalized enough that influences negative impact on the EM constructs and composition. Usage of modified workflow diagrams [10] and formalized questionnaires (by demand free form diagrams can be used as well) solves such shortcomings and properly supports the suggested method. Conceptual MDA architecture enhanced with Knowledge Base Subsystem is depicted in Figure 1.

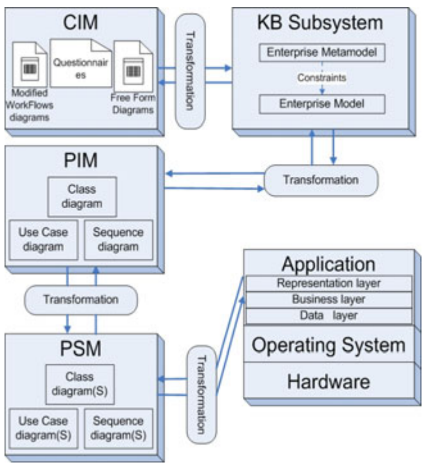


Fig. 1. Conceptual MDA architecture enhanced with Knowledge Base Subsystem

The set of modified workflow models [11] specifies the major CIM constructs. After CIM model is constructed, iterative verification process is started and is repeated until all the incorrect or missing CIM's elements are updated and corresponds to internal structure of EMM. The process leads to the creation of the EM. The following step is the transformation of EM to PIM. PIM consist of the following UML diagrams: Class, Use Case and Sequence. According to IBM [15] these diagram types are essential to PIM and PSM. PSM is the data source for the Application code generation. The following layers of application internal structure can be generated from PSM: representation layer, business layer, data layer.

3.2 Mappings among the Internal Models of PIM and Enterprise Meta-Model

During EM to PIM transformation process the following mapping types exists: EMM to Class Meta-Model (CMM), EMM to Use Case Meta-Model (UCM), EMM to Sequence Meta-Model (SMM). Further tables describe mappings among EMM and PIM model's elements.

Table 1. Mapping among Enterprise Meta-Model and Sequence Meta-Model

CMM elements / EMM elements	Object	Data	Method	Messag.	Constr.
Function			Mapped	Mapped*	
Process			Mapped	Mapped*	
Material Flow		Mapped			
Information Flow		Mapped			
Actor	Mapped				
Event			Mapped	Mapped*	
Information Activity			Mapped	Mapped*	
Business Rule					Mapped

Table 2. Mapping among Enterprise Meta-Model and Class Meta-Model

CMM elements / EMM elements	Class	Method	Attribut.	Relation	Interf.	Package
Function		Mapped				
Process		Mapped				
Material Flow			Mapped		Mapped	
Information Flow			Mapped		Mapped	
Actor	Mapped				Mapped	
Event		Mapped			Mapped	
Information Act.		Mapped				
BusinessRule				Mapped		
GroupingElement						Mapped

Table 3. Mapping among Enterprise Meta-Model and Use Case Meta-Model

CMM elements / EMM elements	Actor	Use Case	Exten. Point	Additi UC	System Bound.	Pack.
Function		Mapped				
Process		Mapped				
Actor	Mapped				Mapped	
Event		Mapped				
Information Act.				Mapped		
Business Rule			Mapped			
GroupingElement					Mapped	Mapped

3.3 The Main Steps of Knowledge Base MDA Approach

The following steps are depicted in figure 2:

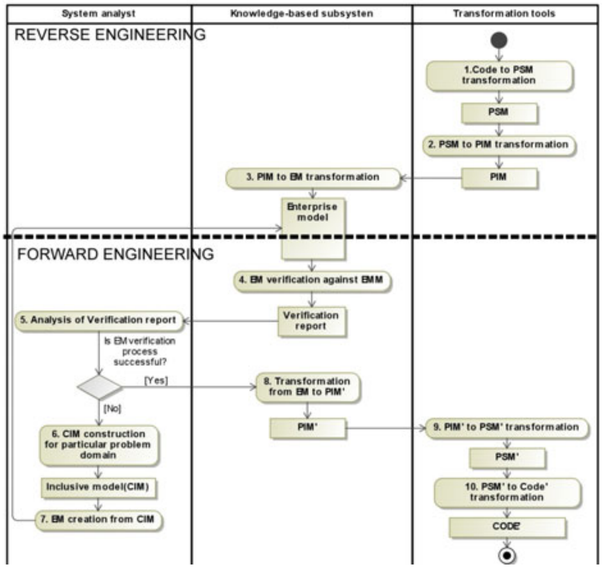


Fig. 2. Main steps of Knowledge Base MDA approach

1. Code to PSM transformation. Particular MDA tool performs transformation from programming code to PSM; *2.PSM to PIM transformation*. Particular MDA tool performs PSM to PIM transformation process, removing or transforming platform related constructs to higher abstraction (PIM) level; *3. PIM to EM transformation*. Knowledge- Based subsystem transforms PIM to EM of particular problem domain; *4. EM verification against EMM*. EM is verificated against Control Theory based EMM internal structure. Missing or incorrect EM data elements that do not correspond to EMM internal structure are determined during this step; *5. Analysis of Verification report*. System analyst evaluates Verification report and approves transformation from EM to PIM' process in case of success or defines necessary actions in order to solve EM inconsistency against EMM internal structure issue in case of failure; *6. CIM construction for particular problem domain*. Problem domain knowledge acquisition and CIM composition are performed at this step. This step is empirical in nature thus heavily depends on the system analyst's experience and qualification; *7. EM construction from CIM*. CIM is transformed to EM; *8. Transformation from EM to PIM*. XMI standard compatible PIM is constructed according to EM knowledge. It ensures PIM conformation to EMM defined formal constraints. PIM; *9. Transformation from PIM to PSM*. Particular MDA tool performs transformation from PIM to PSM, adding to PIM platform specific information; *10. Transformation from PSM to CODE*. Particular MDA tool performs transformation from PSM to programming code as well as to other artifacts such as executable files, direct link libraries, user documentation etc.

4 Conclusions

Knowledge Based Subsystem, which improves traditional MDA conception with best practices of problem domain knowledge and user requirements acquisition methods, is presented in this article. It ensures the problem domain knowledge verification against EMM internal structure. Such usage of Knowledge Based Subsystem along MDA will improve the consistency of software artifacts and will reduce IT projects dependency on empirical processes.

The EMM is intended to be formal structure and set of business rules aimed to integrate the domain knowledge for the IS engineering needs. It is used as the "normalized" knowledge architecture to control the process of construction of an EM. Some work in this area has already been done [12], [14]. The EM is used as the main source of Enterprise Knowledge for discussed MDA approach.

References

1. Asadi, M., Ramsin, R.: MDA-based Methodologies: An Analytic Survey. In: Schieferdecker, I., Hartman, A. (eds.) ECMDA-FA 2008. LNCS, vol. 5095, pp. 419–431. Springer, Heidelberg (2008)
2. Ellis, K.: The Impact of Business Requirements on the Success of Technology Projects. Benchmark, IAG Consulting (2008)
3. Ambler, S.W.: Agile Modeling, `http://www.agilemodeling.com/essays/inclusiveModels.htm`
4. OMG: MDA Guide Version 1.0.1, `http://www.omg.com`
5. ENV 40 003. Computer Integrated Manufacturing Systems Architecture - Framework for Enterprise Modelling, CEN/CENELEC (1990)
6. Vernadat, F.: UEML: Towards a Unified Enterprise modelling language. In: Proceedings of MOSIM'01, Troyes, France, April 25/27 (2001), `http://www.univ-troyes.fr/mosim01`
7. Gupta, M.M., Sinha, N.K.: Intelligent Control Systems: Theory and Applications. The Institute of Electrical and Electronic Engineers Inc., New York (1996)
8. Gudas, S., Lopata, A., Skersys, T.: Approach to Enterprise Modelling for Information Systems Engineering. Informatica 16(2), 175–192 (2005)
9. Lopata, A., Gudas, S.: Enterprise model based computerized specification method of user functional requirements. In: International Conference 20th EURO Mini Conference, EuroOpt-2008, Neringa, Lithuania, May 20-23, pp. 456–461 (2008) ISBN 978-9955-28-283-9
10. Lopata, A., Gudas, S.: Workflow- Based Acquisition and Specification of Functional Requirements. In: Proceedings of 15th International Conference on Information and Software Technologies IT 2009, Kaunas, Lithuania, April 23-24, pp. 417–426 (2009) ISSN 2029-0020
11. Gudas, S., Pakalnickas, E.: Enterprise Mangement view based Specification of Buisness Components. In: Proceedings of 15th International Conference on Information and Software Technologies, IT'2009, Kaunas, Technologija, pp. 417–426 (2009) ISSN 2029-0020
12. Leondes, C.T., Henry, R., Jackson, F.: Manufacturing and Automation Systems: Techniques and Technologies, p. 97. Academic Press, London (1992)
13. Skersys, T.: Business Knowledge-Based Generation of the System Class Model. Informatica 37(2), 145–153 (2008)
14. Lloyd, A.: How to migrate from code-centric to model-centric development using Rational Software Architect, `http://www.ibm.com/developerworks/rational/library/04/r-3247/index.html`

Incomplete Information within Relative Pairwise Comparisons as Utilized by the AHP

Edward W.N. Bernroider[1], Konradin Maier[2], and Volker Stix[2]

[1] Royal Holloway, University of London, School of Management, Egham TW20 0EX, United Kingdom
edward.bernroider@rhul.ac.uk

[2] WU, Vienna University of Economics and Business, Institute for Information Business, Augasse 2-6, 1090 Wien, Austria
konradin.maier@ai.wu.ac.at, volker.stix@wu.ac.at

Abstract. Motivated by a rising demand for feasible and practical decision aids this article relaxes the need for complete pairwise comparisons in Analytic Hierarchy Process (AHP) based approaches regularly applied in multiple criteria decision making practice. We propose a method based on the k-walk procedure from Graph Theory to extract a weighted ranking profile from an incomplete pairwise comparison matrix offering increased flexibility and practicability to the user. Our simulation tests show relative advantages over Harker's method. The connection into Graph Theory offers an understandable interpretation of weight vectors derived from either complete, incomplete, consistent or inconsistent pairwise comparison data.

Keywords: Paired Comparisons, Incomplete Information, Analytical Hierarchy Process, Decision Science, Simulation.

1 Introduction

Research into decision support seeks to improve human decision making, either from a normative, descriptive or prescriptive perspective [1]. While these views provide many different approaches in decision support, a common component in either perspective is to provide a ranking of alternatives within a decision problem based on the decision maker's (DM) preferences. In normative models the DM relies on axiomatic representations from rational thinking. One dominating model of rational choice is the expected utility model von Neumann and Morgenstern [2]. The descriptive view focuses on how real decision makers actually think and act, which often deviates from the normative ideal. One of the most known concepts are Prospect Theory models [3]. The prescriptive view lies in the middle and seeks to combine rational models with descriptive realities. Those models most likely include relaxations, approximations, less rigid assumptions and short cuts that make the model easier to formulate and execute. Given the

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 39–50, 2010.
© Springer-Verlag Berlin Heidelberg 2010

limited application of decision models in management practice (for example for Information System decision making see [4]), contributions in this area are most needed from a practical viewpoint. An example for a very prominent decision support methodology used in practice is the Analytic Hierarchy Process (AHP) proposed by Saaty [5], which uses a pre-defined structure with criteria and alternatives to be assessed with paired relative comparisons, which are seen to be more understandable in the involved assessment problem than absolute value judgments. From psychology we know that there are two different fundamental ways to measure preferences [6]. Using absolute measurement, the decision maker rates his preference towards a single stimulus (e.g., an alternative) on an absolute scale (like school grades). In relative measurement the decision maker judges on how much a stimulus A is preferred to another stimulus B on a relative scale. This is, however, more time consuming and contains the risk of incorporating inconsistencies. In the case of relative measurement with n alternatives, there are $n(n-1)/2$ possible paired comparisons (every alternative is compared with each other). Most of the comparisons are, however, redundant, since for constructing a weighted ranking we technically only need $n-1$ (specific) paired comparisons. A major challenge of redundant paired comparisons is to consolidate all relative preferences since the human judgements are not necessarily consistent. For example a DM could judge that an alternative A is twice preferred to B, B twice preferred to C, and A three times preferred to C (which should be, however, four times to due transitivity $A \to B \to C$). A DM could rank even prefer C over A. This inconsistency can be explained by the principle of bounded human rationality [7]. Therefore it is important to understand that we need redundant measurements of one and the same object, to overcome this imprecise human perception, i.e. more comparisons than only the technical sufficient $n-1$ (even though we might not need all possible $n(n-1)/2$). Referring back to AHP, the method needs all paired relative comparisons resulting in its major disadvantage of not being able to handle incomplete information. Thus, using the AHP we generally must perform all $n(n-1)/2$ paired comparisons. In some situations, the inability to deal with incomplete information renders the AHP inapplicable.

Harker lists three reasons, why incomplete information can occur in a paired comparison approach [8]. First, it can be a time-consuming task to complete $n(n-1)/2$ paired comparisons. Thus, the DM could wish to skip some of the comparisons to reduce the effort. The second reason why incomplete data can occur is that the DM does not want to openly compare two sensitive criteria; e.g., cost of treatment vs. expected life-extending effect in the case of public health insurance. The last reason listed by Harker is that a DM is simply not able (i.e. does not have enough information) to compare two elements. Instead of forcing the DM to make a guess and to spend much time on one comparison, he could simply omit such questions. We found at least a fourth reason, why incomplete information can occur. This is the comparison of two elements exceeding the relative scale of subjective measurement; e.g. using AHP on a 1 to 9 scale, an alternative A is eight times more important than B, B is twice as important as C, the comparison of A and C exceeds the scale. Here the comparison cannot properly reflect the DM's preference and furthermore, the DM is forced to judge inconsistent. Therefore, it would be advantageous to omit such comparisons as well.

Consequently, decision making practice benefits from strategies to deal with incomplete and inconsistent paired comparisons available at decision time. We add to literature by presenting a method to extract a weighted ranking profile out of an incomplete pairwise comparison matrix, which offers more flexibility to the user (e.g. for the AHP). In the next section we shortly introduce the AHP as a prominent candidate which utilizes pairwise preferences and we review other strategies dealing with incomplete information. For a better understanding we interpret in Section 3 the construction of the weight profile and we subsequently generalize the process for incomplete pairwise comparison of objects. In Section 4 we analyze the behavior of the proposed algorithm and compare it with Harker's method. Finally, we summarize the findings and identify further research issues in the last Section.

2 Basic Methods, Definitions and Applications

The Analytic Hierarchy Process (AHP) is one of the most popular multi criteria decision making methods. The AHP is based on three main principles [5]: the hierarchical decomposition of the decision problem into a goal-criteria hierarchy, the usage of complete relative paired comparisons and the measurement of preferences on a 1 to 9 ratio scale. This means the DM can prefer A over B 1, ..., 9 times or B over A which is equivalent to 1, 1/2, ..., 1/9. Given n objects, the decision maker (DM) estimates the $n \times n$ pairwise comparison matrix A. This is done for each level of the hierarchy by comparing each children nodes (sub-goals) with respect towards their common parent node (higher goal). After that for each matrix a weight-vector is calculated (as described below) which is used to build a weighted sum of all sub-goals. This is done recursively until the top level of the hierarchy is reached and a final weighted ranking is achieved.

Definition 1. *A pairwise comparison matrix A is called* consistent *if it fulfills the transitivity condition $a_{ij} = a_{ik}a_{kj}$ for all i, j, k.*

Definition 2. *A pairwise comparison matrix A is called* incomplete *if there exist at least on entry which is undefined. We mark such entries with zero values.*

It is worth mentioning, that the AHP is at least partly able to deal with some problems mentioned before. For example, Saaty proposed a workaround for situations where preferences exceeding the 1 to 9 scale are required by expanding the scale to $[1; \infty)$ [9]. Nevertheless, this requires a clustering of elements into sets which do not exceed the scale. The number of necessary comparisons can be reduced, as well, on a limited scale by restricting the number of child nodes of each parent node of the hierarchy. Nevertheless, this leads to a more complex task of hierarchy construction and the AHP does not offer a generally accepted methodology to build such hierarchies. The matter of inconsistent matrices is dealt by Saaty by allowing A an acceptable deviation and introducing a consistency index measure CI. Matrices which do not meet the required CI need to be reevaluated. Nevertheless, all these procedures do not solve the problem of our concern here namely the capability to deal with incomplete information. For a recent and more detailed introduction to the AHP, the reader is referred to [10].

Clearly A is reciprocal, i.e. $a_{ij} = {}^1/_{a_{ji}}$ and represents only the judgements of a decision maker. Let $w^T = (w_1, \ldots, w_n)$ be the relative weight vector, then the corresponding consistent pairwise comparison matrix A need to be $a_{ij} = w_i/w_j$ and

$$Aw = \begin{pmatrix} 1 & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & 1 & \cdots & \frac{w_2}{w_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \cdots & 1 \end{pmatrix} w = \begin{pmatrix} w_1 + w_2\frac{w_1}{w_2} + \cdots + w_n\frac{w_1}{w_n} \\ w_1\frac{w_2}{w_1} + w_2 + \cdots + w_n\frac{w_2}{w_n} \\ \vdots \\ w_1\frac{w_n}{w_1} + w_2\frac{w_n}{w_2} + \cdots + w_n \end{pmatrix} = nw. \quad (1)$$

Thus w is a eigenvector of A with the eigenvalue $\lambda = n$. This is the weight profile we are interested in, given a consistent matrix of pairwise comparisons. It is known that for these types of matrices this is the maximal eigenvalue, thus we are looking for the principal eigenvektor (EV) $w_{\lambda_{max}}$ of A.

Several approaches have been developed to deal with incomplete information within the AHP. In general, we can distinguish three different strategies. The first one is to replace missing comparisons a_{ij} by some estimation, based on transitive information available $a_{ij} = f(a_{ik}, a_{kj})$, and a subsequent calculation of the principal EV (e.g., consistency based estimation of missing preferece information [11]). The second strategy is to replace the EV-method by some other method, which can deal with incomplete information (e.g., [12]). As sketched above and showed by Saaty, the EV is required [13], therefore the second class of strategies is not very appealing. The probably most popular strategy is to reformulate the incomplete matrix A, as introduced by Harker [8]. The Harker method (HM) replaces missing comparisons with zeros, transforms the complete matrix A by incrementing the values of the main diagonal a_{ii} by one for each zero in row i and subsequently calculates the principal EV of the transformed matrix. For a deeper analysis of the HM using Monte Carlo Simulations see [14]. In the next Section we will present another strategy to deal with incomplete information by generalizing the approximation of the eigenvector method to incomplete comparison matrices.

3 A Method for Deriving a Weight Profile with Missing Information

The principal eigenvector can be approximated by the k-th power of the respective matrix A. This is done by averaging the row sums of A^k:

$$w_i \approx \frac{(A^k e)_i}{e^T A^k e} \quad (2)$$

where $e^T = (1, \ldots, 1)$. As $k \to \infty$ the approximation converges to the principal eigenvalue $w_{\lambda_{max}}$ [15].

Let A be a consistent but incomplete pairwise comparison matrix, where we replace unknown paired comparisons by zeros. Table 1 gives an example of such a matrix, where $a_{13} = 0$ and $a_{31} = 0$ indicate missing values. This Table also shows A^2 and the respective estimation of $w_{\lambda_{max}}$ using (2).

Table 1. Example of EV approximation for incomplete matrices

A

	E1	E2	E3	E4
E1	1	1/2	0	1/8
E2	2	1	1/2	1/8
E3	0	2	1	1/2
E4	8	4	2	1

A^2

	E1	E2	E3	E4
E1	3	3/2	1/2	3/8
E2	6	4	3/2	1
E3	8	6	3	3/2
E4	24	16	6	4

≈EV

E1	0.062
E2	0.145
E3	0.214
E4	0.579

From a_{43} we know that the DM prefers E4 over E3 two times. The approximated weight w does not reflect this preference, according to w E4 is ${}^{0.579}/_{0.214} = 2.7$ times preferred over E3. One could argue that A^2 is only a rough estimation for w. But let us have a closer look at e.g. a^2_{23} and a^2_{24}:

$$a^2_{23} = a_{13} \cdot a_{21} + a_{23} \cdot a_{22} + a_{33} \cdot a_{23} + a_{43} \cdot a_{24} = 0 + 2 + 2 + 2 = 6$$
$$a^2_{24} = a_{14} \cdot a_{21} + a_{24} \cdot a_{22} + a_{34} \cdot a_{23} + a_{44} \cdot a_{24} = 4 + 4 + 4 + 4 = 16$$

It can be seen that the zeros in matrix A lead to zero-sum sub-terms in a^2_{23}, but not in a^2_{24}. In general, when we calculate A^k, with increasing k the values of a_{ij} grow with different relative pace, depending on how many missing values (i.e. zero values) are present in A. This differences in relative pace biases the estimation of w for the DM. Therefore we cannot simply use the naive strategy of replacing unknown values with zero values to deal with incomplete information. This result is not very surprising since setting $w_i/w_j = 0$ and $w_j/w_i = 0$ implies that $w_i = w_j = 0$ this, however, is a contradiction within the relative importance of the two respective objects.

Lets have a different more intuitive look at the given problem at hand. We interpret A as a directed graph and the approximation of the principal EV as a k-walk problem in graph-theory. For a more detailed discussion on the EV-method and its graph-theoretic interpretation the reader is referred to [16]. We can think of any paired comparison matrix A as a directed graph, where a_{ij} represents the dominance of the edge from node i to node j. Figure 1 visualizes the incomplete matrix A given in Table 1 as a graph. The intensities specified in A reflect the one-step dominance of adjacent nodes. Similar, dominance can occur in two steps (e.g. in Table 1 node E1 dominates node E3 via E4), in three steps, in four steps, and so on. In graph-theoretic terms, the k-step dominance of a node over another node can be defined as follows. A k-walk is a sequence of k adjacent edges. The dominance of a k-walk is the product of the weights (dominances) of the edges in the walk. Then within the matrix A^k, a^k_{ij} is the sum of the dominance of all k-walks from node i to node j. Therefore, we can

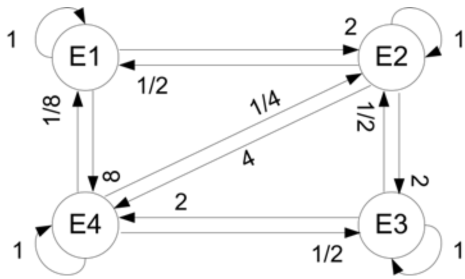


Fig. 1. Matrix A from Table 1 visualized as a graph

interpret the weight vector w derived by (2) in the following way: w_i reflects the normalized k-walk dominance of node i over all other nodes.

By interpreting A in graph-theoretic terms, it becomes clear what happens when simply replacing missing values with zeros. In Figure 1 apparently, there are four 2-walks from node E2 to E4: (E2, E1, E4), (E2, E2, E4), (E2, E3, E4), (E2, E4, E4), but only three 2-walks from node E2 to E3: (E2, E2, E3), (E2, E3, E3), (E2, E4, E3). Therefore, a_{24} increases faster than a_{23} as we raise matrix A to the k-th power, which biases the averaged weight vector w. To adjust to this effect, we need a weighted average of the number of k-walks. For this purpose, we define an adjacency matrix $B^{n \times n}$:

$$b_{ij} = \begin{cases} 1, & \text{if } a_{ij} > 0 \\ 0, & \text{else.} \end{cases} \tag{3}$$

Thus, $b_{ij} = 0$ for missing values and one else thus B^k represents the number of k-walks. We generalize (2) for incomplete paired comparison matrices, where unknown comparisons are marked with zeros:

$$w_i \approx \frac{(C^k e)_i}{e^T C^k e}, \text{ where } c^k_{ij} = \frac{a^k_{ij}}{b^k_{ij}} \ . \tag{4}$$

Doing so, we accept that not all nodes are directly connected and thus that not all walks of length k are possible. We average, however, the dominance between two nodes over all k-walks available. It is intuitive to see why a k-walk converges to the our weight profile w and that it is the logical extension, when pairwise comparisons (i.e. adjacent connections) are missing. Clearly there needs to be at least one walk from any node i to any node j. It follows that the graph needs to be connected. This condition is reasonable, since we need at least one (even vague) link between any two objects we are seeking to rank.

4 Behavior of the Proposed Method

While Section 3 introduced an algorithm to deal with incomplete pairwise comparison matrix, the purpose of this section is to discuss more general observations. In the following, we analyze the behavior of the proposed algorithm in three different cases: (i) complete comparison matrices, (ii) incomplete but consistent comparison matrices and (iii) incomplete and inconsistent comparison matrices.

4.1 Complete Matrices

For complete matrices, clearly our proposed estimation for w summarized in (4) is equal to (2), the approximation of the principal EV.

Proof. If A is complete than $a_{ij} > 0$ for all i, j. Therefore by (3) $b_{ij} = 1$ for all i, j and thus for all i, j the $b^k_{ij} = b^{(k)}$ are equal, namely n^{k-1}. Thus, (4) can be reformulated as follows:

$$w_i \approx \frac{\frac{1}{b^{(k)}} A^k e}{\frac{1}{b^{(k)}} e^T A^k e} \ , \tag{5}$$

which clearly reduces to (2).

In other words, in a complete graph the number of k-walks from node i to node j are identical for all i, j and therefore, we can interpret the number of k-walks as a scalar. Since (2) is invariant under scalar multiplication of the matrix A, (2) and (4) yield the same results. Note that here it is irrelevant whether the matrix is consistent or not.

4.2 Incomplete Matrices in the Case of Consistent Information

If A is consistent but incomplete with m values missing, and there exists at least one k-walk from node i to node j for all i, j, than our proposed algorithm (4) is able to reconstruct the underlying matrix and therefore (by the precious sub section) (2) and (4) yield identical weight vectors.

Proof. Since A is consistent it follows from Definition 1 that all k-walks from node i to node j have the same dominance. While a^k_{ij} represents the sum of these dominances, b^k_{ij} represents the number of all these k-walks, therefore

$$a_{ij} = \frac{a^k_{ij}}{b^k_{ij}} \ . \tag{6}$$

Thus after at most m iterations, where m is the number of missing values, using (6) the algorithm reconstructs all missing values and thus (2) and (4) yield identical weight vectors.

4.3 Incomplete Matrices in the Case of Inconsistent Information

Of course, the most interesting and most common case is, that some comparisons are missing, while the complete matrix is inconsistent as well. To analyze the behavior of the proposed method in this case, we will discuss the outcomes of some simulations within this sub section. For this purpose, a JAVA program was implemented using the IEEE 754 double precision standard for all numerical operations. One simulation-run, consisting of 112.000 simulations, lasted 8 hours executed on a 64-bit dual core processor with 1.83 GHz. We performed two runs to analyze two different scenarios. In the first scenario, we assumed a DM who is able to judge consistent. Nevertheless, since we used the 1 to 9 ratio integer scale, the assumed DM is forced to judge inconsistent to express intermediate preferences (e.g., 7.5 is rounded to 8). To generate matrices fitting this scenario, one row of each matrix was filled randomly based on the 1 to 9 scale (surely including 1 to 1/9 as well) and then the remaining values were calculated by transitive information and rounded to nearest-integer. In the second scenario, we assumed the other extreme: a DM who does not know any preferences. Therefore each pair was randomly generated, also based on the 1-9 scale. Within each of these two scenarios, we generated matrices of $n = 5, \ldots, 10$ and replaced randomly $m = 1, \ldots, (n(n-1)/2) - n$ pairs by zeros, indicating unknown values. A rule was used, which ensured that the resulting incomplete matrices were still connected graphs and therefore reasonable. For each type of matrix $A_m^{n \times n}$, 1000 matrices have been generated leading to 112.000 random matrices in total for each of the two runs. We compared our proposed algorithm with the Harker method to provide comparable results. The convergence criterion was set to 1×10^{-7} to analyze whether these two approaches lead to similar weight vectors. Nevertheless, A^k was restricted to $k < 200$, since the simulation showed unstable numeric results for $k > 200$, due to the limitations of the IEEE 754 double precision standard.

In the following, we will discuss the convergence behavior as well as the ability to estimate the weight vector for both, the Harker method and the method proposed in this paper. The following Figures only illustrate the outcomes of the near consistent scenario, since the observed patterns are similar to those of random matrices but the near consistent simulates reality much more appropriate, since DMs are not supposed to judge completely random. Nevertheless, we will mention interesting differences in the outcomes of the two scenarios.

Figure 2 shows the convergence behavior of the method proposed in this paper, while Figure 3 illustrates this for the Harker method. In both Figures the x-axis shows the number of missing values and the y-axis the number of iterations (k) to converge. As it can be seen, the proposed method converges for all n, m much faster than the Harker method. In more detail, in the case of the proposed method the iterations needed to converge increase rather linear, while in the case of the Harker method the iterations increase rather over-linear. Therefore, the Harker method did not converge frequently within the numerical limit of $k = 200$, while our proposed method never exceeded the limit of iterations. Figure 4 shows the number of matrices where the Harker method did not converge. Especially highly incomplete matrices (sparse ones) were hard for this method.

The matrices where the Harker method converged showed nearly identical results compared to our proposed method. The averaged absolute difference of the weight vectors is smaller than 0.1% for all n, m as illustrated in Figure 5. Not surprisingly, random matrices led here to increased differences, up to 2.5% (not illustrated).

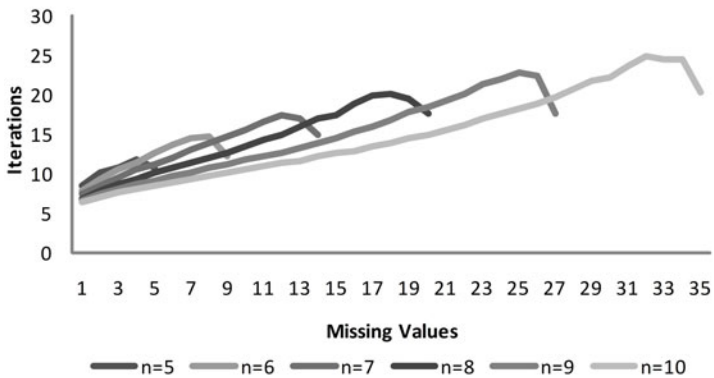


Fig. 2. Convergence Behavior of the proposed Algorithm

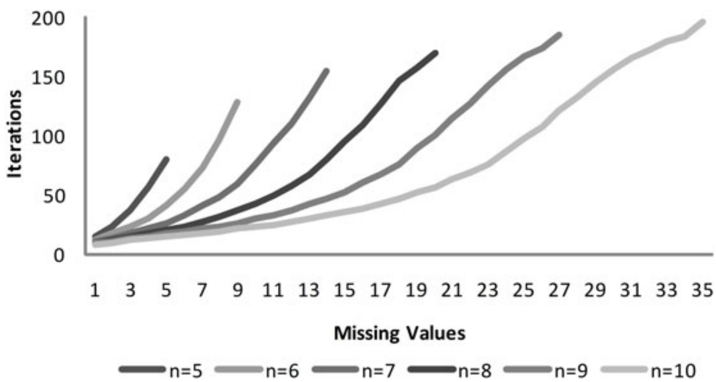


Fig. 3. Convergence Behavior of the Harker Algorithm

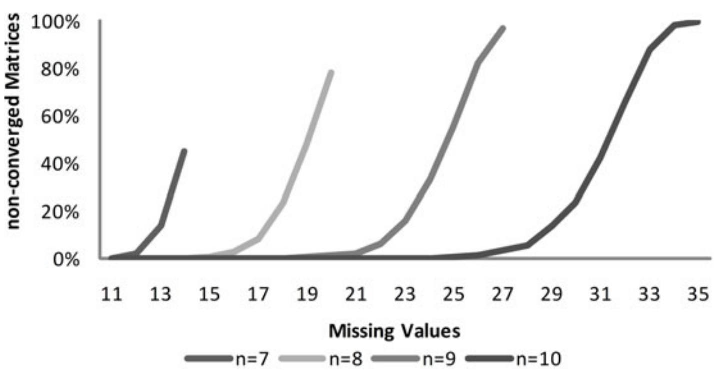


Fig. 4. Percentage of Matrices, where the Harker Method did not converge

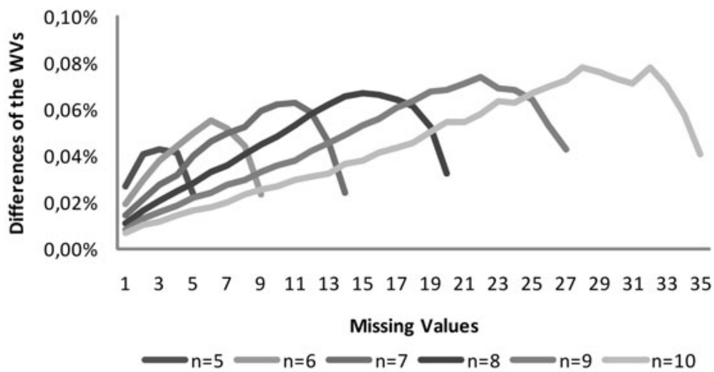


Fig. 5. Absolute Differences of the Weight Vectors derived from Harker Method and the Method proposed in Percentage (near consistent Matrices)

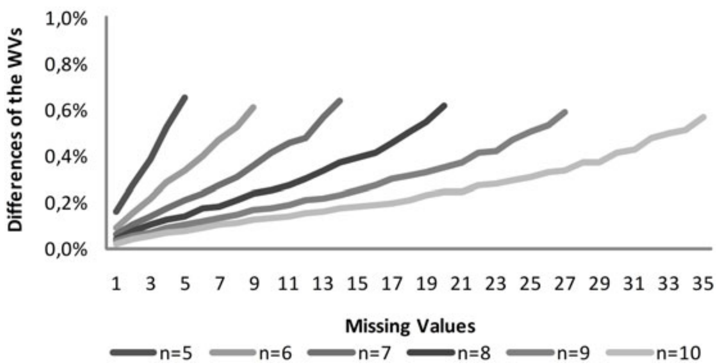


Fig. 6. Absolute Differences of the Weight Vectors derived from complete Matrices and from incomplete Matrices (near consistent Matrices)

As can be seen, the differences of the weight vectors slightly increase with size of the matrix and number of missing values. Nevertheless after reaching a peak the differences reduce. This is because there is little freedom left of estimating missing values (as in the beginning). However, the differences of the weight vectors are small $< 0.1\%$, thus the results are of both algorithms are consistent. Of course, the most interesting output of the simulation is the differences of weight vectors based on complete information and based on incomplete information. Figure 6 illustrates this only for our proposed method, since the differences towards the Harker method are less than 0.1% anyway. As expected, the averaged differences towards the weight vector derived from the complete matrix increase, as lesser information is available while the matrix size does not affect the weight deviation. The differences are quite low $< 0.8\%$. This small deviation is not very surprising, since the underlying matrices are only slightly inconsistent. Of course, in the case of random matrices the averaged difference between the weight vectors is much higher, namely up to 10%. Although one could think that this deviation of 10% is not that much, it is enough to completely change the rank order of the weight vector (note that a difference of 100% means, that one weight vector is the exact opposite of the other).

5 Conclusion

Answering the need for relaxations and less rigid assumptions in models for decision making practice, we have proposed a new approach coping with incomplete pairwise comparisons in the context of the AHP. By interpreting the Eigenvector problem in graph-theoretic terms, we developed a method to derive meaningful weight vectors from incomplete matrices based on a k-walk procedure. We showed that in the case of complete information and in the case of incomplete but consistent information this procedure leads to the same weight vectors as the Eigenvector Method. From the simulation we can conclude, that at least in the case of near consistent matrices the deviations of the weight vectors are negligible. In summary, the weight vectors derived by both methods are similar. However, the simulation showed that the Harker method needs more iterations to converge, especially if the incomplete matrices are sparse. This can lead to numeric instabilities which could render exact implementations more complex.

One main advantage of the method proposed in this paper is that it offers a common interpretation of weight vectors derived from complete, incomplete, consistent or inconsistent pairwise comparison data. In more detail, we can interpret the resulting weight vectors in graph-theoretic terms - the weights represent the k-walk dominance, relative to the number of walks and to the total sum of dominance. This is exactly in the line with the Eigenvector method suggested by Saaty.

Further research may include empirical comparisons of different approaches to deal with incomplete pairwise comparisons. Also testing whether the original AHP or incomplete information designs yield more accurate weight vectors in the case of large matrices or in the case of alternatives exceeding the relative scale would be valuable. The method proposed offers several issues for future research. For example, a detailed analysis of the method in the context of the rank reversal problem would be interesting, as well as to study the behavior of the method in dependence of different levels of inconsistency.

References

1. Smith, J.E., von Winterfeldt, D.: Decision Analysis in Management Science. Management Science 50(5), 561–574 (2004)
2. von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton (1944)
3. Kahneman, D., Tversky, A.: Prospect theory: An analysis of decision under risk. Econometrica 47, 263–291 (1979)
4. Smithson, S., Serafeimidis, V.: Information systems evaluation as an organizational institution - experience from a case study. Information Systems Journal 13, 251–274 (2003)

5. Saaty, T.L.: How to make a decision: the analytic hierarchy process. European Journal of Operational Research 48(1), 9–26 (1990)
6. Thurstone, L.L.: A law of comparative judgement. Psychological Review 34, 273–286 (1927)
7. Simon, H.A.: The New Science of Management Decision. Harper and Brothers, New York (1960)
8. Harker, P.T.: Alternative modes of questioning in the analytic hierarchy process. Mathematical Modelling 9(3-5), 353–360 (1987)
9. Saaty, T.L.: The Analytic Hierarchy and Analytic Network Processes for the Measurement of Intangible Criteria and for Decision-Making. In: Figueira, J., Greco, S., Ehrgott, M. (eds.) Multiple Criteria Decision Analysis: State of the Art Surveys, vol. 78, pp. 345–405. Springer, Berlin (2005)
10. Saaty, T.L.: Decision making with the analytic hierarchy process. International Journal of Services Sciences 1(1), 83–98 (2008)
11. Alonso, S., Chiclana, F., Herrera, F., Herrera-Viedma, E., Alcalá-Fdez, J., Porcel, C.: A consistency-based procedure to estimate missing pairwise preference values. Int. J. Intell. Syst. 23(2), 155–175 (2008)
12. Wang, K., Takahashi, I.: How to select paired comparisons in AHP of incomplete information-strongly regular graph design. Journal of the Operations Research Society of Japan 41(2), 311–328 (1998)
13. Saaty, T.L.: Decision-making with the AHP: Why is the principal eigenvector necessary. European Journal of Operational Research 145(1), 85–91 (2003)
14. Carmone, F.J., Kara, A., Zanakis, S.H.: A Monte Carlo investigation of incomplete pairwise comparison matrices in AHP. European Journal of Operational Research 102(3), 538–553 (1997)
15. Quateroni, A., Sacco, R., Saleri, F.: Numerical Mathematics. Springer, Berlin (2007)
16. Saaty, T.L.: Rank According to Perron: A New Insight. Mathematics Magazine 60(4), 211–213 (1987)

Short Term Trading Strategy Based on Chart Pattern Recognition and Trend Trading in Nasdaq Biotechnology Stock Market

Saulius Masteika

Vilnius University, Faculty of Humanities, Muitines 8, 44280 Kaunas, Lithuania
saulius.masteika@vukhf.lt

Abstract. The main task of this paper is to show the results of stock market trading strategy based on short term chart pattern. Proposed short term chart pattern is a trend following pattern and is relative to fractal formations and chaos theory. The proposed trading strategy consists of two steps: on the first step the stock screening algorithm is used to select volatile stocks in Nasdaq Biotechnology market; on the second step technical analysis and mathematical calculations for selected stocks are applied and profitability of strategy is calculated. The proposed trading strategy based on short term chart pattern was tested using historical data records from the USA Nasdaq Biotechnology market (2008-2010). The trading strategy applied in Biotechnology stock market had given higher returns if compared to the main USA stock market indexes (Dow Jones, S&P, Nasdaq).

1 Introduction

When taking decisions in the stock markets, investors rely on various criteria. Two main directions of market analysis are distinguished. One is fundamental analysis, based on financial indicators the other is technical analysis, based on historical data of price changes [8, 10]. With the increasing use of computerized technologies and diminishing trust in financial reports of enterprises, technical analysis is increasingly applied for market analysis [12].

Chart pattern is a part of technical analysis. A chart pattern is a distinct formation on a stock chart that creates a trading signal, or a sign of future price movements. There are two types of patterns within this area of technical analysis, reversal and continuation. A reversal pattern signals that a prior trend will reverse upon completion of the pattern. A continuation pattern, on the other hand, signals that a trend will continue once the pattern is complete [15]. Continuation pattern is used in volatile and unstable markets where there can be many overreactions, which cause the market movement trends to last longer than anticipated [9, 11, 13].

Typically the time required to build chart pattern varies from 5 till 60 periods [1, 7]. In these days' unstable economic environment and volatile markets the decision was made, that it is less risky to use short term trend trading strategies which in case of a sudden change in the market behavior can cause the adequate reaction.

The main goals of this article is to present a newly build trading strategy based on short term continuation pattern, to formalize it and to show the results of experimental

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 51–56, 2010.
© Springer-Verlag Berlin Heidelberg 2010

investigations in a volatile stock market. The method of formalization of chart pattern was chosen from previous research [7].

2 Basic Concept of the Proposed Short Term Trading Strategy

The proposed short term trading strategy is based on trend following chart pattern. Proposed chart pattern is relative to fractal indicators, presented by Ph.D. Bill M.Williams [2, 3]. The fractal indicator is build from chart bars and is used to determine when there is continuation in the market. There are many forms of fractals. That fact makes many uncertainties when using this indicator. The proposed chart pattern with 3 periods lets increase prediction sharpness because of ignoring older data information. Generalizing the proposed short term trading strategy one can be stated that the trading system looks for securities with the biggest price increases; if the trend of increase is confirmed, such securities are bought.

2.1 Chart Pattern Used for Trading Strategy

The block scheme of the identification algorithm of the chart pattern is provided in the next figure:

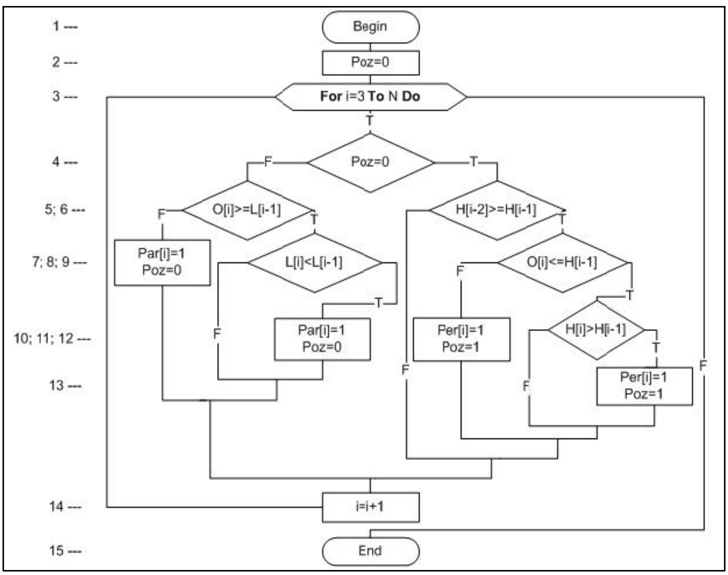


Fig. 1. Block scheme of the identification algorithm of chart pattern

The algorithm uses the *FOR* cycle, which allows to identify the formed signals in the period i. The cycle starts from $i=3$, because the data of the previous periods $i-2$ and $i-1$ are necessary for the identification of the signal. The cycle ends with variable N, which determines the last period of experimental investigation. If there are no open positions and the variable *Poz* is equal to zero (block 4), a signal of buying is sought.

The signal of buying is formed by checking whether the highest price *H* of period *i-2* is higher or equal of the highest price *H* of period *i-1* (block 6). If so, further check is carried – whether the period's *i* opening price *O* is lower or equal to the highest price *H* of period *i-1* (block 9). If the condition *O[i]<=H[i-1]* is not met, buying operation is carried out, the variable *Per* being attributed the value of one in the period *i*. Also, the value of one is given to the variable *Poz* (block 11). If the condition *O[i]<=H[i-1]* is met, a check is carried out whether the highest price *H* of the current period *i* is higher than highest price *H* of the previous period *i-1*. If so, a buying signal is identified (block 13). The same way the selling signal is identified by observing the lowest prices *L* of periods.

Short term chart pattern based on trend trading is characterized in the following way: after a large price increase of a share, price stabilization is awaited. After the price stabilization, one awaits the moment when the highest price of the previous day will be reached. At this point, shares are bought. With the application of this indicator, one expects that the price rising trend will remain strong and will continue for some time.

2.2 Quality Coefficient and Stock Ranking

In order to rank securities, there must be build a stock quality coefficient [7]. Proposed short term trading strategy is looking for securities, which during *n* day's period increased in price the most. The quality of the stock is measured using the following equation:

$$Q = \frac{C_{i-1} \times 100}{C_{i-1-n}} - 100 \tag{1}$$

In the equation (1) *Q* represents the quality coefficient, *C*- the price of the period *i-1*, *n*- the number of days used in calculation of price change. The biggest increase of a particular stock in price during *n* day's period means the best quality. When ranking the stocks there are applied some rules: the price of the stock in period *i-1* must be bigger than $2 and the volume must be at least 100000 stocks traded per day.

3 Experimental Investigation

The efficiency of the proposed short term trading strategy was tested using historical stock market data. The data set comes from USA Nasdaq Biotechnology stock market and represents daily *open*, *high*, *low*, *close, Adj.Close* prices and trading *volumes* from February 1, 2008 till February 1, 2010. The Nasdaq Biotechnology stock market was chosen not by a chance but rather than looking for volatile and variable markets. The conclusion, that Nasdaq Biotechnology market is one of the most volatile and variable was made after doing some calculations of standard deviation. The calculations showed that average standard deviation of price changes of companies that belong to Nasdaq Biotechnology index is much greater if compared with other indexes. More detailed comparison of calculations is shown in the next table:

The calculation shows that price changes of the companies belonging to Nasdaq Biotechnology 125 index tend to be at least 50% more variable than price changes of

Table 1. Standard deviation of price changes (average of index components)

Index	Standard deviation
Nasdaq Biotechnology 125	5.00
Nasdaq100	3.26
S&P100	3.24
DowJones30	2.86

Nasdaq100, S&P100 or DowJones30 components. The conclusion can be made that there are more companies with exceptional behavior of stock price changes in Nasdaq Biotechnology market in comparison with the other markets.

A portfolio of the trader was filled with one stock with the highest rank when applying the proposed short term trading strategy. Only one stock with the highest rank at that particular period was tested for chart pattern presence. If there were no chart pattern signals generated (Fig.1) no action were taken at that period. The transaction costs of 0.35 % were withdrawn from the returns per contract. Taking into consideration the rules applied to formalization of technical analysis indicators specialized standardization of historical data was used [7]. Specialized standardization of historical data consisted of the following parts:

- Recalculation of prices, taking into account share splitting and dividends paid. Following equations were used to recalculate the prices:

$$Delta = \frac{Adj.Close \times 100}{Close},\ Open(new) = \frac{Open \times Delta}{100},\ High(new) = \frac{High \times Delta}{100}, \quad (2,3,4)$$

$$Low(new) = \frac{Low \times Delta}{100}, \quad (5)$$

  where *Adj.Close*, *Close*, *Open*, *High* and *Low* are historical data of a particular stock at concrete time period. *Yahoo!®* *Finance* was chosen as data provider.
- Elimination of lack of data. The securities with interruptions in their data sequences were eliminated from experimental investigation.
- Recalculation of absolute values into percentage values
- Buying/selling imitation of securities;
- Calculation of the results of trading.

The detailed functioning of the proposed short term trading strategy can be described as follows:

- Obtaining and specialized standardization of historical data (Eq.2-5);
- Calculating the quality of stocks at particular period; the parameter of $n=2$ was chosen (Eq.1). Selecting the stock with the highest quality coefficient;
- Application of short term chart pattern algorithm (Fig.1);
- Buying/selling imitation of securities;
- Calculation of the results of trading.

The total returns when using proposed short term trading strategy are presented in Figure 2 and Figure 3.

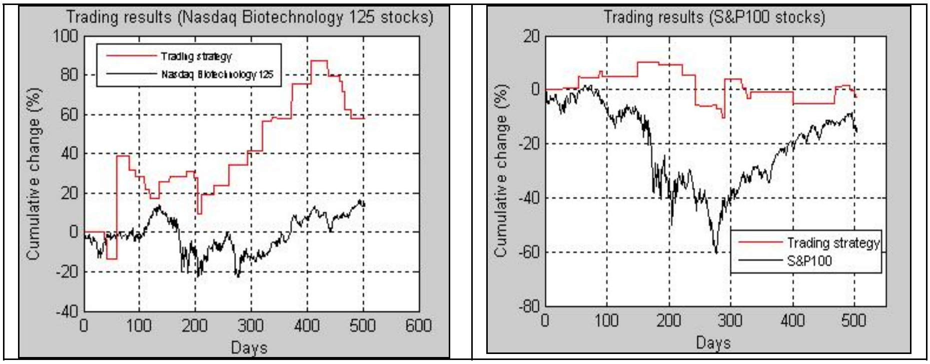


Fig. 2. Profitability of trading strategy in Nasdaq Biotechnology and S&P100 markets

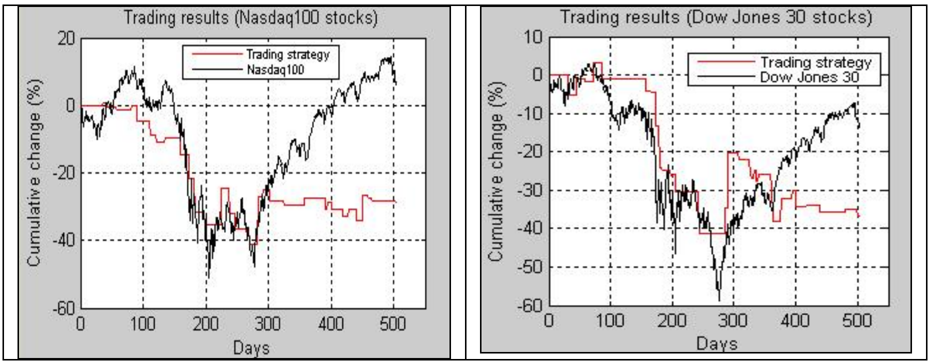


Fig. 3. Profitability of trading strategy in Nasdaq 100 and Dow Jones 30 markets

The results of experimental investigation have shown that the proposed short term trading strategy serves the purpose when trading in companies bellowing to the indexes of Nasdaq Biotechnology and partly S&P100. There are a low number of transactions when using proposed trading strategy. Capital fluctuations are medium.

The comparison of results with other trading strategies presented by other researches is not presented because the comparative analysis faces difficulties with accuracy and compatibility. It is so, because the indicators of technical analysis which are used in other trading strategies are not formalized completely or formalized differently. While doing experimental investigations other researches treat differently the quality of historical data, precise description of indicator's parameters and trading signals, the strength of the signals and etc. This leads to the absence of unanimous rules concerning specific indicators which affects analysis inaccuracies. In order to compare different trading strategies and to eliminate possible heuristic deliberations about indicators, the same formalization method of technical analysis indicators must be applied for all the strategies that take place in analysis [7]. Generalizing the experimental investigation one can state that the proposed trading strategy has proved that the best results were gotten in volatile and variable market. The strategy has infrequent trading signals which lead to periods when the investment capital is not used. In order do deepen the investigation further experiments with other trading strategies using the same method of formalization of technical analysis indicators should be done.

4 Conclusions

Preliminary experimental tests have shown that the proposed short term trading strategy tested with the companies that belong to Nasdaq Biotechnology index outperforms the main USA stock indexes. Less promising results were got with stocks that belong to less volatile and variable indexes Nasdaq100 and DowJones30. The conclusion can be made that the proposed strategy can be attractive to individual traders who are intended to trade using short term strategies especially when volatile movement appears in the markets. Despite of good preliminary results in Nasdaq Biotechnology market, more extensive tests with the proposed methodology may be done *(different markets in other countries, time periods, portfolio diversification)* to evaluate more precisely the whole domain of possible applications of this strategy.

References

1. Alan, N.: The complete guide to investing in short term trading: How to earn high rates of returns safely, p.288. Atlantic Publishing Company (FL) (2008)
2. Williams, B.: New trading dimensions, p. 288. John Wiley & Sons, New York (1998)
3. Williams, B.M., Williams, J.G.: Trading chaos: maximize profits with proven technical techniques, 2nd edn., p. 228. John Wiley & Sons, Inc., New Jersey (2004)
4. Masteika, S., Simutis, R.: Stock trading system based on formalized technical analysis and ranking technique. In: Alexandrov, V.N., van Albada, G.D., Sloot, P.M.A., Dongarra, J. (eds.) ICCS 2006. LNCS, vol. 3994, pp. 332–339. Springer, Heidelberg (2006)
5. Ned, G.: Winning Edge trading: successful and profitable short and long term systems and strategies, p. 224. Wiley trading, Chichester (2009)
6. Turner, T.: Short-term trading in the new market, p. 384. St.Martin's Griffin (2006)
7. Masteika, S.: Formalization and application of technical analysis indicators in securities markets, Doctoral Dissertation, Vilnius, p.147 (2007)
8. Achelis, S.B.: Technical Analysis from A to Z, 2nd edn., p. 380. McGraw-Hill, New York (2000)
9. DeBondt, W., Thaler, R.: Does the stock market overreact? Journal of Finance 40, 793–805 (1985)
10. Chande Tushar, S.: Beyond Technical Analysis: How to develop and implement a winning trading system, 2nd edn., p. 336. Wiley, Chichester (2001)
11. Edwards, R.D., Magee, J.: Technical Analysis of Stock Trends, 7th edn., p. 792. John Magee, Inc. (1998)
12. Carr Thomas, K.: Trend trading for a living: learn skills and gain the confidence to trade for a living, p. 312. McGraw-Hill, New York (2007)
13. DeBondt, W., Thaler, R.: Further evidence on investor overreaction and stock market seasonality. Journal of Finance 42, 557–582 (1987)
14. Hong, H., Stein, J.: A unified theory of underreaction, momentum trading, and overreaction in asset markets. The Journal of Finance LIV (6), 2143–2184 (1999)
15. Janssen, C., Langager, C., Murphy, C.: Technical analysis: chart patterns (February 25, 2010), http://www.investopedia.com/university/technical/techanalysis8.asp

A Model of Employee Selection for SME Based on Innovation Transfer

Justyna Patalas-Maliszewska[1] and Irene Krebs[2]

[1] University of Vienna/Austria, University of Zielona Góra, Polen
j.patalas@iizp.uz.zgora.pl
[2] Brandenburg University of Technology Cottbus, Germany
krebs@iit.tu-cottbus.de

Abstract. Human resource strategy can emerge within a decentralized decision structure that gives managers autonomy to take responsive actions while overall strategic direction is considered within a strategic planning process. This study defines the concept of employee selection (especially strategic employee selection) and hypothesizes on the positive correlation between innovation characteristic in SME and value of strategic employee (the so called personnel usefulness function). An empirical study illustrates the importance of both elements in an integrative human resource strategy formation process particularly for firms operating in the international environments.

Keywords: the personnel usefulness function, innovation characteristics, the employee selection efficiency method.

1 Introduction

Growing international competition necessitates increasing interest in new scientific advancements and new solutions in the field of management based on innovation transfer. Management staff forms a basis for expediting structural changes in the regional economy. Staff is also a decisive factor for enhancing competitiveness. In the knowledge based economy, while trying to maintain their position in the market or access new markets, managers and employees need to improve their skills. This will help in being able to use innovative technologies that exist all over the world.

It is generally recognized that effective human resource strategy formation processes consist of central planning activities orchestrated at the corporate strategic apex as well as emerging strategic decisions influenced by empowered managers within the organization [Hill and Jones, 2000; Johnson and Scholes, 2002; Mintzberg et al., 2003].

Companies operating in a market economy need to introduce changes to their systems organisation and management. In business practice, decision making in a company depends on activity, competition, changing external factors, e.g. technical advancement, and results achieved by research and development departments, knowledge, employees' skills, social relations, know-how, and in particular, effective investment in intellectual are an added value for a company. Companies investing in human capital and work systems acquire competitive advantage thanks to their

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 57–66, 2010.
© Springer-Verlag Berlin Heidelberg 2010

readiness to learn and adopt new qualifications as well as efficient communication and information channels.

Small and medium sizes enterprises (SMEs) that are about to make decision concerning the employee selection tend to make a pre-evaluation of the efficiency of this decision. There is a demand for developing the method that would diminish the risk of an inadequate employee selection and at the same time would allow to solve the problems which otherwise could be missed. A relevant framework to this issue is based on a database referring to:

The SME, with defined the selected functional area, the business processes, the workplaces and the personnel usefulness function for m-the employee in the n-th SME: W_{nm}, $n,m \in N$ is given. Allowing responding to the following question: Whether a existed (given) the employee selection efficiency method based on the innovation transfer in the small and medium enterprises (SME)?

It can be concluded that there is a need to define the criteria of the employee selection (especially strategic employee) efficiency in the SME. Consequently an appropriate reference model of the company should be developed to enable both defining: the needs in the areas of functionality and the success evaluation of employee selection. Such a model should provide a kind of guideline for the future employee selection framework.

The decision as far as the selection of the proper employee should be preceded by the management profitability assessment and the prospect advantage. In order to view the prediction of the described innovation indicators the authorized consulting system supporting decision making enabling to asses and forecast of knowledge in SMEs (software) can be used (an Advisory Computer System for Forecasting the Efficiency of a Strategic Employee Selection in Small and Medium Enterprises). This tool is the implementation of the method of assessment of efficiency strategic employee selection based on innovation indicators in the small and medium enterprises (SME), which in this paper is presented.

This computer program could be a supplement as the module to the for example compute system: Enterprise Resource Planning (ERP). ERP is process oriented software that implementation can improve business processes running through functional areas of a company onto computer system [Klaus at al., 2000]. Implementation of ERP is a strategic decision, that results about efficiency of business processes, charge of resources and generally about compete and management efficiency of the enterprise for the next years.

2 The Employee Selection Efficiency Method - ESE

Knowledge in the paper is understood as improving qualifications and skills among employees by participating in the implementation of innovative projects [Quinn et al., 1996]. The process of managing intellectual capital should consist of two stages: identifying and measuring. Literature distinguishes qualitative measures (e.g. Danish project of IC measurement, ‘Scandia’ navigator, intengible assets monitor, IC model – TM Rating, VCSTM, balanced result sheet, report by Saratoga Institute) and methods of valuating intellectual capital (e.g. MV / MB, q-Tobin, CIV, KCE, VAICTM, economic added value, IAV model, Strassmann’s method, IAMVTM, technology broker)

[Dudycz, T.; 2005], [Edvinsson, L., Malone, M.; 1997; Sokołowska, A.; 2005; Fitz-enz, J.; 2001; Nonaka, L.; Takeuchi, H.; 1995]. However, there are no methods assessing the efficiency of decisions on acquiring knowledge based on innovation transfer.

The author's method was developed on the basis of the defined method for the intellectual capital profitability assessment, including innovation characteristics. It consists of 4 elements that is:

(1) Experience in SMEs as regards investment in knowledge – so called the reference SME model,
(2) Values of personnel usefulness function for m-th employee in the n-th SME: W_{nm},
(3) Innovation parameters (characteristics),
(4) Algorithm that enables values of personnel usefulness function: W_{nm} and innovation characteristics binding – GMDH algorithm.

The name of the method is an acronym that enhances the main operating methods, namely: **E**fficiency – **S**election – **E**mployee.

In order to define the employee selection efficiency method (ESE) it is necessary to find the answer for the following questions:

(1) What kind of SME is to be discussed?

A reference SME model proposed is based on literature and own scientific research [Kasprzak, 2005]. It involved a survey of selected SMEs, in the sales area. The research group consisted of 10 companies. Business processes were defined in detail in each division of the company. The reference model has been developed (an SME according to the regulation dated November 12, 1999, Commercial Law – Dz .U .Nr 101, poz 1178), which includes the following (1): business processes, employees (description of workplaces), and value of personnel usefulness function for m-th employee in the n-th SME. The business processes in each functional areas in SME describe employees activities. The personnel usefulness function is defined for each m-th employee, which realize the determined set of business processes.

(2) What personnel usefulness function is to be considered?

So, the following personnel SME usefulness function W_{nm} for the m-th employee in the n-th SME: is proposed [Patalas – Maliszewska J., 2009]:

$$\mathbf{W_{nm} = f_1(GK) + f_2(PK) + f_3(A) + f_4(E) + f_5(P) + f_6(C) + f_7(P),} \text{ where: } n, m \in N,$$

The linear form of this function W_{nm} is chosen because all elements are equally important to assess effectiveness and efficiency of investment in knowledge.

- $f_1(GK)$ – the general knowledge function for the m-th employee in SME, where: $GK \in R$, and $0 \leq f_1(GK) \leq 5$,
- $f_2(PK)$ – the professional knowledge function for the m-th employee in SME, where: $PK \in R$, and $0 \leq f_2(PK) \leq 5$,
- $f_3(A)$ – the professional abilities function for the m-th employee in SME, where: $A \in R$, and $0 \leq f_3(A) \leq 5$,

- $f_4(E)$ – the experience function for the m-th employee in SME, where: E – is a synthetic index of experience for the m-th employee in SME binding the factors di: E = $\frac{\sum_{i=1}^{3} di}{3}$ where: d1- year of work, d2- age of employee, d3- number of realized project. Each indicator $f_4(E)$ is assessed on the points scale (0 – 5) and $0 \le f_4(E) \le 5$,
- $f_5(P)$ – the patents function for the m-th employee in SME, where: P - synthetic index of patents for the m-th employee binding the factors ei: P = $\frac{\sum_{i=1}^{4} ei}{4}$ where e1- number of patents, e2 - value of investment of new patents, e3 – value of copyright, e4- number of project, which are waiting for patents. Each indicator $f_5(P)$ is assessed on the points scale (0 – 5) and $0 \le f_5(P) \le 5$,
- $f_6(C)$ – the clients function for the m-th employee in SME, where: C - synthetic index of clients for the m-th employee binding the factors ki: C = $\frac{\sum_{i=1}^{3} ki}{3}$ where: k1- number of all clients, k2 - number of permanent clients, k3– number of transactions. Each indicator $f_6(C)$ is assessed on the points scale (0 – 5) and $0 \le f_6(C) \le 5$,
- $f_7(P)$ – the m-th employee's personality in SME, where: $P \in R$, and $0 \le f_7(P) \le 5$.

It is possible to receive indispensable data for account of value personnel usefulness function from companies belonged to reference model of SME by interview in each enterprise.

(3) What innovation characteristics are to be considered?

The notion of innovation is treated flexibly depending on the field in which it is used. The bases of being economically competitive enterprise in the SME sector are directly initiative and innovation. The authors treat enterprise innovation as the ability and motivation to search and commercially use the results of scientific researches. So, following innovation characteristics are proposed [Krebs, I.; Patalas-Maliszewska, J. 2009].:

- Research and development cost.
- Employee's number in SME in R&D.
- R&D centres 's number.
- R&D cost financed by public funds (euro).
- R&D cost financed by private funds (euro).
- Number of cooperation's agreement with R&D company.
- New product's number/product's number.
- Computer's number.
- Patent's number.
- Project's number waiting for patent.
- Consulting cost per year.

- Employee's number: education's level: college.
- Employee's number: education's level: university.
- Employee's number: education's level: post doc.
- Employee's number: education's level: MBA study.
- Employee's number: education's level: post graduate study.

(4) In which way the information concerning value of personnel usefulness function and the innovation parameters is to be accessed?

The development of the employee selection decision model should be started with collecting as much information as possible in relation to the structure and the dynamics of the object (SME) in question. It can be a subjective knowledge which involves the empirical data obtained as a result of the observation on the SMEs functionality. This approach consists of the complete data because it has been obtained in real situation. The elements in decision model (the value of the personnel usefulness function, the value of the selected innovation parameters) were obtained as a result of the research group consisted of 10 companies.

(5) What kind of algorithm can be used to combine value of personnel usefulness function, value selected innovation parameters?

For defined object – the pair: the value of personnel usefulness function W_{nm} for the m-th employee in the n-th SME and the value of selected innovation parameters was making the empirical database of this indicators. The algorithm that enables the both values was defined as GMDH (Group Method Data Handling), that involves the following assumptions [Farlow S.J. (Ed.), 1984; Iwachnienko A.G., 1982]: a precise description of the interdependence between the output and input data (the selected innovation parameters with the characteristics of a given the personnel usefulness function as well as the characteristic of the company – see Table 1) and minimum modeling error. As a result of the algorithm GMDH implementation the best possible polynomial (the polynomial decision-making model) was obtained which was characterized by the lowest value criteria for regularity assigned to the pair object (respectively – the personnel usefulness function and the innovation parameters).

The polynomial decision-making model will allow us to define the value of personnel usefulness function W_{nm} for the m-th employee in the n-th SME for the assessment of the efficiency of the selection of a senior management employee. On the basis of the forecasted values of these indicators, the company's board of directors will make a decision about the employment of a strategic employee.

(6) What kind of decision supporting structure should be used in relation to the employee selection?

Decision model is contracted on the basis of the database. It includes a complex information about all the processes which could be observed while the database was created, so both examples of successful and unsuccessful employee selection (value of personnel usefulness function) are included. The application of empirical knowledge enabled the application of GMDH as a modeling tool. In conclusion the decision model, which was under examination, binds the selected indicators of innovations characteristics with the value of the personnel usefulness function for m-th employee in the n-th company. This restriction makes the decision making process simple and brings it to some kind of pattern of the restriction propagation (chosen decision making indicators of employee selection system under examination). It means that, for some companies, the assessment of the effects which employee would bring can be done on the basis of previously defined so called reference SME model.

In accordance with the date included in the value of the personnel usefulness function database, including the innovations parameters all the variations of the GMDH algorithms were investigated in the computer system supporting decision making enabling to asses and forecast of knowledge in SMEs (an Advisory Computer System for Forecasting the Efficiency of a Strategic Employee Selection in Small and Medium Enterprises).

3 A Model of Employee Selection for SME

The main problem in responding to the question whether a given strategic employee (a product manager) will guarantee to obtain the assumed level of a SME performance index or not – it is presented like the decision problem.

3.1 A Polynomial Decision-Making Model Using GMDH Algorithm

In order to solve the research problem, a polynomial decision-making model using GMDH algorithm (see Formula 3.1) has been designed for an strategic employee selection for an innovative SME. The model compiles all groups of the elements of the method ESE. The generally form of the decision model, which will enable the connection of the value of the personnel usefulness function to the value of the innovation indicators is defined:

$$y(x_n, x_m) = A + B\,x_n + C\,x_m + D\,x_n^2 + E\,x_m^2 + F\,x_n x_m \qquad (3.1.),$$

where:

where: y (x_n, x_m) - SME performance index , x_n - the value of the personnel usefulness function for a strategic employee in SME, x_m - the value of the innovation indicators in SME.

In order to illustrate the possibility of answer let us consider the situation: the problem considered regards of chosen a product manager and of assessment of effects of the strategic employee selection for enterprise A.

So, it is proposed the database: the values of personnel usefulness function for product manager (strategic employee) in the 10 SMEs (x_1), the values of the innovation indicators from 10 SMEs (x_2 – research and development cost per year, x_3 - employee's number in SME in R&D, x_4 - R&D centers ‘s number , x_5 - R&D cost financed by public funds (euro), x_6 - R&D cost financed by private funds (euro), x_7 - number of cooperation’s agreement with R&D company, x_8 - new product’s number/product’s number, x_9 - computer’s number, x_{10} - patent’s number, x_{11}- project’s number waiting for patent, x_{12} - consulting cost per year (euro), x_{13} - employee's number: education’s level: college, x_{14} - employee's number: education’s level: university, x_{15} - employee's number: education's level: post doc, x_{16} - employee's number: education's level: MBA study, x_{17} - employee's number: education's level: post graduate study and the values of SMEs performance index (y), (see Table 1):

Table 1. The database: the values of the innovations characteristics, the values of the personnel usefulness function for a product manager, values of performance index in the 10 SMEs – research results

	y	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}
SME1	183250	16	12500	4	0	2500	10000	2	0	0,06	25	0	0	17500	11	21	0	0
SME2	225000	17	37500	6	0	0	37500	4	2	0,05	30	0	0	22500	10	29	0	0
SME3	262500	18	50000	7	0	10000	40000	5	4	0,06	39	0	3	25000	11	32	0	0
SME4	162500	17	5000	2	0	0	5000	3	3	0,06	20	0	1	5500	8	20	0	0
SME5	225000	15	15000	4	0	10000	5000	5	4	0,05	25	0	0	12000	6	45	0	0
SME6	750000	16	2000	3	0	2000	0	2	8	0,10	40	0	0	3500	5	34	0	0
SME7	125000	19	0	0	0	0	0	0	2	0,00	12	0	0	0	0	76	0	0
SME8	100000	21	20000	4	0	18000	2000	5	3	0,02	32	0	0	1000	3	42	0	0
SME9	195000	15	5000	1	0	0	5000	7	5	0,30	18	0	0	500	2	21	0	0
SME10	195000	17	1000	3	0	0	1000	2	2	0,06	10	0	2	2000	7	12	0	0

The stage of the employee selection variations and prediction testing of certain innovations indicators allows for the introduction of the conclusions concerning the ways in which the prediction is proceeded using ESE implementation method. The main aim for the experts was to determine prediction value for the personnel usefulness function for product manager. The object on which efficiency examination is carried out in relation to the pair: a SME company, which considers the value of the personnel usefulness function for a product manager and a properly defined the value of the innovation indicators. The author's software *Advisory Computer System for Forecasting the Efficiency of a Strategic Employee Selection in Small and Medium Enterprises* [Patalas-Maliszewska J., 2010] facilitates the proceeding of previously described experiment.

So, the best possible polynomial (decision model) binding the selected indicators like: x_1 - the value of the personnel usefulness function for a product manager (employee), x_2 - research and development cost per year (using the author's software) – see Table 1, so-called the model of employee selection for SME is defined:

$$y\ (x_1, x_2) = 186379 - 1{,}37x_1 + 0{,}83x_2 + 0{,}4x_1^2 + 0{,}0000008x_2^2 - 0{,}0000003x_1x_2 \qquad (3.2.)$$

where: $y\ (x_1, x_2)$ - SME performance index , W x_1 - the value of personnel usefulness function for product manager (employee), x_2 - research and development cost per year.

The model of employee selection for SME (see the Formula 3.1), which has been defined, enables us to carry out a assessment of the employee chosen. As a result, on the basis of the obtainable prediction values, it could be recommended for the company to chose strategic employee, which was conditioned by the prediction of the better values for the SME performance index.

The so-formulated decision model (see the Formula 1) on the basis of the data from the Table 1 is built and can be changed depended on the implemented data.

3.2 The Scheme of the SME Procedure in the Process of a Strategic Employee Selection Using the Model of Employee Selection

The decision-making situation, where an SME is considering the employment of a senior management employee, is presented in the diagram:

Such a strategic employee is sought who guarantees an SME a gain of the desired values of qualification criteria for an innovative company?

↓

Step 1: Establishing the current value of the innovation index: research and development cost per year in a company A.

↓

Step 2: Conducting an interview with potential strategic employees – establishing the current value of a m-th strategic employee (the value of the personnel usefulness function W_{nm} for the m-th strategic employee – see chapter 2, question 2)

↓

Step 3: Using the decision model in the process of employee selection for SME (the application of method ESE) according with a formula 3.2.

↓

Step 4: Assessing of the forecast value of the personnel usefulness function W_{nm} for the m-th strategic employee in the n-th SME

↓

Step 5: The comparison of the forecasted value of a strategic employee (the value of the personnel usefulness function W_{nm}) with the current value.

↓

Step 5: The recommendation of the strategic employee (in terms of higher values of the personnel usefulness function W_{nm})

Fig. 1. The scheme of the SME procedure in the process of a strategic employee selection using the model of employee select

On the basis of the decision making model a forecast of a defined indicators value is introduced to the company . As a result, the company must make a decision as far as the product manager selection based on value of SME performance index forecasting is concerned.

4 Concluding Remarks

The enterprises functioning in market economy have to implement changes in systems of organization and management that they use. In economy practice making a decision in enterprise is conditioned by competitors' action, changing factors of environments, eg. technical progress and results of the research works. Added value for SME can be determine as knowledge, employees' skills and abilities, social relation, know-how, and particularly effective investing in intellectual capital. The enterprises which invest in human capital and systems of work are achieved competitive advantage because of their workers' readiness to learning and qualifying themselves and also thanks to effective information and communication transfers.

On the basis of the employee selection efficiency method – ESE the decision model for a strategic employee efficiency chosen has been developed . It was concluded that on the basis of the decision model the company of SME sector will obtain the prediction of the defined strategic employee (the value of the personnel usefulness function W_{nm}).

The method proposed shows just a concept associated with the assessment of the employee selection in order to find their SME's effective chosen. It means that, for some companies, the assessment of the effects which employee would bring can be done on the basis of previously defined indicators and the experience of those companies.

Acknowledgement. The research leading to these results has received funding from the European Community's, Seventh Framework Programme [FP7/2007-2013; FP7-PEOPLE-IEF-2008], under grant agreement n°[235585] "SKnowInnov".

References

1. Dudycz, T.: Enterprises value management. PTE, Warsaw (2005)
2. Edvinsson, L., Malone, M.: Intellectual Capital: realizing your company's true value by finding its hidden brainpower. Harper Business, New York (1997)
3. Farlow, S.J. (ed.): Self-organizing Methods in Modelling: GMDH-type Algorithms. Marcel Dekker Inc., New York (1984)
4. Fitz-enz, J.: Investment in human capital, Krakow, Poland (2001)
5. Hill, C.W.L., Jones, G.R.: Strategic Management: An Integrated Approach, 5th edn. Houghton-Mifflin, Boston (2000)
6. Iwachnienko, A.G.: Induktiwnyj metod samoorganizacji modelej sloznych system po eksperymentalnym danym. Naukowa Dumka, Kijów (1982)
7. Johnson, G., Scholes, K.: Exploring Corporate Strategy: Text and Cases, 6th edn. Pearson Higher Education, London (2002)
8. Kasiewicz, S., Rogowski, W., Kicińska, M.: Intellectual capital, Krakow, Poland (2006)
9. Kasprzak, T.: The reference mode. In: The business processes management, Difin, Poland, Warsaw (2005) (in polish)
10. Klaus, H., Rosemann, M., Gable, G.G.: What is ERP? Information Systems Frontiers, pp. 141–162. Kluwer Academic Publishers. Manufactured in The Netherlands (2000)

11. Krebs, I., Patalas-Maliszewska, J.: An E-Administration Strategy for Innovation Transfer. In: Grasserbauer, M., Sakalauskas, L., Zavadskas, E.K. (eds.) 5th International Vilnius Conference. Knowledge-Based Technologies and Methodologies for Strategic Decisions of Sustainable Development. VGTU Press Technika, Vilnius (2009)
12. Mintzberg, H., Lampel, J., Quinn, J.B., Ghoshal, S.: The Strategy Process: Concepts, Contexts, Cases. Prentice- Hall, Upper Saddle River (2003)
13. Nonaka, L., Takeuchi, H.: The Knowledge-Creating Company. Oxford University Press, New York (1995)
14. Patalas-Maliszewska, J.: The concept of system supporting decision making enabling to asses and forecast of knowledge. In: SMEs – research results, Applied Computer Science, Zilina, Slovak Republic (2010)
15. Quinn, J.B., Anderson, P., Finkelstein, S.: Managing professional intellect: making the most of the best. Harvard Business Review (March-April 1996)
16. Sokołowska, A.: Intellectual capital management in the small enterprise. PTE (2005)

Mining Decision Activity Logs

Razvan Petrusel and Daniel Mican

Faculty of Economical Sciences and Business Administration, Babes-Bolyai University, Teodor Mihali str. 58-60, 400591 Cluj-Napoca, Romania
{razvan.petrusel,daniel.mican}@econ.ubbcluj.ro

Abstract. This paper introduces our work regarding the mining of decision activity logs generated by the users of a decision support system-like environment. We will show that a DSS can be modified in order to become "decision-aware. If the system offers support for all the data and information needs of the decision maker, how the user interacts with the software can provide us with a new perspective over the implicit and explicit knowledge employed in the decision process, as well as the decision patterns and strategies used for that decisional situation. All this valuable information will be stored as activity logs. Those logs need to be mined in order to build a graphical representation of the decision process. As proof-of-concept we focus on the enterprise loan contracting decision situation. We will show some of the models we created using several process mining algorithms and our own approach. Based on those models, we argue the new insights we can provide into the decision making process and the knowledge that is now explained and depicted as diagrams.

Keywords: decision mining, user activity log, process mining, mining algorithm, decision reference model.

1 Introduction

This paper aims to show some meaningful insights into our work regarding decision mining. We define decision mining as the activity that, based on the user interaction and event logs obtained from a software tool (usually a decision support system - DSS), extracts and creates a model of the decision making process. In many ways, what we introduce as "decision mining" is similar to process mining. What is specific to decision mining is that it aims to create a model of decision maker's thinking patterns based on his interaction with decision-aware software. Decision mining is more challenging because, besides developing the actual mining algorithm, it first requires us to develop means for capturing the essence of human thinking expressed during the human-computer interaction.

In the second section, we connect our work to several major research areas and some related papers and argue how our paper builds on existing knowledge, and also what is new and different from existing research. *In the third section* of the paper we provide some insights concerning the proof-of-concept implementation for an experiment regarding the decision process of financing a company by loans. *In the*

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 67–79, 2010.
© Springer-Verlag Berlin Heidelberg 2010

fourth section we shortly review several process mining algorithms (Alpha, Heuristic, Fuzzy, Workflow Patterns and Genetic miners) and their performance on the decision activity logs. Also in the fourth section we introduce a new algorithm that we are developing for decision mining.

2 Theoretical Foundations and Related Work

Our work combines aspects regarding several broad research areas as: decision theory and analysis, decision support systems, process mining, workflow management, and simulation.

As in decision analysis, some of our higher goals include prescribing the recommended course of action in a decisional situation by applying some theorems and algorithms to create and exploit a well-formed representation of the decision process. We aim to create formal representations of decisions and translate them into new insights and recommendations for the decision makers. Our immediate goals include identifying and explaining various aspects of actual decision making processes by using an automated mining approach based on decision maker's interactions with computer software.

Our work draws on decision theory as well as workflows and process management. We are innovating in this area of research because we are arguing a new way of acquiring knowledge (both explicit and implicit) from the expert decision makers, besides the classic methods (interviews, questionnaires, etc). Our approach is based on the assumption that the actions of a person during a process will provide an external observer with a better understanding of a workflow then what the user says in an interview about that workflow. The same assumption is also used in process mining [1], [2].

An approach that aims to simulate an operational decisional situation based on workflows is presented in [3]. This paper shows how, based on observed behavior in an enterprise, captured in event logs, a simulation of the workflows is created. The approach proposed in this paper is based on using pre-existent workflow software as the starting point of the simulation. However, to our best knowledge, in decision making there is no such tool available.

The basic guidelines in creating a process simulation of operational decision making are presented in [4]. The paper provides: a critical analysis of simulation approaches and tools used in process modeling; and a detailed analysis of the effect of resource availability in simulation studies. One of the conclusions is that people prefer to work in chunks [4]. This is of interest to us because it indicates it is possible to divide the decision into several clusters of related activities that are performed with random delays and in batches. Based on this conclusion, we introduce the clustering of activities and the notion of parallel activities in decision mining.

We aim to create a specialized decision mining algorithm. A similar problem of generating a workflow graph out of event logs was originally addressed in [5], then improved in [6] and [7]. Our approach over decision mining algorithm is similar because the logs show sequential mental activities and duplicate tasks are present frequently. Our graph generation technique will start from [8], [5]. Similar to process mining we need to perform three types of mining: discovery, conformance checking and improvement [9].

The term “decision mining” was used before by Rozinat in [10]. The mining algorithms are implemented in a plug-in called Decision Miner that is implemented in ProM Framework (application and documentation is available at: www.processmining.org). This approach uses a derivation of C4.5 algorithm to build decision trees that allow analysis of choices in the decision points of a workflow. Rozinat proposes the use of Petri Nets theory in order to identify the points in which a choice was made and one or other of the branches were followed. However, this is different from our approach. There are differences due to the fact that the properties determined by Rozinat that change the path followed after a decision point can not map to the properties of the mental activities captured in the decision workflow. Also, timing activities in a decisional workflow does not make sense.

3 Logging Decisional Activities

The experimental results are based on a DSS-like software designed to be used by experts and by students in the master degree program at our university. The general approach used in building this system is depicted in Fig. 1.

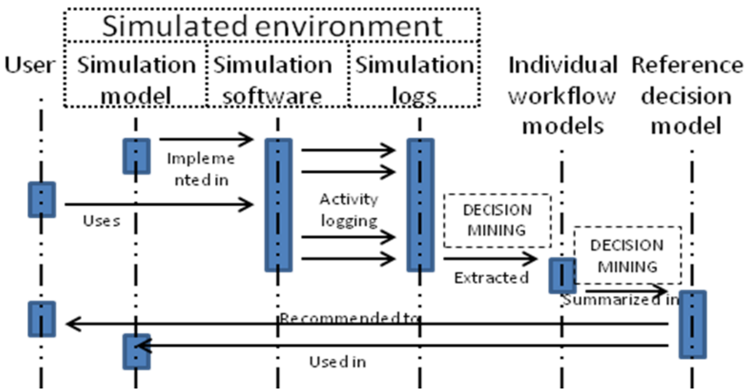


Fig. 1. General approach over decision mining

The same decisional situation is intended to be presented to numerous decision makers that are required to use our decision-aware system in order to reach a decision. The aim of the experiment is to log all activities of the user in order to re-trace the mental activities and to get a better insight into the actual knowledge employed in the decision making process. The ultimate goal of the experiment is to create a “map” of the knowledge necessary when making a decision regarding the loan choice for enterprise financing, depicted as a reference model.

The essential element for the success of our research is our ability to produce software that will allow the decision maker to enact all his mental activities, as much as possible. This is due to the fact that the most important and relevant knowledge for a certain decisional situation is implicit. The major innovation, so that a regular DSS becomes decision-aware, is that no data is shown when the forms are loaded. Only when explicitly requested (a text box containing a piece of data is clicked) the data is displayed and can be used in the decision process. The sequence of the events is recorded by using timestamps for each action. This way, we can see what data was actually used and which is the order of usage. The overall conclusions we reached are:

a) The system must be DSS-like and must provide the decision maker with all the data and information related to the decision at hand, ranging from critical to unimportant.
b) In order to keep trace and log all the activities of the user, it is crucial to split the interactions with the software into basic thinking items. This actually requires a reconsideration of the entire software engineering process.
c) The software must trace and log the actual data used by the decision maker. If the user is deriving data, the source and the formula employed should be logged.
d) The software must be designed so that all the mental activities of the decision maker must be performed inside the software. The user should not need or be allowed to look for data and information someplace else.

In order to perform decision mining we must make sure that it is possible to record user actions within the software such that:

a) each action refers to one mental activity (is a well defined step in the decision making process);
b) each action refers to one and only one well defined instance of a decision process;
c) each decision process is performed by known decision makers;
d) in case of collaborative decision making, each action can be attributed to only one decision maker and each interaction between decision makers can be reduced to basic actions attributed to one user;
e) each action within the software has a time stamp and actions are totally ordered.

The software used for this experiment is divided into four main sections: the decision at-hand research, the enterprise financial position research, the loan market research and the decision making section. The decision research section presents the user with data and information regarding the investment that generates the need for financing. The enterprise financial position research allows the user to review all the figures from the enterprise's financial statements (balance sheet, cash-flow, trial balance, key indicators). The loan market research allows the user to review the financing conditions (interest, commissions, etc) for six different types of loans. The decision making section allows the user to perform what-if analyzes regarding the installments and interest paid. Also in this section, the user needs to input the chosen decision alternative (as decisional variables like credit amount, advance payment amount, loan type and loan duration).

The logs we obtained are classified into three types, according to the level of knowledge exhibited by the users. The logs of the user actions were stored in four tables, as required by the ProM Import Tool (see also Fig. 4).

4 Decision Mining Algorithm (DMA)

The main process mining algorithms are reviewed in the next paragraphs. The framework that allows comparisons of the performance of each algorithm for process models is introduced in [11], [12].

Alpha Miner relies on four basic relations used directly for building the Petri Net: A>B, A→B , A#B, and A||B [1]. There are three steps in the algorithm: a) discovery of causal dependencies in the log (an activity always followed by another activity); b) construction of the input and output expressions; c) search for long distance dependency relations [13]. One of the downsides of this algorithm, when applied to decision mining, is that it does not consider the frequency of traces in the log. Alpha ++ miner is an advancement of alpha algorithm that explicitly captures non-free-choices in the mined model [14].

Heuristic miner aims to mine the control-flow perspective and only considers the order of events within a case (i.e. the order of events among cases is not important) based on the timestamp of the action [13]. There are three mining steps: a) the construction of a dependency/frequency (D/F) table; b) the induction of a D/F graph out of a D/F table; and c) the reconstruction of the workflow-net out of the D/F table and graph [13], [15].

Genetic algorithm performs mining in four steps: a) build the initial population (randomly or using a heuristic approach), b) calculate the fitness of the population, c) generate the next generation using elitism, crossover and mutation, and d) stop when reaching a defined "stop" point [16].

Fuzzy miner was created to be used in conjunction with less-structured real data logs [17]. It intends to show a model using concepts like: aggregation, abstraction, emphasis and customization. There are three phases in the algorithm. The first two phases (conflict resolution and edge filtering) remove edges (i.e., precedence relations) between activity nodes, while the final aggregation and abstraction phase removes and/or clusters less significant nodes [17]. The main benefit of Fuzzy Miner is that it is able to clean up a large amount of confusing behavior, and to infer and extract structure from chaotic behavior [17].

Workflow Patterns Miner focuses on the discovery of local workflow patterns [18] instead of global mining. The algorithm performs well in mining structure, parallelism and non-free choices from delimited areas of the logs.

Based on existing process mining algorithms, we argue a specific DMA needs to resemble more with Fuzzy algorithm (for dealing with less-structured decision making processes and for reducing the high number of activities in one trace).

DMA must be centered on mining the sequence (defined as activities that must be performed one after the other) because it is essential to be able to show to the user which actions can be performed randomly and which must be performed in a certain order. The timestamps of activities, the conclusion that people work in chunks extracted from [4] and some of the formalism in [18] are at the core of the creation of local thinking patterns (activity clusters). A complete workflow model will consist of connected local patterns. Concurrent activity clusters can be performed in parallel or in any order. The conceptual model of DMA is shown in Fig. 2.

The basic activity of decision making can be divided into two major levels, considering the degree of generalization involved. The first one is the generic approach over the decision at hand. This level contains the generic activities (like research the problem at hand, learn new information about the situation and/or some aspect of it, create alternatives, etc.). At this level, the rules and means of clustering the decisional activities is the major concern. The generic activity clusters can also be extracted from experts by using existent methods (interviews, questionnaires) and asking them to describe the overall decision making strategy. The concurrence is the major concern while not forgetting sequence. The second level contains actions that take place one time or several times within each generic activity. This level basically describes each step taken in performing the generic activity. DMA mines for each aspect.

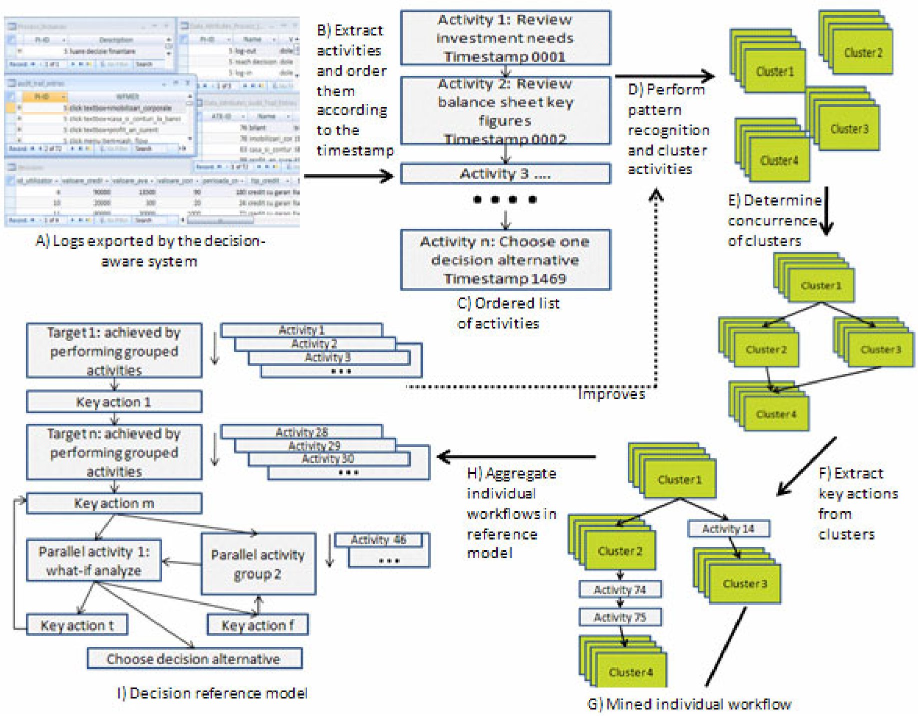


Fig. 2. Overview of the Decision Mining Algorithm

The steps we include in DMA are:

a) logging decision activities expressed as user system interaction,
b) creation of clusters of activities for all traces in the log,
c) creation of individual workflows for each trace in the log,
d) creation of reference models containing clusters and main actions for each cluster,
e) extraction of key actions from the clusters. This is based mainly on the frequency of actions in multiple workflow instances. If the same action is performed by multiple users it will be classified as a key action.
f) creation of the final model.

It is difficult to use machine learning algorithms in decision mining because there are only positive examples in the log and no negative examples. There are no negative examples in decision making because simply there is no “right” and “wrong” when making a decision.

Some means for evaluating the quality of the models discovered by process mining algorithms are described in [12]. The mined results for a decision activity log containing only 13 traces and 649 events are shown in Fig. 3 (results were worse for logs with more traces and events). Even if the actual elements in the model are unreadable, we show the mined models in Fig. 3 in order to provide the reader with an overview for each algorithm's output when mining decision workflows. Therefore, we focus at this stage on the general structure not on the actual content of the model. It is easily observed that even for a small number of traces the models are extensive and the number of connected activities (lines in the model) makes each model difficult to follow. The overall structure of the output models (Fig. 3) is used as one of the criterions in the comparison of algorithm performance on decision logs.

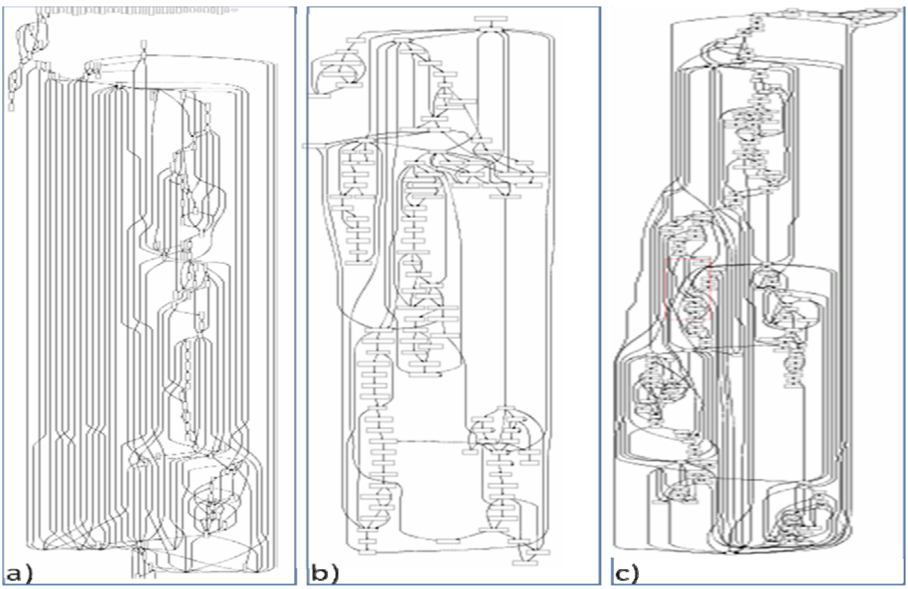


Fig. 3. Overview of models mined by a) Alpha++, b)Heuristic and c)Workflow Patterns Miners

As shown in Fig. 3, the models created by Alpha++, Heuristic and Workflow Patterns miners are quite similar and break the basic guidelines of a good model [19]. The major flaw of Alpha++, Heuristic, Genetic and Workflow Patterns models is that each unique action of the decision maker is represented individually in the model and therefore the model lacks structure and has too many elements depicted. One can notice that Heuristic model (Fig. 3b) is the most intelligible of the three. Fuzzy miner (see Fig. 5b) performs better; extracts a model that uses fewer elements (due to the use of clusters) and is better structured.

Due to the lack of space in this paper, further insight into the content of the mined models is limited to the following conclusions (based on criterions in [12], [19]):

a) Fitness indicates how much of the logged behavior fits the model. Each algorithm performs well at this aspect, except Fuzzy Miner.

b) Completeness quantifies the percentage of traces in a log that can also be generated by the model. All algorithms have a satisfying performance due to the fact that the mined models are large. Alpha Miner performs best while Fuzzy miner provides slightly less complete traces than can be replayed (only 58%).
c) Duplicates denote the number of identical actions present in the model compared to the log. Again, all algorithms have a satisfying performance.
d) Precision addresses the overly general models. For Alpha++, Heuristic, Genetic and Workflow Patterns algorithms, the precision is better than for Fuzzy but still is poor because each model connects activities that should not be connected.
e) Generalization means extracting the essential model from the logs. Fuzzy miner is the only one performing good at this aspect due to clustering. All other models enumerate most actions.
f) Structure denotes the easiness of reading the model. The "spaghetti" models are less structured. Again, Fuzzy Miner is the only one performing good at this aspect because all other algorithms provide a model that is impossible to understand entirely.

Fuzzy Miner has the best overall performance. However, none of the models is providing us with too much knowledge and new insights over the decisional process due to the size of the model and the "spaghetti" type relationships. We argue that the main problem of process mining algorithms, when applied to decision activity logs, is that they do not:

a) show a condensed, easy to read and understand view over the decision process
b) do not emphasize on the key actions.

In the remainder of this section we show how the experimental expert activity log (containing 5 complete and 8 incomplete traces) was mined using DMA approach.

To illustrate *the first step* of DMA we show a partial sample from each table that stores the log (Fig. 4). We used multiple tables in order to create a log that can be imported into ProM Framework. The records represent actions of the user while interacting with the system (e.g. clicking menus, clicking textboxes to see the figures, etc). A complete trace must contain three actions: log-in, log-out and reach decision (e.g. Data_Attributes_Process_Instances PI-ID records 1, 2 and 3). Audit_Trail_Entries stores the actual action in WFMEIt field encoded as action object+object name (e.g. in record 4: click textbox+investment_value). In Data_Attributes_ATE table we can find the actual value stored for an object (e.g. record 4 shows that when textbox.investment_value was clicked it showed the value of 100000).

To illustrate *the second step*, Fig. 5a shows the clusters created for this log. So far, we used for clustering the knowledge acquired from the experts stating the major concerns when making a loan decision. If we compare the actions in the DMA model (Fig 5a) to the model mined by Fuzzy Miner (Fig 5b) we can see that the clusters are mapped to the menu items (due to the implementation of the experimental software).

In *the third step*, we create a workflow model for each complete trace in the log. Basically, the workflow model is a linear model with clusters and ordered actions inside each cluster, similar to the one in Fig. 5a. It can become tree like when repetitive actions are performed (e.g. what-if analyzes that require multiple inputs for the same figure).

Process_Instances

PI-ID	Description
1	loan financing decision
2	loan financing decision
3	loan financing decision
4	loan financing decision
5	loan financing decision
6	loan financing decision
7	loan financing decision
8	loan financing decision
9	loan financing decision
10	loan financing decision
11	loan financing decision
12	loan financing decision
13	loan financing decision

Data_Attributes_Process_Instances

PI-ID	Name	Value
1	log-in	giurgiualexand
1	log-out	giurgiualexand
1	reach decision	giurgiualexand
2	log-in	gavrisionut
2	log-out	gavrisionut
2	reach decision	gavrisionut
7	log-in	costinrazvan
7	log-out	costinrazvan
8	log-out	costinrazvan

Data_Attributes_Audit_Trail_Entries

ATE-ID	Name	Value
1	cash_flow	cash_flow
2	evaluare_investitie	evaluare_inve:
3	tip_investitie	autoutilitara
4	valoarea_investitiei	100000
5	bilant	bilant
6	venituri_prognozate	310000
7	cheltuieli_de_punere	4500
8	Conturi Profit si Pierc	Conturi Profit s
9	cheltuieli_de_exploa	40000

audit_trail_entries

ATE-ID	PI-ID	WFMElt	Timestamp	Originator
1	1	click menu item+cash_flow	12/15/2009 4:42:23 PM	giurgiualexand
2	3	click menu item+evaluare_investitie	12/15/2009 4:43:17 PM	soldereaioan
3	3	click textbox+tip_investitie	12/15/2009 4:43:24 PM	soldereaioan
4	3	click textbox+valoarea_investitiei	12/15/2009 4:43:27 PM	soldereaioan

decisions

id_u	loan_value	advanced_payment	initial_taxes	loan_period	loan_type	installment_type
6	90000	13500	90	180	mortgage	Equal
10	20000	300	20	24	mortgage	Equal
11	80000	20000	1000	72	mortgage	Equal
19	80000	16000	0	60	warranty free	Equal

Fig. 4. Log sample

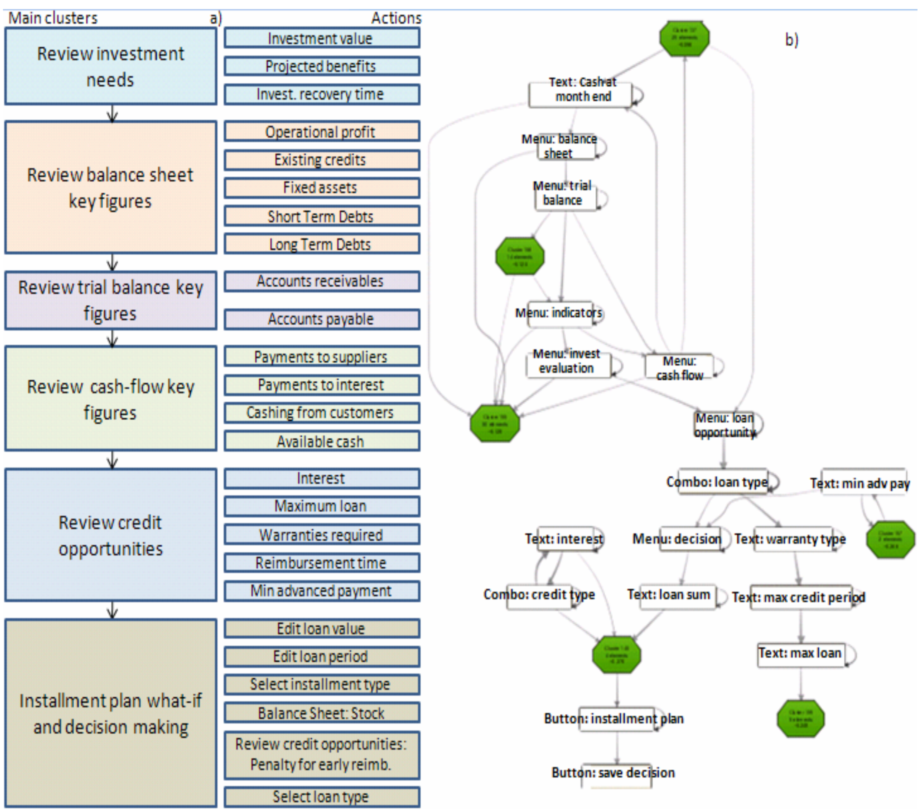


Fig. 5. a) Model mined using DMA, b) Model mined by Fuzzy Miner

In *the fourth step*, the reference model is created starting from the individual models (Fig. 6). The entire formalism for the last three steps will be presented in a separate paper. The basic idea is that each cluster of each individual trace is compared to similar clusters in other traces, performing thus local pattern recognition. In Fig. 5a the actions in each cluster show up in at least 80% of the individual workflows (4 out of 5 complete traces for the experimental log). If we compare the actions in the DMA model (Fig 5a) to the model mined by Fuzzy Miner (Fig 5b) we can see that 9 out of 10 actions (clicks on items other than menus) in Fuzzy model also show up in DMA list of actions. During the interviews for this particular experiment, we identified that the number of actions that should be performed (as identified by at least 80% of the experts) are 27. DMA depicts 93% of those (as shown in Fig. 5a), while Fuzzy model depicts only 36% of expert actions for a 0.38 significance cutoff set when creating the model in Fig. 5b. For lower cutoffs, the number of items in Fuzzy model increases but show up both actions in and outside DMA (e.g. for 0.2 significance cutoff, Fuzzy model shows 59% of the actions in DMA but the total number of items in the model is 42 and it looses generalization and structure).

In *the fifth step,* we define a key action as an action that shows up in clusters in every trace. If a key action is identified, it is "pulled out" of the cluster (Fig. 6b). All other actions are ordered inside the cluster based on a calculated average sequential position in the clusters in each trace. Fig. 6a shows that the credit value is the first action in the cluster because it was the first one in 4 of the 5 complete traces, then it was followed in 4 traces by selection of the credit type, etc). The actions that do not pass a threshold of show ups are discarded (e.g. an action that does not show up in more than 10% of the traces is deleted). We are aware that this approach sacrifices completeness in order to improve generalization and structure.

In *the last step* we unite the clusters and establish the relationships among each other and with key actions.

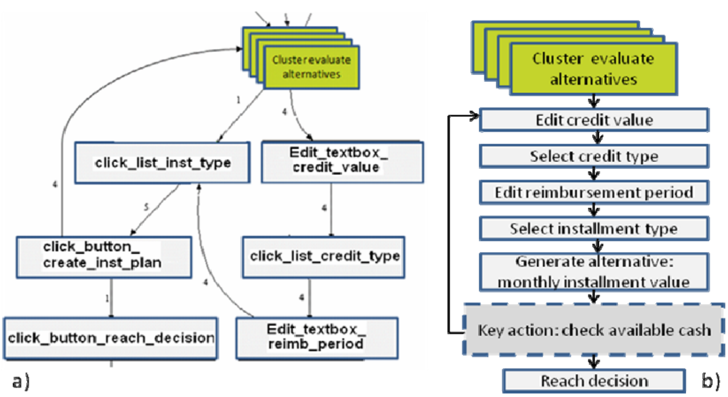


Fig. 6. a) Partial individual decision workflow, b) Partial reference model

We also applied decision mining successfully in evaluating the academic performance of students involved in Virtual Businesses course. The students used the software at the beginning and at the end of the semester. The mined models were compared with each other and then with some "expert" logs. Each student's improvement in knowledge was quantified and graded by calculating a percentage of the actions in the expert workflow that also show in the student model divided to the number of actions that show only in the student workflow.

5 Conclusions and Future Work

We argued in this paper that it is possible to gain an improved insight into the decision making process by mining the logs containing the user interaction with a decision-aware system. We do not intend to replace the classic, well-known knowledge acquisition methods (interviews, questionnaires), but we are arguing that decision mining can be an alternate, sometimes closer to reality, knowledge acquisition tool, that also has the advantage of being far less time and resource consuming. The ultimate goal of decision making is to prescribe the (mental) actions that should be followed and the particular sequence that needs to be followed (e.g. the result of the experiment in section 4 is that, when making a loan decision, at least the actions in Fig. 5a should be followed in the particular order shown in the model).

To make decision mining possible we first need to develop an environment that can force the decision maker to enact knowledge as interactions with the software. In the third section we introduced some features of such a system and some fundaments we argue as mandatory when converting each action into an item in the activity logs.

The second challenge is to create a mining algorithm that extracts new pieces of knowledge from those logs. The decision mining (DMA) approach argued by us relies, at this point, on a pre-existent model extracted from the experts which is improved by user action clustering, sequence and concurrence. It is possible to compare our approach with existing process mining algorithms. In the fourth section, we compared the most important process mining algorithms performance on sample decision activity logs and concluded Fuzzy Miner performs best. Therefore, for a sample of experimental data created for a loan contracting decision, the model created by using interviews, DMA and Fuzzy Miner were compared. We concluded that DMA performs significantly better for this particular situation. However, we do not expect the same performance for other decisions at this stage of DMA development. We replicated and extended the experiment so far only for academic purposes. We will show improvements on specific DMA features in future papers.

We argue that the decision reference workflow model type DMA proposes (a model should contain two levels: clusters and key actions on the upper level and individual actions on the lower level) is better than the process mining models because it provides a better overview and enables the user to easily understand the path that needs to be followed.

We will address in our next papers issues like: presenting the formal foundation for the last steps of DMA, creating an improved framework for decision mining, implementing the DMA algorithm in ProM, creating several new environments for other decisions, extending decision mining to collaborative activities.

Acknowledgement. This work is supported by the Romanian Authority for Scientific Research (CNCSIS) under contract IDEI_573.

References

1. van der Aalst, W.M.P., van Dongen, B.F., Herbst, J., Maruster, L., Schimm, G., Weijters, A.J.M.M.: Workflow Mining: A Survey of Issues and Approaches. Data and Knowledge Engineering 47(2), 237–267 (2003)
2. van der Aalst, W.M.P., Weijters, A.J.M.M.: Process Mining. Computers in Industry 53(3), 231–244 (2004)
3. Rozinat, A., Wynn, M., van der Aalst, W.M.P., ter Hofstede, A.H.M., Fidge, C.: Workflow Simulation for Operational Decision Support. Data and Knowledge Engineering 68(9), 834–850 (2009)
4. van der Aalst, W.M.P., Nakatumba, J., Rozinat, A., Russell, N.: Business Process Simulation: How to get it right? In: vom Brocke, J., Rosemann, M. (eds.) International Handbook on Business Process Management. Springer, Berlin (2009)
5. Agrawal, R., Gunopulos, D., Leymann, F.: Mining Process Models from Work-flow Logs. In: Schek, H.-J., Saltor, F., Ramos, I., Alonso, G. (eds.) EDBT 1998. LNCS, vol. 1377, pp. 469–483. Springer, Heidelberg (1998)
6. Schimm, G.: Process Miner: A Tool for Mining Process Schemes from Event-based Data. In: Flesca, S., Greco, S., Leone, N., Ianni, G. (eds.) JELIA 2002. LNCS (LNAI), vol. 2424, pp. 525–528. Springer, Heidelberg (2002)
7. Herbst, J., Karagiannis, D.: Integrating Machine Learning and Workflow Management to Support Acquisition and Adaptation of Workflow Models. International Journal of Intelligent Systems in Accounting, Finance and Management 9, 67–92 (2000)
8. Wen, L., Wang, J., van der Aalst, W.M.P., Huang, B., Sun, J.: A Novel Approach for Process Mining Based on Event Types. Journal of Intelligent Information Systems 32(2), 163–190 (2009)
9. van der Aalst, W.M.P.: Using Process Mining to Generate Accurate and Interactive Business Process Maps. In: Abramowicz, A., Flejter, D. (eds.) BIS 2009 Workshops. LNBIP, vol. 37, pp. 1–14. Springer, Berlin (2009)
10. Rozinat, A., van der Aalst, W.M.P.: Decision Mining in ProM. In: Dustdar, S., Fiadeiro, J.L., Sheth, A.P. (eds.) BPM 2006. LNCS, vol. 4102, pp. 420–425. Springer, Heidelberg (2006)
11. Rozinat, A., de Medeiros, A.K.A., Gunther, C.W., Weijters, A.J.M.M., van der Aalst, W.M.P.: The Need for a Process Mining Evaluation Framework in Research and Practice. In: ter Hofstede, A.H.M., Benatallah, B., Paik, H.-Y. (eds.) BPM Workshops 2007. LNCS, vol. 4928, pp. 84–89. Springer, Heidelberg (2008)
12. Rozinat, A., de Medeiros, A.K.A., Gunther, C.W., Weijters, A.J.M.M., van der Aalst, W.M.P.: Towards an Evaluation Framework for Process Mining Algorithms. BPM Center Report BPM-07-06, BPMcenter.org (2007)
13. van der Aalst, W.M.P., Weijters, A.J.M.M., Maruster, L.: Workflow Mining: Discovering Process Models from Event Logs. IEEE Transactions on Knowledge and Data Engineering 16(9), 1128–1142 (2004)
14. Wen, L., Wang, J., Sun, J.G.: Detecting Implicit Dependencies Between Tasks from Event Logs. In: Zhou, X., Li, J., Shen, H.T., Kitsuregawa, M., Zhang, Y. (eds.) APWeb 2006. LNCS, vol. 3841, pp. 591–603. Springer, Heidelberg (2006)

15. Weijters, A.J.M.M., van der Aalst, W.M.P.: Rediscovering Workflow Models from Event-Based Data using Little Thumb. Integrated Computer-Aided Engineering 10(2), 151–162 (2003)
16. de Medeiros, A.K.A., Weijters, A.J.M.M., van der Aalst, W.M.P.: Genetic Process Mining: An Experimental Evaluation. Data Mining and Knowledge Discovery 14(2), 245–304 (2007)
17. Gunther, C.W., van der Aalst, W.M.P.: Fuzzy Mining: Adaptive Process Simplification Based on Multi-perspective Metrics. In: Alonso, G., Dadam, P., Rosemann, M. (eds.) BPM 2007. LNCS, vol. 4714, pp. 328–343. Springer, Heidelberg (2007)
18. Gaaloul, W., Baïna, K., Godart, C.: Towards mining structural workflow patterns. In: Andersen, K.V., Debenham, J., Wagner, R. (eds.) DEXA 2005. LNCS, vol. 3588, pp. 24–33. Springer, Heidelberg (2005)
19. Mendling, J.: Metrics for Business Process Models: Empirical Foundations of Verification, Error Prediction, and Guidelines for Correctness. In: Mendling, J. (ed.) Metrics for Process Models. LNBIP, vol. 6, pp. 103–133. Springer, Heidelberg (2008)

Applications of Voice Based Informative Public Services: Lithuanian Case

Vytautas Rudžionis[1], Rytis Maskeliūnas[2],
Ramūnas Šablinskas[3], and Algimantas Rudžionis[2]

[1] Department of Informatics, Vilnius university Kaunas faculty, Lithuania
[2] Kaunas university of technology, Lithuania
[3] UAB Omnitel, Lithuania

Abstract. The paper presents analysis of the possibilities to implement some voice based informative services in public sector of Lithuania. The paper reveals the applications in the public sector where the implementation of voice based modality may contribute significantly to the convenience and the popularity of the service. Possible solutions for the cases when Lithuanian speech will be used as the main speech in the service are evaluated both from technological and economical point of view. Several solutions how the obstacles related with the insufficient recognition accuracy of the Lithuanian voice commands may be lessened are revealed also. It is shown that implementation of properly trained foreign language speech recognizer may achieve recognition accuracy of Lithuanian voice commands which is necessary for applications in public sector.

Keywords: voice based informative services, speech recognition, speech synthesis, speech engine, phonetic transcriptions.

1 Introduction

Speech based interfaces used as a basic modality for the control of telephony based services has enormous potential and advantages. In recent years we saw significant number of new voice based services and system prototypes presented and developed in various countries and in various areas of application. One of the main forces that caused increased popularity of voice based automated telephony services was the advent of speech engines – economically sensible and at the same time assuring high quality solution for voice interface based applications. The speech engines created the opportunities to implement voice based interfaces into information systems even for the small and medium enterprises since they significantly lowered the costs of the development. But speech engines have one very important aspect: they are not language independent and depend on the language spoken by the users of the system. Microsoft and other major speech technologies providers develop and distribute speech engines for widely spoken languages having many native speakers (English, Spanish, French, etc.). Many other less widely used languages remain out of the scope of interest for the major speech recognition solution providers.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 80–89, 2010.
© Springer-Verlag Berlin Heidelberg 2010

In such situation businesses and state institutions in countries were such less popular languages are used as a main source of spoken language communication, face a challenge of development of their own speech recognition tools. Two major ways for solution are as follows [1]:

- to develop completely new speech recognition engine oriented to the recognition of that language;
- to adapt the foreign language based engine for the recognition of your native language.

The first approach has potentially higher facilities to exploit peculiarities of selected language and hence to achieve higher recognition accuracy. But the drawback of such approach is the same why the providers of major speech technologies avoid implement these languages in their products – high costs of the development in the general sense of this word.

The second approach [5] has the potential to achieve some practically acceptable results faster than when developing entirely new speech recognition engine. Another advantage of this approach is the potential to achieve compatibility with the existing technological platforms faster and cheaper. This advantage is often important for business customers, since they need to follow various technical specifications in order to guarantee consistent functioning of the enterprise. But this approach also requires careful investigation of the ways of adapting speech engine and optimizing adaptation algorithms. Our previous activities [4] showed that proper adaptation of the foreign language based speech engine may achieve the recognition accuracy which is close to the requirements set for the commercial applications and the other applications oriented to the general audience.

This paper deals with the attempts to identify the areas of the applications in the Lithuanian public sector which has the biggest and most urgent need to implement voice based user interfaces in their information systems. Then we will examine the main requirements to the design of the system in general and to the voice interface in particular of the some of the perspective applications. The accuracy drawbacks and the possibilities to overcome them are presented also. Experimental evaluation showed that the use of restricted set of digits for the identification of person in voice based information systems may guarantee accuracy level which is necessary to achieve success and to avoid complaint between the customers of the service.

2 Some of the Potential Applications of Voice Based Interfaces for the Public Sector in Lithuania

The requirements that are more and more often set on the public organizations such as governmental, municipal, healthcare or educational institutions emphasize the necessity to provide the necessary for the user information in a convenient and ubiquitous way. It should be noted that such requirements are often set as judicial requirements and public institutions are forced simply to follow them. It is important to note that such a trend will continue in the upcoming years since it perfectly fits to the general trends of "e-democracy" and openness of the public institutions. Such trends are particularly encouraging to look at the public sector as an area were first and the most difficult stage of the practical implementation of voice based services should occur.

Let us consider some of the potential voice commands controlled applications that could have significant market potential. In principle such applications could include the abilities to provide or to correct income declarations by voice, to register for the upcoming elections, to get the results of the entrance test to the university and many others. One of the newest examples of such kind in Lithuania is the move to the electronic form of sick-lists that are issued by the healthcare institutions to the state social security foundation. Instead of the traditional paper sick-list physicians will be forced to fulfill electronic declaration form and to send this electronic form to the state social security foundation. The benefits of the electronic forms are clear: those forms will reduce the time necessary to check and to evaluate them at the state institutions and will reduce the possibilities of unintentional errors, etc. In this way it is expected that overall administrative costs of the social security foundation will be reduced also. Such electronic service could be realized using traditional graphic user interface and the desktop computer with Internet connection (e.g. eligible physicians signs in into the website of the social security foundation and fulfills the appropriate fields of the predefined form of the sick-list). And such Internet based realizations are the most common way to move the service to the electronic media. In practice this approach has several drawbacks. These days there are about 12 thousand doctors that have the rights to issue and are issuing sick-lists to the patients in Lithuania. Of course not all of them are familiar with the computers and some estimates show that only 2 thousand of them have the computer at their working place. In this situation some healthcare institutions are establishing special departments where trained personal will fulfill electronic sick-lists using the data provided by the doctors in that institution. It is not difficult to understand that lowering the administrative costs of the state institutions such approach will increase the costs of the healthcare institution. Voice commands using interface could provide alternative which could be more efficient: doctor calls to the special server (e.g. using mobile phone – the device used by the nearly each doctor these days) and tells the necessary data by voice – identifies patient, illness and says the day of the start and the duration of the sick-leave. The careful investigation of the vocabulary used in this application show that most of the commands could be constructed from the digit names or digit strings, e.g. the person could be identified from the personal code. This identification means that having personal code we could identify the name of the patient, declared place of residence, social security card number and other necessary data uniquely. It is also clear that the command controlling the duration of the sick-leave could be composed of digit names as well as the start of that period (today – zero, tomorrow – one, etc.). The very important component of the sick-list – the code of disorder – itself is composed from one letter and three digits. So for such application the proper recognition of the digit names is of crucial if not decisive importance.

Another potential application of voice controlled interface in the public sector could be the electoral system and particularly the problems arising with the advent of the e-voting. Many countries are trying to introduce electronic voting as one of the options of their electoral system and Lithuania isn't exception (it could be noted that Lithuania belongs to the group of pioneering countries due to the specific pressures to the country's political system such as the large number of the voters abroad due to the job emigration and the political passivity of the younger generations). But no one country is using and it is hard to expect that it will use in the future electronic voting as a single option. The traditional voting system will live in parallel with the electronic system. And in this situation new problem to secure that each person could vote only once (or electronically, or traditionally) arises. Voters who want to vote electronically must register in advance but no one could stop such voters from changing their choice and going to the polls on the election's day. In such situation members of the local electoral commission need to nullify electronic vote in the central electoral commission's database. This could be done using GUI but the commissions at the local vicinities do not have the computers and the Internet connection. In some vicinage it could be even difficult to organize Internet connections from the technical point of view (local electoral commissions aren't permanent bodies and work only for a short period before the elections). So another solution may look attractive and be viable in practice: the member of the local electoral commission calls to the server of the central commission, identifies by voice the person which vote should be nullified, gets the confirmation that the person has been identified properly and finally confirms the cancellation of the electronic vote. The main factor causing the success of such system is the proper recognition of the person's identity by the automatic recognition system.

Another example of the increased accessibility to the information of public institutions may be the results of the entrance to the university. It was observed that candidates to the future students want to know the results of the entrance competition as soon as possible. All higher educational establishments (universities and colleges) provide such information on their websites these days. But for many candidates access to the traditional Internet could be unavailable immediately since they may be on a trip or in other places (enrolment to the universities is done during summer vacation period and most of the candidates are young and mobile persons). The ability to get this information using mobile phone and speech recognition may provide increased accessibility. In this approach the user needs to call to the universities datacenter and to identify himself or herself giving the name or some other identification code and system automatically generates and provides to the candidate results of the enrolment.

It should be noted that voice based user interfaces may increase informational accessibility and provide new level of the user convenience in practically each service provided by the public institutions.

From a system designer's point of view the main question arises: how complicated speech recognition system should be in such services. Despite the fact that in principle the more natural system will be the higher user convenience level will be achieved in practice we may restrict the vocabulary used in such services with a rather limited set of words. This restriction may lead to the substantial increase in the feasible recognition accuracy and hence lead to the appropriate customer satisfaction level. It has been shown in many different studies that users are more willing to limit themselves in vocabulary but to get higher robustness of the system performance than to have more freedom in conversation but at the same time to have less stable recognition.

Looking at the general requirements to the content of such services we may observe that the essential component of them is the necessity to identify the person (the user, the provider, the patient, etc.). The identification may be done providing person's name or some pre-assigned code. The identification using names isn't convenient approach due to the unlimited set of the names that may be used for the identification. It should be noted that names themselves don't have semantic information and this fact increases the level of complexity: surname should be recognized exactly on a phoneme-to-phoneme basis independently on how similar those phonemes are (e.g., system designers can't give any preference to the name Gražys versus the name Grašys, the name Kavaliūnas versus the name Karaliūnas, etc.).

In this situation recognition using codes formed from the digit strings could be better approach. First of all such approach will enable the system designer to form vocabulary of limited size. In many cases it is possible to organize pronunciation of the personal identification number using spelling of the digit numbers. In this situation we obtain limited set of ten digit names that could be used to carry out core voice information for the public service. Using this approach we also could implement error – correction codes in some situations to detect false recognitions and to use necessary measures to avoid misrecognition (e.g., to ask the user to repeat once more one digit instead of asking to repeat whole string of digit names).

Despite the fact how we will organize public service the essential question remains the same – the quality of speech recognition. For several years we are working with the possibilities to adapt foreign language based recognizers for the Lithuanian spoken commands recognition. Our previous attempts were oriented to the recognition of proper names and some voice commands. Those attempts showed that it is feasible to achieve recognition accuracy close to the requirements set for the commercial systems using foreign language recognition engine and properly selected transcriptions of Lithuanian voice commands. Our next step was to try to achieve high accuracy level for the recognition of Lithuanian digit strings using foreign language speech recognition engine as economically feasible solution for many customers of voice based information services.

3 Speech Corpora

There were two speech corpora collected for this study. The first one is called initial corpora (denoted ICT). Here 35 non-professional speakers (17 male and 18 female speakers) pronounced each of the ten digit names 20 times during single recording session. Different speaker's data were used in a training session and in a testing session. Training session in our case was the selection of the better suited transcriptions for the recognition. Testing session in our case was the procedure of testing the accuracy of selected transcriptions.

The characteristic property of the initial corpora was relatively large number of pronunciation inaccuracies. Since all speakers were inexperienced a lots of pronunciation errors were observed – hesitations, improper stressing, mistakenly pronounced phonemes, loss of the first and the last phonemes and similar acoustic events. So the analysis of the speech corpora was done in order to remove the most persistent errors. Analysis showed that about 12.5% of all utterances contained pronunciation errors in the initial corpora.

Modified and corrected speech corpora (denoted MCT) has been obtained in such a way: one male speaker replaced all the corrupted utterances of male speakers while single female speaker replaced all the corrupted utterances of female speakers.

4 Digit Number Recognition Accuracy

As it was shown above digit number recognition may be used in the variety of public services as a tool for convenient identification of a person (both user and person of interest). The efficiency and the quality of the service and consequently customer's satisfaction will be determined by the recognition accuracy.

Since Microsoft speech server recognition engine is commercial recognition system we can't modify nor the acoustic models, nor to retrain the acoustic models. The single parameter that we could modify (or being more precise to select more appropriate one) is the phonetic transcription of the voice command. In these experiments we used combined selection methodic of multiple transcriptions proposed in [3]. This method showed is superior to the method used in [2]. Those results formed the basis for the belief that also in the case of the recognition of digit names we could obtain the results that will be close to the recognition accuracy necessary for the commercial and public sector applications.

We obtained 70 different transcriptions for ten Lithuanian digits using combined transcription selection method. The number of the transcriptions varied from two transcriptions used for such short digit names as "two", "three" and "no" (in Lithuanian du, trys, ne) up to 12 different transcriptions used for the word "eight" (in Lithuanian "aštuoni") which also contain diphthong. It should be noted that word "yes" (in Lithuanian "taip") also has relatively big number of phonetic transcriptions. This word had 7 different transcriptions since it also contains diphthong.

The results of the recognition of ten Lithuanian digit names are presented in the Table 1.

It could be seen that average recognition accuracy for ten digit names and two control words was about 81%. In principle such recognition accuracy isn't acceptable for the commercial applications or for the applications in the public sector. Another observation was that recognition accuracy of different digit names varied significantly. While the recognition accuracy for such digit names as „šeši“ (six), „aštuoni (eight) , „septyni“ (seven), „vienas“ (one) was high the average recognition accuracy has been degraded by the recognition of such digit names as „keturi“ (four), „du“ (two), „devyni“ (nine).

Those observations led us to the conclusion that higher recognition accuracy could be achieved if we would be able to limit vocabulary to the seven digit names. Implementation of error correcting codes showed that it could be possible to construct reliable identification of a person using only seven digits and in this case we would be able to avoid the use of those digit names that showed to be less reliable in the pilot study. In this case we could expect also significantly higher overall recognition accuracy.

The next group of experiments was performed trying to investigate the possibilities to improve recognition accuracy of digit names that were difficult to recognize with the transcriptions used in experiments above. One of the difficult to recognize digit

Table 1. The recognition accuracy of 10 Lithuanian digit names using the adapted transcriptions for the Microsoft Speech Recognition engine

	Recognition accuracy, %		
Comand	All transcr.	Two transcr.	Single transcr.
Nulis	41.5	40.0	67.0
Vienas	97.5	98.0	97.5
Du	73.5	83.0	83.0
Trys	62.0	77.5	67.5
Keturi	31.5	39.5	45.5
Penki	78.0	72.0	77.0
Šeši	99.5	100.0	100.0
Septyni	96.5	96.5	96.5
Aštuoni	98.5	97.0	97.0
Devyni	64.0	66.0	66.5
Taip	85.5	87.5	87.0
Ne	85.5	90.0	90.0
Average	76.1	79.0	81.1

name was the word „Keturi“ (four). In these experiments we used method proposed in [3] to construct phonetic transcription of the word from transcriptions of the syllables.

Initially, the syllable transcriptions were constructed out of 7 consonants (P, T, K, B, D, G, R) and all possible 16 vowel and diphthong combinations. The recognition accuracy of the corpus, containing the syllables, constructed using the consonants in the open, semi open and closed vowel context (word "Keturi" syllables „Ke, Tu, Ri"), was analyzed. The results of this experiment are further noted as "C_7V_{16}". The recognition result analysis proved, that the syllable "Ke" was recognized as some transcription variation 98,1 %, syllable "Tu" – 94,3 %, syllable „Ri“ – 66,3 %. The most different answers out of 112 possible transcriptions were selected for the syllable "Ri", although its recognition accuracy was the worst of all three.

A similar experiment was made using all 24 consonants. The 384 syllable transcription variations were created in this case.

The detailed result analysis proved that the recognizer is capable of correctly recognizing the syllables, based on the open / semi open vowel: "Ke" and "Tu". Syllable "Ke" was recognized ~ 98 % in both types of the experiments. The number of omissions was larger for the syllable "Tu". The syllable "Ri" was the most poorly recognized syllable: ~ 66,3 % and 59,1 %. The recognition accuracy was a bit higher in the case of "C_7V_{16}" possibly due to a more limited set of the transcriptions used.

Most of the frequently recognized syllable transcriptions were quite phonetically different from their Lithuanian counterparts. The phonetically similar alternatives were selected very rarely.

The value of the combined transcription selection method may be illustrated in the next experiment using the digit name "Keturi" (four). At the first stage the transcriptions were created, using the syllable recognition results and the best of those were selected using the iterative selection method. Trying to further improve the recognition accuracy, the most dissimilar syllable transcriptions were replaced with a more phonetically close, though less frequently recognized transcriptions. For example, the second syllable's transcription from the word "Keturi" - D OW (pronounced *dau*) was replaced with a more phonetically similar alternative - T UH (pronounced *tu*). The best transcriptions from that modified set were iteratively selected. The overall recognition accuracy achieved (91 %) was almost 3 times higher than for the LBT based transcriptions. The results are presented in Table 2.

Table 2. The best iteratively selected word "Keturi" transcriptions

Transcription		Recognition accuracy, %
G EH T UH D IH		40,5
G EH T UW D IY		20,5
G EH T UW D IH		19,0
G EH T OW G EY		3,5
K EH T UH D IY		7,5
	All:	**91,0**

Similar transcription development method has been applied to get transcriptions of all ten digit names and the words "Taip" (yes) and "Ne" (no). The results are presented in the Figure 1 together with the recognition accuracy obtained using the earlier method:

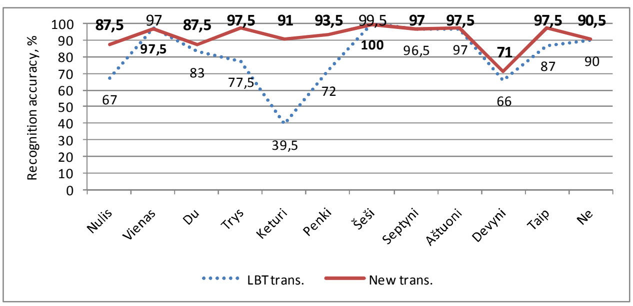


Fig. 1. Recognition accuracy of ten Lithuanian digit names and the words "taip" and "ne" using transcriptions obtained with combinative method (solid line) and most similar transcriptions method (dashed line)

It could be seen that combinative method allowed significant increase in overall recognition accuracy. It could be seen that recognition accuracy of 9 words from 12 exceeded 90% and the recognition accuracy of 6 words exceeded even 95% (this accuracy level is important since many research on customer satisfaction emphasizes 95% recognition level as the level from which customer satisfaction is enough to use the service and to be ready to pay money). We may assume that some further optimization is possible but since the number of the digit names recognized worse isn't big we performed another group of experiments using restricted set of digit names.

The words Taip" and „Ne" were removed from the recognition vocabulary, as these voice commands are usually modeled using separate semantic rules and recognition grammars. Table 3 shows recognition accuracy of seven digit names: there were used 3 different digit names sets. Data sets were obtained removing three digit names from the total set of 10 digit names. The number in the table means the Lithuanian pronunciation of the denoted digit name. In those experiments we used some additional optimization of the transcription selection procedure in order to evaluate how much the recognition accuracy could be improved using optimization procedures.

Table 3. The recognition accuracy of seven digit names in three sets

	Recognition accuracy of voice commands, %							
1 set	1	2	3	5	6	7	8	All
	99.5	99.5	97.5	97.0	99.0	97.5	100	98.5
2 set	1	2	3	4	6	7	8	All
	99.5	98.0	98.0	93.5	100	98	100	98.1
3 set	1	2	4	5	6	7	8	All
	99	97.5	95.5	97.5	98.0	99.0	100	98.0

The results in the table show that using all three sets of seven digit names average recognition accuracy exceeded the targeted recognition accuracy of 95%. The best results were achieved using first data set which contained the names of digits 1,2,3,5,6,7,8. In this case average recognition accuracy was 98.5%. Also it is important to note that in this set recognition accuracy of each individual digit name exceeded 95%. The same is true for the third set of digit names but in this case average recognition accuracy was slightly lower. Interesting situation has been observed with the second digit names set. Here the recognition accuracy of six digit names was the highest if we compare with two other sets but the average has been degraded with the relatively poor recognition of the word "penki" (five). It should be checked additionally if we could optimize the recognition accuracy of this digit name in this particular set of digit names.

The general observation is that recognition accuracy for the seven Lithuanian digit names using English speech recognition engine exceeds target accuracy of 95%. We could expect that further increase in the recognition accuracy could be achieved implementing additional optimization procedures. Also we could expect that implementation of such foreign language recognition engines as Spanish (more similar acoustics to the Lithuanian) could lead to further increase in the recognition accuracy of Lithuanian voice commands.

5 Conclusions

Many services in public sector require the identification of the user. This could be done giving users name/surname or some nickname as well as giving some code (personal number) in the form of the digit string. The second solution may be more appropriate since it enables the development of services using limited vocabulary for the automatic speech recognition.

Experiments with the recognition of seven digit names out of the ten possible digit names showed that in this case it is possible to achieve overall recognition accuracy of more than 95% which is the target level for many applications in public sector. We could suggest that further improvements in the recognition accuracy could be achieved using better optimization procedures for the selection of phonetic transcriptions and implementation of several foreign language speech engines as an approach for the recognition of different Lithuanian voice commands.

These experiments show that implementation of Lithuanian voice command recognition in the public sector of the country is feasible. We see the public sector as one of the most perspective areas for the development of information services with the voice controlled user interfaces in the country. Those services will extend access to the public sector information and will improve the relations and interaction between public enterprises and the society in general.

References

1. Lindberg, B., et al.: Noise Robust Multilingual Reference recognizer Based on Speech Dat. In: Proc. of ICSLP 2000, Beijing, vol. 3, pp. 370–373 (2000)
2. Kasparaitis, P.: Lithuanian Speech Recognition Using the English Recognizer. Informatica 19(4), 505–516 (2008)
3. Maskeliūnas, R.: Recognition of Lithuanian voice commands uisng on Multiple Phonetic Transcriptions. Ph. Thesis, Kaunas University of Technology (2009)
4. Rudzionis, V., Maskeliunas, R., Rudzionis, A., Ratkevicius, K.: On the Adaptation of Foreign Language Speech Recognition Engines for Lithuanian Speech Recognition. In: Abramowicz, W., Flejter, D. (eds.) BIS 2009 Workshop. LNBIP 37, pp. 113–118. Springer, Heidelberg (2009)
5. Zgank, A., et al.: The COST278 MASPER initiative – crosslingual speech recognition with large telephone databases. In: Proc. of 4th International Conference on Language Resources and Evaluation LREC'04, Lisbon, pp. 2107–2110 (2004)

Structure and Levers for Computational Evaluation of Customer Capital

Virgilijus Sakalauskas and Dalia Kriksciuniene

Department of Informatics, Vilnius University, Muitines 8, 44280 Kaunas, Lithuania
{virgilijus.sakalauskas,dalia.kriksciuniene}@vukhf.lt

Abstract. The article presents conceptual model and empirical research for evaluation of customer capital of the enterprise. The method is aimed not only for forecasting of turnover and return, but also for evaluating potential response to the enterprise efforts targeted for scoring customers and evaluation of stability of the customer capital. The compound index of customer capital is designed and tested by experimental research.

Keywords: Customer capital, evaluation, compound index, financial indicators.

1 Introduction

The success of the enterprise depends on numerous compounds, which seem of different importance to the managers of the enterprise in order to find most targeted levers which could affect them precisely by the activities of the enterprise. Among these important compounds which are extensively covered in the scientific research literature are various sorts of enterprise wealth, including financial, technological, human, informational, intellectual and customer capital. This list of various types of capitals affects not only functioning of the enterprise but also reflects the opinion of market experts, investors and lenders when the decisions for acquisition of the whole company, evaluation of application for bank loan and audit conclusions are made.

The analysis of market values of the companies as compared to their book values, registered in the enterprise information systems reveal that the biggest part of the enterprise resources cannot be registered in the balance sheets of the enterprises due to lack of adequate methods in the traditional bookkeeping for measuring the value of corporations, whose value lies mainly in their so-called intangible assets.

This problem of inadequate valuation is transferred to the area of business information systems as well, as these systems appear to be unable to reveal the adequate value of the wealth which is used by the enterprise. This leads to inefficient usage of the business information systems, despite to their high capacities of data recording and high availability of intelligent analysis and forecasting methods.

The extent of the problem of valuation of the intangible assets can be illustrated by the difference of enterprise capitalization (total market price value of shares of the listed companies) and book values. In 1982 about 62% of assets in USA companies were registered by enterprise accounting systems. Currently most advanced companies account only 6,2% (Microsoft), 4,6% (SAP), 6,6% (Coca Cola) of their

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 90–101, 2010.
© Springer-Verlag Berlin Heidelberg 2010

assets, leaving the other part of their value to the category of intangible assets, not registered in the accounting information systems [1].

The main problem which is still unsolved in this area is lack of methods and frameworks which could reveal value of customer capital and highlight variables with the highest impact to its changes. Some of the customers are not worth the time or financial investment of establishing learning relationships [2]. Data integration between departments, data warehousing and other intelligence systems empower enterprises with the capability to link customer data, improve functionality of information systems aimed to increasing customer equity by proper management of its levers and metrics.

The article presents conceptual model for evaluation of customer capital of the enterprise based on design of compound index by combining several methodologies: scoring, time series analysis of turnover and return, measurement of customer loyalty influence and forecasting stability of the customer capital. The method is applied for the experimental research for evaluating customer capital of the trading enterprise.

2 Methods for Evaluation of Customer Capital

Customer capital (or customer equity) is differently defined in various research works in marketing management, finance and computational analysis areas. In the studies of intangibles assets sphere customer capital is a structural part of intellectual capital of the enterprise [3,4].

The research of customer capital valuation from marketing perspective is oriented to define its structural parts, finding indicators or factors, which influence its value. In [5] it is advised to calculate the customer equity as index of weighted compounds: customer advocacy, preference, customer price premium, satisfaction and reach. Lemon, Rust and Zeithaml [6] suggest that an enterprise should pay attention to three drivers of customer equity: brand equity, value equity and relationship equity. Within each of these drivers are actions the enterprise can take to increase its overall customer equity. Brand equity is built through image and meaning, it has never been calculated as a monetary amount in any satisfactory way [2]. Brand acts as a ‘magnet’ that can attract new customers to the enterprise. Brand equity can be defined broadly as the extensive set of attributes that influence customer choice above and beyond its objectively perceived value. Value equity is basically the amount of value a customer perceives in the company‘s product or service. The main levers for managing value equity are quality, price and convenience. The key levers controlled by the enterprise for managing relationship equity include: loyalty programs, special recognition and treatment, affinity programs, community building programs and knowledge building programs. These levers can maximize the likelihood of customer repurchase, maximize the value of customer’s future purchases, and minimize the likelihood of the customer switching to a competitor.

As the compounds are highly subjective metrics, their value is calculated by empowering various methods of expert judgment and is further applied to make impact in the area of customer capital by managing separate variables. Therefore the total value of the customer capital remains highly biased and difficult for monitoring progress of the enterprise or change in its value.

Variables, derived from transactions data (which are objective and quantitative) are included to the model by Reinartz and Kumar [7]. They suggested including data of purchase propensity, past customer spending level, cross-buying behavior, purchase frequency and recency, and marketing contact by firm.

These variables were used as proxy variables for defining financial value of the customer, namely customer lifetime value (CLV), which leads to applying the financial approach to customer capital evaluation in two main areas used for evaluation of the enterprise. These are financial assets and flow. Numerous studies presented in the scientific research this area define customer capital by measuring customer lifetime value in several ways and in defining its monetary values based on variety of assumptions. The financial evaluation methods are mainly applied for analysis of the listed companies, as their market capitalization value can be used for testing the ability suggested methods of measuring customer equity and explain the difference of market value from the book value in the accounting system.

Customer equity can be thought of as the sum of the lifetime values (LTVs) of all the customers in the customer base. Customer equity can describe the effectiveness of customer strategies and implementation because it is primarily determined by the total value of all customer relationships of the enterprise [2].

In Blattberg and Deighton [8] customer equity is defined as the total of the discounted lifetime values of all the customers. The enterprise views the total value of its customers in terms of their current profitability and also with respect to the net discounted contribution flow from the customer over time. In [9] customer lifetime value (CLV) is the present value of all future profits obtained from a customer over the life of the relationship with the firm.

The variables included into calculation of customer lifetime value are presented in the formula (1):

$$CLV = \sum_{t=0}^{T} \frac{(p_t - c_t) r_t}{(1-i)^t}, \tag{1}$$

where p_t - price paid by customer at time t; c_t - direct cost of servicing the customer at time t; r_t - probability of customer buying from the firm at time t; *i*-discount rate of cost of capital to the firm; *T*- time horizon for estimating CLV.

This concept is similar to the discounted cash flow approach used in finance, except for two key differences, as CLV is calculated for each individual by recognizing differences between customers by their importance and profitability, and it also has in mind the possibility, that the customer can switch to the competitors in future.

Gupta, Lehman, and Stuart [10] provide the following model for calculating customer equity (CE) and substantiate the possibility to further modeling of its compounds:

$$CE = \int_{k=0}^{\infty} \int_{t=k}^{\infty} n_k m_{t-k} e^{-ik} e^{-(\frac{1=i-r}{r})(t-k)} dt dk - \int_{k=0}^{\infty} n_k c_k e^{-ik} dk, \tag{2}$$

where n_k is the number of newly acquired customers for cohort *k*; *m* – margin; *r*-retention rate; *i*- discount rate; *c*-acquisition cost per customer.

The negative side of the model is that the customers are analysed individually, the variables include various types of uncertainty, and the attitude of the company to customers vary not only because of their profitability, but other variables as well. Kumar and Shah [12] suggest evaluating CLV by including variables of predicted probability that the customer will purchase in a certain period, predicted contact with the customer, and also index for future time periods. This model is oriented to forecast customer capital value for the three year period suggested by the authors.

The computations of the models include variety of forecasting, modeling and optimization methods. The high level of uncertainty included to defining values of varibles encourages applying artificial intelligence methods. The fuzzy Analytic Hierarchy Process (fuzzy AHP) is proposed in [11].

These methods of CLV are oriented to forecasting value of customer capital, and bring negative outcomes due to the subjective judgments applied for defining their value, based on volatile assumptions and market prices, which cannot be applied to non-listed companies, and attempts to make evaluation by converting customer related data to the monetary equivalents.

The application of customer capital evaluation in forecasting potential of the enterprise of developing current business relationships and enhance them in future suggests that the monetary value of these characteristics can be defined only in subjective way. Therefore the approach of designing compound customer capital index is selected for further research. The suggested index includes quantitative information available in the accounting systems and documented during the relationships with the customers and reflects the structure of customer capital by defining variables responsible for tracking its changes.

3 Customer Capital Evaluation Method

The suggested method for evaluating customers allows to make comparative evaluation of change in customer capital by creating customer capital index. In this chapter the compounds of the index are suggested by taking into account the quantitative information of the sales department, and then their influence to the index value is substantiated and explored.

The compound index of customer capital is calculated as the weighted average of the indicators, which reflect changes in the customer capital from three perspectives; financial, marketing and enterprise image. The number of variables which belong to the three perspectives is optional and can be varied by the enterprise, but we have defined two requirements for their selection: we can define variable only in case if it is related to the enterprise goals and is it is of quantitative origin. All the indicators have to be regularly available from the enterprise accounting system and have no subjective values. Therefore ten indicators were selected by taking into account the enterprise goals in the selected areas and the importance of the indictors, expressed by weights (Table 1). Each indicator, included into the customer capital index is evaluated not by its monetary value, but by its change. All the indicators have maximum value 1. The core idea of the method is to explore the employment of the

Table 1. Set of indicators selected for research of customer capital index

Indicator	Measure	Variable
Financial indicators:		
Size of the customer	Assigned to category according to annual turnover	KD
Profit margin	Profit margin calculated per customer	KAP
Annual order quote	Annual order volume quoted by customer	MKP
Number of customers	All customers of the enterprise, including present and lost customers	IKS
Marketing indicators:		
Life cycle	Life cycle of the customers is measured in number of days since first visit	KGC
Frequency of visits	Frequency of usage of services	PND
Intensity of visits	Time interval between visits, measured in number of days	LTA
Enterprise image indicators:		
Satisfaction	Customer satisfaction level derived from customer remarks (positive or negative) registered in the accounting system and level of customer order fulfilment	KPL
Share of long-term customers	Rating of customers according to their life cycle	IKD
Acquisition ratio	Ratio of new customers to lost customers	N/P

customer capital in the enterprise according to the defined goals: the index value 0 indicates that the customer capital is not employed as the asset, and its maximum value 1 means, that the enterprise fully uses the potentials of the customer capital.

One of the problems of customer capital evaluation is finding common basis for evaluation, whether we have to take into account influence of each customer or to analyse influence of the customer groups which are addressed with different strategy of the enterprise. The influence of each particular customer could be difficult to notice and interpret. Therefore the customers were grouped according to their influence. The influence of the group can be understood in different way for various types of enterprises, in most cases they can be divided by turnover or profit margin, depending on different level of services required by the group. Number of customer groups can be explored and selected by each enterprise by applying computational analysis (clustering or other grouping method) or subjectively defined by the competent experts of the enterprise.

The weight of each group for defining customer capital value can be assigned by applying the cumulative ‘whale curve‘ method, where the share of the group in the enterprise annual turnover or profit margin is calculated.

We denote m groups with the corresponding weights $w_1, w_2, ..., w_m$, and the number of customers in each group $n_1, n_2, ..., n_m$. In order to calculate the corresponding value of each indicator the enterprise has to have a set of hypothetic (optimal) indicators, which describe ideal results of the enterprise aimed towards each sphere of its performance. These indicators of ideal performance form the set of targeted values of goal indicators ‘*goal( )*', where the variable in the brackets shows the group indicator;

e.g. $goal(n_i)$ denotes the ‘ideal‘ value of number of customers in group i. The indicators are calculated as ratio of current values to the values from the goal set.

The variable “Size of the customer” (KD) is defined according to the annual turnover generated by the customer (3):

$$KD = \sum_{i=1}^{m} w_i \cdot \frac{n_i}{goal(n_i)}, \tag{3}$$

The goals of the enterprise are related to keeping big customers, as the biggest share of enterprise profit is generated by them. Information about specific features of customer orders, their history and conditions can be optimized according to the enterprise goals.

The variable “Profit margin” (KAP) is calculated for each customer enables to identify most important customers.

$$KAP = \sum_{i=1}^{m} w_i \cdot \frac{1}{n_i} \sum_{j=1}^{n_i} \frac{annual\ return(j)}{goal(annual\ return\ \ group\ \ i)} \tag{4}$$

The goal of the enterprise is to thoroughly watch this group of customers, analyse their requirements, formulate special offers. The main indicators for analysis of customer profitability include predictive information about possible churn and level of reaching target profit estimated for each customer. The target profits for customer groups are set among the goals of the enterprise. The profit share of each customer inside the group is calculated and is used for deriving the value of KAP indicator as weighted average of all customer groups.

The “Annual order quote" variable (MKP) can be calculated if the enterprise uses the quoting system, where the customer makes preliminary order for the forthcoming year.

$$MKP = \sum_{i=1}^{m} w_i \cdot \frac{1}{n_i} \sum_{j=1}^{n_i} \frac{order\_quote(j)}{goal(order\_quote)} \tag{5}$$

The ordering system allows to apply the „blanket ordering“ method which enables optimizing supply by scheduling partial orders, which could minimize logistic costs and be compatible to production, seasoning, analysis of demand for various products or other processes. The goal defines order quote for each customer and group. This indicator can be undermined, if the enterprise cannot fulfil the customer orders during the financial year.

Management of “Number of customers” (IKS) is one of the most important goals of the enterprise.

$$IKS = \frac{annual\_change\_of\_customers}{goal(annual\_change\_of\_customers)} \tag{6}$$

The customer capital and its potential to generate financial yields increases not only by expansion of total number of customers, but increasing the share of bigger customers. This indicator shows the change of number of customers at the start and end of the period. The goal is to reach optimal number of customer which could enable highest performance.

The group of indicators, related to marketing efforts of the enterprise include variables of strength and intensity of the relationships with the customer. These are variables of customer life cycle, frequency and duration between visits.

Customer "Life cycle" variable (KGC) reflects the constant collaboration between the customer and the enterprise.

$$KGC = \sum_{i=1}^{m} w_i \cdot \frac{1}{n_i} \sum_{j=1}^{n_i} \frac{life\_cycle(j)}{goal(life\_cycle(i)} \tag{7}$$

This indicator is diminished if the share of single-visit customers is high. The goal of the enterprise is to increase share of long-term customers and create background for continuous financial flow by order quoting and fulfilment. The customer life cycle can be measured in number of days since the first visit of the customer. Naturally, each enterprise has part of migrating customer, newly acquired and loyal customers. The weighted average of customer life cycle within groups can be compared to the enterprise goals related to customer loyalty and extension of duration of cooperation.

Frequency of using services (PND) shows the dependency of the customer and the enterprise.

$$PND = \sum_{i=1}^{m} w_i \cdot \frac{1}{n_i} \sum_{j=1}^{n_i} \frac{frequency(j)}{goal(frequency(i)} \tag{8}$$

On the one hand, high frequency expresses strong customer relationship, but on the other hand the enterprises need to optimize this indicator by introducing minimum order, as high frequency of services increases operation costs.

"Intensity of visits" (LTA) is measured by the number of days between usage of services by the customer.

$$LTA = \sum_{i=1}^{m} w_i \cdot \frac{1}{n_i} \sum_{j=1}^{n_i} \frac{time\_between\_visits\,(j)}{goal(time\_between\_visits(i)} \tag{9}$$

It is the indicator, which denotes the tightness of the relationship of the customer to the enterprise, as the shorter intervals between visits usually can show tendency for the customer to become constant and easier to forecast his requirements. The variable is optimized according to the ability to arrange work processes of enterprise.

The "Satisfaction" variable (KPL) expresses level of customer satisfaction derived from customer remarks (positive or negative) registered in the accounting system and also the level of fulfilment of customer orders.

$$KPL = \sum_{customer} 0.8 \cdot \frac{fulfilled\_orders}{quoted\_orders} + 0.2 \cdot \frac{\sum_{\text{acknowledgement}} 1 - \sum_{\text{reprimand}} 1}{\sum_{\text{acknowledgement}} 1 + \sum_{\text{reprimand}} 1} \tag{10}$$

The first variable is calculated for each customer by summarizing all positive acknowledgements and subtracting number of reprimands (applying weight 0.2). In order to increase this indicator it is very important to provide service in the most convenient way for the customer. That is very important for keeping long-term customers. The second variable is calculated as ratio of fulfilment of the quoted customer order (multiplied by weight 0.8). The corresponding weights of 0.2 and 0.8 express that it is very easy to lose long-term customers, but it is very difficult to replace them by new customers and to make them loyal.

The "Share of long-term customers" is the variable (IKD) which enables us to forecast long-term profit (11).

$$IKD = \sum_{i=1}^{m} w_i \cdot \frac{1}{n_i} \sum_{j=1}^{n_i} relationship\_coefficient(j) \tag{11}$$

The value of IKD is calculated as the weighted average of all customers' indicators. The bigger share of long-term customers, the better is stability of enterprise financial flow. Each customer is assigned the coefficient (see table 2) according to the duration of his relationship with the enterprise:

Table 2. Customers coefficients

Relationship duration (years)	<1	[1; 2]	[2; 3]	[3; 4]	> 4
Coefficient	0.2	0.4	0.6	0.8	1

The variable "Acquisition ratio" (N/P) is calculated as the ratio between newly acquired and the lost customer during analysis period.

$$N / P = \frac{\sum_{New} 1 - \sum_{Lost} 1}{\sum_{New} 1 + \sum_{Lost} 1} \tag{12}$$

The negative value of the variable shows, that the enterprise loses more, than acquires new customers. This indicator is related to the enterprise strategy not only in the area of attracting new customers, but also in optimizing number of customers in different groups. Some enterprises put more value in finding new customers despite their small share or risk that they will make only one-time purchase. But if the strategy is based on promoting long term contracts, then the loss of customers in the most valued groups brings more impact.

The selected list of indicators is included to the compound index of customer capital by calculating the weighted average and can reflect changes of customer capital in the selected areas of the enterprise. The selection of weights has quite big impact for the value of index.

If the method is applied for tracking changes of customer capital within one enterprise, the weights can be selected and approved by the experts and management. The analysis of groups of enterprises can lead to designing the customer capital index structure for comparative evaluation between ability of companies to manage their customer capital, and their attractiveness for investment.

Further we present the experimental research for investigating expert and quantitative approaches for defining weights of the index and exploring the impact of the variables of the customer capital. The experimental data was applied for computational analysis, interpretation and application of the suggested method.

4 Experimental Research

The experimental database was collected in middle size trading enterprise during 2008. The data of 325 customers, 5630 transactions was analyzed. The experimental research aimed to apply suggested method of the customer capital index and its ability to explain the customer– related changes of the enterprise value.

The research was performed in two stages: the weights of compound index were explored; the structure of enterprise customers was analyzed. The compound variables were calculated and explored by their impact to the enterprise strategy. The possibilities to apply the method of customer capital index were researched and evaluated.

The weights denoting importance of the variables were firstly defined subjectively by surveying the experts of the enterprise [13]. The opinions of the experts were rather different; they tended to assigned higher weight to the variables, which reflect their work responsibility: finance, trade or marketing. The average values of the weights, calculated after the survey of enterprise experts are presented in the Fig. 1.

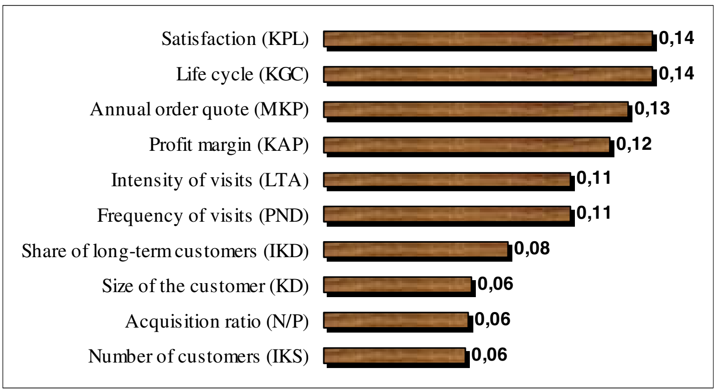


Fig. 1. Weights of the variables included to the compound customer capital index

The biggest weight was assigned to the customer satisfaction and the duration of the life cycle. The smallest weight was assigned to the customer size, number of customers and customer acquisition ratio. These weights were used for calculating the customer capital index value (CCI).

Table 3. The values of the variables

CCI	IKS	KD	MKP	KAP	LTA
0,6189	0,9121	0,4205	0,4747	0,5060	0,8761
	KGC	**PND**	**KPL**	**IKL**	**N/P**
	0,9370	0,5823	0,6948	0,9429	-0,4286

The values of the variables (Table 3) were calculated by using the trading data of the enterprise and the algorithm described in section 3. The value of customer capital index (CCI) 0,6189 means, that in 2008 the potential of the customer capital was used only by approximately 62%, as related to the enterprise goals.

The result shows that the main reason of this low index value was in fulfillment of annual order quotes only by 47% and annual profit goals were met only by 51%. This year was also marked by negative value of N/P variable, which meant that there more customers are lost than newly acquired. The variable KD shows that the most profitable customers make only small share of the total number of customers. The separate analysis of each indicator can give insights and advice for improving work of managers. Usually this type of work is done in all enterprises, but it does not reflect the customer capital as the asset and its change in value. Therefore the index value CCI reflects the size of unfulfilled potentials, and attractiveness for investors.

As the customers are very heterogeneous by size of their generated annual turnover, profitability and duration of their life cycle the related variables are difficult to interpret. Therefore the research of the influence of the variables included to index was analysed by grouping customers. By applying K-means Clustering five groups of customers were defined according to their annual turnover. The number of customers' included to each group are presented in frequency table (Table 4).

Table 4. Frequency table by the customer annual turnover

From To	Frequency table: Code (daznis4.sta)			
	Count	Cumulative Count	Percent	Cumulative Percent
0<x<=10 000	188	188	57,85	57,85
10 000<x<=50 000	76	264	23,38	81,23
50 000<x<=100 000	27	291	8,31	89,54
100 000<x<=200 000	23	314	7,08	96,62
200 000<x<=300 000	11	325	3,38	100,00

The weights were assigned to groups according to the clustering results. The fifth group consists of VIP customers, who accumulate 63,44 % of enterprise turnover. The smallest turnover is generated by the group of small customers, which is highest in number of participants (Fig. 2).

Fig. 2. The weights of the groups of customer

The further investigation of changes in customer capital CCI was done by experimenting with different customer groups. The CCI value reached its maximum when the 1st group was eliminated. This group is biggest in size, therefore the increase in CCI (from 0,6189 to 0,7030) is kind of paradox, as large number of customers is lost. The increase of CCI is related to changes in quality of the customer base. As the 1st group consists mainly of one-time purchase customers, the relationship with them is missing or is very weak and unmanageable, therefore their influence in the customer capital is related to high efforts with the results, which are difficult to forecast. After eliminating this group, the customers' base consists only of best valuated customers, which directly impacts CCI index.

The impact in CCI achieved by eliminating 2nd, 3rd and 4th groups decreases CCI value. By eliminating these groups, the enterprise loses more important customers not only by their turnover which has already been generated but by losing their future potential as well.

The elimination of the 5th group caused decrease in CCI by almost 50 percent. This group, which is smallest in members, but most highly ranked by the length of their relationships, constant purchases and turnovers, as expressed by the values of variables included to the CCI. The results in changes of CCI are presented in Table 5.

Table 5. CCI and other indicators calculation results by eliminating groups of customers

	Eliminated group of customers				
	1	*2*	*3*	*4*	*5*
CCI	**0,7030**	**0,5669**	**0,5241**	**0,4672**	**0,3593**
IKS	0,9236	0,9010	0,9124	0,9154	0,9122
KD	0,9691	0,6504	0,3180	0,3050	0,2283
MKP	0,4567	0,4191	0,3900	0,3475	0,2856
KAP	0,4888	0,4592	0,4301	0,3833	0,2625
LTA	0,8335	0,7761	0,7261	0,6261	0,5426
KGC	0,8992	0,8536	0,8029	0,7048	0,487
PND	0,5520	0,5034	0,5022	0,4602	0,3114
KPL	0,6570	0,6217	0,5788	0,4986	0,4231
IKL	0,9028	0,8605	0,8126	0,7022	0,4929
N/P	0,6000	-0,6129	-0,4286	-0,4118	-0,4286

The experimental research revealed that the method of customer capital index can be applied for analysis of enterprises for their internal use and for expert evaluation for the external purposes. The main advantages are the availability of objective and consistent information, based on operational transactions of the enterprise.

5 Conclusions

High capacities of data recording, analysis and forecasting are not used to their full extent, and the investment to business IS projects do not meet expected return on investment and high performance in the area of clearly presenting information about the real value of the enterprise and its performance.

The evaluation methods should be applied for valuation of the intangible assets of the enterprise, where customer capital is one of its parts. The basis for evaluation include customer lifetime value and its modifications related to forecasting of its future value, and series of customer related variables which could be used for presenting insights to the changes of customer values.

The main drawbacks of the existing methods are caused by the qualitative variables, subjective judgments applied for defining their value, forecasting of future values, based on volatile assumptions and market prices, which cannot be applied to non-listed companies, and attempts to make evaluation by converting customer related data to the monetary equivalents.

The method suggested in the article is based on designing compound index of customer capital (CCI). The variables included into index serve as levers which define change in its value. These variables are calculated by using customer data consistently registered in the enterprise information systems and the enterprise goals defined in each area.

The experimental analysis explored aspects of interpretation of changes in CCI, defining variable weights and customer grouping, and managing influence of the compound levers. The results reflect changes in customer capital, possibilities to modify method according to the enterprise goals, specific customer related features of business, and information stored in the business IS.

The suggested method can enable enterprise to exploit its potential of customer capital and to precisely target efforts of human capital and marketing expenditures. The modeling and forecasting calculations of the CCI can support successful management of customer capital.

References

1. Ryan, B.: Corporate finance and valuation. Thomson learning, London (2007)
2. Peppers, D., Rogers, M.: Managing customer relationships. John Wiley & Sons, Inc., New Jersey (2004)
3. Sveiby, K.E.: The new wealth: intangible assets (1997), `http://www.knowledgecreators.com`
4. Skandia: Skandia's intellectual capital Annual report (1996), `http://www.skandia.se`
5. Hull, J.: IPA and marketing society (2008), `http://www.ipa.co.uk`
6. Lemon, K., Rust, R., Zeithaml, V.: Driving customer equity: how customer lifetime value is reshaping corporate strategy. The Free Press, New York (2000)
7. Reinartz, W.J., Kumar, V.: The impact of customer relationship characteristics on profitable lifetime duration. Journal of marketing 67(1) (2003)
8. Blattberg, R.C., Deighton, J.: Manage marketing by the customer equity test. Harvard business review 74 (July-August 1996)
9. Gupta, S., Zeithaml, V.: Customer metrics. The past, the present, and the Future in Academia and practice. In: Marketing Science institute. Special report No.05-200 (2005)
10. Gupta, S., Lehmann, D.R.: Managing customers as investments. Wharton School publishing, New York (2005)
11. Chen, H.H.: Measuring intellectual capital using fuzzy analytic hierarchy process. International Journal of Innovation and Learning 6(1) (2009)
12. Kumar, V., Shah, D.: Expanding the role of marketing: from customer equity to market capitalization. Marketing Science institute. Working paper ser. Issue 3. No. 08-003 (2008)
13. Timofejeva, T.: Application of intellectual capital concept for evaluation of enterprise customers. In: Proceedings of 14th Annual Conference 'Information Technologies 2009'. Vilnius University (2009)

Data Mining for Small Organizations

Auksė Stravinskienė and Saulius Gudas

Kaunas Faculty of Humanities of Vilnius university, Muitinės st. 8, Kaunas, Lithuanian
saukse@gmail.com,
saulius.gudas@vukhf.lt

Abstract. The article analyses the use of data mining technology for small businesses, which are not able to invest in information technologies but seeking to optimize business processes and use accumulated data effectively, is presented. A comparison of data mining tools is very important for choice an appropriate tool.

Keywords: Data mining technologies, enterprise knowledge, decision support.

1 Introduction

A problem of extraction and operation of knowledge is one of the urgent issues in a business world. An ability to operate and use this knowledge in order to implement objectives of an enterprise is one of the guarantees of successful business development. Big organizations have applied BI tools perfectly well, however, small businesses frequently encounter a problem, since majority of data mining tools costs a lot comparing with received income of an enterprise, therefore small businesses, by necessity to analyse data, use simple spreadsheets. However, most often a gained result does not give such benefit as it could be obtained from the usage of data mining (DM). The article gives some suggestions how this problem might be solved.

The aim of the work is to recommend the data mining tools, which would be suitable to use for small businesses evaluating convenience and opportunities of their usage. In this article a comparison of two tools, which might be effectively applied in small businesses, is presented. The advantage of main tools is that they both are integrated into Excel spreadsheet of Microsoft Office packet.

2 The Importance of Data Mining Tools for Small Organizations

Data mining in small businesses is little studied and the majority of the offered products are oriented to big or at least medium size enterprises. A few years ago Melissa Solomon made a research about the willingness of small businesses to use high technologies. He distinguished a decade of IT fields in which small businesses planned to invest their money. After gathering the information about what they do and what they are interested in, this information might be used to analyse work of an enterprise and plan new suggestions regarding demand. (Solomon M., 2005)[1].

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 102–105, 2010.
© Springer-Verlag Berlin Heidelberg 2010

However, how important data mining is for small businesses is greatly revealed in Barny Ritholz article ``Intro to Data Mining for Small Businesses". The main idea of the article is that data mining can help enterprises to react beforehand to the market changes and to survive easier even during financial crises. It is stated that if small businesses used data mining tools by themselves they could notice the market changes, their tendencies and various anomalies themselves.(Ritholz, 2009)[2].

Small businesses ought to pay more attention to data analysis. One of the main and most important functions of DM for small businesses would be a prognostication. After evaluating enough indexes it would be much easier for enterprises to plan their activities, future income and expenditure. Subject to activities of an enterprise it can be predicted its income, sale, a number of orders and so on [3].

3 A Comparison of Data Mining Tools

As it has already been mentioned above small businesses often do not have enough finances to purchase DM tools, however, they seek to optimize activities processes and efficient use accumulated data [4]. Further on the table is presented (see table 1) which gives a survey of main DM functions of two analysed tools– XLMiner and Microsoft SQL Server 2008 Data Mining Add-Ins.

Table 1. A comparison of Data mining tools

	XLMiner	Microsoft SQL Server 2008 DM
Prediction	Multiple Linear Regression, K-Nearest Neighbours, Regression Trees, Neutral Network.	Time Series (forecasting), Logistic Regression, Decision Tree algorithm.
Classification	Discriminant Analysis, Logistic Regression, Classification Trees, Naive Bayes, Neutral Network, K-Nearest Neighbours.	Clustering (setting categories), Decision Tree algorithm, Logistic Regression (completing according the example), Logistic Regression.
Affinity	Association rules	Naive Bayes, Association Rules.
Data Reduction	Principal Component Analysis, k – Means Clustering, Hierarchical Clustering.	-

The analysed tools have functions both for DM and modelling. One of the most important DM functions is data partition which allows grouping the data. One group is used for teaching a model and the other one is used for testing. If the data is not full, for example some records are missing or there is a wish to change one part, for this purpose tools have many functions. Enterprises should estimate that the more definite and comprehensive data is analysed the better outcomes are gained. One of the most frequently used functions in enterprises is prognostication which helps to evaluate future prospects easier and plan actions of an enterprise more reliably.

A wide choice of classification functions allows a user to process the same data applying various ways, to observe similarities and differences of outcomes and to draw consistent conclusions. Also Microsoft SQL Server 2008 Data Mining Add-Ins gives other different possibilities to use classification, it allows modelling particular settings. Such modelling functions allow evaluating different possibilities without experiencing loss.

As we could see both tools have very similar functions, however, their settings and particularity a little bit differ. Microsoft SQL Server 2008 Data Mining Add-Ins might be liked by users due to its wizards which function quite by the light of nature and a user only need select the main settings. Whereas XLMiner would be more appropriate to users who have more knowledge or want to gain it.

4 Results

The data summary of a small logistic enterprise was used in order to conduct the experiment. The main goal of the enterprise is to optimize its activities. The enterprise seeks to reject orders which are loss-making for the enterprise and strengthen relationships with the enterprises which make the best orders.

In order to implement the data mining process these products such as XLMiner and Microsoft SQL Server 2008 Data Mining Add-Ins were used. The data used for the experiment: a customer, a carrier, a date of ordering, a date of delivery, a country from which a load is transported, a type of a load, and a price of an order.

There is a graphic showing distribution of order prices (picture 1). It is apparent that most frequently made orders are those which value is up to 9715.66 Lt. Orders of higher value are rarer, consequently a primary conclusion could be drawn that it is unrewarding to implement them. Our aim is to clarify which main factors influence a price of an order.

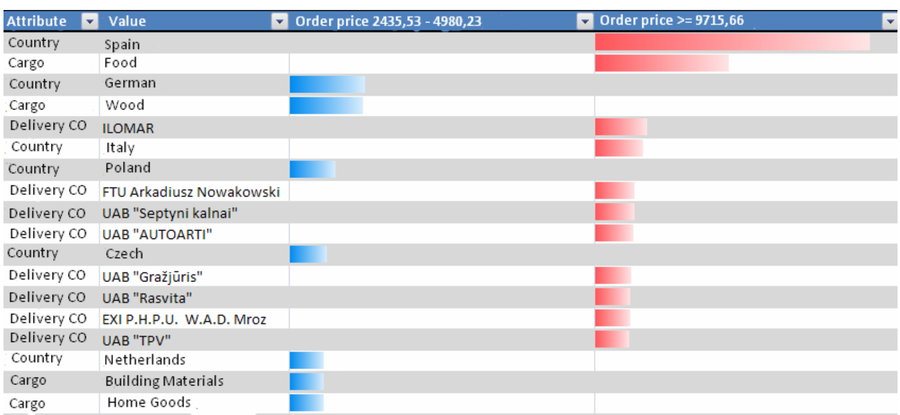


Fig. 1. Distribution of order prices

After using algorithm of neutral networks for a classification problem, an outcome has shown what makes the biggest influence on a high price. Different results obtained from the classification using decision tree algorithm (picture 2).

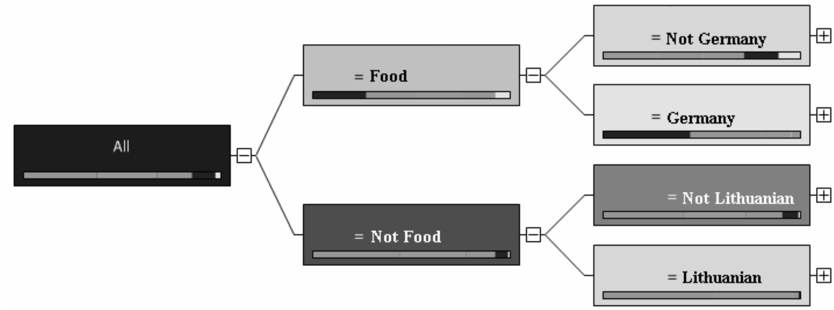


Fig. 2. Result of Decision Tree algorithm

5 Conclusions

All things considered, we could emphasize that data analysis and data mining is talking point not only for big enterprises. Small businesses often don't have DM tools. They seek to optimize business activities by efficient usage of accumulated data.

Using appropriate tools it might be applied to small businesses which are considering about future prospects as well. Employees of the enterprise can independently evaluate outcomes of activities of the enterprise and make decisions easier. It's very important to use DM tools for data analysis, because organization could predict the market changes, their tendencies.

Two tools of data mining have been analysed - XLMiner and Microsoft SQL Server 2008 Data Mining Add-Ins, which do not operate as separate applications but as supplements of MS Excel spreadsheet. Both tools are perfectly suitable to use in small businesses because they have a huge variety of functions. Various functions of the tools can help to analyse data, after evaluating gained outcomes small businesses can optimize activities' process and effectively use accumulated data.

References

1. Solomon, M.: Ten High-End Technologies You Can Afford, CDW Corporation (2005), `http://www.biztechmagazine.com/article.asp?item_id=87`
2. Ritholz, B.: Intro to Data Mining for Small Businesses. American Express Company (2009)
3. Piróg-Mazur, M., Setlak, G., Szajnar, W., Kożak, T.: The Application of Data Mining Methods to the Classification of Finished Products. International Book Series Information Science and Computing (2007)
4. Larose, D.T.: Data mining methods and models, p. 344. IEEE Computer Society Press, Los Alamitos (2006) ISBN-13 978-0-471-66656-1

BITA 2010 Workshop Chairs' Message

Ulf Seigerroth and Kurt Sandkuhl

Jönköping University, Jönköping, Sweden
{Ulf.Seigerroth,Kurt.Sandkuhl}@jth.hj.se

A contemporary challenge for enterprises is to keep up with the pace of changing business demands imposed on them in different ways. There is today an obvious demand for continuous improvement and alignment in enterprises but unfortunately many organisations don't have proper instruments (methods, patterns, best practices etc.) to achieve this. Enterprise modelling and business process management are two areas belonging to a tradition where the mission is to improve business practice and business and IT alignment. In this tradition the alignment process usually is manifested in taking a business from one state into another improved state, i.e. a transformation of the business and its supporting IT into something that is regarded as better. A challenge in business and IT alignment is to move beyond this narrow focus on one tradition or technology. We need to be aware of and able to deal with a number of dimensions of the enterprise architecture and their relations in order to create alignment. Examples of such dimensions are: organisational structures, strategies, business models, work practices, processes, and IS/IT structures. Among the concepts that deserve special attention in this context is IT governance. An effective IT governance aligns IT investments with overall business priorities, determines who makes the IT decisions and assigns accountability for the outcomes. There are ordinarily three governance mechanisms that an enterprise needs to have in place, 1) decision-making structures, 2) alignment process, and 3) formal communications.

This workshop aimed to bring together people who have strong interest in business and IT alignment. We have invited researchers and practitioners from both industry and academia to submit original results of their completed or ongoing projects. We have encouraged broad understanding of possible approaches and solutions for business and IT alignment, including IT governance subjects. Specific focus was on practices of business and IT alignment, i.e. we have encouraged submission of case study and experiences papers.

The workshop received 11 submissions. The program committee selected 4 submissions for presentation at the workshop.

We thank all members of the program committee, authors and local organizers for their efforts and support.

May 2010 Ulf Seigerroth, Kurt Sandkuhl

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, p. 106, 2010.

Design Thinking: Towards a Unified View of Organizational and Technological Realms

Dimitris Bibikas and Eleni Kargioti

South East European Research Centre (SEERC) of the University of Sheffield and CITY College, 24 Proxenou Koromila Str.,
54622, Thessaloniki, Greece
{dbibikas,ekargioti}@seerc.org

Abstract. When is it appropriate to discuss the divergence of organizational and technological realms and based on which assumptions? Could the social and technical disconnection be merely a constructed phenomenon, and if so, what does design thinking has to offer towards the creation of aligned organizational and technological views? Drawing from these questions this paper argues that the two notions of organization and technology are both human artifacts with an intertwined and unfolding ontology. We base our argument inspired by i) structuration theory according to which organizational structures and technologies are mutually constituted and co-evolving elements, and, ii) a conceptualization of organizations and technologies as 'epistemic objects' with a lack in completeness of being. We finally discuss design thinking principles as a basis for an integrated reflection of the two notions hopping to invoke creative and constructive dialogue among communities of technologists, managers and academics alike.

Keywords: Design thinking, organizations, technologies, structuration theory, epistemic objects.

1 The Socio-technical Gap Is Not 'Out There'

1.1 Motivation

As the bold title of this introductory section implies, our motivation lies on the tendency of both academic and practitioner communities to treat the disconnection between the organizational and technological realms as an established and inevitable situation. We think that such an approach tends to objectify the divergence between the *social* and the *technical* and give rise to the so-called 'socio-technical gap'. Contrary to this view, we believe that such a disconnection between organization and technology can be particularly narrow-focused and does not do justice to neither of the two notions. Therefore, our paper avoids the inclination of much of the literature to reify such a parting and examines critically the nature of this relationship.

1.2 Aim

Drawing from the body of literature investigating the social construction of both *organization* as well as *technology* [1], we discuss the alleged socio-technical gap not

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 107–118, 2010.
© Springer-Verlag Berlin Heidelberg 2010

as a fixed situation, but rather as a product of our own thinking and (inter)action. Viewed this way, the disconnection between the social and the technical emerges through our daily enactments [2] as organizational theorists, computer scientists, technologists, managers, IT staff and knowledge-workers inside work settings.

Under this light, we investigate the creation of this disconnection through the lens of different schools of thought about the nature of technology and organization. Our aim is to present a mutually constitutive view of organizational structures and technological artifacts and to examine both as an unfolding accomplishment of everyday situations. We discuss design thinking principles as an alternative lens for a unified view of organizational structures and technological artifacts. Defining a specific methodology for applying design thinking or prescribing a technique for organizing and developing technology is beyond our intention. Nevertheless, we argue that by employing design thinking to view organization and technology as one 'intertwined object of knowing', characterized by an 'unfolding ontology' being constantly in flux, rather than fully formed [3], the idea of a persisting gap is averted.

1.3 Approach

We base our conceptualization on *technology* and *organization* to the theory of structuration [4], whereby technology-in-use and organizational orderings interact recursively. Human activity is mediated by information technology, and actors interpret and appropriate technology in order to produce and reproduce organizational arrangements [1,2,4]. Interpretive schemes, moral norms and understandings of power from the broader institutional context are enacted during technology use by actors; actors appropriate technology in action and select to retain or change social structures. Similarly, we avoid conceiving the notions of organization and technology as a static property of specific procedures or relations between individuals and groups, and see them as a situated and ongoing accomplishment that emerges in everyday work [1].

Organization and technology are viewed as having a reflexive relationship: technology-in-use is *constitutive of* and *constituted by* changing organizational orderings, while, temporal organizational structures are *constitutive of* and *constituted by* technological artifacts in-use. Therefore, organization and technology are an ongoing and situated action with capacity to unfold indefinitely. Following this combined and unfolding ontology of organizations and technologies, we are liberated from the stable disconnection between them. The two notions are conceptually integrated and explored based on the design thinking principles as incomplete socially constructed objects with both intended and unintended consequences.

1.4 Structure of the Paper

In the next sections we discuss some of the main arguments of the socio-technical gap and argue that socio-technical theory may have nurtured certain conceptualizations of the organizational as well as the technological realms that sustain this gap. We then explore these conceptualizations and justify their restrictive focus. Examples of these views, based on secondary data from literature, are given to support our argumentation. Design thinking is presented next as a promising theoretical foundation to overcome views of alienated social and technical realms. Finally, we conclude this paper with implications for further research and practice.

This section introduced the motivation and aim of the paper. Next, we present the main viewpoints of the alleged ‘socio-technical gap’, and argue that it emerges from false assumptions that concern *organizational structures* and *technological artifacts* as stable and fixed phenomena.

2 Main Opinions on the ‘Socio-technical Gap’

The socio-technical gap is perceived within a large proportion of the literature as the result of a misfit between social arrangements and the design and implementation of information technology solutions. According to this view, over emphasizing on the technical requirements of IT infrastructures (i.e. hardware and software components) often results in diminished attention to organizational structures [5,6]. As many theorists argue, the technical sub-system leaves a significant part of the social sub-system unsupported. Therefore, the socio-technical gap indicates a social sub-system that is weakly supported by the technical infrastructures of the organization.

Socio-technical theory was proposed to address the abovementioned divide. It focuses on the *joint optimization* of both organizational structures and technical infrastructure that constitute the total work system [7]. The technical sub-system comprise tools, hardware and software systems and codified knowledge assets, necessary to produce certain outputs [8], while the attitudes, beliefs, relationships and results of work arrangements constitute the social sub-system. The main premise of socio-technical studies is the belief that the outcomes of organizations can be managed by jointly ‘optimizing’ the social and technical elements of the workplace.

Without dismissing the contributions of the socio-technical theory to constitute technology responsive to social structures [9], this paper is opposed to its basic grounding that renders the entities of social structures and technological artifacts as dichotomized that can be jointly optimized in order to achieve alignment. Socio-technical theory relies on analyzing the relational properties of the social and technical subsystems in order to acquire a better understanding of the total work place. Nevertheless, such a view has been criticized to have a high degree of formalism, much like a rule governed approach, reinforcing the separation of the social and the technical [11]. The presumption that there is a finite ‘optimized’ state in which organizational structures and technological infrastructures are aligned could be restrictive within a complex and dynamic environment. Similarly, the social construction view of technology seeks to explain how individual and collective perspectives give meaning to technologies [12]. While this view provides a rich picture concerning the ways technologies are conceptualized, it is reported that it has “…yet to examine how agency shapes the way technologies influence work practices and organizational structures once the technology is deployed and used in organizations” [13, p.149].

Technologies and organizations are both at the same time social and physical artifacts. If we focus primarily to a strictly constructionist or materialist view, we neglect the continuous change, negotiation and inter-play between agency and materialism; a procedure present in both organizational structures and functionalities of technological tools, as these are being mutually (re)formed by one another in practice. We argue that the main cause of diminished attention towards a reciprocally constitutive relationship between the social arrangements and technical artifacts lies on specific conceptualizations of technology and organizing within work settings.

In this section we briefly presented the main arguments of the socio-technical theory and also argued that aiming at their mutual optimization tends to objectify organizational arrangements as well as technological functionalities. Next we describe the main conceptualizations of technology and organizing within institutions that lead to diminished attention towards a reflexive relationship between organizational and technological artifacts and then attempt to provide an alternative view of both.

3 Conceptualizations of Technology and Organizations

In this section we describe the conceptualizations of technology and organizations in literature and demonstrate how they do not explain well all aspects of technologies and organization. These views are accompanied with secondary data from a relevant case study found in the literature [14]. The example examines the organization of the development, use and integration of a particular information technology system in a large, multi-national software consulting firm. The information technology refers to 'productivity tools' that were designed and created by the company's developers, based on a requirement by management, to better support the firm's consultants during the systems development process.

3.1 First Technology and Organization View: 'Effectual' to Social Structures

The first view of technology and organization reflects the properties of an abstract object with measurable and predictable impacts. In terms of technology, the emphasis is placed to its foreseen consequences on organizational properties, such as size, performance, effectiveness and productivity [14]. A similar approach to organizing guided by the functionalist paradigm views organizations as entities privileged by decision-making processes and strategic planning capabilities based on a priori rationality [10]. The order of organizational performances is then largely reified and taken for granted. Given this view, the interactivity between social systems and technology can be predictable and measurable [15].

An example of this objectified view of both technology and organization can be seen in the motivation of the company's managers behind the development of the productivity tools. Managers sought to "maintain [the company's] profitability ratio and beat the competition [by] increasing productivity [through] decreasing the length of systems development, and reducing the number of consultants required on each project" [14, p.412]. Under this light, social arrangements are seen as inherently stable and orderly, while the functionalities of technological artifacts are treated as independent of human actions and interactions.

This approach can be useful when attempting to study the normative power of organizations' structural characteristics and technologies' features based on the choice of individuals and teams, and to generalize results to various industries and contexts. However, this view of both technology and organization provides limited insights on the interaction between the social aspects of organizations and the functions performed by technological artifacts [14]. Thus, the unintended consequences of technology and organization during everyday practice within different contexts are ignored. For example, Poole and DeSanctis [16] have found that the appropriation of technology inside work settings can be *faithful* or *ironic*. The former behaviour refers to the use of technology as intended by its initial design, while the latter reflects the disaccordance or the inconsistency of the technology-in-use and the designers' intentions. Other authors discuss the varying outcomes of the same technology within different organizations [17] and the contradicting conceptualizations of the same technology by different work groups [18].

3.2 Second Technology and Organization View: 'Embedding' Social Structures

The second approach emphasizes the malleable nature of both technology and organization. It does that by acknowledging that technology can have different consequences in various contexts [19] and by minimizing the rational calculations of efficiency in organization [10]. In terms of technology, such a view attempts to bring into light the motivations, objectives and interests of key players as the major elements that shape technology inside organizations. The underlying principle is that technology inside work settings can be structured so as to depict the specific organizational characteristics, and that the social structures of the organization can be faithfully modeled within technology. In terms of organization, the institution is treated as an entity proceeding by trial and error, gradually building its own experience from the interaction with its external environment. Coordination and function vary according to particular contexts and relations among different work-groups [10]. According to such view, technology can embed organizational arrangements. This is because this particular approach promotes organizational entities like roles, relationships and work routines as the main drivers of coordination on the one hand and the shaping of technical infrastructures in conformity to those organizational arrangements on the other.

In the example, the company's developers perhaps influenced by the above view of technology – embedding social structures and repertoires – set out to accurately model the institutional properties and social structures by "articulat[ing] and rationaliz[ing] the […] procedures that […] consultants utilized daily in their work" [14, p.414].

The shortcoming of such an approach, which provides primacy to the agency of technology and actors in the organization, is two-fold. First, it moves the attention away from the changes that the new technology introduces in every-day work practices. Second, the assumption that the way '*things get done*' within every-day work practices can be clearly described and explicitly depicted to be embedded in technological artifacts is often faulty. In the example, the tools were built to work "in a standardized, structured […] manner, which leaves little discretion in the hands of individual consultants" [14, p.415].

Both of the above views make the assumption that technology and organization have primarily either functionalistic or agentic characteristics. Managers are inclined to functionalist views of organization and technology: the tools are going to be used '*in the right way*' so as to have controllable impact on the organization, that is, to advance the productivity of knowledge workers. They largely ignore the reciprocal changes between technology and social structures during the use of the tools. Developers on the other hand, tend to provide primacy to the agency power of technology and organization, and think that the tools are going to be used '*as expected*', since they were designed to embed the roles and relationships of individuals and groups inside the organization. As a result, the emerging changes on functional requirements are ignored. Functional changes are often considered as an unpleasant thing, a fault in gathering initial requirements or as a threat to the budget and delivery time of the project.

In the example case study, the productivity tools were thought to meet the motivations, objectives and intentions of managers by evoking manipulable and intentional consequences. Managers expected that "the tools [would give] an opportunity to push standardization further, and more importantly, to enforce it in practice, which had not been feasible before" [14, p.414] and also that "tools [would] contribute to the company's structure because the knowledge embedded in them directs the manner in which problems are interpreted and work is conducted" [14, p.417]. Driven by the above expectations developers assumed that there is one way of interacting with the tools, and as a result "the process of systems development is rational, sequential, and unambiguous, [and] consultants should interact passively with tools" [14, p.415].

The views described above treat technology either as an abstract artifact with deterministic impacts on organizational aspects, either as a socially constructed object with stable functionalities, neglecting modifications by changing needs while in use. Similarly, both views treat organizational orderings either as controllable, neglecting the social dynamics in everyday work, or as relationally defined, not taking into account the material aspects of organizing. This view is justified in the case study. The tools were considered to "discourage reflectiveness" [14, p.418]; nevertheless some actors deviated from the norm and created unexpected situations. So, some "consultants choose not to use the tools in the authorized manner or choose to bypass the tools altogether" [14, p.418]. Others "manipulated their access to the underlying computer system in a way that allowed them to surreptitiously bypass the tools" [14, p.419]. Finally there were consultants that "would often circumvent the tools in order to get on with the work they wanted to do" [14, p.419].

It becomes evident that by following these views technology and organizing is regarded as a project with finite end, developed in an optimal way for the given time and context but in danger of becoming obsolete due to continuous change. These approaches prevent us from exploring a view in which technology shapes social activity and roles inside work settings, as well as it is continuously shaped by the complex social structures and interactions of human actors during use. Such a view is presented next.

3.3 Technology and Organization: A Relationship Enacted in Practice

With this view, human (social) and non-human (technical) structures are understood as both shaping and being shaped by each other within an on-going activity stream. Neither social nor technical arrangements can be seen as entities independent of or superior to one another. Organizations and technologies are not perceived in this paper as being embedded in objects or routines (i.e. objectively existing 'out there') nor as inscribed in human brain or communities (i.e. subjectively existing ‘in here’). Through a reciprocal constitution of organizational orderings and technological infrastructures in-use, the social and the technical are approached as temporal entities with provisional status, bridging the gap between objectivity and subjectivity. They are seen as ongoing social accomplishments, constituted and reconstituted in everyday practice [3]. The social realm is defined through its material instantiations, and accordingly, technological artifacts are conceived as only temporarily stable objects of human knowledgability. We view the enactment of technology and organization as an ongoing accomplishment that is always in the making [14] dependent on human and material agency which cannot be fully planned or prescribed.

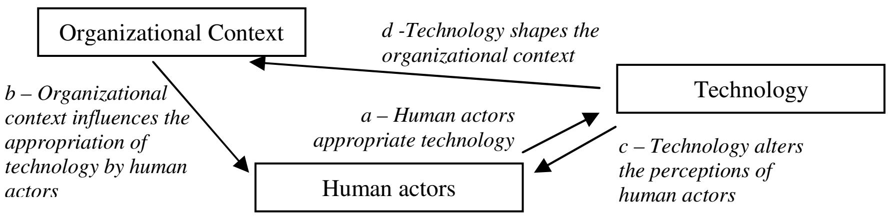


Fig. 1. Dialecting cycle of interactions between human actors, technology and institutional properties (adopted from [14])

Given this reflexive relationship, technology is expected to influence and be influenced by the institutional properties and its appropriation by human actors. This is evident by the outcome of the case study, where “consultants managed to convince their project managers that tools were unduly restrictive, and a few technical consultants were assigned to modify the functionality of the tools” [14, p.419]. A reflexive relationship is based upon a “dialectic cycle of interactions” [14, p.420] between human actors, technology and institutional properties that persists over time, depicted in Fig. 1. During their work consultants appropriate the tools (arrow-a) in a way that is influenced by their socialisation within the institutional context (arrow-b). While using the tools their work reality is facilitated or constrained by the norms and assumptions embedded in the tools (arrow-c), since the tools typically model and reaffirm the institutional characteristics (arrow-d). The cycle is repeated when any deviation to the appropriation of tools by consultants (arrow-a) transforms the company's institutional properties and triggers a change in management strategy, so that managers may authorize the developers to modify the tools (arrow-b). This would temporarily change the form or functioning of the tools (arrow-a), but once deployed, tools would again become institutionalized and support the work of the functional consultants (arrow-c), while also reproducing the company’s institutional system (arrow-d).

To obtain a unified and aligned view of organization and technology, one that will allow us to be flexible, responsive and adaptive to emerging perspectives, is a challenging endeavor. This alignment is further impeded by the complex dynamics of the social and material contexts, which are mostly governed by indeterminacy in the goals and requirements. Given this wicked profile of the above outlined situation, we discuss next design thinking and its core principles in our attempt to reflect and account for both social and material and embrace the respective communities.

4 A Design Perspective of Organization and Technology

4.1 Defining Design

While the creation of Design concept can be traced back to the 1920s within objectivity and rationality influences, during the 1960 political and economical challenges radically changed the perceptions of design. The idea of design was promoted by Simon [20] in his book *The Sciences of the Artificial* (1969), who saw design mainly as a problem-solving paradigm differentiated from natural sciences – which were perceived to be mainly concerned with how things are – in that design is focused on how things ought to be. Literature offers numerous references in defining the epistemology of design and the debate on the clarification of the relationship between design and science is still prevailing. Much work on design stems from the engineering domain and more specifically, from architecture. Indeed, Lawson [21] commented that engineers tend to follow a solution-focused strategy towards the desired outcome, while natural scientists follow a problem-based strategy, focusing on first discovering the fundamental underlying rules that would allow the solution. Yet, design is a much broader concept which transcends the limits of a specific domain as noted by Simon, who proposed that the *design activity* is endemic in all professions, since “anyone designs who devises courses of actions aimed at changing existing situations into preferred ones” [20, p. 111]. Simon’s initial view of design was influenced by positivism and could thus be considered ‘technical’ in the sense of an approach aiming to solve well defined problems. Schon [22] on the other hand in his work *The Reflective Practitioner* proposed design as a reflective process which can adequately address messy, problematic situations of instability, uncertainty and value conflict. Such practice encompasses the intuitive processes of a “reflective conversation with the situation” [22, p.76], where the perception of the initial situation evolves and changes during the design process with not always intended consequences. The reciprocal and iterative process of design is reflected in the latest edition of Simon’s book, which included a section with the seemingly paradox heading “Designing without final goals” [23, p. 162], one that could be viewed as an alliance to Schon’s ideas.

The perception of design as a process as well as an outcome within an ever-changing environment is evident in literature with references to the constant (re)creation of goals and the (re)definition of task constraints [24], the changing nature of the ‘problem’ at hand and the difficulty of remaining open to emerging perspectives [25], the mutually constituted relationship between the entity being designed and the context within it is being designed [26] among others. Design has a dual meaning and is simultaneously the process of developing an object, as well as the outcome of this process. As such, design focused the attention equally to both the design outcome and the value of the reflective journey towards realizing that particular result. Both the process and the outcomes of design can be conceived as “objects of knowledge” which “[…] are characteristically open, question-generating and complex[;] [t]hey are processes and projections rather than definitive things” [3, p.190]. Thus, design can be employed to address not necessarily ‘problematic’ occurrences but rather continuously unfolding situations in which there is a range of goals, requirements and constraints. These temporal structures can bind experts to produce an outcome adhering to the notion of *satisficing* [20, 23], where the aim is to meet criteria for adequacy, rather than pursue an optimal solution. As apparent from the above discussion, our perception of design is that of a reflective, hermeneutic process [27] allowing us to address unique, indeterminate situations within an ever-evolving environment.

4.2 Design Thinking Principles

The discussion above accommodates us to define the following principles governing design thinking:

- Design addresses ill-defined, ambiguous and unfolding situations as it is not set out to literally solve a problem. The purpose is to meet a range of goals, requirements and constraints with a *satisficing* attitude rather than pursuing optimal solutions.
- A synthetic rather than an analytic strategy is followed during the design process. Designers avoid to begin with excessive data gathering in order to fully decompose the perceived phenomenon. Thus, they are able to think beyond the frame of the current situation and recognize new ways for achieving the desired outcomes [28].
- Design is a continuous activity that signifies the value of incompleteness. Outcomes of a design process are only partially and temporally stable and could be considered merely as a pause in an ever-evolving phenomenon.
- The design scope transcends the production of the desired outcome and embodies also the advancement of the understanding of the phenomenon within which is being developed and whereby the set of goals and requirements are continuously transformed and enriched.
- Design encompasses both the notions of a process as well as an outcome. Thus, design is conceived as being reciprocal and having reflexive relationship with the context and the result.

We suggest that these principles could advance our understanding and offer a suitable framework of mind towards an integrated and co-evolutionary ontology of organization and technology. Furthermore, design thinking can be used to guide our reflections of organizational arrangements and technology-in-use as unfolding social objects characterized by a lack of completeness of being. Organizations and technologies can be viewed as intertwined objects of design activity with the capacity to unfold and extend indefinitely: they are materially defined, yet they continuously acquire new properties and alter the ones previously possessed [3]. Organizations and technologies through their on-going co-evolution constitute representations of ill-defined situations – never fully attained – forming a wicked nexus with material instantiations of for-ever unfolding arrangements [3].

4.3 Organizations and Technologies through Design Lens

We argue that organizations and technologies can be conceptualized as 'epistemic objects' that "frequently exist simultaneously in a variety of forms. They have multiple instantiations, […] and other representations to material realizations" [3, p.191]. Given this wicked nexus of the above outlined condition, we suggest that design thinking could guide our reflections. A design thinking approach is suitable for addressing ill-defined and wicked situations [29], meaning unique occurrences that cannot have a definitive formulation and which are characterized by unstable requirements. Complex and reflexive interactions between the context and the solution are observed, while the ramifications to the overall context can be confusing. Identifying a finite and true solutions within these situations is impossible, since there can be multiple solutions, each depending critically on the human cognitive (creativity, intellectual perspective) and social (teamwork, participation) abilities [30].

Being within an intertwined and co-evolutionary ontology of organizational orderings and technological artifacts in-use can be perceived as a reflective and hermeneutic process adhering to the design thinking principles. Organizational theorists, computer scientists, technologists, managers, IT staff and knowledge-workers are all active participants to this dynamic process, which is conducted in a collaborative fashion, intentionally involving and affecting everybody's expectations, agendas and views [31]. As outcome of this process we bring to fore a renewed view that positions technology as integral part of organizational structures and our daily enactments within work settings; one that offers an alternative perspective where organizational norms and technological products can remain modifiable by their interactions with human actors throughout their existence. Incomplete objects of organizational arrangements and technology-in-use pose further questions, and only by considering these notions as incomplete we can move our understandings forward. Such a view is achieved through iterative design circles where the sense of permanent structure is transcended and unexpected consequences are welcome. The diversity of the various social perspectives is considered as source of valuable feedback that can shape new potentials and advance not only the ontological stance of the artifact but also the ever deeper interpretation of the phenomenon.

5 Concluding Remarks

We have argued that the role of technology within organizations is not well explained by existing views. We have also explained that organizational structures and technical infrastructures have a mutually constitutive relationship. This is because it is the interaction between people and technology within institutions that continuously creates both social orderings as well technological artifacts. Therefore, organizational orderings and technological artifacts in-use are infinitely unfolding objects creating a nexus of ill-defined circumstances. This is why we have proposed design thinking as a suitable framework of mind addressing the 'wicked problem' of organizational and technological mutual connections. The design view of these two entities guides us to treat them as incomplete, in flux, materially instantiated social accomplishments.

This is a theoretical paper without primary data and empirical grounding, which consists an important limitation of our study. We intend to further explore the relationship between technology and organizations and pursue filed studies in specific work settings, investigating the role of technology in organizational work practices and the consequences of social structures to technology appropriation.

Acknowledgments

Part of this work was funded by the research project OrganiK, EC 7th Framework Programme, Research for the Benefit of SMEs, under Grant Agreement 222225.

References

1. Orlikowski, W.J.: Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations. Organization Science 11(4), 404–428 (2000)
2. Orlikowski, W.J., Iacono, C.S.: The Truth Is Not Out There: An Enacted View of the "Digital Economy". In: Brynjolfsson, E., Kahin, B. (eds.) Understanding the Digital Economy: Data, Tools, and Research. MIT Press, Cambridge (2000)
3. Knorr Cetina, K.: Objectual practice. In: Schatzki, R.T., Knorr Cetina, K., von Savigny, E. (eds.) The Practice Turn in Contemporary Theory. Routledge, London (2001)
4. Giddens, A.: The Constitution of Society: Outline of the Theory of Structure. University of California Press. Berkeley (1984)
5. Bibikas, D., Kourtesis, D., Paraskakis, I., Bernardi, A., Sauermann, L., Apostolou, D., Mentzas, G., Vasconcelos, A.C.: Organisational Knowledge Management Systems in the Era of Enterprise 2.0: The case of Organic. In: Flejter, D., Grzonkowski, S., Kaczmarek, T., Kowalkiewicz, M., Nagle, T., Parkes, J. (eds.) Proceedings of the 2nd Workshop on Social Aspects of the Web (SAW 2008), 11th International Conference on Business Information Systems, BIS 2008 (2008)
6. Bibikas, D., Kourtesis, D., Paraskakis, I., Bernardi, A., Sauermann, L., Apostolou, D., Mentzas, G., Vasconcelos, A.C.: A Sociotechnical Approach to Knowledge Management in the Era of Enterprise 2.0: the Case of Organic. Scalable Computing: Practice and Experience Scientific International Journal for Parallel and Distributed Computing 9(4), 315–327 (2008)
7. Keating, C.B., Fernandez, A.A., Jacobs, D.A., Kauffmann, P.: A methodology for analysis of complex sociotechnical processes. Business Process Management Journal 7(1), 33–49 (2001)
8. Fox, W.M.: Sociotechnical system principles and guidelines: past and present. Journal of Applied Behavioral Science 31(1), 95–105 (1995)
9. Trist, E.L., Bamforth, K.W.: Some Social Psychological Consequences of the Longwall Method of Coal Getting. Human Relations (4), 3–38 (1951)
10. Gherardi, S.: Organizational Knowledge: The Texture of Workplace Learning. Blackwell Publishing Ltd., Oxford (2006)
11. Horton, K., Davenport, E., Wood-Harper, T.: Exploring sociotechnical interaction with Rob Kling: five "big" ideas. Information Technology & People 18(1), 50–67 (2005)
12. Bijker, W.E., Pinch, P.: The Social Construction of Facts and Artifacts. In: Bijker, W.E., Hughes, T., Pinch, T. (eds.) The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology. MIT Press, Cambridge (1987)
13. Orlikowski, W.J., Bariey, S.R.: Technology and Institutions: What can research on Information Technology and research on Organizations learn from each other? MIS Quarterly 25(2), 145–165 (2001)
14. Orlikowski, W.J.: The Duality of Technology: Rethinking the Concept of Technology in Organizations. Organization Science 3(3), 398–427 (1992)
15. Attewell, P., Rule, J.: Computing and Organizations: What We Know and What We Don't Know. Communications of the ACM 27(12), 1184–1191 (1984)

16. Poole, M.S., DeSanctis, G.: Understanding the Use of Group Decision Support Systems: The Theory of Adaptive Structuration. In: Fulk, J., Steinfield., C. (eds.) Organizations and Communication Technology, pp. 173–193. Sage Publications, Newbury Park (1991)
17. Barley, S.R.: Technology as an occasion for structuring: Evidence from observation of CT scanners and the social order of radiology departments. Administrative Science Quarterly 31, 78–108 (1986)
18. Iacono, S., Kling, R.: Computerization Movements: The Rise of the Internet and Distant Forms of Work. In: Yates, J., Van Maanen, J. (eds.) IT and Organizational Transformation: History, Rhetoric, and Practice. Sage Publications, Newbury Park (1999)
19. Davis, L.E., Taylor, J.C.: Technology, Organization and Job Structure. In: Dubin, R. (ed.) Handbook of Work, Organization, and Society, pp. 379–419. Rand McNally, Chicago (1986)
20. Simon, H.A.: Sciences of the Artificial, 1st edn. MIT Press, Cambridge (1969)
21. Lawson, B.: How Designers Think. Architectural Press, London (1980)
22. Schon, D.A.: The Reflective Practitioner: How Professionals Think In Action, 1st edn. Basic Books, New York (1984)
23. Simon, H.A.: Sciences of the Artificial, 3rd edn. MIT Press, Cambridge (1996)
24. Akin, O.: An Exploration of the Design Process. Design Methods and Theories 13(3/4), 115–119 (1979)
25. Jones, J.C.: Design Methods. Wiley, UK (1970)
26. Garud, R., Sanjay, J., Tuertscher, P.: Incomplete by Design and Designing for Incompleteness. Organization Studies 29(03), 351–371 (2008)
27. Purao, S.: Design Research in the Technology of Information Systems: Truth or Dare. GSU Department of CIS. Working Paper, Atlanta (2002)
28. Nadler, G., Hibino, S.: Breakthrough Thinking. Prima, Rocklin (1990)
29. Rittel, H.J., Webber, M.M.: Planning Problems are Wicked Problems. Developments in Design Methodology, New York (1984)
30. Buchanan, R.: Wicked Problems in Design Thinking. Design Issues 8(2), 5–21 (1992)
31. Jelinek, M., Romme, G., Boland, R.J.: Organisation Studies as a Science for Design: Creating Collaborative Artefacts. Organisation Studies 29, 317–329 (2008)

Interaction between Heterogeneous Autonomous Systems: Principles and Practice

Mats-Åke Hugoson[1], Ulf Larsson[1], and Vladimir Tarasov[2]

[1] Jönköping International Business School, Jönköping University, Jönköping, Sweden
{mats-ake.hugoson,ulf.larsson}@ihh.hj.se
[2] School of Engineering, Jönköping University, Jönköping, Sweden
vladimir.tarasov@jth.hj.se

Abstract. Today, many enterprises use information systems to support their activities but as a rule these systems are heterogeneous and incompatible. When enterprises need to cooperate with each other, they have to overcome this heterogeneity and establish interaction between their information systems. Based on a case study from international business development, this paper analyzes principles and conditions for creating interaction between autonomous systems. Furthermore, an approach to interaction between heterogeneous information systems in the case study is presented, which is based on a service-oriented architecture and Web services. In this context, a business process model describes how the distributed autonomous systems should be used by companies in different countries to support establishment of business relationships. These systems have different data models, and use different industrial classifications. Our main contribution to research in the field of interaction between autonomous systems are (1) a proposal for principles to be taken into account when federating autonomous systems and (2) an experience report when putting these principles into practice in the case study presented.

Keywords: Autonomous systems, Information systems interaction, Service-oriented architecture, Structures of heterogeneous systems, Web Services.

1 Introduction

Nowadays IT-systems supporting business operations or the sharing of information are inherently heterogeneous and geographically dispersed. Many such systems emerge as a result of cooperation between organizations in one country or even in several countries. Bridging information systems that supports businesses of different organizations imposes a challenge because of diverse underlying technologies, interfacing techniques and data structures used. To solve this problem, service-oriented architectures on the basis of Web Services are often employed, being a modern approach for computerized interaction between heterogeneous, complex information systems.

Systems Architecture is a well established, central area within Informatics. Results from research in Integrated Information Systems have been presented gradually for many decades, spanning from early works by [3] and [1], via principles in the field of

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 119–130, 2010.
© Springer-Verlag Berlin Heidelberg 2010

enterprise architecture (e.g. [4]), to modern approaches for enterprise application integration as proposed by [2] and others. These principles and approaches are dealing with design of socio-technical systems, basically for one single complex system, e.g. an organization or an enterprise. These systems can be considered as autonomous systems as they are independent of control from an external entity. Possibly, the same principles can be used for development of a structure of interacting information systems of the same kind, when all the systems in the structure can be centrally controlled, and dependencies between different systems in the structure are accepted (or aimed at). However, when the systems in the structure are heterogeneous and there is a demand for autonomy between systems, different principles must be applied. This is a research area which is not fully matured. Some results are published (e.g. [5], [6]) but the basic differences between demands on interaction in the two alternatives are seldom analyzed.

In this paper, work in the field of interaction between autonomous systems is motivated by a case from international business development presented in section 2. Principles for computerized interaction between heterogeneous and fully autonomous systems are analyzed in section 3. In section 4 of the paper a service-oriented architecture based on Web services for systems interaction is presented and applied in the case study. Furthermore, applicability of the presented principles is discussed.

As the main contribution to research in the field we offer (1) a proposal for principles to be taken into account when federating autonomous systems and (2) an initial empirical study of putting these principles into practice.

2 The Case

This part of the paper is closely linked to a project sponsored by SPIDER (The Swedish Program for Information and Communication Technology in Developing Regions), a development agency promoting the development of ICT infrastructure, human capacity and relevant ICT content. The main intention of this project (runtime: September 2005 - December 2006) was to support establishment of business relationships between enterprises in Sweden and Vietnam by providing ICT tools for finding appropriate partners and utilizing diaspora (i.e. Vietnamese living outside their homeland) competence.

2.1 Business Process Model for the Use of Information Systems

Fig.1 presents the overview of the business process model, which was created based on the SPIDER project case. The following main actors can be identified in this process.

In Vietnam:

- Companies interested in establishing a new relationship with enterprises abroad. The main purpose is probably export of products and/or services, but it could also include import from other countries, e.g. machinery for the production line raw material or know how.
- VCCI – The Vietnamese Chamber of Commerce and Industry. VCCI has developed the database VNemart (http://www.vnemart.com.vn/). It can be

assumed that companies interested in a new business relationship abroad have registered in VNemart and have entered both general company information and information about products and offers (so called trade leads). The actual number of Vietnamese companies registered in VNemart was about 6000. VNemart also contains information on companies abroad. The foreign country with the largest number of companies registered was China. The number of European companies was low.

- Swedish (and other European) individuals and organizations in Vietnam that may have possibility and interest to support development of the business relations between Vietnamese companies and companies in their own country.

In Sweden:

- Companies in Sweden (and in other European countries), interested in finding partners in Vietnam.
- Chamber of Commerce. Eight Swedish Chambers of Commerce has developed and owns the database Chamber Trade (http://www.chambertrade.com/). In this database a high number of both Swedish and European companies have registered. The web based form (in English) for entering of company and actual product information is accessible worldwide, but very few Vietnamese companies have registered.
- Vietnamese individuals and organizations in Sweden (and other European countries) that possibly can support development of business relations with Vietnamese companies. It is of utmost interest to enroll competence within the Vietnamese diaspora in this process.

It must be assumed that a Vietnamese company, if using an IT-based support to search for a partner abroad, will turn to VNemart and not to a foreign database, such as Chamber Trade. Likewise, a Swedish (or European) company will probably use Chamber Trade (or other European business database) to search for a partner in Vietnam in the most cases as VNemart is not known in Europe. This ends up with low probabilities for companies in Vietnam and Sweden (or other European countries) of finding each other through the IT-based support.

Focusing on VNemart and Chamber Trade, the current situation is thus a kind of dead-lock as there is no direct interaction between the two databases.

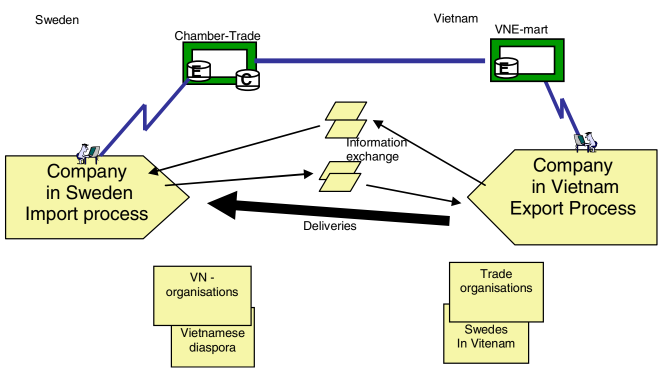


Fig. 1. Overview of the business process to establish business relationship

2.2 System Landscape

In this section we describe the key concepts of the model in Figure 1.

Enterprise databases (E-DB). Information on companies and products (E in Fig. 1) is currently stored in both Chamber Trade and VNemart with different classifications and different database structure. A proposal is to establish interaction between these two databases through the concept of Web Services. The approach will allow the user of either VNemart or Chamber Trade to transparently search both databases simultaneously and get aggregated results. This will not interfere with the internal structure of the databases, and duplication of data will be avoided. Several software components, which provide support for distributed search and translation between the industrial classifications, would need to be installed on both VNemart and Chamber Trade, see Fig 3.

The Competence Database (C-DB). Once the possible partner is found in the E-DB there is a need to find information on possible competences to support establishment of business relationships with the partner. It is proposed that a dedicated Competence database is developed separated from the E-DB bases. The C-DB should support development of the Business process between companies in Sweden and companies in Vietnam, and be directly connected to Chamber Trade (C in Fig. 1). The Competence database must be made available for users of VNemart.

Both contents and interface for C-DB are directly dependent on the different steps in the business processes. These processes must be analyzed in more detail.

Analysis of the business processes. In Fig. 1 an overview of the business processes is shown. Both for the company in Sweden and for the company in Vietnam, the first step is to find possible partners, supported by use of E-DB. Establishment of a business relationship then implies exchange of many different documents, and ends up in deliveries of products or services. The task for the Competence database C-DB is to offer support during establishment of the business process Based on information from C-DB, additional competence from individuals or organizations can be involved in the process (not shown in Fig. 1.)

In the SPIDER project interviews with companies in Sweden and in Vietnam were carried out, in order to specify different activities in the business processes and to indicate what can be supported by the Competence database.

3 Principles

3.1 Heterogeneous Autonomous Systems

Heterogeneity within a system of systems generally means that the systems have different origin. From a design point of view this means that the inner structure differs, and it also means that the inner structure in different systems cannot be centrally controlled. Different systems in the structure may be developed at different points of time, in different projects, using different methodology and technology.

The concept of Autonomy is very much a question of independence between systems in the structure. Autonomy means primarily that every system can support its users, independent of other systems. It must be possible to develop an autonomous system (heterogeneous or not) and implement it independently of other systems. *Changes in the inner structure of one system, should not affect other systems in the structure*.

Another dimension in autonomy is time independence. *An autonomous system must be capable to function and support its users, even if other systems in the structure are not up and running for a period of time*. A consequence of autonomy is that disturbances, or even break downs, can be isolated to the system that malfunctions. Concepts from the field of distributed systems related to time independency are robustness or failure transparency (see [7]). However, autonomy does not mean that relations to other systems in the structure can be neglected. Demands on interaction (with maintained autonomy) must be considered, that's what really creates the interoperability problem.

An example related to the case study may clarify this as shown in Fig. 2. The database and information systems **Chamber Trade** and **VNemart** were developed completely independent of each other. As a consequence they have different inner structure, even if they contain mainly the same type of information on companies and products. The third system in Fig. 2, **PDM** (product data management) is supposed to be the system that supports product development in a company, which is interested to use the service from Chamber Trade. The company seeks for a partner in Vietnam to produce some components. All the three systems are evidently heterogeneous.

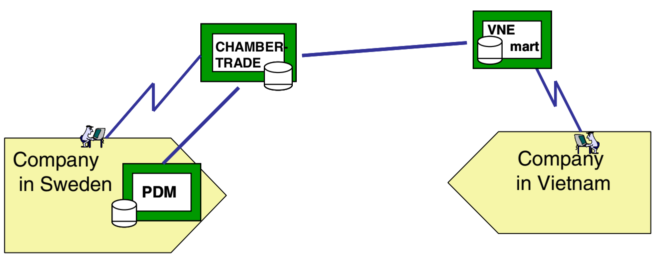


Fig. 2. A structure of heterogeneous systems, with high degree of autonomy

The demand for computerized interaction between the systems could be analyzed in the following way.

- Between Chamber Trade and VNemart.

Information on Vietnamese companies, their products and their production capacity, is probably entered only in VNemart, while Chamber Trade contains mainly information on European companies and their products. The European company naturally turns to Chamber Trade (it is perhaps not even known that VNemart exists). Some kind of interaction between Chamber Trade and VNemart will improve the possibility for the company to find a partner in Vietnam.

- Between Chamber Trade and PDM.

As the specification of products and components represents a considerable amount of data it could be of interest to establish a computerized transfer from PDM into

Chamber Trade. Chamber Trade has a well defined data entrance for this. In the other direction information on the Vietnamese partner could be captured and via Chamber Trade to be entered into PDM.

Autonomy in this structure of systems means that every system can change its inner structure, or even be replaced, without any change in any other system, as long as the specified interaction is unchanged. Time independence means that each system can work even if other systems are down. For PDM this can be established if the transfer of information can be buffered. Time independence between Chamber Trade and VNemart is more crucial, as each system cannot fully support its users if the other system is not up and running. Robustness can however be maintained also for these two systems.

How can interoperability then be established with maintained autonomy?

3.2 Inner and Outer Structure of Systems, Different Demands

Each system in a system of systems has an ***Inner Structure*** as a result of design work for that system, whichever architectural principles are used. In this paper the interaction *between* the different systems in the structure is referred to as the ***Outer Structure***.

Demands on the inner and outer structure are different in a system of heterogeneous autonomous systems. For the inner structure fast interaction and tight integration is needed in order to support the user of the system, perhaps in a real time situation. For the outer structure the demand is to maintain autonomy, which means that interaction must be specified and implemented with a high degree of independence. *The principle is to specify the outer structure based on demands (and restrictions) for interoperation in the enterprise where the systems are operating, not by the data storage in the different systems.*

Another difference between inner and outer structure is the long term perspective. Every system with its inner structure has a lifecycle, which means that the inner structure will be replaced (or deleted) sooner or later. The outer structure is continuous as long as the system of systems exists. The outer structure may be changed, perhaps extended when new systems are inserted, but it survives particular systems.

3.3 Responsibility Aspects

Responsibility for the inner structure of a system can always be pointed out. The designer has the right to decide on the inner structure.

For the outer structure the responsibility is not always so clear. In a company or organization perhaps there is a possibility for the CIO to decide on the outer structure top down, and direct how the outer structure should affect different systems, but when the systems in the structure are developed and operated by different companies or organizations, it is difficult to point out responsibility for the outer structure.

A bottom up approach can be used for this type of Inter Organizational Structures (IOS), where interaction demands are picked up for different systems system in the structure through bilateral discussions between owners of different systems in the structure. Anyhow, *the principle is to define the outer structure through interaction agreements; it cannot be specified from one system, if autonomy is to be maintained.*

3.4 The Principle of Independence between Inner and Outer Structure

The analysis of the inner and outer structure so far leads to the principle of independence: *The outer structure must be described independent of the inner structure if autonomy is to be maintained*. This principle, which is well-established in software and systems engineering, has to be applied for all parts of the socio-technical system under consideration in order to create/maintain autonomy between the parts.

The desired consequence is then evident: Changes in the inner structure will not affect the outer structure, which was the basic condition for autonomy. As long as the outer structure is unchanged, every system can be changed or even replaced. The only condition is that the changed or replaced system can perform the specified interaction with other systems in the structure.

There will however be dependencies the other way round: Changes in the outer structure may affect the inner structure for some systems, but only for systems that have changed relations to other systems.

The next conclusion is then that it would be preferable to specify the outer structure before the design of the inner structure. This is possible when a new system is inserted in the system of systems, and when a system is replaced, but in many cases heterogeneous and autonomous systems are developed first and interaction is established later on. This means that development of interoperability between heterogeneous and autonomous systems generally cause changes in existing systems. The real crucial point is then to apply the principle of independence between inner and outer structure, otherwise long term autonomy can not be maintained.

In software and systems engineering, the subject of inner and outer structure is addressed in architecture development. Different architecture styles [8] or reference architectures [9] were proposed in order to identify principal elements of larger systems and facilitate the specification of interfaces between sub-systems. Furthermore, architecture description languages (ADL) [10] contribute approaches how to describe components, which can be considered sub-systems, and connectors for interfacing between components. These approaches can be used for defining both inner and outer structure. Howerver, ADL do not contribute to defining correspondences between inner and outer structure on information model, process and organization level.

3.5 Implementation of the Outer Structure

There is a wide range of alternatives for implementation of the outer structure, from Electronic Data Interchange (EDI) solutions, sometimes with dedicated nets, to web based technology. *Every solution calls for some kind of infrastructural support, and this means of course that autonomy is to some degree limited*.

The way to implement interaction can, however, cause some dependence by itself. This can be a conscious choice in order to use an efficient solution. In part 4 the case is presented in more detail, and it shows that changes in the inner structure of a system may call for minor changes in other systems. The principle of independence between inner and outer structure can be a tool to understand the consequences of a certain solution in a long term perspective.

How will implementation of an independent outer structure affect the possibility to use heterogeneous systems? Sometimes there is a demand to develop the inner structure fully independent of other systems in the structure. If autonomy is achieved, this can be done. In other cases it is of interest to unify the inner structure in some systems in the structure, for instance through reuse of systems or components. Also in this case autonomy in the outer structure brings advantages. The principle of an independent outer structure promotes reuse.

4 Putting Principles into Practice

The principles for interaction between heterogeneous autonomous systems presented in the previous chapter will be used to assess the case presented briefly in chapter two. In this chapter, a detailed description of the interaction between the systems of the case will be presented, and in the next chapter the analysis and assessment according to the principles are done.

4.1 Web Services Architecture for Interacting Information Systems

The web services architecture included a new competence database and a web service connecting together the databases used in the SPIDER project as described in chapter 2. The user interface for each database is provided by its own web application running on the same web server as the database. There was no interconnection between the two web applications and a search was only made in Chamber Trade (or in VNemart respectively). The web service would allow a Chamber Trade or VNemart user to make a distributed search in both databases simultaneously. The databases describe products/services produced by companies using industrial classifications in different ways, so a mapping between the classifications was needed. The competence database is to store information about individuals and organizations that could help in establishing business relationship between two companies. The web service would also allow inclusion of the competence database in a distributed search, either simultaneously with enterprise search or independently.

The proposed architecture contains the following major components. The competence database itself, created and located on the Chamber Trade server. A user interface component to update and search the competence database locally (by users in Sweden), this component is integrated with the existing user interface of the Chamber Trade enterprise database. A search component making it possible for the VNemart web application to remotely retrieve data from the Chamber Trade enterprise and competence databases, either simultaneously or independently, as shown in Fig. 3. A translation component mapping the Chamber Trade industrial classification onto the VNemart classification before sending retrieved data back.

Although not formally part of the web service, a web service consumer component is to be added to the Chamber Trade web application to make a remote search in VNemart when a search in Chamber Trade is initiated, and to integrate data remotely retrieved from VNemart with data locally retrieved from Chamber Trade before presenting them to the Chamber Trade user.

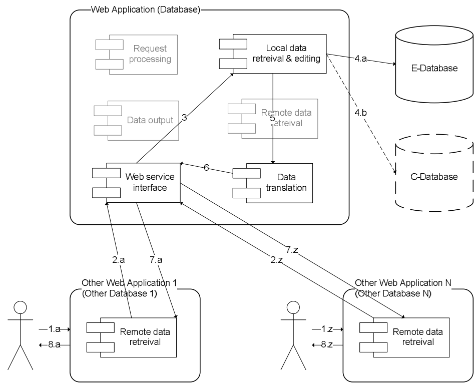


Fig. 3. Remote user searches the database (Web service provider)

4.2 Mapping of Industrial Classifications Used in the Databases

As soon as the mapping between the classifications was needed, we identified the similarities and differences between Chamber Trade and VNemart databases. The review of databases was based on the import or export industry, the kind of operations, the size of the firms and the number of levels used in each category for classification.

VNemart provided two classifications, one was initially used in VNemart website, the other was under development to be launched soon. The following similarities between the existing Chamber Trade and VNemart databases were found. The number of categories used to classify the businesses is the same (12). There are only several similar categories used in both databases. In general, although using 12 categories, each database clusters the businesses differently. Chamber Trade combines two or more categories such as foodstuffs and agricultural products, but VNemart focuses on one typical type of product. The service sector in VNemart is not clearly described and VNemart database totally lacks the categories industrial like machinery or transport.

The new VNemart database was in the developing phase. The number of categories was 28. Some new categories were added. The new database mostly uses three levels to cluster the businesses: category, sub-category, and sub-sub-category. Each category is broken down into very detailed and specific type of products, product tools and equipment.

Chamber Trade and VNemart categorize the businesses mostly based on the outputs (the products and services that organizations offer), and the only principal difference between the two databases is that Chamber Trade uses general terms to describe the businesses, while VNemart tries to break them down as specific as possible. In order to map the two databases, one approach is identifying similar categories of VNemart's categories and then allocating them into Chamber Trade's ones. The mapping can be drawn as presented in Table 1. Another approach could be to use an international classification, like [11], as an intermediate one and to do translation from the Chamber Trade or VNemart classification to the intermediate one and vice versa on each side. This allows adding more databases in the interaction. In this case, the mapping would also look like the one presented in Table 1.

Table 1. Fragment of the Mapping between Chaber Trader and VNemart

CHAMBER TRADE	VNemart
CT0103 Confectionery, Flavourings, Additives, Spices	524 Food & Beverage 611 Candy . . . 1364 Food Ingredients 1365 Chocolate Ingredients 1366 Cultures, Enzymes & Yeast . . .

5 Conclusion

5.1 Validity of the Principles in the Web Service Architecture

The web service architecture that is applied in this case study for interaction between the two systems, Chamber Trade and VNemart, leads to some conclusions related to the basic principles.

It is evident that the systems are heterogeneous, but what are the conditions for maintained autonomy when interaction is established? This question was of particular interest as VNemart was changing the inner structure during 2006. Can this be done without any changes in Chamber Trade?

The answer to the question is that

- if the classification of each database is mapped to a new classification (for instance an international standard, which exists) and
- this standard is used for the web service interaction and
- mapping back to the receiving systems classification is done in the receiving system,

Then each system can be changed without any impact on the other system. The case study shows that this is possible through the applied web service architecture.

If, on the other hand, in each system the web service architecture were based on direct transformation from classification in sending system to the classification in receiving system, then a change in the inner structure in one system would affect the other system, and a change in an extended structure of systems affect all other systems.

The conclusion is that it is not the web service technology as such that creates and maintains autonomy, it is the possibility to describe interaction independent of the inner structure in each system

Time independence is achieved only for updating in this case study. Support to users through interaction between the two systems using the applied web service architecture is possible only when both systems are up and running at the same time.

Analysis of responsibility aspects in this case study shows that the web service architecture can be settled through bilateral agreements. It is then possible that the responsible party for each system develops their software, necessary for interaction. This means that this is not a federal structure; there is no need for any common federation above participating systems. The structure can be extended with new systems (for which other parties are responsible) without any changes in the existing systems, if the defined web service architecture is accepted for additional systems.

5.2 Comparison to Other Autonomy Principles

In the field of federated database systems, autonomy has been discussed intensively. Conrad proposed to distinguish between three different aspects of autonomy [12], which roughly can be summarised as follows:

- Design autonomy concerns the design of database schema and information models for the information systems under consideration. For autonomous systems it has to be possible to make changes in the information model without affecting other systems
- Communication autonomy demands that all systems involved in an interaction decide independently, when interaction with other systems is performed
- Execution autonomy includes that an autonomous system decides, which operations are executed on their local database, and in which order and at what point in time this is done.

Comparison of the above listed aspects of autonomy and the proposed principles from chapter 3 leads to similarities and differences. Conrad's design and execution autonomies are both related to our principle of separation between inner and outer structure. Conrad explicitly focuses only on (federated) database systems and thus does not include general system architecture or organisational aspects. However, separation between design aspects and execution of operations is a valuable contribution to discussing autonomy of socio-technical systems.

Execution and communication autonomy are both related to the proposed principle of time independence. Communication in federated database systems is more complex as communication between two peers, as entering and leaving a federation has to be communicated thoroughly in the federated system.

Responsibility aspects are not explicitly mentioned in Conrad's proposal, but indirectly integrated in communication and execution autonomies, as both include a clearly defined responsibility for inner and outer structure as a prerequisite for achieving these autonomy aspects.

Acknowledgments. We would like to thank Bengt Henoch from Jönköping University, Sweden, Marianne Dott and Maria Karlberg from Chamber Trade Business to Business AB, Sweden, and Thao Nguyen Van and Tuan Bui Duc from Vietnam Chamber of Commerce and Industry (VCCI), Vietnam, for valuable input to the SPIDER project.

References

1. Scheer: Architecture of Integrated Information Systems: Foundations of Enterprise Modelling. Springer, New York (1992)
2. Linthicum: Next generation application integration: From simple information to Web services. Addison-Wesley, Reading (2003)

3. Zachman: A Framework for Information Systems Architecture. IBM Systems Journal 26(3) (1987)
4. Vernadat: Enterprise Modeling and Integration. In: Principles and applications. Chapman & Hall, Boca Raton (1996)
5. Greiner, Lippe, S., Kahl, T., Ziemann, J., Jäkel, F.: A multi-level modeling framework for designing and implementing cross-organisational business processes. In: Manolopoulos, Y., Filipe, J., Constantopoulos, P., Cordeiro, J. (eds.) ICEIS 2006 - Proceedings, Paphos, Cyprus, May 23-27 (2006) ISBN 972-8865-41-4
6. Vernadat: Interoperable Enterprise Systems: Architectures and Methods. In: Dolgui, A., Morel, G., Pereira, C. (eds.) INCOM 2006 Proceedings, St. Etienne, France, May 17-19, IFAC (2006)
7. Tanenbaum, van Steen, M.: Distributed Systems. In: Principles and Paradigms. Prentice Hall, Englewood Cliffs (2002)
8. Shaw, Garlan, D.: Software Architeture – Perspectives on an Emerging Discipline. Prentice-Hall, Englewood Cliffs (1996)
9. Fowler: Patterns of Enterprise Application Architectures. Pearson Education, London (2003)
10. Medvidovic, Taylor, R.: A classification and comparison framework for software architecture description languages. IEEE Transactions on Software Engineering 26(1) (January 2000)
11. United Nations Statistics Division. Central Product Classification Version 1.1, CPC Ver. 1.1 (February 2, 2002), `http://unstats.un.org/unsd/cr/family2.asp?Cl=16`
12. Conrad: Föderierte Datenbanksysteme: Konzepte der Datenintegration. Springer, Heidelberg (1997)

Business/IT Alignment: The Executive Perspective
New-Generation EIS Architecture as a Case Example

Jörg H. Mayer[1] and Bernhard Krönke[2]

[1] University of St.Gallen, Institute of Information Management
Müller-Friedberg-Strasse 8, 9000 St. Gallen, Switzerland
joerg.mayer@unisg.ch
[2] Continental AG, Corporate Controlling
Vahrenwalder Strasse 9, 30165 Hanover, Germany
bernhard.kroenke@conti.de

Abstract. How are companies managed today and what part does state-of-the-art business/IT alignment play? Executive information systems (EIS) should support top management in "managing a company." But many executives complain that EIS bear little relevance to their management task. This article examines the role of business/IT alignment within a new-generation EIS architecture and contributes a business-driven model. A prototype demonstrates an instantiation of this approach. The findings are synthesized to seven recommendations on business/IT alignment that should lead to better EIS architecture.

Keywords: corporate management, IT support for top management: executive information systems (EIS), business/IT alignment in IS architecture.

1 Introduction

Companies must operate in an increasingly dynamic context. Due to their overall responsibility, executives are particularly affected by this situation [1].

Solutions meant to help top management are known as *executive information systems* (EIS). Their IT side (IT in general) is often unable to keep pace with the changes in business. Thus, *business/IT alignment* is a topic of information systems (IS) research gaining more and more importance [2].

Directly (point-to-point) linking business and IT artifacts often fails to meet flexibility and agility[1]. To provide flexible links, researchers often propose using enterprise services, IT capabilities, (logical) applications, and domains in service-oriented architectures (SOA). Virtual decoupling is one recent approach [5]. Services are encapsulated functional capabilities – known in software development as encapsulated functionalities – with published interfaces [6].

[1] Flexibility is the cost-adequate, qualitatively good capability to implement new but anticipated requirements [3]. Agility additionally covers responses to unexpected requirements as well [4]. Hereafter, we use flexibility to cover both senses.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 131–143, 2010.
© Springer-Verlag Berlin Heidelberg 2010

While practitioners recognize the increasing need for flexibility, they evaluate approaches such as virtual decoupling as difficult to understand and little applicable in practice [7]. This article assesses the *potential business/IT alignment has for EIS design from an executive perspective*—using a new-generation EIS architecture as an example.

Research Method

The paper takes a *design research* approach: according to Hevner et al. [8] and March, Smith [9], the outcomes of a construction process can be classified as constructs, models, methods, and instantiations. The to-be-developed architecture of a new-generation EIS that respects business/IT alignment can be categorized as a *model*. It aims at balancing the needs of practitioners without sacrificing scientific rigor. While the focus is on EIS architecture, the findings should contribute to making IS more effective in general.

To develop artifacts, several construction processes have been proposed [8, 9, 10, 11]. The process described by March, Smith [9], which distinguishes between 'build' and 'evaluate' activities, is predominant in the literature [8]. The article follows this approach, while also using findings from a current empirical survey.

- *Identify a need*: The outcome of a recent survey among Financial Times "Europe 500" companies show what business/IT alignment issues currently exist in the field EIS architecture (section 2). A review of the state-of-the-art completes this picture.
- *Build*: section 3 examines what an architecture for new-generation EIS should look like. The key objective is a business-driven model that allows better IT support of executives than the state-of-the-art.
- *Evaluate*: The prototype at an automotive supplier demonstrates the relevance of the new model (section 4) by showing various types of business value business/IT alignment brings.
- *Learn and theorize:* Finally, the key findings are formulated as recommendations to start future research and EIS implementation (section 5).

2 Business/IT Alignment from Executive Perspective

The term "*business/IT alignment*" is used to indicate that a company's IT is designed in terms of its business value [12]. The alternative approach, in which a company's business functions are determined based on IT capabilities, is less common.

The first definition should make clear that IT should significantly contribute to the company's success. One purpose of business/IT alignment is thus to ensure that a company and its individuals have the information needed to fulfill their tasks.

The amount of data available within a company has continually grown in recent years thanks to business intelligence (BI) [13, 14, on state-of-the-art, 15]. This approach combines all system components that either prepare and store operational data used to generate information or perform the functionalities to analyze or present this information. *EIS* represent the "top" of such an architecture [16, 17]. They can be defined as systems explicitly designed to support executives with information tailored to their managing a company.

State-of-the-Art Review

A study by Henderson/Venkatraman [12] defines business/IT alignment as a means to link the company's business strategy, organizational infrastructure, and processes to IT strategy, IT infrastructure, and processes. Looking at the state-of-the-art, using the potential of business/IT alignment has barely played a role in EIS architecture design [18, 19]. One could assume, that financial reporting is relatively stable, given that it is governed by legal regulations and quasi-legal guidelines. However, demand for more operational information shows that such clusters as management accounting, compliance, or program management is subject to *increasing changes* [1].

The performance dashboards described by Eckerson [20] or Few [21] or the "management reporting" in Gleich et al. [22] do not discuss the potential of business/IT alignment during the design process. N*o end-to-end system architecture* could be identified that makes use of business/IT alignment to handle changing processes. Finally, the focus of IS architecture design is not design research, but "*theory building*"—descriptive, empirically validated statements about users and sociotechnical issues [23], no artifacts for a "better" world.

More Operational Responsibility at the Corporate Center

Our survey (appendix) within the Financial Times "Europe 500" report started by examining executives' processes for managing their companies. A first question covered the split between strategic leadership by the executives and operational management at the division level.

During the period of growth from 2003 to 2007, executives devoted most of their time to "pure" strategic leadership. Yet, one-third of the executives said that they currently intervene frequently and extensively in day-to-day business. A further 22 percent even characterized their involvement as very frequent and very extensive. This response shows that executives have significantly expanded *their involvement in operational business*—in parallel to their strategic leadership. Therefore, executives are asking for direct access to more operational data.

Biased IS Objects

A second question observed executives' demand for flexibility. Almost 50 percent of them answered that they are operating in a market environment that is more aggressive and dynamic than ever. As a result, they evaluate *flexible IT* as "high" to "very high." While executives have viewed IT as a "cost pool", the survey showed that flexibility of IT has achieved similar importance to executives as their cost (Fig. 1).

A third question was about the areas in which executives forsee the greatest need for change. In addition to changes in the organizational structure (arithmetic mean of 3.1) and the data model (arithmetic mean of 3.6), the need to respond flexible to changing *management processes* was identified (arithmetic mean of 4.0).

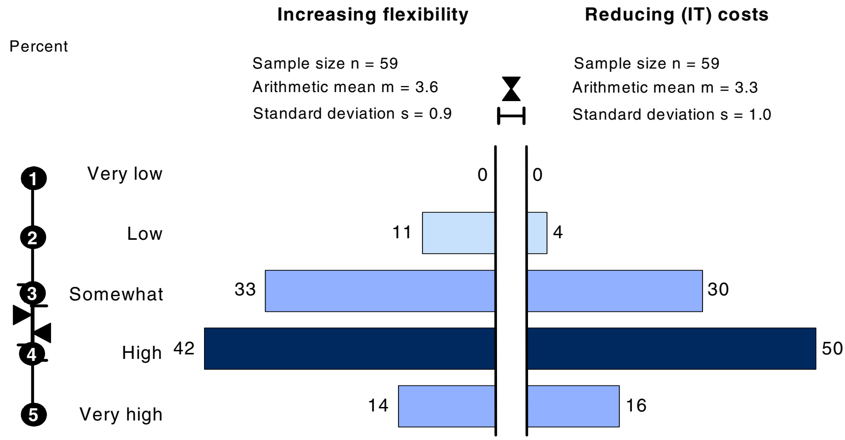


Fig. 1. Executives' perspective of current IS architecture objectives

3 Business/IT Aligned Architecture Design for EIS

Architectures can be defined as "the fundamental organization of a system, embodied in its components, their relationships to each other and the environment, and the principles governing its design and evolution" [24]. So, they are universal models, in the sense that they represent abstractions from specific characteristics. This section describes the "*construction plan*" of new-generation EIS architecture, specifying their most important components. (Sociotechnical) design principles for the artifacts of each layer are not in the focus of the article.

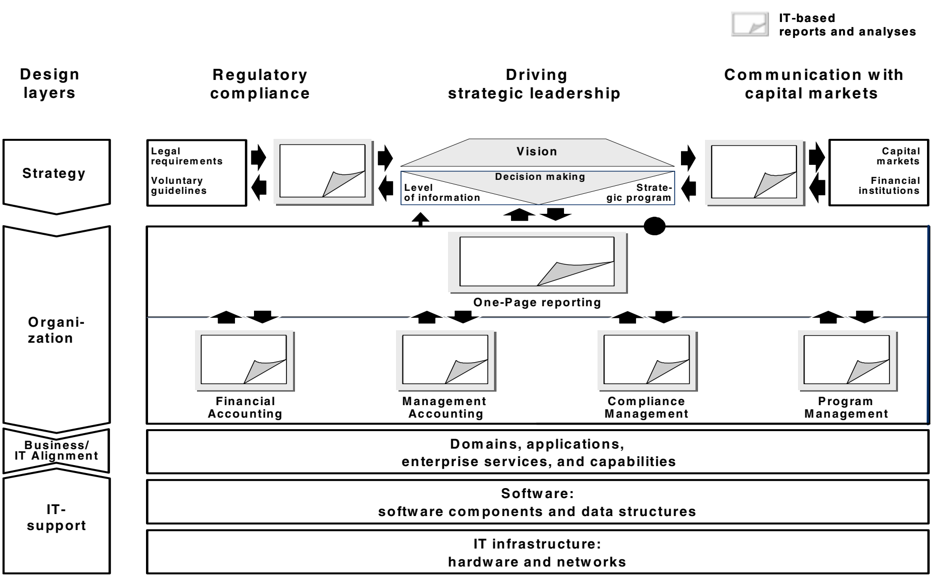


Fig. 2. Four-layer architecture of new-generation EIS

Several approaches exist in both academia and practice that claim to support proper IS architectures [25]: architecture of integrated IS [26], the trading-H model [27], and the semantic object model [28]. Following Henderson/Venkatraman [12], we propose an integrated top-down model based on four layers[2]: strategy (corresponding to Henderson/Venkatraman's strategy perspective), organization (corresponding to their organizational infrastructures and processes), alignment, and IT support (corresponding to their IT infrastructures and processes). Each layer can be specified as follows:

Strategy

The strategy layer specifies the purpose of EIS and should be oriented toward executives' key tasks. The empirical survey showed that *internal management* dominates in the day-to-day work of the executives surveyed (64 percent). Within this area, almost half (48 percent) of their working time is devoted to *strategic leadership*. Given the preponderance of this task, the objective of new-generation EIS can thus be defined as "driving (internal) strategic leadership."

A second task of executives was about external communication (36 percent). On the one hand, these activities include communication with capital markets. Despite the increased "regulatory compliance", 59 percent of executives consider *communicating with capital markets* as equally important. Furthermore, 38 percent spend more time on capital market communication than on regulatory compliance. Supporting executives in this area is therefore defined as an accessory objective of new-generation EIS. The fulfillment of legal requirements and quasi-legal guidelines ("*regulatory compliance*") is a prerequisite (the design layer "strategy" of Fig. 2).

Organization

Herein, the objectives of EIS from the strategy layer are executed using appropriate organizational structures (design layer 2 "organization"). Key processes to support executives should be derived based on the tasks executives actually perform. The most important has been identified in the survey as *strategy execution*. New-generation EIS should do this by breaking down strategic targets and communicating these targets to the organization ("strategy steering") as well continually monitoring the extent to which they are met ("strategy tracking").

IT Support

IT support has two aspects: software and IT infrastructure (Fig. 2). The *software layer* comprises software components and data structures. Our implementations have shown that new-generation EIS need heterogeneous aggregates, but these can predominantly come from standardized business software packages. In order to provide flexibility at acceptable cost (Fig. 2), the design for the software layer of new-generation EIS should combine (mostly common) standard software products in a tailored way and help to configure distinctive data structures.

The *IT infrastructure* consists of hardware and networks. No special issues regarding EIS are anticipated. Guaranteeing scalable, stable, and resource-efficient hardware and networks would be an appropriate maxim for the IT infrastructure layer.

---

[2] The approach is based on the St. Gallen Management Model. It offers a system-based perspective that serves as an integrative framework for various business disciplines [29].

Business/IT Alignment

Since development cycles are generally shorter for business artifacts than for IT (section 2), gaps develop over time. Therefore, the business/IT alignment layer has the role of bringing the business requirements (artifacts) together with IT capabilities. This decoupling is traditionally supported by domains, (logical) applications, and capabilities. Enterprise services represent logical clusters of IT functionalities that can ideally be reused by different processes.

In the accounting field, for example, *enterprise services* for EIS represent activities such as "perform payment" or currency conversion. *Applications* can cover accounts payable, accounts receivable, or the general ledger (financial accounting), value-driver trees or margin calculations (management accounting). *Business domains* represent a further level of abstraction, characterized by a large number of connected applications and enterprise services (e. g., financial or management accounting).

Creating Synergies

We consider the above approach to be more strongly business-driven than the state-of-the-art (section 2), contributing to better IT support for executives. This assessment assumes that the inherent flexibility of this approach largely comes from the extent of integration across design layers in the modelling.

To do so, *modelling chains* should encompass the relevant (management) processes at the organization level, their mapping with IT capabilities by domains, applications, enterprise services, and capabilities in the alignment level, and finally, the generation of code that can be executed automatically to guide the implementation. The next section shows the findings an instantiations delivered so far.

4 Implementing the New-Generation EIS Architecture

Implementation of the EIS architecture at an automotive supplier with, as of December 31, 2009, approximately 140,000 employers has had two dimensions: modelling the financial accounting domain for "pure" financial reporting (section 4.1, in detail, Fig. 3) and new-generation EIS architecture for drill-through issues (section 4.2).

4.1 Consolidation Process within a Business/IT System Architecture

Legal group consolidation shows the impact of using the proposed four-layer model to design EIS with an alignment of business and IT.

Modelling at the Organization Layer

Based on the legal requirements (strategy layer), a group consolidation process was designed at the organizational layer. BPMN was used for documentation [30]. The process has *six steps*: first, the unconsolidated divisional statements are captured ("reporting data"). Then the data are converted to group currency ("currency conversion"), and, third, these data are consolidated. Finally, the results are validated. If any failure appears, adjustment bookings on line item must be performed. The final stage in legal consolidation is to design the report and approve it (Fig. 3).

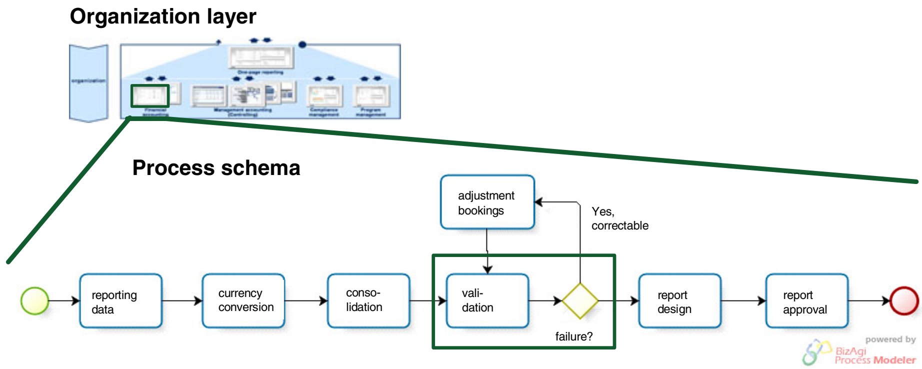


Fig. 3. Designing of legal group consolidation in the organization layer

Translation to the Business/IT Alignment Layer

Applications serve to refine the designed process. Whether each activity is automatic or manual must be specified. (a) With the exception of unanticipated failures, automated activities have no interaction with human users and can be performed separately within IS. Currency conversion of foreign division statements is such an associated service (layer 2, Fig. 4).

(b) For manual activities, authorities must be provided with information for decision making. "Adjustment bookings" activity is detailed in the alignment layer with a string of enterprise services such as, "failure report" (automatic), "require adjustment booking" (automatic), and "check adjustment booking" (manual, Fig. 4).

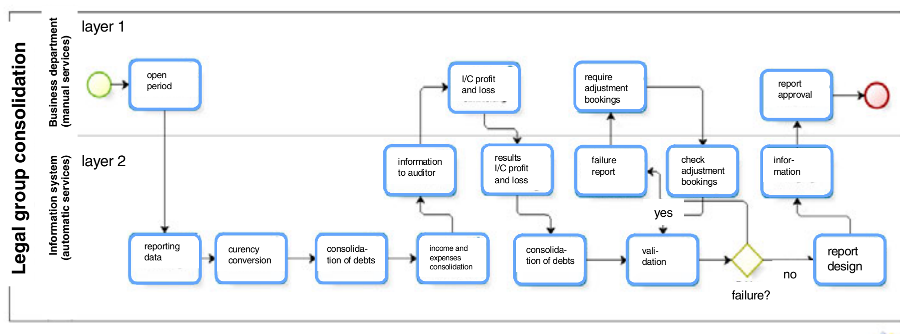


Fig. 4. Design of legal group consolidation in the alignment layer

Translation to the Software Layer

A last step of the alignment process covers linking the services of the alignment layer with existing IT functionalities (software layer). For that matching the business perspective of IT architecture can be used (Fig. 5).

To identify the IT functionalities, open interfaces in the standard software SAP Business Consolidation (BCS) and SAP Financials (FI) were used. In a SOA

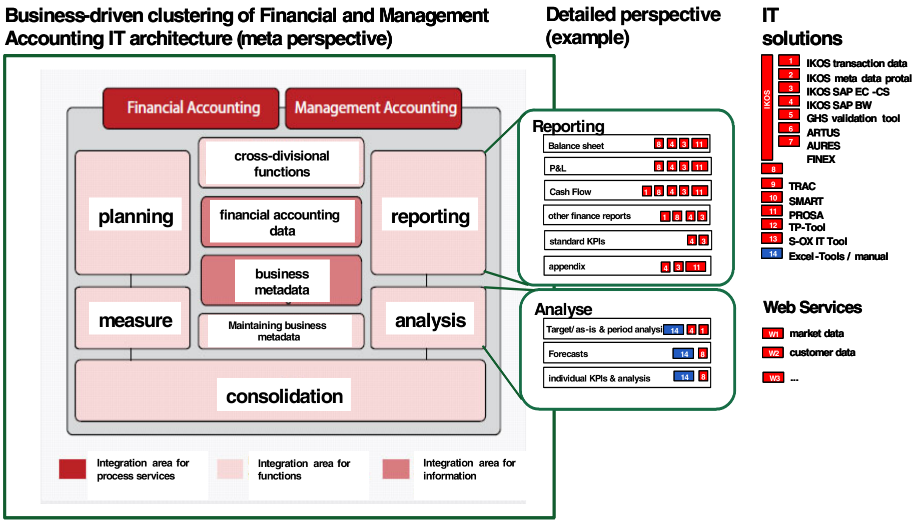


Fig. 5. Business perspective on IT architecture

architecture, web services are appropriate because they allow technology- and location-independent referencing. Such interfaces or web services are not limited to the legal group consolidation process and can be reused in other contexts as well.

4.2 Generating "Drill-through" Functionalities

The strategy layer pointed out the new significance of *operational information* for executives. EIS should provide a way to drill through to underlying operational data.

The first drill-through function implemented provided access to financial reporting. If the EIS provides condensed versions of the balance sheet, P&L statement, and cash-flow calculation, the drill-through should allow direct access not only to overviews of accounts receivable and payable, but also entire balance-sheet, P&L, and cash-flow statements. The procedure for generating such a *flash report* in our business/IT aligned model is based on the previously modelled consolidation process.

In the legal group consolidation, much time is devoted to the consolidation activities (activity 3, Fig. 3). From an internal perspective, just a *"simple summation"* of the separate division statements is sufficient for a first management view. To allow this, the legal consolidation process can be modified easily in the new-generation EIS architecture by including an additional activity.

Changes made in Organization Layer

The separate division statements on group currency must be captured and then a summarized P&L must be compiled and presented for approval. To avoid delaying the consolidation processes, reporting data should be duplicated (Fig. 6).

Changes made to the Alignment and the Software Layer

In our instantiation, two new services – "duplicate reporting data" and "create flash report" – were added and aligned to the process flow. Both services could be classified as (automated) activities.

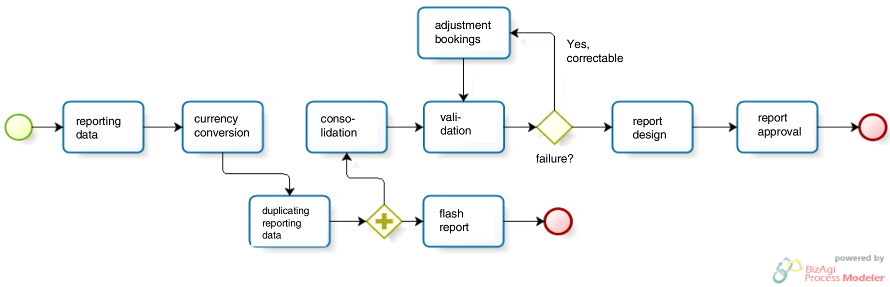


Fig. 6. Flash report design in conjunction with the legal group consolidation

Finally, a link was created with the software interface SAP BCS to allow the new "create flash report" process to be performed automatically. The function of the SAP NetWeaver Business Warehouse could perform the "duplicate reporting data" function.

5 Synthesis

The findings of the case study lead to seven business/IT alignment recommendations to drive future research and implementations.

1. In executives' perspective, flexibility currently is the biggest challenge business/IT alignment has to face

The surveyed executives pointed out that flexibility is one of their most important objectives for IT today (Fig. 1). Business/IT alignment has to manage changing *processes* in particular. In the domain of EIS executives ask for *additional analytical capabilities*, especially direct access to more operational information.

In our case example, the distinctive separation and loose coupling of processes (organization layer), domains, applications, enterprise services, and capabilities (alignment layer) with software components (software layer) represents a step towards managing changing processes.

2. IS flexibility was identified as an important EIS architecture design principle, but not yet implemented due to current short-term pay-off focus

Due to the economic crisis, investments are selected more than ever on short pay-off time. IT flexibility is a long-term project. Political opposition and a lack of capacity can exacerbate this resistance. But in the longer view, flexibility should gain importance as an architecture design principle: On the one hand, vision, strategic program and companies' organization will change more often. On the other hand, business/IT alignment should help to reuse functionalities, such as planning and consolidation, by identifying such functionalities in several IS and make their "recycling" possible.

In our case example, the business perspective on IT allowed to identify the "currency conversion" in several financial software solutions. It was centralized and maintenance was leveraged for just a single functionality. As a result, cost and time were reduced by 30 percent.

3. Measuring the business benefits of business/IT alignment is a strong lever for its future success

The implementation of our EIS architecture showed that a conceptual model of business/IT alignment is important, but to demonstrate its power, the model's *business benefit* must be measurable.

Our case revealed that the alignment can be measured in several dimensions, such as the *degree of integration* into business processes on the one hand or ease of handling and cost-efficiency on the other hand. Furthermore, the business-driven domain model of IT gave the business side a better understanding of IT capabilities. Because of its acceptance, the *principle of economic efficiency* can be the lead architecture principle. Under this standard, business/IT alignment does not have to achieve the "best" conceptual or technical solution, but the most efficient in economic terms.

4. A top-down approach seems to be the best business/IT alignment model

The scope of executive tasks is continually expanding. That will increase the *functional design work* in business/IT alignment projects. These efforts should aim to more strongly reflect a company's vision and strategic program and translate these targets into action.

Our case example showed that reports and the associated analyses can be used to align business requirements with IT capabilities. EIS architecture design should be a business-driven *top-down project* and should be communicated that way.

5. Conceptual information logistics assure that information flows from several upstream systems and is a newer determinant of business/IT alignment

Required data for dispositive IS come from many sources, often in different formats. For successful business/IT alignment, the *path this data follows through the organization* must be determined. In addition to the EIS data model and IT infrastructure, business-/IT alignment must cover this new topic of information logistics both within the company and from external sources.

6. Further domains for business/IT alignment

Besides EIS, other domains should be feasible for using business/IT alignment capabilities. *Business metadata* is an example. These data about ownership relations or addresses, customer data about order information like cash discounts, or customer liquidity rankings seems to be business/IT alignment topics. Such data are needed by several IS and must be consistent, e.g., for group consolidation and one face to the customer. A good approach would be to outsource them to a separate IS, maintain it centrally, and provide them as the single-point-of-truth to company authorities and IS.

7. Executives must champion the business/IT alignment process and devote time to clarifying their (information) requirements, while the project team must deliver the first prototype quickly

The case example indicates that executives must take an *active role* in the business/IT alignment process to ensure its success—not only to clarify their information needs, but for the first prototype review as well. The latter is an area in which they are often unfamiliar.

From the project side, executives are best served by providing them hands-on "look & feel" with a first *prototype* and then enhancing this prototype. Our case example showed that will help to maintain executives' interest in the project.

6 Outlook and Future Research

The objective of this article was to develop a business-driven model to contribute to better EIS architecture design. The proposed business/IT aligned model consists of four design layers: strategy, organization, alignment, and IT support.

In terms of timing, the present moment seems fortuitous: today's executives grew up with IS and should have a positive attitude toward IT [1]. Furthermore, technical progress has been made in the domain of modelling IT from a business perspective.

Looking ahead, we expect that the younger management generation will be interested in new, more communicative ways to perform their day-to-day work. In response, IT should, for example, integrate individual desk research, collaborative problem solving and joint decision making in board meetings, with communicating the results into the organization in close to real time—tasks that until now have been separated. Such development on the front end side could push business/IT alignment from the top ranking in executives' agendas or give it a new direction.

Empirical Survey

A comparative field survey (cross-section analysis) was selected as the research method for this project. Large, international companies were defined as the population. As they are predominantly share-based, the survey was limited to those listed in the Financial Times "Europe 500" report on April 1, 2009.

To collect the data, we conducted a written survey among CEOs and CFOs of the corporate center. The survey consisted of 58 questions, divided into four categories: questions about the company and its organizational structure, functional requirements, design requirements (to-be profile), and the current instrument used by the companies (as-is profile).

Fifty-one companies responded with a total of 59 questionnaires, a population-based response rate of 23.6 percent. Representativeness was tested using the chi-squared test of homogeneity. Frequencies were given for the metrically scaled values. These were analyzed using the arithmetic mean and corresponding standard deviation.

References

1. Mayer, J.H., Winter, R.: Emerging from the economic crisis with a clearer view—new-generation executive information systems: the CFO perspective, research paper. University of St.Gallen, St.Gallen (2010)
2. Luftman, J.N.: Key Issues for IT Executives 2004. MIS Quarterly Executive 4(2), 269–285 (2005)
3. Schelp, J., Winter, R.: Integration Management for Heterogeneous Information Systems. In: Desouza, K.C. (ed.) Agile Information Systems—Conceptualizations, Construction, and Management. Butterworth-Heinemann, Oxford (2007)

4. Yusuf, Y.Y., Sarhadi, M., Gunasekaran, A.: Agile manufacturing: the drivers, concepts and attributes. International Journal of Production Economics 62(1-2), 33–43 (1999)
5. Aier, S., Winter, R.: Virtual Decoupling for IT/Business Alignment – Conceptual Foundations, Architecture Design and Implementation Example. Business & Information Systems Engineering (BISE) 1(2), 150–163 (2009)
6. Dietsch, A., Goetz, T.: Nutzenorientiertes Management einer Serviceorientierten Unternehmensarchitektur. Wirtschaftsinformatik 47(2), 10–21 (2005)
7. Friedrich-Alexander-University Erlangen-Nürnberg/Detecon: Die geschäftsfokussierte Informationstechnologie—Business-IT-Alignment als zentrales Steuerungsinstrument zur strategischen Ausrichtung der IT, Erlangen-Nürnberg und Bonn (2009)
8. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly 28(1), 75–105 (2004)
9. March, S.T., Smith, G.F.: Design and natural science research on information technology. Decision Support Systems 15(4), 251–266 (1995)
10. Rossi, M., Sein, M. K.: Design Research Workshop: A Proactive Research Approach, `http://www.cis.gsu.edu/~emonod/epistemology/Sein%20and%20Rossi%20-%0design%20research%20-%20IRIS.pdf` (accessed 2009-04-29)
11. Peffers, K., Tuunanen, T., Gengler, C.E., Rossi, M., Hui, W., Virtanen, V., Bragge, J.: The Design Science Research Process: A Model for Producing and Presenting Information Systems Research. In: Proceedings of the First International Conference on Design Science Research in Information Systems and Technology, DESRIST, Claremont, CA (2006)
12. Henderson, J.C., Venkatraman, N.: Strategic Alignment. A Framework for Strategic Information Technology Management. Center for Information System Research, Working Paper Nr. 190, Massachusetts Institute of Technology, Cambridge (1989); Reprint: IBM Systems journal 38(2&3), 472–484 (1993)
13. Greene, R.M.: Business intelligence and espionage. Dow Jones-Irwin, Homewood (1966)
14. Luhn, H.P.: A Business Intelligence System. IBM Journal of Research and Development 2(4), 314–319 (1958)
15. Watson, H.J., Wixom, B.H.: The Current State of Business Intelligence. Computer 40(9), 96–99 (2007)
16. Ackhoff, R.: Management misinformation systems. Management Science 14(1), 147–156 (1967)
17. Rockart, J.F., Treacy, M.E.: The CEO goes on-line. Harvard Business Review, 82–88 (January-February 1980)
18. Gluchowski, P., Gabriel, R., Dittmar, C.: Management Support Systeme und Business Intelligence, Berlin (2008)
19. Schulze, K.-D., Besbak, U., Dinter, B., Overmeyer, A., Schulz-Sacharow, C., Stenzel, E.: Business Intelligence-Studie 2009, Steria Mummert Consulting AG, Hamburg (2009)
20. Eckerson, W.: Performance Dashboards—Measuring, Monitoring, and Managing Your Business. John Wiley & Sons, Hoboken (2006)
21. Few, S.: Dashboard Design: Beyond Meters, Gauges, and Traffic Lights. Business Intelligence Journal 10(1), 18–24 (2005)
22. Gleich, R., Horváth, P., Michel, U.: Management Reporting—Grundlagen, Praxis und Perspektiven, München (2008)
23. Watson, H.J.: The BI Ideal: A Single Enterprisewide Data Access Tool. Business Intelligence Journal 10(2), 4–5 (2005)
24. The Institute of Electrical and Electronics Engineers, Inc.: ANSI-IEEE 1471. Recommended Practice for Architectural Description of Software-Intensive Systems. IEEE Computer Society, New York (2000)

25. Schekkerman, J.: How to survive in the jungle of enterprise architecture frameworks: Creating or choosing an enterprise architecture framework. Trafford, Victoria (2006)
26. Scheer, A.-W.: Wirtschaftsinformatik: Referenzmodelle für industrielle Geschäftsprozesse, 7. Aufl, Berlin (1997)
27. Becker, J., Schütte, R.: Handelsinformationssysteme, 2. Aufl., Frankfurt/Main (2004)
28. Ferstl, O.K., Sinz, E.J.: Grundlagen der Wirtschaftsinformatik, 6. Aufl., München (2008)
29. Dubs, R., Euler, D., Rüegg-Stürm, J., Wyss, C.: Managementlehre—Einführung in die Managementlehre, Bern et al (2004)
30. OMG, Business Process Modeling Notation (BPMN) Specification (2006), http://www.bpmn.org/Documents/OMG%20Final%20Adopted%20BPMN%201-0%20Spec%2006-02-01.pdf (accessed 2008-07-29)

Enterprise Architecture Design Principles and Business-Driven IT Management

Mats-Åke Hugoson, Thanos Magoulas, and Kalevi Pessi

IT University of Göteborg
P.O.B. 8718, SE-412 96 Göteborg, Sweden
Jönköping International Business School
P.O.B. 1026, SE-551 11 Jönköping, Sweden
Mats-Ake.Hugoson@ihh.hj.se, thanos@ituniv.se, pessi@ituniv.se

Abstract. The strategic role of IT and its significance throughout the organization increases in environments characterized by complexity, variety and change. Hence IT management must deal with uncertainties and with different, conflicting and ever changing demands. In this sense Enterprise Architecture is playing an increasingly important role in improving IT management practice. This paper presents architectural design principles that strive to create and maintain alignment in a dynamic and uncertain business environment. Two crucial architectural aspects will be highlighted: (1) the delineation of the information systems, (2) the choice of interoperability principle. A case from Swedish Industry is used to demonstrate the application of the principles. In the paper we argue that the choice of principles has an significant impact on operational as well as strategic alignment. The larger, heterogeneous, and dynamic the business are, the more crucial are the issues of information systems delineation and interoperability.

Keywords: Enterprise Architecture, Strategic Alignment, Operational Alignment, IT management, Interoperability.

1 Introduction

During the last decade IT has received a more central role in transforming organizations in order to meet the challengers of the 21th century. The focus on IT management has changed from an emphasis on mastering the technology, developing information systems and controlling the costs of the IT department to seeing IT as an essential means to achieve business value as well as to create new organizational forms with an increased ability to innovate and change. The business environment is characterized by greater dynamics and complexity. Contemporary organizational boundaries are no longer clearly easily defined or delineated. Today's IT has opened the way to the creation of many different types of influences on organizations as well as cooperation between different organizations. One consequence of this is that IT management must deal with uncertainties and with different, conflicting and ever changing demands. In this sense enterprise architecture is playing an increasingly important role in improving IT management practice. Architecture matters are more

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 144–155, 2010.
© Springer-Verlag Berlin Heidelberg 2010

and more critical for the creation and maintaining of operational (short term) alignment as well as strategic (long term) alignment between the business and its information systems. Enterprise architecture design and management should therefore be guided of principles that contribute to a suitable and sustainable alignment.

This paper presents architectural design principles that strive to create and maintain alignment in enterprise architecture in a dynamic and uncertain business environment. In the paper we discuss two crucial architectural aspects that have significant impact on alignment between business and information systems: (1) the delineation of the involved information systems and their ability to change and operate independently of each other and thus achieve operational alignment, (2) the choice of interoperability principle and its impact on strategic alignment. A case from Swedish Industry is used to demonstrate the application of the approach. Through the use of the case we discuss the above two aspects as well as the impact of the design principles on the possibilities to attain agility through business oriented IT management.

In the paper we argue that the choice of principles for delineation and interoperability has a significant impact on operational as well as strategic alignment. The larger, heterogeneous, and dynamic the business and its environment are, the more crucial are the issues of information systems delineation and interoperability.

The paper is organized as follows; first in chapter 2 we discuss Enterprise Architecture and the concept of alignment. Then the following two chapters present principles for information systems delineation and interoperability. In chapter 5 the case study will be introduced and in chapter 6 we will analyze the case with respect to the architectural principles and their impact on alignment. Last, in chapter 7 we summarize the paper and draw some conclusions.

2 Research Method

The research has been inspired by collaborative practice research [14] and its inside/outside perspectives. One author of this paper has been working several years with the case companies and provide “inside perspective”. The other two authors are full-time academic researcher and provide “outside perspective”, which allows for deeper assessment and reflection.

The research methodology is essentially interpretive case study [15], [16]. We carried out data collection primarily through observations, semi-structured interviews, and workshops with CEO, CIO, Process Managers, Enterprise Architects, Project managers and Users. Observations and activities were recorded in a research diary. Research rigor is a question in any case study; typical critiques target the validity of generalisation or the lack of self-criticism. However, the main objective in these case studies is to increase the understanding of the relationship between Enterprise Architecture and Management of IT Investments.

3 Enterprise Architecture and Alignment

Since the 1970’s, the fast growing number of systems and in many cases the ad hoc manner in which the systems were integrated have exponentially increased the cost and complexity of information systems. At the same time organizations are finding it more and more difficult to keep these information systems in alignment with business need. In light of this development, a new field of research and practice was born that soon came to be known as Enterprise Architecture [1][2][3][4].

Enterprise Architecture is usually divided into different categories or domains or architecture types. For example Aerts et al. [5] identify three domains in which architecture matters: (1) The business architecture, (2) The application architecture (or information systems architecture), (3) ICT platform architecture (or IT architecture).

The developments in the various domains influence each other and the increasing need of business for agility to cope with changes may be provided by architectures supporting reflectivity. Alignment practices must take into consideration the relation between the various architectures.

Enterprise Architecture is an expression of workable alignment between business requirements and systems assets that have been organized in a particular way as a response to the nature of business environment. Therefore, Enterprise Architecture Frameworks should give guidelines for the choice of architectural principles for information systems [4]. Such guidelines and principles should take into account:

- How to delineate information systems and create “systems of systems”.
- How the different systems should interoperate in order to satisfy the expectations of the business and its environment.

Enterprise Architecture Frameworks should also give guidelines for how to manage the requisites of alignment between business architecture and information systems architecture [12]. Unfortunately, very few existing Enterprise Architecture Frameworks give clear guidance about how to architect and manage these critical issues. Thus, there is a need for more knowledge about architectural principles that holds in harmony 1) the ever changing nature of business and 2) the capabilities provided by information systems to respond both quickly and in a long term perspective to these changes.

Architectural principles are statements that express how your enterprise needs to design and deploy information systems across the enterprise to connect, share and structure information. The value of such principles can be given in terms of decision guidance. In this paper we focus on principles for delineation of information systems and principles for interoperation between systems.

4 Information System Delineation Principles

Delineation of information system can be done during different principles. One principle is based on different types of information models i.e. Information driven principle. One other principle is based on responsibilities in the enterprise i.e. Responsibility driven principle. In this paper we present these two principles.

The Information Driven Principle

The Information driven principle is applied in several different architectural frameworks and is widely used [9][11]. Allen and Boynton [13] called it “The high road alternative". The information driven principle focus on information as basis for delineation of systems, for example, all product data should be gathered in one system, the "Product Data System". The core IS activities of the business are centralized and core information systems are designed to be organizationally independent – that is, immune to restructuring in business.

The Responsibility Driven Principle

The responsibility driven principle also referred to as "The low road alternative" [13] is based on the principle that the responsibility for an area of operations in the enterprise should comprise all resources (including information system) used to fulfil objectives and tasks specified for this area. Each information system should therefore support only one area of responsibility, and can then be developed for efficient support to users in that area. Necessary interaction and information exchange between different areas of responsibility in the organization and with external partners must however be considered. Every system should respond to two tasks: to support users and to support computerized information exchange with other systems in the structure. Similar areas of responsibility can use the same type of solution, but still each instance of information system can operate (and be changed internally) independent of other systems. Independence in development, operations and change is an important issue when delineation of information systems is based on responsibilities in the enterprise.

5 Interoperability Principles

One of the major issues in designing and developing enterprise architecture is which degree of interoperability should there be between the various business units of the enterprise and how this should be reflected in the integration of their information systems. Interoperability is usually defined as "the ability of two or more systems or components to exchange and use information" [7]. TOGAF [3] defines interoperability as: (1) the ability of two or more systems or components to exchange and use shared information and (2) the ability of systems to provide and receive services from other systems and to use the services so interchanged to enable them to operate effectively together. Much work has been done in defining the concept of interoperability from Organizational as well as from Information Systems perspective [8][9][10][11]. In this paper we present two interoperability principles: Intersection and Interlinking.

Intersection Principle

The objective of the intersection principle is to improve the quality of information and the efforts of information management through the elimination of redundancies. Accordingly, the intersection principle (or overlapping) takes place when the structure of each of the participating systems has one or more elements whose properties are identical (or sufficiently similar) or one of the participating systems has elements which can be used in some of the other systems [10]. Intersection strives to eliminate duplicates.

The intersection principle creates a shared information space, which can be exemplified by a situation where the participating information systems share some of their constituent parts, i.e. the conceptual base, the information base, or the rule base.

In this sense, the participating information systems can be coordinated either (1) by keeping the previously redundant part in one system and allowing the other systems to access that part, or (2) by unifying the redundant local parts into one shared and common part (like a common database). Thus, a shared information environment might be seen as the result of a coordinating process aiming either to eliminate the overlapping parts of the participating systems or the globalization of some of their parts (unification)

The shared parts become a common property because any kind of change in those parts may have effects on all of the participating systems. Locally, the right to alter the system is limited to unshared parts only. In this sense the effects of intersection might be described in terms of expectations for better quality and availability of information services, as well as in terms of dependencies between systems and acceptance of a limited freedom of alterations.

Interlinking Principle

Interlinking principle represents a different concept for systems interoperability. Computerized interaction between different systems is carried out through the exchange of messages which are based on business demands. A significant feature of interlinking is that interoperability takes place without substantial interference with the inner structures of the participating systems and without substantial limitation of their independence [10]. Only information exchange is federated in the interaction agreement, which means that the inner structure in each system can be specified and developed quite independent of the inner structure in all other systems in the total structure. Each system has to automatically produce specified messages according to interaction agreements, either as a planned task or on request, and has to handle incoming specified messages.

Interlinking implies a change from access or sharing thinking to messaging. If independence is to be preserved, it is not enough to ´open the books´ and to show the data structures in each system. Interaction is instead bridged through the defined messages based on business relations. It is not a question of understanding data in other systems; it is a question of understanding what information is to be transferred between the systems. The different data structures are local and connected to the messages through mapping mechanisms, which must be developed and maintained for each system. If the inner structure in one system is changed, then the mapping mechanism may need to be changed in that very system in order to maintain proper interaction, but there will be no change in any other system as long as the interaction agreement stands.

In the long perspective interlinking creates the possibility to replace a system without any changes in interoperability, as long as the new system fulfills interaction agreement. This has a major impact in that it reduces migration problems and facilitates the sustainability of complex structures of systems.

6 Case Study

In order to illustrate how delineation principles and interoperability principles may be applied in a business context, we use a case from industry. The case is a Swedish company that constructs, manufacture and sells submersible pumps on the European market. Sale companies are established in 14 countries. Head office, product development and production is situated in Sweden. Smaller standardized pumps to be delivered off the shelf are manufactured in one plant, big pumps are manufactured to customers order in another plant.

In order to improve efficiency in the organization a Business Process Reengineering (BPR) project, headed by the Chief Administration Officer (CAO) has been carried out. An Enterprise Model describing the new processes has been develop, but the business processes are not yet implemented. An overall Enterprise Model of the business processes and their interaction is shown in fig 1.

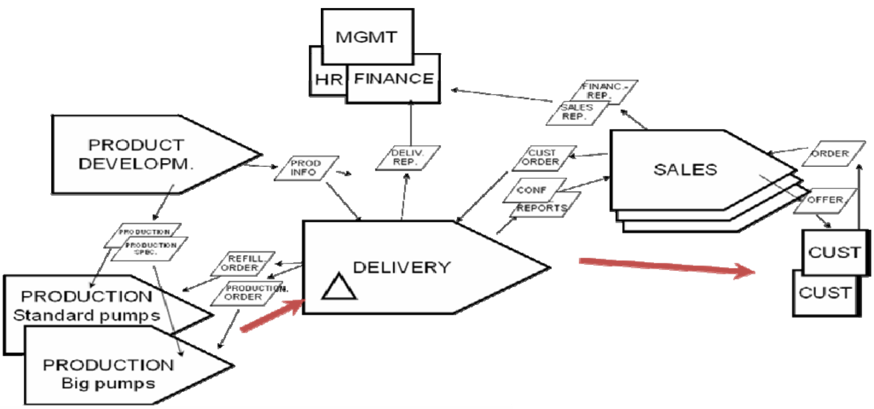


Fig. 1. Process Oriented Enterprise Model

The SALES process has in each country the full responsibility to fulfill customer demands. Standardized pumps should be delivered within 3 days, using the central DELIVERY process. Central resources can be used to handle special demands and give customers offers for big, one off pumps. SALES process interacts with the delivery process through specified Customer Orders, and receives Confirmation and Delivery Reports. There is an official price list for standardized pumps, but pricing in offers to customers is a responsibility for each SALES process, which buys the pumps at an off factory price. Each SALES process must give both different types of Sales Reports on business performance and Financial Reports to central management. Reporting to authorities in each country is seen as a local responsibility.

SALES is , through the BPR project, a centrally defined business process, but the process orientation has to be carried out in each country, headed by the local Managing Director (MD). Even if the intention is commonality in the SALES process, this decentralized implementation allows for some local adaptations, which is seen as an opportunity to be competitive in each market.

The DELIVERY process is, on the other hand, highly centralized. DELIVERY handles logistics, including ware-housing and transports, for all pumps from final testing in production line out to customers. For pumps manufactured to order a ´walk through´ flow should be applied in order to shorten delivery time. For pumps off the shelf a distributed net of warehouses may be used (some of them now existing in local sales organizations) but the total flow of pumps ready for delivery will be controlled by the DELIVERY process. Production Orders are sent to the two PRODUCTION processes. For standardized pumps Refill Orders are calculated to maintain delivery ability. Reports on Performance in delivery and warehousing are to be sent to central management.

The two PRODUCTION processes differ, both in lay out of the production and in equipment/tools. Standardized pumps are manufactured in batches. Modularity supports different capacities and specifications. Production Specifications for assembly along the productions lines are fixed and well defined. Manufacturing to customers order in the other PRODUCTION process is carried out in single units, using a number of options to satisfy different customer demands. Special components must be used for customization. Production Specifications for big pumps contain a number of options to be used for different customer orders. Production planning must be specific for each pump.

The PRODUCT DEVELOPMENT process is highly project oriented. For a new variant of standardized pumps the PRODUCT DEVELOPMENT process involves activities from market analysis to test of a number of pumps before Product Specifications can be released for marketing, and Production Specifications can be given to the PRODUCTION process. For big pumps a number of options have to be designed and tested. All combinations of different options can however not be tested before launching to market and to production, and therefore there must be an information exchange between the PRODUCT DEVELOPMENT process and the PRODUCTION process also during the production of big pumps.

In addition to the structure shown in this enterprise model, the business strategy opens for future changes. Additional SALES processes, perhaps through Agents, for new markets are to be entered without major changes in existing SALES processes. New PRODUCTION processes in other countries or by outsourcing can be added in order to decrease production costs, but deliveries to customers will still be centrally controlled.

Current IT Support

It is apparent that the current set of information systems as shown in fig 2, is not aligned with the enterprise model. The main information system is composed of a number of subsystems integrated through a central data base. This system is also integrated with system for product structures. These systems are rather old and have gradually grown out of date. There are strong dependencies between different subsystems; a change in one will cause undesired side effects in others. The central Accounting system is however rather new. It interacts with the central system through transactions, but financial reports from sales companies in different countries are manually updated.

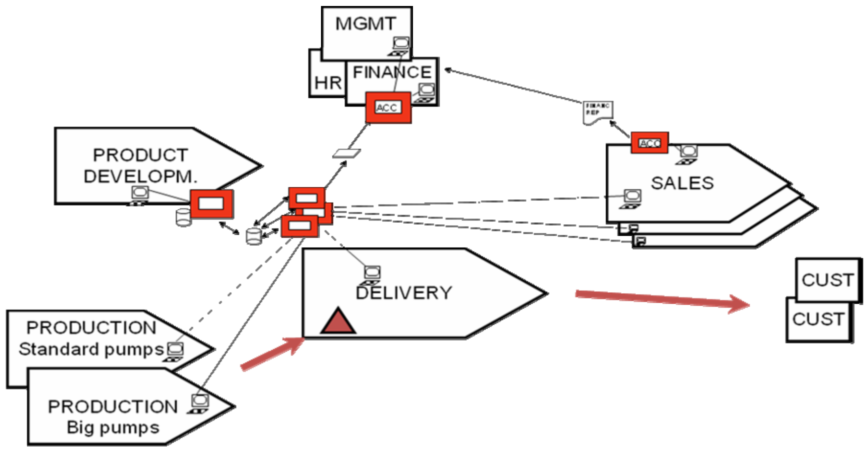


Fig. 2. Current IT Support

The main problem is that the process orientation will not be supported by these systems. Order Entry, which is the starting point for the central system, is the end of the SALES PROCESS, sales activities, offerings etac are not supported. Projects in the PRODUCT DEVELOPMENT process are not supported, only final results are entered into the system. All production orders are handled the same way, even if the two production processes are different, etc. As the systems in use are instable, they cannot form the basis for development; they must be replaced.

The New IT Support

The following directives are given by the central management:

- Development of new information systems must be carried out step by step.
- The Business strategy with future structural changes, new business processes, possible outsourcing etc, must be considered in the strategic IT plan.
- Implementation of the business processes should be coordinated with implementation of new IT support.
- CAO is given central responsibility for this coordinated development, but Managing Directors and Process Managers for central processes are responsible for the development within each business process.

Two main architectural principles were applied during the strategic IT planning, in order to satisfy the given directives:

Information systems are delineated according to business responsibilities. A separate information system (the ***CRM*** system) is delineated for the SALES process. As the sales process can be implemented in different ways in different countries, each country may have a separate instance of the ***CRM*** system. This gives freedom of action for step by step development, and for future insertion of new systems to support structural changes (new, or new type of SALES process), without major changes in existing systems. A central information system, the ***LOG***-system supports the central DELIVERY process. Future changes in logistics through new warehouses may call for local WH-systems. Two different systems (***MPS-S*** and ***MPS-B***) for material and production planning are delineated, as the two PRODUCTION processes are different. If new production plants are established, each plant will have a separate system. Outsourcing of production will then be possible without changes in remaining systems. For the PRODUCT DEVELOPMENT process a certain system ***P-DEV*** is delineated. The existing ***ACCOUNT*** system fits in to structure. An additional ***PERFORM*** system is inserted as an executive support system.

Systems interaction should be based on business transactions, and performed through exchange of computerized messages. This principle allows every information system to form an inner structure independent of other systems in the structure. Systems from different origin can interact.

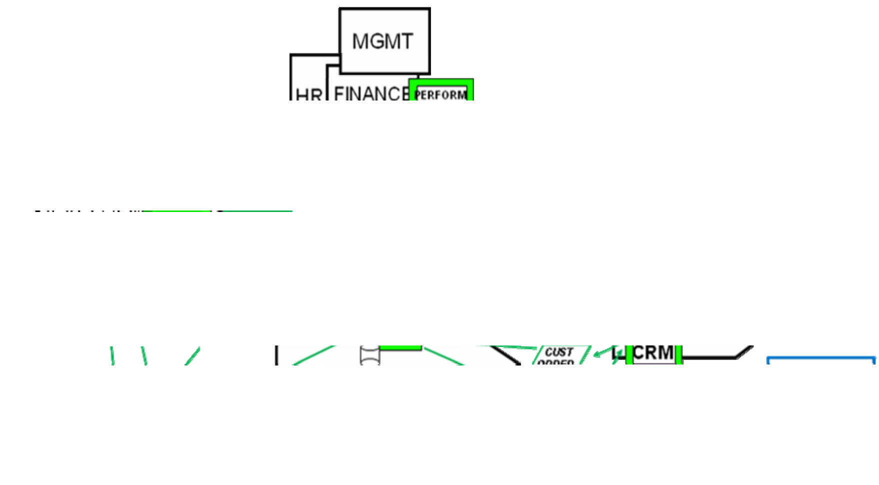


Fig. 3. The Strategic Future Structure of Systems

The strategic structure is used for coordination of different development projects. Standardized software packages will be used and a central license contract is signed with a software house. CAO controls that every proposed system interacts with other systems according to the strategic structure, but features in each system are decided locally. The IT manager is responsible for the infrastructure, including Message Handling System to support systems interaction.

Managing Director in Italy is eager to start implementation of the new SALES process, and step 1 is to specify and develop a CRM system for Italy. Vendor proposes a solution with a common database for all SALES processes, but CAO just states that he controls the structure, and as there are no common customers between different SALES processes, the system for Italy shall be quite independent from other CRM systems. Each CRM system must send ***Customer Orders*** to the central system and receive ***Confirmation*** and ***Delivery*** messages. The CRM system must also interact with central systems for accounting and for performance control, as specified in the Enterprise model. The new SALES process and the system ***CRM-IT*** is implemented within 6 months. The infrastructure is financed centrally, but the CRM system is financed locally, which gives incentive for a rather simple system. Only features that have a business value are chosen in the CRM package.

On demand from the MD in The Netherland, step 2 is to specify and implement a CRM system for the Dutch SALES process. The vendor proposes an extension of the Italian system with a common data base, but again the CAO gives clear directives that an independent ***CRM-NL*** shall be developed. The same specification should be used for systems interaction, but the Dutch system may use some other features suitable for the Dutch SALES process. The long term ability to use different versions and even install systems from other origin is to be maintained.

Gradually other CRM systems are inserted in the structure in parallel with the business process implementations. The business value from investments in each project can be evaluated and the time to value (TTV) is short as the implementation time for each project is rather short.

Development of the P-DEV systems started early, but it took some time to establish the new logistics for DELIVERY, and therefore the system LOG was delayed. The two systems for the PRODUCTION processes were developed in parallel with the LOG system in order to fully replace the old system.

7 Analysis of the Case

A partitioned approach is used in the case. Systems are delineated according to the business structure (areas of responsibility), which differs from other strategic principles for Enterprise Architectures, such as Zachman [1], TOGAF [3] etc where the structure is based on some kind of information model. The structure of information systems reflects the business structure, which can be referred to as Strategic Alignment.

For systems interaction (interoperation) the principle of interlinking is chosen, which means that defined computerized messages are exchanged between systems in the structure. There is no common data storage between different systems in the structure. As messages for interoperation in the case study are derived from business transactions they can be defined and described independent of the data storage and inner structure in each system.

Every system must be able to send and receive the specified messages (using some kind of infrastructural message handling system), but each system can then be developed fully independent of other systems. The implication is that a change in the interoperability structure can cause changes in some systems, but a change in the inner structure in one system will not cause changes in any other system. A system can even be fully replaced without structural changes if the new system fulfils the specified interoperability.

In the case this ability is used for adaptation of each system to local demands, which could be refereed to operational alignment. The combination of these principles for delineation and interaction results in a structure where only information for interoperability must be controlled centrally. The big amount of information within each local system can be handled in a more decentralized way. Three major principles for IT governance are applied in the pump company:

1. Central responsibility for the strategic structure and for strategic choices.
2. Local responsibility for adaptation and implementation of each information system.
3. Coordinated business and IT development. As the strategic structure of information system is directly aligned with the business structure, and the coordinated development is seen as a part of business responsibility, the model can be characterized as Business-driven IT management.

The necessary condition for this model is that each system really can be developed and implemented without dependence on other systems. On the other hand a well defined and maintained interoperability structure must be implemented in order to avoid anarchy and ´information islands´. In this case study it is shown that architectural principles that promote systems independence will enable alignment and agility and make business oriented IT management possible.

8 Summary and Conclusion

Enterprise Architecture should be an expression of workable alignment between business requirements and systems assets that have been organized in a particular way as a response to the nature of business environment. Alignment is any form of fitness or workable harmony between the requisites of business units and the capabilities of information systems to satisfy these requirements.

In order to extend our knowledge of these issues we have in this paper discussed the following two crucial architectural aspects in the context of alignment and IT management: (1) the delineation of the involved information systems and their ability to change and operate independently of each other, (2) the choice of interoperability principle and its impact on agility. A case from Swedish Industry has been used to demonstrate the application of Business-driven architectural principles.

The case has demonstrated that the enterprise architecture makes sense because it is able to hold a strong alignment between the capabilities of information systems architecture and the ever changing demands of the business architecture. Furthermore, the capability of the enterprise architecture is grounded on the strong agreement between the enterprise architecture and with the right choice of principle of delineation, and principle of interoperation

In the paper we argue that the choice of principles for delineation and interoperability has a significant impact on operational alignment as well as on strategic alignment. The conclusion is that the choice of a suitable architectural principle is a main condition for aligning information systems to business demands and to create benefits. If the wrong principles are chosen, there is a considerable risk for misalignment and expensive consequences.

References

1. Zachman, J.A.: A Framework for Information Systems Architecture. IBM Systems Journal 26(3), 276–292 (1987)
2. Sowa, J.F., Zachman, J.A.: Extending and Formalizing the Framework for Information Systems Architecture. IBM Systems Journal 31(3), 590–616 (1992)
3. The Open Group: The Open Group Architecture Framework, Version 8.1.1, Enterprise Edition (2007)
4. CIO Council: A Practical Guide to Federal Enterprise Architecture. Chief Information Officer Council, Version 1.0 (February 2001)
5. Aerts, A.T.M., Goossenaerts, J.B.M., Hammer, D.K., Wortmann, J.C.: Architectures in context: on the evolution of business, application software, and ICT platform architectures. Information & Management 41, 781–794 (2004)
6. van der Poel, P., van Waes, R.: Framework for Architectures in Information Planning. In: Falkenberg, I.E., Lindgreen, P. (eds.) Information Systems Concept: An In-depth Analysis, Elsevier Science Publishers, B.V., Amsterdam (1989)
7. IEEE STD 610.12. Standard Glossary of Software Engineering Terminology. IEEE, Los Alamitos (May 1990) ISBN: 155937067X
8. Clark, T., Jones, R.: Organisational Interoperability Maturity Model for C2 (1999), `http://www.dodccrp.org/events/1999_CCRTS/pdf_files/track_5/049clark.pdf`
9. Alter, S.: Information Systems. Cummings Publishing Company, Inc., The Benjamin (1996)
10. Solotruk, M., Kristofic, M.: Increasing the Degree of Information System Integration and Developing an Integrated Information System. Information & Management 3(3) (1980)
11. Magoulas, T., Pessi, K.: Strategic IT Management. Doctoral Dissertation Department of Informatics. University of Gothenburg, Sweden (1998) (in Swedish)
12. Hugoson, M.-Å., Magoulas, T., Pessi, K.: Interoperability Strategies for Business Agility. In: Dietz, J.L.G., et al. (eds.) CIAO! 2008 and EOMAS 2008. LNBIP, vol. 10, pp. 108–121. Springer, Heidelberg (2008)
13. Allen, B.R., och Boynton, A.C.: Information Architecture: In Search of Efficient Flexibility. MIS Quarterly 15(4) (December 1991)
14. Mathiassen, L.: Collaborative Practice Research. Scandinavian Journal of Information Systems (14), 57–76 (2002)
15. Walsham, G.: Interpretive Case Studies in IS Research: Nature and Method. European Journal of Information Systems (4), 74–81 (1995)
16. Klein, H.K., Myers, M.D.: A set of principles for conducting and evaluating interpretive field studies in information systems. MIS Quarterly 23(1), 67–94 (1999)

ILOG 2010 Workshop Chairs' Message

Wolfgang Deiters[1], Anita Mirijamdotter[2], and Kurt Sandkuhl[3]

[1] Fraunhofer ISST, Dortmund, Germany
deiters@do.isst.fhg.de
[2] Linnæus University, Växjö, Sweden
anita.mirijamdotter@lnu.se
[3] Jönköping University, Jönköping, Sweden
Kurt.Sandkuhl@jth.hj.se

For many enterprises, it is of decisive strategic importance to optimize the internal information flow and to implement an efficient reuse of existing knowledge. Especially in knowledge-intensive industry and service sectors, information is a major factor in production processes, and knowledge reflects an important asset of the enterprise. Similarly, public organizations and governmental bodies are dependent on accurate and timely information supply for efficient and high quality processes and services. Intelligent information supply has become an important issue that is characterized by just-in-time, demand-oriented and context-sensitive information.

Experience shows that successful solutions for intelligent information supply involve several ingredients: a sound business case with clearly defined benefits and returns for the (networked) organization, a clear understanding of the user's demand including the organizational context, and the right use of "enabling technologies" like semantic technologies, knowledge management or ubiquitous computing. The workshop on Information Logistics and Knowledge Supply (ILOG 2010) has the ambition to create a forum bringing together researchers and practitioners from both industry and academia with a strong interest in the above topics.

ILOG 2010 is the 3rd workshop on this subject following events in 2004 and 2007. In 2003, the seminar on "Knowledge Gap in Enterprise Information Flow" (Ljungby, Sweden) resulted in the conclusion that information logistics and knowledge supply in networked organizations was a topic interesting and attractive for both industry and academia. As a consequence the first ILOG workshop was organized in conjunction with "Multikonferenz Wirtschaftsinformatik 2004". The 2nd event in this series was organized at Jönköping University in November 2007. Both workshops had roughly 25 participants.

At ILOG 2010, we encouraged a broad understanding of possible approaches and solutions for information logistics and knowledge supply. Specific focus is on practices of, i.e. we encouraged submission of case study and experiences papers, and of contributions bringing together business cases and enabling technologies.

The workshop received 16 submissions. The program committee selected 7 submissions and agreed to invite one additional paper for presentation at the workshop.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 156–157, 2010.

Acknowledgments

We thank all members of the program committee, authors and local organizers for their efforts and support.

Wolfgang Deiters
Anita Mirijamdotter
Kurt Sandkuhl

May 2010

Ubiquitous Microblogging: A Flow-Based Front-End for Information Logistics

Martin Böhringer[1] and Martin Gaedke[2]

[1] Fakultät für Wirtschaftswissenschaften, Chemnitz University of Technology
Thüringer Weg 7, 09126 Chemnitz, Germany
martin.boehringer@wirtschaft.tu-chemnitz.de
[2] Fakultät für Informatik, Chemnitz University of Technology
Straße der Nationen 62, 09111 Chemnitz, Germany
gaedke@informatik.tu-chemnitz.de

Abstract. The success of information supply strongly depends on successful user adoption. This especially is the case for the integration of non-human information sources deriving from ubiquitous computing. To allow ordinary users to participate, there is a clear need for simple but yet powerful front-end technology. Therefore, we suggest leveraging existing and proven application patterns rather than building new concepts out of scratch. Especially Web 2.0 applications are designed for the management of millions of (human) networked information nodes and could be very useful in the context of Information Logistics. Beyond the Web 2.0 tool family, in particular microblogging could show a perfect match with Information Logistics scenarios due to its ad-hoc character and its simplicity. This paper discusses these possibilities and presents the vision of Ubiquitous Microblogging, which means a Twitter-like front-end for information from human and non-human information sources.

1 Introduction

Never has there been more information available in history of mankind than today. The Web is a huge and very fast source of knowledge and information. The world-famous example of the plane accident in the Hudson river which has been reported first not via news agencies but on Twitter [23] was only a final demonstration of the development towards a world with pervasive information availability and unprecedented publishing possibilities, which very soon will reach also the last offline areas in our towns and countryside [10].

However, especially researchers in Information Logistics have to consider the problem of information overload, which has been identified as a problem related to information systems more than three decades ago [9]. The claim 'data smog' came up more than one decade ago and it seems clear that in the meantime the amount of information noise has not been reduced [27]. A common answer to this problem is abstraction and hiding underlying information complexity from the user. However, in a near-future world with ubiquitous sensors everywhere it seems clear that 'hiding' something from users might strongly enforce the fears of an Orwellian scenario.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 158–167, 2010.
© Springer-Verlag Berlin Heidelberg 2010

Further, due to increasing complex information demands, modeling them becomes a nearly impossible task, which leads to the need for bottom-up solutions.

We argue that the key to user acceptance is transparency. Further, full access to all information is a fundamental condition for enabling collective intelligence and serendipity effects [29]. So the question and the problem which we are going to discuss in this paper is 'How can we enable full access and transparency to pervasive information without producing the ultimate data fog machine?'.

This paper presents a pragmatic and simple approach. We argue that especially in the context of ordinary users we have to leverage existing and widely well understood concepts from which we know that they do work. The Hudson River example already gave a hint to a possible solution: Twitter-like microblogging applications. They provide an extremely simple user interface to millions of information sources while not producing information overload due to the underlying publish-subscribe mechanism. Hagel & Brown [15] consider this pull principle as groundbreaking trend in information technology and business. In addition, microblogging is mobile, microblogging is near-real time and it supports collective intelligence, e.g. via its re-tweeting functionality [7].

We start our discussion with a summary of existing works and concepts for user access to Information Logistics in the context of pervasive computing. Afterwards, we provide an overview on microblogging and the status quo of research on this type of application and go on with introducing our concept of 'ubiquitous microblogging'. It will be especially discussed, what advantages microblogging qualify for pervasive information and what shortcomings and needs for further development are. After providing this conceptual overview section 4 includes a survey of technological implications of microblogging usage, before section 5 presents our research agenda towards realising the vision.

2 Information Logistics, Ubiquitous Computing and the User

The need for pervasive information sources to be public available is a known issue. Therefore, past works focused on opening pervasive computing to a wider audience. However, many of the existing concepts still have IT experts in mind and aim to create a programming middleware rather than a user front-end ([24], [17]).

Though, from a theoretical point of view we already have a clear idea of suitable user interfaces. Edwards & Grinter [12] argued that user front-ends have to be extremely simple and should hide the underlying technical complexity, especially if they are aimed to support collective intelligence and rich interaction with all kinds of data. This is important both for consuming and producing as "the simplicity of sharing this information is key" and "easy data discovery, subscription, and republication will be crucial" [10]. This clearly leads to Web 2.0 tools – in the sense of read/write Web tools, which can be used by everyone as it describes the typical use case for social applications.

Without knowing microblogging, early in 2006 sensorbase.org has been presented which uses 'slogging', an analogy for blogging, as mechanism for accessing sensor data [8]. This contribution meets most of the discussed objectives for a user-ready front-end for pervasive Information Logistics. However, especially for the re-use and re-publication objective, there is a lack of critical mass and the integration with other information sources. Therefore we suggest taking into account recent developments and best practices for 'blogging' small and time-based information pieces. The best example to learn from might be the million-user platform Twitter, which introduced the concept of microblogging. Although being a very new technology, microblogging has already been applied to enterprise scenarios ('enterprise microblogging'). Analysts forecast a rapid distribution of such tools into existing information systems [14]. Besides of numerous products for stand-alone enterprise microblogging, there are already vendors of 'classic' enterprise software integrating microblogging as mechanisms for information allocation, e.g. Salesforce Chatter.

3 Ubiquitous Microblogging

3.1 Twitter-Like Applications

The principle of microblogging is best known via Twitter, its most famous application. Users have their own public microblog where they can post short updates. Other members can be 'followed' by adding them to one's personal network. An aggregated view of all updates by followed microblogs appears in chronological order on the user's start page. Microblogging services often support a wide range of contribution possibilities. For example, messages to Twitter can be posted via mobile text messages, desktop clients or several third-party applications.

Microblogging can be considered a world-wide trend today. While early adopters mostly came from America and Japan [22], other parts of the world followed soon later. In Germany, for example, the first broad media coverage of Twitter could be observed during the civil unrest in Iran in summer 2009 [3]. Compared with other social networking services like Facebook, Twitter still has a small user base but its basic principles should be known by a broad range of people today.

Parallel to the adoption beyond internet users and interestingly faster than previous Web 2.0 applications, microblogging in general became a subject of research. Due to its unique approach, researchers from different disciplines are interested in the topic. Two kinds of research works relevant for our research context can be identified in the existing body of knowledge. First, a large number of works searches for explanations of microblogging usage. A significant group focuses on statistical descriptions of Twitter (e.g. [18], [21], [20], [19]).

The most notable result of these works is evidence for the user-driven shaping of the microblogging technology. Several works describe that Twitter users only partly answer the service's question 'What are you doing?' but leverage Twitter for information sharing, conversations and other usage patterns. In the same context belongs the user-driven feature enhancement. The well-known @-reference was a short code originated by the users as well as the re-tweeting codex [21]. Twitter added support for this text code-based functionality in its front-end later on. Another group of research works in the first category discuss microblogging from a social point of view. Several researchers discussed the phenomena as tool for building awareness, which is a concept with lots of implications on pervasive computing [13].

The second class of research on microblogging is technology-related especially addressing shortcomings of Twitter. These shortcomings are a lack of semantics [25] as well as the central and proprietary infrastructure [26]. Especially interesting and a great approach for an intuitive access to pervasive information is the concept of visual microblogging presented in [2]. Visual microblogging means the enrichment of the textual medium microblogging with graphical information where possible. In a context of sensor data this could be diagrams or sparklines. Such visualisation could provide context in the flow of microblogging information where only new postings stand on top and are soon displaced by newer ones.

3.2 Terminology

The concept of 'ubiquitous microblogging' leans on the well-known term of 'ubiquitous computing'. The latter understands ubiquity in the sense that artificial computing sensors and devices are everywhere integrated into everyday things and tools in the real world [1], therefore microblogging as front-end to these devices must be 'ubiquitous', too. In our definition, *ubiquitous microblogging means a microblogging system, which enables information streams of and between human and non-human information sources within the scope of the system.*

Weiser [30] in his vision of ubiquitous computing stated that 'the most profound technologies are those that disappear'. Figuratively speaking, this is also true for ubiquitous microblogging as the goal behind this approach is to hide the complexity of information access in the real world by providing a flat information space accessible by the easy following-mechanism. This leads to the objectives of a solution for ubiquitous microblogging: following the core microblogging approach it should support a maximum of ease of use with nothing more necessary for usage than to understand the natural principle of following and un-following other people or information sources. The main task for an information system for ubiquitous microblogging therefore is to provide this simple access to manifold information sources in the required information space (i.e., an organisation) and make the underlying complexity 'disappear'.

3.3 Conceptual Description

Microblogging structures information towards its origin. While this approach seems to be rather natural (speaking humans are a trivial example; books are labelled with their author), it is surprisingly quite uncommon in the context of IT systems which are normally organised towards taxonomies or folksonomies. It can be assumed, that especially this unique mechanism is a strong reason for the success of blogging and microblogging, with microblogging providing even more interaction possibilities between the information streams [5]. Therefore, we suggest adopting this principle and using *one microblog per information source* respectively sensor.

An open point is how it can be assured that the usage of a simple text-based medium does not reduce information quality. Furthermore, if pervasive data sources use ubiquitous microblogging, it is an obvious approach to use it not only for human end users but as a general middleware for mashups and automated bots. However, to fit this scenario, posting content should be accessible for non-human actors which would presume the possibility for automated understanding.

Figure 1 shows different approaches of how to transfer sensor data into microblogging postings. In 1a raw sensor data is posted regularly in the same structure. This might be easy to be automatically interpreted. 1b shows a microblog where the numbers are interpreted and turned into natural language sentences before being posted. And in 1c the raw data (time) is being converted into, well, nearly useless 'bongs'. The point is, that even if postings are well structured like 1a, nonetheless it has to be crawled, which needs manually programming. Automated discovery and composition of microblogging data would need some kind of semantics.

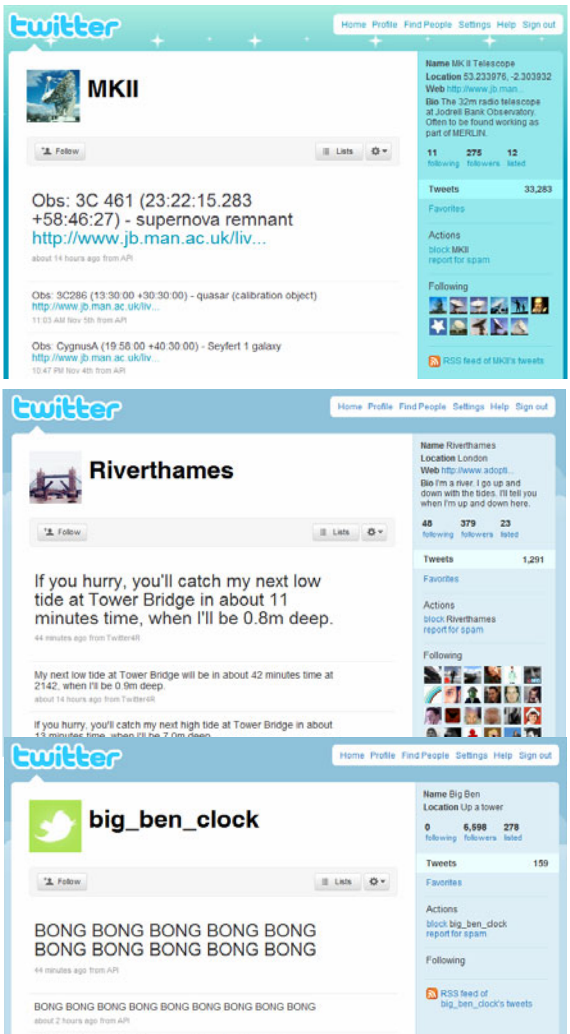


Fig. 1. Information encoding in Twitter (a-c)

Given the availability of sensor microblogs, their management could use standard microblogging functionality. Aggregation could be done by bots (e.g. looking for strange sensor data and re-tweeting them) which could be followed by end users. With its new list feature, Twitter shows another interesting approach for organising such microblogs.

While we discussed above that the information-source-oriented information struturing of microblogging differs from taxonomies/ folksonomies, the latter is also possible in microblogging using the user-invented hashtags. The concept of hashtags is to put a '#' in front of a word in the posting, which can be seen as tag. Therefore this is an interesting alternative structuring possibility. A good example for hashtag-structuring is their use at conferences, e.g. "#BIS10" [11].

Microblogging includes further standard features which could be used in Information Logistics scenarios. Figure 1a shows the usage of GPS coordinates in the location meta data. The networking functionality (following/un-following) can map sensor networks including hierarchical structures. The standard link in the microblog's profile can lead to static information. And, finally, especially around Twitter and similar services (like identi.ca) exists a rich spectrum of third-party-applications and clients.

To sum up, microblogging comes with a simple and small but yet powerful set of functionality which answers a wide range of the requirements for ubiquitous computing front-ends. Its concept is known to many users; it is very easy to learn and – in the case of Twitter – comes with an existing critical mass of users.

4 Technology

Since Twitter's launch in 2006, there has been remarkably little development on the platform from a user's point of view. Twitter leaves functional improvements to a large number of third party suppliers, which are encouraged by the open and rich Twitter API. Even such core functionality as an effective search has not been included in Twitter until 2008, when the microblogging service bought a third-party search application. While Twitter develops with a 'trial and error' strategy on the costs of its third-party programmers, a system for ubiquitous microblogging has to provide a reliable application stack right from the beginning in order to guarantee core functionalities like security and to achieve a sustainable adoption. There are open research issues at all levels of this stack.

Beginning with the user front-end, there is significant potential in supporting users in searching the 'haystack' for important information. For Twitter there exist a number of small applications for visualisation and text mining. However, they are prototype implementations and are not integrated to each other or well-documented. From a scientific point of view, little research has been done in this area.

Further, the user would benefit from support by automated agents understanding the semantic meaning of the information. For this reason, Passant et al. [25] present a concept for semantic microblogging. However, the main question for the enrichment of text with semantic information is user acceptance, which is an important point in microblogging as a medium that is mainly based on the ability to publish information fast and easily. An alternative to semantic markup of microblogging postings is to understand the meanings of natural language using Natural Language Processing (NLP). However, it can be questioned how effectively NLP could work on the very short, often informal and possibly flawed microblogging postings. First experimental prototypes have demonstrated the basic feasibility of the approach (see, for example, the service akibot.com).

An important research issue is the realisation of a suitable architectural concept for ubiquitous microblogging. We have already discussed the possibilities of a centralised and decentralised option. Passant et al. [25] argued for a decentralised architecture, so did Sandler and Wallach [26]. One can intuitively agree with the advantages of a decentralised RESTful architecture, since it is proven in the WWW and the disadvantages of Twitter's centralistic approach are publicly visible with server failures and, in particular, DDoS attacks, which regularly lead to closing of the platform.

However, it is questionable whether an architecture similar to the blogosphere with self-hosted, distributed microblogs allows the fast and easy connectivity that is typical for Twitter. Twitter enables addressing of other microblogs by using the syntax

@<microblogname> (e.g. @boehr). Without a centralistic application stack or at least a centralistic directory, this simple addressing might not be possible. In this case, full URI-based addressing would have to be used (basically, @boehr can be seen as a abbreviation for http://www.twitter.com/boehr which is possible due to its centralistic approach). Due to the objective of ease of use, it can be questioned whether a URI-based addressing mechanism would be accepted by the users. Further, a suitable architecture has to integrate with existing Information Logistics applications and therefore would have to be designed with respect to existing standards.

A technical challenge of ubiquitous microblogging is not only the storing of its data but also the communication layer. A central characteristic of microblogging is its real-time approach. In a ubiquitous microblogging system with thousands of globally distributed users and tens of thousands of non-human information sources, this is a special requirement. Recently, there has been some development in this area with a number of approaches aiming to combine advantages of a persistent storage (with the user's possibility to pull information) and active real-time notification (push principle). Such recent initiatives are Pubsubhubhub, Simple Update Protocol and RSSCloud. However, none of them has been tested in the described scenario.

5 Future Work

The presented concept of Ubiquitous Microblogging as front-end for multiple information sources is a research vision. The short discussion of open issues shows that there are lots of open questions on the way towards its realisation. Therefore, we have two main tasks in our research agenda.

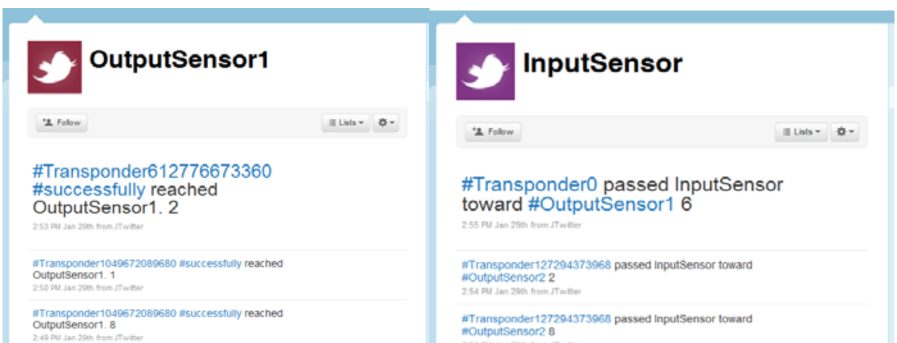


Fig. 2. Prototype of tweeting RFID sensors

First, we are going to test the vision in different use cases. We are currently running projects with RFID sensors in a LEGO-based factory and Arduino-based mobile sensors (see figure 2). Further, in laboratory experiments we examine microblogging's capabilities of supporting emergent ad-hoc processes which include human and non-human postings (see figure 3). In addition to that, in case studies and experiments in Twitter and enterprise microblogging applications we are trying to understand how microblogging is applied at the user side (e.g. [4]).

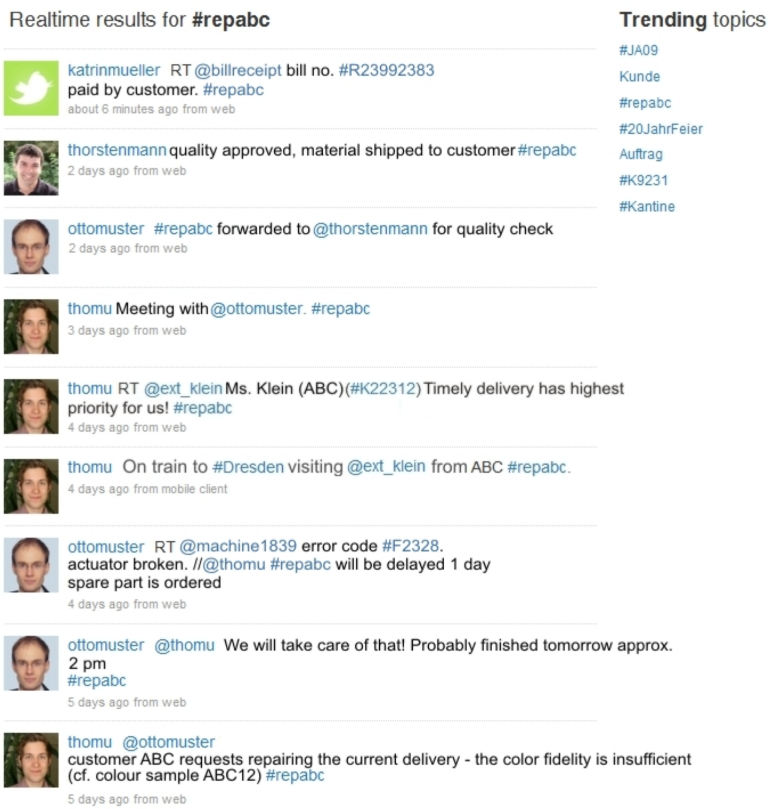


Fig. 3. Monitoring of ad-hoc processes via hashtag usage

Second, we are working to technological foundations for Ubiquitous Microblogging. Our key project in this area is the REST-based Data Grid Service (DGS) as backbone for the integration of multiple information sources ([16], [28]).

6 Conclusions

This paper presented an overview of Ubiquitous Microblogging, a research concept which leverages Twitter-like functionality as pragmatic solution for information allocation. The approach could provide a light-weight front-end for Information Logistics. However, an implementation of the approach still faces some conceptual and technical problems, which have been described in more detail in [6].

The community around the idea of Ubiquitous Microblogging is organised around http://ubimic.org. We encourage other researchers to contribute and discuss their ideas there or on Twitter using the Hashtag #ubimic.

References

1. Abowd, G.D., Mynatt, E.D.: Charting past, present, and future research in ubiquitous computing. ACM Transactions on Computer-Human Interaction 7(1), 29–58 (2000)
2. Assogba, Y., Donath, J.: Mycrocosm: Visual Microblogging. In: Proceedings of the 42rd Hawaii International Conference on System Sciences, HICSS-42 (2009)
3. Barnes, S.J., Böhringer, M.: Continuance Usage Intention in Microblogging Services: The Case of Twitter. In: Proceedings of the 17th European Conference on Information Systems (ECIS), Verona, Italy (2009), `http://www.ecis2009.it/papers/ecis2009-0164.pdf`
4. Barnes, S.J., Böhringer, M., Kurze, C., Stietzel, J.: Towards an understanding of social software: the case of Arinia. In: Proceedings of the 43rd Hawaii International Conference on System Sciences (HICSS-43), Koloa, Kauai, Hawaii, January 5-8 (2010)
5. Böhringer, M.: Really Social Syndication: A Conceptual View on Microblogging. Sprouts: Working Papers on Information Systems 9(31) (2009), `http://sprouts.aisnet.org/9-31/`
6. Böhringer, M., Gluchowski, P.: The Beauty of Simplicity: Ubiquitous Microblogging in the Enterprise. In: Proceedings of the 12th International Conference on Enterprise Information Systems (ICEIS), Funchal, Madeira – Portugal, June 8-2 (2010)
7. Boyd, D., Golder, S., Lotan, G.: Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. In: Proceedings of the 43rd Hawaii International Conference on System Sciences (HICSS-43), Koloa, Kauai, Hawaii (2010)
8. Chang, K., Yau, N., Hansen, M., Estrin, D.: Sensorbase. org-a centralized repository to slog sensor network data. In: Proceedings of the International Conf. on Distributed Networks (DCOSS)/EAWMS (2006)
9. Chervany, N.L., Dickson, G.W.: An Experimental Evaluation of Information Overload in a Production Environment. Management Science 20(10), 1335–1344 (1974)
10. Cuff, D., Hansen, M., Kang, J.: Urban Sensing: Out of the Woods. Communications of the ACM 51(3) (2008)
11. Ebner, M., Reinhardt, W.: Social networking in scientific conferences – Twitter as tool for strengthen a scientific community. In: Proceedings of the 1st International Workshop on Science 2.0 for TEL at the 4th European Conference on Technology Enhanced Learning, EC-TEL'09 (2009)
12. Edwards, W., Grinter, R.: At home with ubiquitous computing: Seven challenges. In: Abowd, G.D., Brumitt, B., Shafer, S. (eds.) UbiComp 2001. LNCS, vol. 2201, pp. 256–272. Springer, Heidelberg (2001)
13. Erickson, I.: The Translucence of Twitter. In: EPIC 2008, Ethnographic Praxis in Industry Conference (2008)
14. Gartner: Gartner Reveals Five Social Software Predictions for 2010, and Beyond (2010), `http://www.gartner.com/it/page.jsp?id=1293114`
15. Hagel, J., Brown, J.: From push to pull—emerging models for mobilizing resources. Journal of Service Science, Third Quarter 1(1) (2008)
16. Heil, A., Gaedke, M.: WebComposition/DGS: Supporting Web2.0 Developments With Data Grids. In: IEEE International Conference on Web Services (ICWS 2008), pp. 212–215 (2008)
17. Holloway, S., Stovall, D., Lara-Garduno, J., Julien, C.: Opening pervasive computing to the masses using the SEAP middleware. In: Proceedings of IEEE International Conference on Pervasive Computing and Communications (2009)

18. Honeycutt, C., Herring, S.C.: Beyond Microblogging: Conversation and Collaboration via Twitter. In: Proceedings of the 42nd Hawaii International Conference on System Sciences, HICSS-42 (2009)
19. Huberman, B.A., Romero, D.M., Wu, F.: Social networks that matter: Twitter under the microscope. First Monday 14(1) (2009), http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2317/2063
20. Hughes, A.L., Palen, L.: Twitter Adoption and Use in Mass Convergence and Emergency Events. In: Proceedings of the 6th International ISCRAM Conference, Gothenburg, Sweden (May 2009)
21. Java, A., Song, X., Finin, T., Tseng, B.: Why we twitter: understanding microblogging usage and communities. In: WebKDD/SNA-KDD '07: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis (2007)
22. Krishnamurthy, B., Gill, P., Arlitt, M.: A few chirps about twitter. In: WOSP '08: Proceedings of the First Workshop on Online Social Networks, pp. 19–24 (2008)
23. Krums, J.: Status Posting 1121915133, There's a plane in the Hudson (2009), http://twitter.com/jkrums/status/1121915133
24. Luckenbach, T., Gober, P., Arbanowski, S., Kotsopoulos, A., Kim, K.: TinyREST: A protocol for integrating sensor networks into the internet. In: Proceedings of REALWSN (2005)
25. Passant, A., Hastrup, T., Bojars, U., Breslin, J.: Microblogging: A Semantic Web and Distributed Approach. In: Proceedings of the 4th Workshop on Scripting for the Semantic Web (2008), http://ceur-ws.org/Vol-368/paper11.pdf
26. Sandler, D.R., Wallach, D.S.: Birds of a FETHR: Open, Decentralized Micropublishing. In: 8th International Workshop on Peer-to-Peer Systems, IPTPS '09 (2009)
27. Shenk, D.: Data smog. MIT's Technology Review 100(4) (1997)
28. Sommermeier, R., Heil, A., Gaedke, M.: Lightweight Data Integration using the WebComposition Data Grid Service. In: First International Workshop on Lightweight Integration on the Web (Composable Web'09) in Conjunction with the 9th International Conference on Web Engineering (ICWE 2009), pp. 30–38 (2009)
29. Vinoski, S.: Serendipitous reuse. IEEE Internet Computing 12(1), 84–87 (2008)
30. Weiser, M.: The Computer for the 21st Century. Scientific American 265(3), 94–104 (1991)

Information Reuse and Interoperability with Ontology Patterns and Linked Data

First Experiences from the ExpertFinder Project

Karl Hammar, Feiyu Lin, and Vladimir Tarasov

School of Engineering at Jönköping University
P.O. Box 1026
551 11 Jönköping, Sweden
{haka,life,tavl}@jth.hj.se

Abstract. Semantic web technologies show great promise in usage scenarios that involve information logistics. This paper is an experience report on improving the semantic web ontology underlying an application used in expert finding. We use ontology design patterns to find and correct poor design choices, and align the application ontology to commonly used semantic web ontologies in order to increase the interoperability of the ontology and application. Lessons learned and problems faced are discussed, and possible future developments of the project mapped out.

Keywords: semantic web, ontology design patterns, ontology alignment, linked data.

1 Introduction

The semantic web extends the current web, using semantic web ontologies to solve the problem of defining common meanings of terms and concepts. This enables automatic integration of information between different platforms and information systems utilizing the common definitions. Furthermore, it allows for reasoning agents to collate such information and from it infer knowledge that is greater than the sum of its parts.

In the context of information logistics the benefits of semantic web technologies are many. By allowing users to query for concepts rather than keywords, the precision of information queries can be substantially improved. Common concepts allow for services to be created in which users can query for information, rather than simple data, originating from a multitude of sources, interlinked using standardized vocabularies.

The potential gains of using semantic web technologies for information logistics purposes are not limited to enterprises or large organizations. In [1] a picture is painted of a future in which the use of semantic web technologies can realize ubiquitous computing scenarios, allowing individuals immediate access to information they need in order to plan and execute their every day life with a minimum of fuzz.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 168–179, 2010.
© Springer-Verlag Berlin Heidelberg 2010

The contribution of this paper is an experience report on applying ontology design patterns and connecting existing linked data sources when developing a semantic web ontology for use in an expert finding system. The paper offers a perspective on two aspects of reuse in ontologies: reusing modeling knowledge via design patterns, and designing for reuse by aligning to commonly used standard ontologies.

The rest of this paper is structured as follows. First, we introduce the theoretical background of the areas and the ExpertFinder project. In section 3, we discuss how reuse is actually performed within our project. Our experience and lessons learnt from this are presented in section 4. Section 5 summarizes the work and gives an outlook on future work.

2 Background

In this section, we discuss the theoretical background of the technologies and methods touched upon in the paper, and the project to which these were applied.

2.1 Linked Data Interoperability

Linked data refers to data and information that has been published online using semantic web standards (ontologies and semantically marked up XHTML). For such linked data to be interoperable and truly form one "web of data" rather than multiple incompatible "islands of data", the ontologies used for defining concepts need to be aligned to one another. Automatically matching and mapping ontology concepts is an unsolved problem as yet, meaning that ontology alignment is still a rather labor intensive process if one does not design one's ontologies with interoperability in mind from the very beginning.

If, however, one can align with the commonly used ontologies of the semantic web, the potential interoperability and integration benefits achieved without specialized integration software or technologies can be significant. Some examples of semantic web integration and interoperability are mentioned in [2], such as the integration of the various types of heterogeneous datasets in the life sciences, and the enabling of exploitation of public data in different formats by citizens. Another example including the manufacturing industry and efficient supply chain management is discussed in [3].

2.2 Ontology Design Patterns

Ontology design patterns have been proposed (by for instance [4]) as a method of simplifying ontology development. The term "ontology design pattern" has multiple definitions in literature. In this paper we adopt the definition used for the similar term *ontology pattern* in [5], which states:

An ontology pattern is a set of ontological elements, structures or construction principles that intend to solve a specific engineering problem and that recurs, either exactly replicated or in an adapted form, within some set of ontologies, or is envisioned to recur within some future set of ontologies.

Recently a lot of work has been done on such patterns, both with regards to construction and to the various uses of them. A general method for constructing and reusing content ontology design patterns are presented in [6]. The use of a web portal to hold a repository of patterns is presented in [7], an idea that has been realized in the creation of the site OntologyDesignPatterns.Org. Ontology patterns and case-based reasoning is used in [5] to perform semi-automatic construction of ontologies from input texts.

Experiments on using ontology design patterns are detailed in [8], which concludes that the use of patterns improve the quality and coverage of the resulting ontologies, and that they help modelers avoid common mistakes.

2.3 The ExpertFinder Project

ExpertFinder is an internal research project at Jönköping University. The driving motivation behind the project is the possibility of using semantic technologies to help university staff in finding experts within various fields. Potential experts are to be found among all of the university researchers and teachers, whose profiles are stored in a semantic web ontology. Two main usage scenarios have been used to focus the project:

- **Internal Expert Finding.** A course manager needs a guest lecturer on IT in health-care and wants to find the best suited person in the university. The course manager writes a simple keyword query, which is matched against the ontology, and information about the persons related to the matching concepts are retrieved. Rules are used for ranking the results, for instance weighing prior teaching experience in the subject higher than being a co-author of a paper remotely concerned with the subject.
- **External Expert Finding.** A company needs help to understand how to setup an ad-hoc sensor network in their facilities. A company representative can use the system for finding relevant competencies in the university. The user writes a simple keyword query, and if terms are not present in the ontology, semantically similar terms are suggested by the system. If no ontology term matches a certain keyword, the keyword is used by a standard IR system, as a complement to the ontology-based query. To help the user to see if the experts are relevant for the task, a motivation of why they were retrieved is also presented in the ranked result list.

The ExpertFinder ontology is implemented in the OWL language. Both the data import logic that populates the ontology and the query interfaces for accessing it are implemented in Java, and the intention is to deploy the latter as a web application on the university web site in the future.

3 Reuse in the ExpertFinder Ontology

This paper deals with reuse in the design of the ExpertFinder ontology. Apart from this, software components and frameworks have also been reused successfully within the actual application that populates and queries the ontology, but such reuse is considered to be outside of the scope of this paper.

During the development of the ExpertFinder ontology, the methodology for ontology construction proposed in [9], was followed. This methodology consists of four distinct phases: requirements analysis, building, implementation, and evaluation and maintenance. The ontology is in accordance with this methodology being developed iteratively.

The ontology resulting from the first iteration did not make use of any ontology patterns, and was not aligned to any standard semantic web ontologies. For the purpose of increasing interoperability and compatibility as well as finding suboptimal design choices made in the first version, these features were implemented in a second iteration. It is primarily the experiences in performing this second iteration that this paper covers, although the section on content reuse refers to the initial design.

In the current version, the domain has been restricted to researchers and teachers of the department of computer and electrical engineering. However, the ambition is that the general design should be general enough that future iterations will be able to accommodate knowledge about staff, projects, courses and programs of all schools and departments of the university.

3.1 Linked Data Use

In the initial iteration, ontology search engines such as Swoogle, Ontosearch, and Watson were used to find academic ontologies describing universities and research. However, after analyzing the ExpertFinder requirements, these ontologies were deemed too complex and not suitable for the project, and the initial construction was therefore performed from a blank slate without any reuse. When the second iteration began, the first task was thus to find semantic web ontologies to align that had a reasonable scope for this project, i.e. that were general enough to further interoperability but still specific enough to be applicable to the application. Also, since aligning ontologies can be rather time consuming, the ontologies found had to be easy to tie in to the first iteration version.

BIBO. The first such ontology that was found to be appropriate was the Bibliographic Ontology, commonly known as BIBO ([10]). Since the ExpertFinder ontology covers scientific publications, standardizing how such publications are represented makes a lot of sense. It further allows easier implementation of import from public data sources that use the BIBO standard (including for instance the US Library of Congress). Implementing BIBO compatibility consisted of importing the ontology and connecting the defined ExpertFinder classes and properties to their BIBO equivalents by defining them to be linked via owl:equivalentClass statements. Many of the concepts were found to map easily in this manner, but some were not found to have any natural counterparts in BIBO. In these cases, the most common superclass was mapped to instead, providing at least some interoperability. For instance, the concept "Conference Papers" had no direct equivalent, and was therefore mapped as a subclass of bibo:Document which is a more general term.

Implementing BIBO compatibility enabled us to notice and fix shortcomings in our own initial design. One such shortcoming was that our ontology made frequent use of datatype properties to define values of publications where using object properties could provide better reasoning capacity and semantics. To take one example, the term "peer reviewed conference paper" was often used as a comment or status string for individual publications, whereas BIBO instead models publication review statuses semantically, making it possible to query for all peer reviewed items in the knowledgebase in a much simpler manner.

FOAF. BIBO itself depends upon the Friend-of-a-Friend vocabulary, FOAF ([11]), that was also imported. FOAF provides a standard way of talking and reasoning about people, groups and organizations. This is of course useful in many situations, and very appropriate for a system such as ExpertFinder that deals primarily with finding people. FOAF compatibility was implemented in the same manner as BIBO, by connecting existing classes and properties to their equivalents or nearly related concepts in the foaf: namespace. We were in this manner able to connect both our concepts of experts (as persons) and universities (as organizations) to standard classes. For a listing of the classes (not including object or data properties) aligned in this manner, see table 1.

Table 1. ExpertFinder classes mapped to BIBO and FOAF

FOAF/BIBO Class	ExpertFinder Class	Mapping relation
bibo:Document	ef:ConferencePaper	Subclass of
bibo:Chapter	ef:BookChapter	Equivalent
bibo:AcademicArticle	ef:JournalArticle	Equivalent
bibo:Report	ef:Report	Equivalent
foaf:Organization	ef:UniversitySchool	Subclass of
foaf:Person	ef:Expert	Subclass of

W3C Time. The third standard ontology that was aligned to was the W3C Time ontology originally presented in [12]. While this standard is formally listed as being retired, it is nonetheless still in active use in a number of ontologies, including BIBO. Discussing time is something that falls within the domain of ExpertFinder (for instance, modeling the times when a certain expert participated in a certain project or course), and being able to do it in a standardized interoperable way motivates the use of this standard, in spite of its retired status.

Aligning to W3C Time required more work than aligning to BIBO and FOAF, since the previous date representation was not compatible with the imported time: namespace definitions. New properties had to be defined, using the W3C Time definitions, and all existing time references from projects, courses, etc. in the ontology had to be manually updated to use the new properties. Since the amount of data in the ontology was relatively small this was not a huge undertaking, but for larger ontologies such a refactoring could be quite time consuming if not possible to automate.

3.2 Pattern Use

Having aligned to some common ontologies, the next task of the second iteration of development was to try to implement ontology design patterns. While the motivation of using such patterns often are to speed up initial development of ontologies from scratch, in this case the motivation was rather to find bad design choices made in the first iteration.

In order to find applicable patterns the OntologyDesignPatterns.org repository was studied. A number of patterns matching needs in the ExpertFinder domain were identified, and out of these a subset were implemented in the ontology, solving design issues that existed. The patterns that were selected, based on ease of applicability and description quality, were:

- **Information Realization**[1]
  One key concept in the ExpertFinder domain is courses. Courses can be held multiple times, each time having different teachers, participants, schedules, etc. but still conceptually being the same course, giving the same number of study credits, covering the same topics, and having same course description. In the initial version of the ExpertFinder ontology, the differentiation between a course as such a static concept and a course as a recurring event was not modeled. This led to an unnecessary duplication of information, as each time the course was run was modeled as a whole new course. This design also makes it unnecessarily difficult to query for information that involves multiple runs of the same course. Implementing the Information Realization pattern solved this problem. While Information Realization is documented as being intended to model information objects and their physical realizations (i.e. a written work and the actual printed book), it works equally well when describing a concept and the performance or implementation of that concept, in this case a course and the actual instance of that course.
- **Partition**[2]
  The partition pattern is a logical pattern published in [13] that models a concept that is composed of several disjoint subconcepts. In terms of OWL and classes, it can be useful as a way of dividing a classification into subclassifications. In the ExpertFinder ontology, two instances of this pattern were applied, dividing the Project class into various types of projects (EU Project, National project, Industry Project, etc.), and dividing the ScientificDegree class into fitting subclassifications (BachelorDegree, MasterDegree, PhDDegree, etc.). Previously this same information about the classification of projects and scientific degrees were present in the ontology, but only as simple data properties that could not be as easily reasoned about.
- **Nary Participation**[3]
  The nary participation pattern models time indexed participation of objects in events. In the ExpertFinder domain, the participation of teachers

[1] `http://ontologydesignpatterns.org/wiki/Submissions:Information_realization`

[2] `http://ontologydesignpatterns.org/wiki/Submissions:Partition`

[3] `http://ontologydesignpatterns.org/wiki/Submissions:Nary_Participation`

in courses or experts in projects match this pattern very well. However, the available reusable pattern building block makes extensive use of other design pattern building blocks from the OntologyDesignPatterns.org repository (events, situations, time intervals, etc) that are not compatible with the previously aligned semantic web ontologies. The general idea of the Nary Participation pattern was thus deemed appropriate for ExpertFinder, but the implementation details of the pattern as it was available made it less well suited. For this reason, the pattern implementation was partially recreated based on concepts already existing within or imported through the aligned ontologies, as displayed in figure 1. This was a successful design and provides the application with the possibility of modeling time indexed expert participation in projects, which is suitable for projects with extensive run times or different phases.

3.3 Content Reuse

One key requirement developed in the first iteration was the need to model fields of education and research areas. Three existing classifications were found that could potentially be reused: the international standard classification of occupations (ISCO-08)[14], the Eurostat classification of fields of education and training (1999) [15], and the ACM computing classification system (1998) [16]. Since the ISCO and Eurostat classifications are not detailed enough, the ACM classification was selected. [17] is a topic hierarchy in rdfs based on the ACM computing classification.

However, the ACM computing classification system has not been updated since 1998 and thereby lacks some important concepts, such as the semantic web domain. Therefore, we introduced an extended taxonomy of semantic web topics [18] (available as a SKOS vocabulary and also available in OWL as an extension of the SWRC ontology) into the ExpertFinder ontology.

The knowledge content of the ontology, the actual facts used to fill the knowledgebase, are imported from the data sources available at the university and in this sense reused from personal web pages (experts' profiles), DiVA (the database for storage of research publications and student theses used at many Swedish universities), Neverlost (the software used for class scheduling), course syllabi, and project description spreadsheets. Since these data sources are all published in different formats, various tools for importing the information needed to be constructed. Since common keys were available to correlate the data in many cases (login names, email addresses, course codes, etc) little integration efforts were necessary, and since the common format of an OWL ontology was already defined, no specific information integration metalanguage needed to be implemented, leading these tools to be relatively simple.

4 Experiences

This section reports the experiences learned from the ExpertFinder project with regard to information modeling and reuse in ontologies.

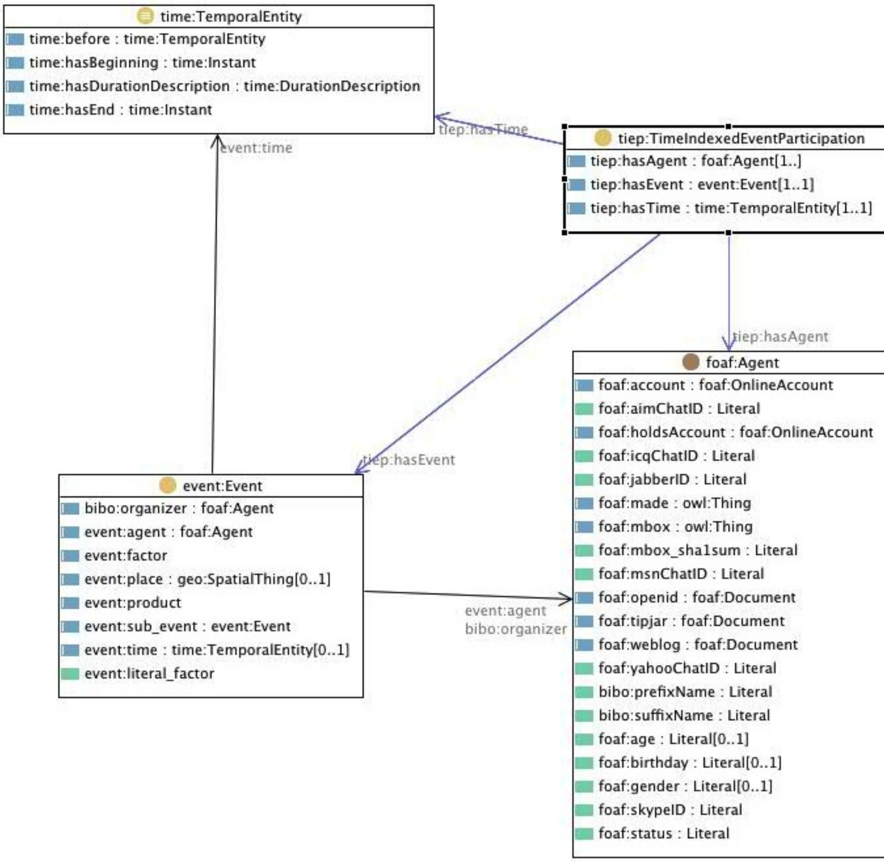


Fig. 1. Alternate implementation of Nary Participation pattern

The modifications resulted in changes to six former top level classes within the class subsumption hierarchy, and the addition of 13 new classes at various levels of the hierarchy. Compared to the 190 classes in the default ontology namespace this is rather few (3,5 %) and the impact of the performed changes may therefore be interpreted as being very small. However, the changes performed concern key concepts within the ontology, making this low number somewhat misleading - many of the 190 classes are in fact research classifications imported as described in section 3.3, having less importance for the overall structure and reasoning capacity of the ontology. The changes made to the properties hierarchy are likely more indicative of the impact of changes performed: 42 % of all properties within the ontology have been changed either in range, domain, or inheritance, indicating a significantly larger change.

4.1 Ontology Design Patten (ODP) Experience

The way of documenting and providing the patterns in the ODP portal is sound, although a lot of patterns currently listed there have rather poor documentation. Our perception is that we actually could save time and increase quality of our ontology by using ODPs. However, patterns related to our area of work, in particular content patterns in the Academy or University domains, do not exist in the ODP portal. Instead we partly reused the ACM topic and semantic web topics hierarchies, which are not patterns in themselves, but provide commonly agreed upon vocabularies within the domain. This improved the quality of our ontology and saved time but to substantiate this statement we need to do more experiments and consider more patterns. We also experienced a lack of tool support for integrating ODPs into an ontology.

One fault in the design of the ontology that has been experienced when applying the partition pattern is the frequent occurrences of cases when strings are used to describe what class a certain individual belongs to, when in fact it would make more sense to model this using the subsumption hierarchy. This was common enough in the ExpertFinder ontology to warrant further study, perhaps being indicative of an antipattern.

Another problem that was experienced in applying the patterns from OntologyDesignPatterns.org is the situation that arose with the Nary Participation pattern, that is, that while the basic idea of a certain pattern is sound, the implementation provided as a reusable OWL building block is not compatible with the ontology, leading one to have to reconstruct the implementation. This has in our case only been a minor problem. If more mapping patterns were developed to connect the concepts within the pattern repository to other well used semantic web ontologies, the occurrences of this type of problems could be minimized.

4.2 Experiences in Aligning Ontologies

The experience of aligning ontologies reinforces our opinion that aligning at the outset is probably a more efficient solution than doing so at a later time, since doing the latter requires both finding the appropriate ontologies and appropriate ways of connecting them, which can necessitate a lot of refactoring of the design. In the ExpertFinder case only a certain subset of the features were aligned in this iteration, due to the complexity of the task.

On the other hand, one important experience is the truth that lies in the commonly used expression that even a little semantics goes a long way - even though the ExpertFinder is not perfectly compatible with every imaginable web ontology, it does at this time allow for querying using BIBO and FOAF syntax, allowing automatic integration of bibliographic and personal information via linked data. This is quite a feat compared to traditional relational database-backed applications or SOA web service type solutions that tend to necessitate quite a lot of forethought and design in order to achieve good interoperability.

Finding the right semantic web ontologies to use was also experienced to be a somewhat tricky task. Since the question of objective quality metrics for ontologies is still an open issue, there are no simple easily quantifiable ways of measuring what is a good ontology and what is not, when picking possible ontologies to work with. Instead, one has to look at factors like how well the ontology in question covers the problem space, how well it is used, and how much community backing it has.

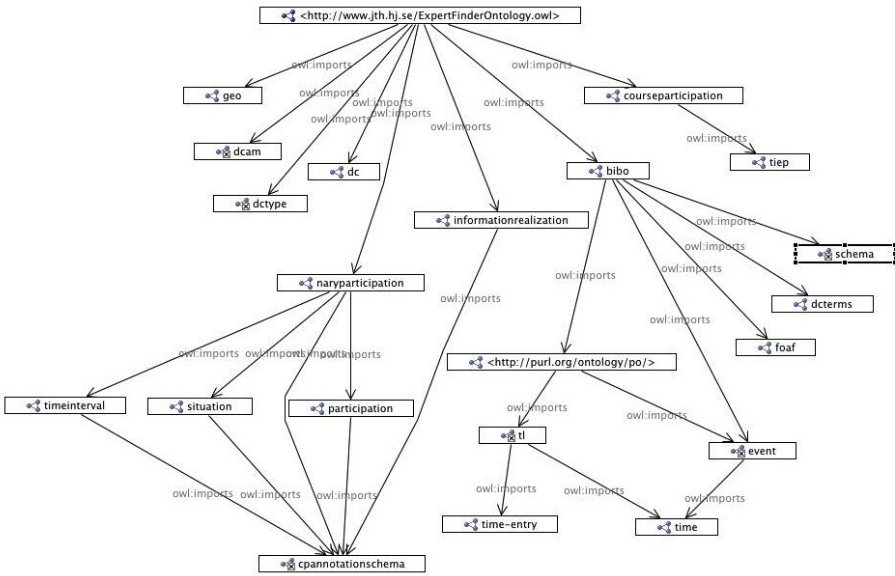


Fig. 2. ExpertFinder ontology import graph. Note that this figure is intended to illustrate the size and complexity of the import closure graph, not what individual ontologies are imported and labelled.

One problem that was experienced in performing alignment was that a lot of semantic web ontologies import and reuse other base ontologies. Since the OWL language has no feature for performing partial imports, and since imports are transitive, this gives the result that the sum total of concepts and properties within the import closure of one's ontology rapidly becomes very large, as displayed in figure 2. This growth in complexity makes alignment more difficult as the possible choices of how to align concepts also grows very rapidly, and understanding the pros and cons of each choice requires extensive knowledge of all of the imported ontologies and how they are designed, not just the first imported one. Another problem with this situation is that the ontology development tools available are not well suited for dealing with this situation when very large numbers of classes and properties are imported. One possible solution to this problem is the RDFex solution proposed by [19], whereby an import proxy service can be used to filter out only the necessary classes and properties from a certain namespace.

5 Summary and Future Work

In this paper, we have discussed our experiences of reusing knowledge and information in the ExpertFinder project, by way of ontology design patterns and the reuse of existing university data sources. We have also looked at designing for reuse, by aligning to ontologies already available on the semantic web.

Our experiences indicate that while there is still work to be done in developing both ontology design patterns and tools for using such patterns, the general idea of reusing design knowledge in this manner seems to be sound. As for ontology alignment, the experiences have also been positive, though we note that there is as of yet no clear consensus on or objective measurement of what ontologies are the most useful to align to, leaving room for future studies.

The ExpertFinder project is still ongoing and several future challenges were discovered while implementing these technologies. To begin with, the way that research areas and research topics are modeled within the ontology will need an overhaul in order to provide better support for reasoning and retrieval with SPARQL queries. Since no documented design patterns covering this area have been found, it is likely that we will have to design our own solution, possibly publishing it as a pattern if found to be general enough.

Another future research task deals with version management of ontologies - a lot of now redundant properties could be pruned from the ontology after the second development iteration, but it must first be studied what the consequences of such pruning might be. Research on this problem becomes even more important when one considers that other ontologies might import the ExpertFinder ontology, making any changes performed cascade throughout the semantic web.

Acknowledgements

Andreas Billig, Michael Ricklefs, and Annika Öhgren have provided valuable contributions to the ExpertFinder project.

This work is partly funded by the project DEON (Development and Evolution of Ontologies in Networked Organizations) based on a grant from STINT (The Swedish Foundation for International Cooperation in Research and Higher Education); grant IG 2008-2011.

References

1. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Scientific American 284(5), 34–43 (2001)
2. Shadbolt, N., Hall, W., Berners-Lee, T.: The Semantic Web Revisited. IEEE Intelligent Systems 21(3), 96–101 (2006)
3. Hendler, J., Berners-Lee, T., Miller, E.: Integrating Applications on the Semantic Web. Journal of the Institute of Electrical Engineers of Japan 122(10), 676–680 (2002)

4. Svatek, V.: Design Patterns for Semantic Web Ontologies: Motivation and Discussion. In: Proceedings of the 7th International Conference on Business Information Systems, Poznan, Poland (2004)
5. Blomqvist, E.: Semi-automatic Ontology Construction based on Patterns. PhD thesis, Linköping University, Department of Computer and Information Science (2009)
6. Presutti, V., Gangemi, A.: Content Ontology Design Patterns as Practical Building Blocks for Web Ontologies. In: Li, Q., Spaccapietra, S., Yu, E., Olivé, A. (eds.) ER 2008. LNCS, vol. 5231, pp. 128–141. Springer, Heidelberg (2008)
7. Presutti, V., Gangemi, A., David, S., de Cea, G., Suárez-Figueroa, M., Montiel-Ponsoda, E., Poveda, M.: D2.5.1: A Library of Ontology Design Patterns: reusable solutions for collaborative design of networked ontologies. NeOn Project Deliverable (EU-IST-2005-027595) (2008)
8. Blomqvist, E., Gangemi, A., Presutti, V.: Experiments on Pattern-based Ontology Design. In: Proceedings of the Fifth International Conference on Knowledge Capture, Redondo Beach, CA, USA, pp. 41–48 (2009)
9. Öhgren, A.: Towards an Ontology Development Methodology for Small and Medium-sized Enterprises. Licentiate Dissertation 1401, Linköping University, Department of Computer and Information Science (2009)
10. D'Arcus, B., Giasson, F.: Bibliographic Ontology Specification, http://bibliontology.com/specification (accessed: 2010-03-01)
11. Brickley, D., Miller, L.: FOAF Vocabulary Specification 0.97, http://xmlns.com/foaf/spec/20100101.html (accessed: 2010-03-01)
12. Hobbs, J., Pan, F.: An Ontology of Time for the Semantic Web. ACM Transactions on Asian Language Information Processing (TALIP) 3(1), 66–85 (2004)
13. Noppens, O.: Concept Partition Pattern. In: Proceedings of the Workshop on Ontology Patterns, Washington D.C., USA (2009)
14. International Labour Organization: ISCO - International Standard Classification of Occupations, http://www.ilo.org/public/english/bureau/stat/isco/isco08/index.htm (accessed: 2010-03-02)
15. Eurostat: Classification of fields of education and training (1999), http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LS%T_NOM_DTL&StrNom=EDU_TRAINI (accessed: 2010-03-03)
16. Association for Computing Machinery: The ACM Computing Classification System (1998 Version), http://www.acm.org/about/class/1998 (accessed: 2010-03-02)
17. Universität Karlsruhe: Acm topic hierarchy, http://www.aifb.uni-karlsruhe.de/WBS/pha/bib/acmtopics.rdf (accessed: 2010-03-04)
18. Diederich, J.: Semantic Web Topic Hierarchy, http://semanticweb.org/wiki/Semantic_Web_Topic_Hierarchy (accessed: 2010-03-04)
19. Knublauch, H.: Rdfex - partial imports of common rdf namespaces, http://sparqlpedia.org/rdfex/ (accessed: 2010-03-04)

MAPPER Collaboration Platform for Knowledge-Intensive Engineering Processes

Håvard Jorgensen[1], Svein G. Johnsen[2], Adam Pawlak[3], Kurt Sandkuhl[4], Till Schümmer[5], and Peter Tandler[6]

[1] Commitment AS, Lysaker, Norway
[2] SINTEF, Oslo, Norway
[3] Silesian University of Technology, Gliwice, Poland
[4] School of Engineering, Jönköping University, Jönköping, Sweden
[5] FernUniversität in Hagen, Hagen, Germany
[6] Teambits GmbH, Darmstadt, Germany

Abstract. The paper presents a collaboration platform with model-based configurable services. It is a part of a more general approach developed by the EU project MAPPER that enables distributed collaboration over adaptive knowledge-intensive engineering processes. In a model-configured infrastructure, active knowledge models (AKMs) of products and processes become the basis for user-composed services, tools and methodologies available to users in customisable workplaces. A multi-layered architecture that offers a powerful means for configuration of services for particular application areas has been described. The presented approach and the collaboration platform have been experimented with in a collaborative engineering use case in electronic industry that integrated two dispersed SMEs. Experiences learned are summarized.

Keywords: enterprise modeling, collaboration platform, distributed collaborative engineering, task patterns.

1 Introduction

More and more complex products are being produced these days, like new mobile phones or vehicles. Engineering processes involved in design and manufacturing of such products are sophisticated, heterogeneous, and very often distributed among dispersed enterprises. Difficulties are consequences of essentially different design techniques required for designing heterogeneous components, being: software, digital or analogue hardware, optical, mechanical or even biological components, as well as in their integration and verification. Designing and integrating such diverse components require collaboration of expert designers that often are available in remote locations only.

In such knowledge intensive industries, like automobile and electronic ones, design and manufacturing processes are heavily based on design and product reuse. Electronic industry demonstrates additionally, new management and control of demand-supply chain with global collaborative networks of involved enterprises [12], [13]. This enables the partner companies to compete against others in terms of time to market, product flexibility and shared knowledge. Information flow in enterprise networks, as

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 180–191, 2010.
© Springer-Verlag Berlin Heidelberg 2010

well as enhancing design processes through effect reuse of engineering knowledge are definitely important aspects of this agile engineering style.

Enterprises in response to global market challenges re-engineer their business models, services and internal structures. The ability to quickly adapt to changing market needs and to cooperate flexibly has become a key element for maintaining competitiveness. This is reflected in the increasing importance of concepts and technologies that support distributed collaboration of enterprises. Industries are moving towards networked organisations, such as collaborative networks and dynamic business ecosystems. Products must be rapidly adapted to customer needs, leading to faster innovation cycles and more complex distributed collaborative engineering. In many industry sectors, a demand for increased value creation on the supplier and sub-suppliers side of the networks is observed, emphasising knowledge content and services. This requires significant growth in inter- and intra-enterprise knowledge reuse and sharing, management services and in new collaborative work environments.

Collaborative environments commonly support communication, collaboration and coordination aspects, which contributes to implementing information logistic principles. The research field of information logistics aims at optimizing information flow and providing demand-oriented information supply. Only the “right information” should be provided to a user engaged in a collaboration situation “just-in-time”. Information logistic especially contributes to coordination support and improves awareness for changes within the distributed work group.

The paper presents the MAPPER project[1] knowledge model-based approach [7] that enables adaptive product and process engineering [8] and supports knowledge-intensive innovative design processes. The approach comprises active knowledge models [1], a collaboration platform, customisable workplaces, a participative engineering methodology and a set of practices defined as task patterns [14].

This article is organised as follows: the MAPPER approach is presented firstly, while the collaboration infrastructure and its components have been detailed in the following section. Main services of the infrastructure, namely: knowledge models, asynchronous and synchronous communication services, as well as secure invocation of remote tools are presented. Finally, the use case with distributed engineering has been addressed.

2 The MAPPER Approach

The main technologies of the MAPPER project are illustrated in the centre of Fig. 1. Configurable *active knowledge models* (AKMs) [1] [9] of products, processes and other enterprise aspects are used for configuring and coordinating both the human and technical aspects of collaborative design. *Customisable workplaces* give different stakeholders access to the information and services they need for performing their tasks. The *secure collaboration platform* enables enterprises to use each other’s engineering tools and product data in a collaborative, yet secure manner, while *participative engineering methodologies* guide joint product and process design, interdisciplinary and inter-organisational collaboration throughout multiple product lifecycles. Each of these main components will be explained in the following sub-sections.

---

[1] MAPPER - Model-based Adaptive Product and Process Engineering, FP6-2004-IST-NMP-2 Project No. 016527, 09.2005-02.2008, `http://mapper.eu.org`

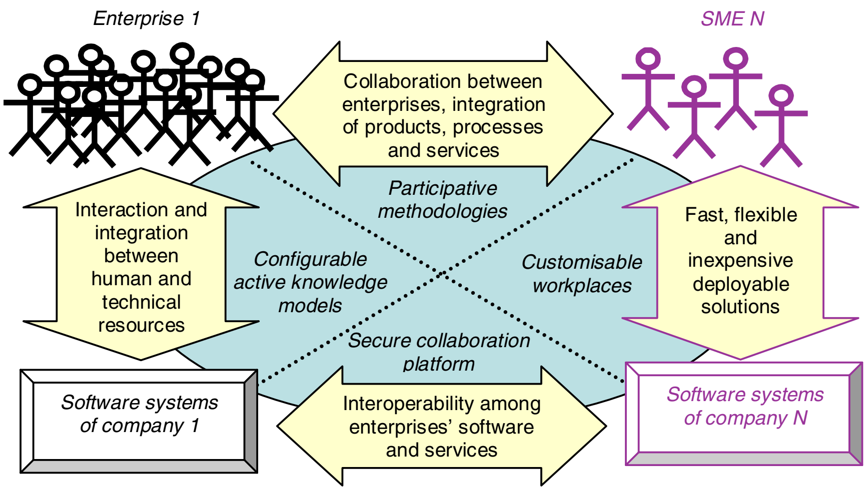


Fig. 1. Overview of the MAPPER approach

2.1 Configurable Active Knowledge Models

Through visual models that reflect evolving business knowledge, all the other elements can be configured and composed. The models are thus used *actively* to customise and adapt the IT infrastructure and the models are executed through process enactment and rule engines. Active knowledge models differ from conventional model-driven architectures (MDAs) in that they primarily capture user knowledge about business realities, rather than technical information about how the computerised support systems work. AKMs thus let evolving business needs directly control the IT infrastructure. Its interactive model execution paradigm [8] creates IT support for the innovative design processes at the core of the business' competitive advantage, not just for the administrative support processes.

Product design today typically involves a multitude of engineering and other disciplines, often from different companies and nationalities. When such a project is to use the MAPPER approach, methodologies and infrastructure, the first step is to build a *scaffolding* model for the project. This model should contain the information needed for setting up a *collaboration space* for the project on the Web, e.g.:

- *Product model:* What kind of product is to be designed, what are the constraints and requirements, which components should it use etc.?
- *Organisation model:* Which companies and people are involved in the project?
- *Process model:* which overall project work plan and general procedures should be followed?
- *System and infrastructure model*: which IT applications and services should be used in the project? Which are the overall content models for managing and navigating project information, etc.?

2.2 Customisable Workplaces

The collaboration platform supports model-configurable workplace components in order to provide each user with the services, information and communication channels needed for performing his or her tasks. These workplaces are available in two main formats:

- *Web workplaces,* easily accessible interfaces for performing well-defined tasks and conventional information processing, e.g. through forms and reports.
- *Visual workplaces,* displaying, organising, modifying and interacting with active knowledge models, defining and doing work at the same time, handling complex and poorly understood problems in innovative ways.

2.3 Collaboration Platform of Model-Configured Services

An ICT infrastructure for model-configured, user-composed services has been developed in MAPPER. These services span different layers:

- *Infrastructure services*, accessible through application programming interfaces, as socket-based communication services (Web services, with explicit XML interfaces and interface definitions (WSDL), accessible over SOAP.
- *User services*, such as interactive portal services and portlets, accessible e.g. as URLs with parameters, pluggable inside HTML frames. Tools and applications, consisting of a number of related software functions, are often used by a specific role to perform a job.
- *Business services*, that are offered, sold and delivered by a company, like products, including generic services (ERP, PLM, application service provisioning etc.), industry sector specific services (e.g. chip design) and company specific services.

The IT industry delivers services across all of these layers. Enabled by a model-configured infrastructure, higher level user and business services can be more easily adapted and composed.

On top of the infrastructure layers, more complex functionality can be model-configured for individuals, groups, projects, companies, business networks and communities. In particular, generic solutions for methodologies were operationalised by configuring lower level user and infrastructure services into coherent task patterns. Each layer filters, combines and contextualises services from the layers below to construct increasingly customized services for business users. In the service teams that perform this customisation, ordinary users and super users are supported by experts on solution and platform modelling. The platforms can thus be extended on different levels.

The system and infrastructure model represents the Web services offered by the components of the MAPPER infrastructure, or by other systems used in each company.

2.4 Participative Engineering Methodology

The objective of the *participative engineering methodology* is to facilitate participative engineering in a networked manufacturing enterprise. The knowledge modeling approach has been used for the development of a methodology consisting of reconfigurable and customizable models. Interdependent sub methodologies are integrated in the overall methodology knowledge model to support the following challenges:

- *Collaboration among humans in a networked enterprise:* How should we go about forming and retaining fruitful collaboration relationships with other people, inside and outside a company?
- *Organisational learning in a networked enterprise:* How should we provide for maximum learning in a networked enterprise? This also involves learning across organisational borders.
- *Multi-project portfolio management:* How to coordinate several parallel projects, where each one may have participants and resources from several companies.
- *Modelling a networked enterprise:* How to plan and perform a modelling effort, covering the whole life-cycle of models, from planning and development to application and management.

3 The Collaboration Infrastructure

3.1 Overview of the MAPPER Infrastructure

The modelling tool Metis [4] which provides services for capturing both the users' data and the configuration models that customise how the users' data should be interpreted constitutes the core of the infrastructure. The Metis Enterprise repository is used for storing the models, and the Configurable Web Portal consists of a set of model-configured components integrated with the Metis Enterprise portal. The visual front-end of the Metis client has been extended with information and view management and task execution to create *configurable visual workplaces* for various design roles involved.

From a group interaction perspective, one can classify the needs for group interaction in distributed virtual enterprises according to the dimension of time. *Synchronous support* is mainly required in the context of virtual meetings. The team needs communication facilities and means for interacting on shared material such as design documents or source files that make up the final product. In MAPPER, this kind of interaction is supported by the Concert Chat system developed at Fraunhofer IPSI.

Asynchronous interaction support is needed to help the team between the different design meetings. It requires that team members are able to share and modify material, discuss and coordinate issues that come up during the design task and stay aware of other users' activities that take place in parallel to the own activities. The CURE platform developed at the Fern Universität in Hagen supports such kinds of interaction.

In both contexts, the users need to have access to the required domain-oriented tools such as simulators or compilers. The tools may in addition require special facilities such as chip test beds or even larger physical installations, like a wind tunnel in the case of car manufacturing. The TRMS system developed at the Silesian University of Technology offers means for securely invoking domain-specific tools. The Tool Registration and Management System [10] enables distance-spanning tool

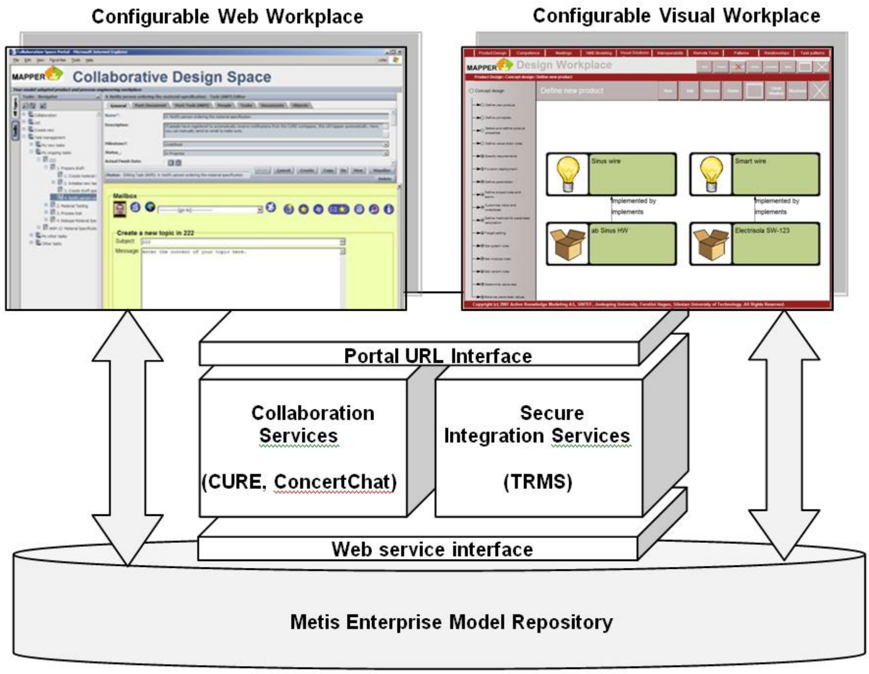


Fig. 2. Overview of the MAPPER infrastructure

integration. Once registered, tools provide their service for distributed design teams. The goal of this component is to allow designers to use tools that are installed at remote sites. All the above tools need to be combined to support task patterns for collaborative engineering. On a technical level, the task patterns are mapped to Metis models that represent the plan for the concrete interaction. Basically, the Metis model configures the user-defined portal. In the portal, the users find their tasks and related tools. The tools are configured by the AKM platform model enactment service that uses the tools' Web service interfaces. If the task requires, e.g., the creation of a set of pages in CURE, the AKM platform will create the pages using Web services. Additionally, a discussion space at the Concert Chat server can be established. Finally, if the task requires the execution of a design tool, the AKM platform will set up a tool configuration in TRMS and trigger the tool execution.

3.2 The Architecture and Its Components

The architecture of the MAPPER collaboration platform is organised as loosely interacting web services, thus the individual components stay as independent as possible while still allowing the users to experience a coherent usage experience. A diagram of the architecture is shown in Fig. 3, the details are explained below. The most important part is the Metis Enterprise Portal that provides the model enactment engine. This engine remote-controls the configuration of the other collaboration support tools. It reads an input from a METIS model describing the workflow between different participants. The METIS model also includes information on the required documents and the expected results of each step in the modelled organization's processes.

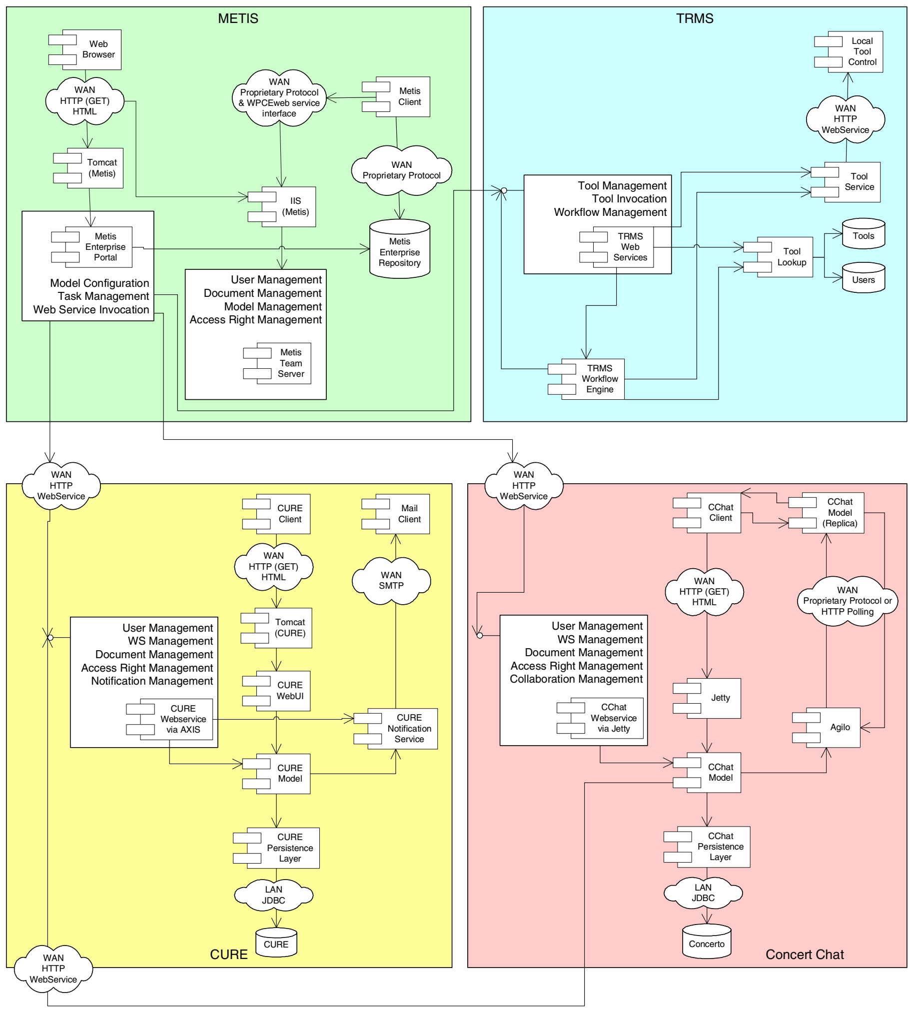


Fig. 3. The architecture of the MAPPER collaboration platform

At each step, the workflow engine detects the required tools, the interacting users, and the input documents. Based on this information, the workflow engine can configure the tools in order to create a collaboration environment that is tailored to the current step and provide this environment to the interacting users. In steps where collaboration support is required, the Workflow Engine will configure TRMS (for secure tool invocation), CURE (for asynchronous interaction), Concert Chat (for synchronous discussions), and other services referred to in the model. The configuration is achieved by submitting a set of web service requests. Below, the components of the platform are shortly explained.

Knowledge modelling tool – METIS [4]. It comprises three main components:

- the *Metis Enterprise* (ME) platform which provides data collection, management, and delivery services required to build any enterprise modelling application,
- *Metis Team* is a file repository used for storing Metis models (views and data) as files. It may also be used for document management. In addition to the web interface used in the MAPPER project portal, Team content can be accessed and managed directly through the repository browser of Metis modelling clients.
- *Metis Client* brings advanced visualization, modelling and analysis to commercial enterprise and government users. This tool enables organizations to support the needs of additional users to enhance their existing Enterprise Architecture (EA) activities.

Remote tool invocation – TRMS [5]. Tool Registration and Management Services system [10] enables distance-spanning tool integration. The architecture (Fig. 4) of TRMS system comprises three main components: the Global Tool Look-up Services (GTLS) server with data bases, the Tool Servers and the Client application.

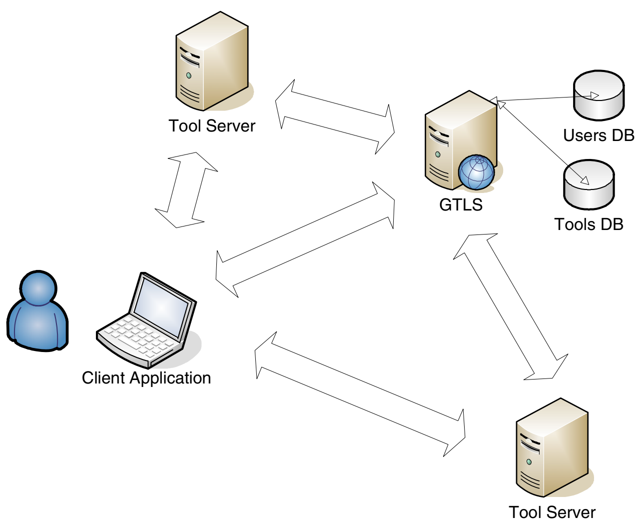


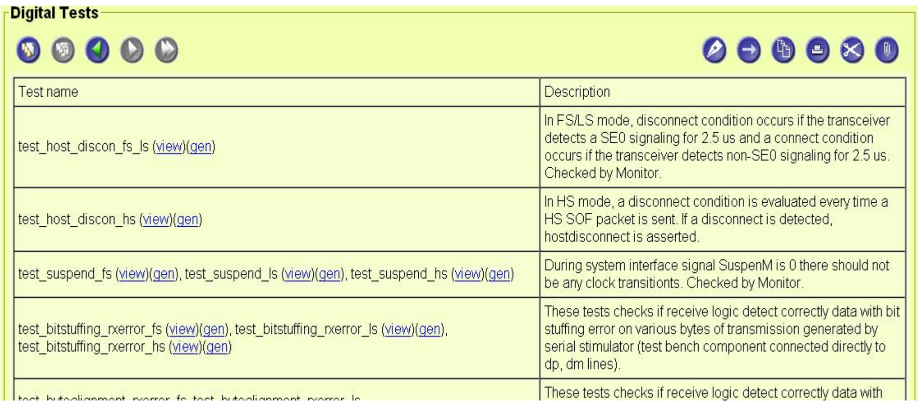
Fig. 4. The TRMS architecture overview

In a general case, all three TRMS system components run on separate machines connected to the Internet. All sensitive data are encrypted (ciphered) and digitally signed by a sender. The Client applet has a simple GUI that allows for login to the system, its administration and usage of available tools. The sequence of tasks to be performed is represented visually in the TRMS Client as a simple workflow. This workflow can be executed by the TRMS Client.

The Tool Server (TS) is responsible for controlling users' access to tools and their execution. A Client invoking a tool does it through the Tool Server. Its additional task constitutes of brokerage in user authentication. The Tool Server queries GTLS whether a user who invokes a tool has sufficient privileges.

Collaboration workspace – CURE [2]. CURE is a web-based system that facilitates collaboration over the Internet using standard browsers [6]. From a user's perspective, CURE is based on a *room* metaphor combined with WIKI ideas, and communication tools. To use CURE, users need a CURE client, which can be any browser that is capable of JavaScript.

To build up structured collaboration environments, a room may be connected to adjacent rooms, thus forming a virtual collaboration infrastructure. Every collaboration infrastructure is represented by a designated entry room and all rooms recursively connected to it. In fact, a whole virtual organization can be constructed by creating an entry room of the network of partners, which links to collaboration spaces for each participating partner and teams formed by representatives of the partners.


Digital Tests

Test name	Description
test_host_discon_fs_ls (view)(gen)	In FS/LS mode, disconnect condition occurs if the transceiver detects a SE0 signaling for 2.5 us and a connect condition occurs if the transceiver detects non-SE0 signaling for 2.5 us. Checked by Monitor.
test_host_discon_hs (view)(gen)	In HS mode, a disconnect condition is evaluated every time a HS SOF packet is sent. If a disconnect is detected, hostdisconnect is asserted.
test_suspend_fs (view)(gen), test_suspend_ls (view)(gen), test_suspend_hs (view)(gen)	During system interface signal SuspenM is 0 there should not be any clock transitions. Checked by Monitor.
test_bitstuffing_rxerror_fs (view)(gen), test_bitstuffing_rxerror_ls (view)(gen), test_bitstuffing_rxerror_hs (view)(gen)	These tests checks if receive logic detect correctly data with bit stuffing error on various bytes of transmission generated by serial stimulator (test bench component connected directly to dp, dm lines).
	These tests checks if receive logic detect correctly data with

Fig. 5. Asynchronous communication of engineers in CURE

A room contains pages, resources and communication tools, which are created, manipulated, navigated and read by users of the room. Users can add/remove/view resources associated with the room. Each room may have its own mailbox that is kept persistent. All users in a room can view and send mails through the room's mailbox.

Distributed engineers may communicate on the design documentation that is shared among design team members over the CURE room, as demonstrated in Fig. 5.

Concert Chat [3]. It provides synchronous collaboration services to the users by offering user awareness, text-based synchronous communication, a shared whiteboard, and referencing functionality using a virtual room to realize a shared workspace.

On the left side in Fig. 6, the shared whiteboard is shown. It provides a *content space* that can be used to share material among all collaborating users providing a spatial layout. The content can be obtained from asynchronous rooms in CURE. This way, Concert Chat adds synchronous interaction services to the set of asynchronous services of CURE. The top right of the figure shows the users that are currently participating in this session. Below is the *interaction space* that offers text-based communication. In order to be able to directly refer to a shared material; chat messages can refer to regions of the shared whiteboard, but also to previous messages.

The Concert Chat server can be configured by the Metis Enterprise model enactment engine to set up collaboration structures for the tasks defined in the project model.

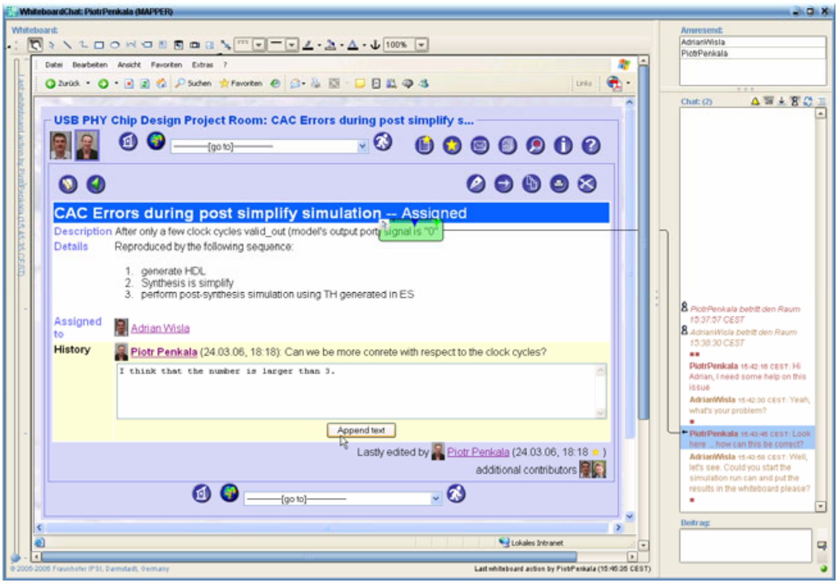


Fig. 6. Synchronous collaboration services provided by Concert Chat

4 Use of Active Knowledge Modelling in Engineering

Active knowledge models (AKMs) can be used *actively* to customize and adapt the IT infrastructure, and the models are executed through process enactment and rule engines. They can serve as systematically updateable knowledge base on the design environment with: relevant design processes including those used for quality assurance, required human resources, and tools. This can simplify transfer of design knowledge around a particular design phase or design task to new employees. AKM defined as a design workflow, when executed can support (semi-) automation of design processes;

AKMs in a natural way can model *electronic products* at various levels of abstraction, from simple Intellectual Property components to complex heterogeneous systems. Configurability of active knowledge models supports design of product families.

4.1 Use of AKMs to Collaborative Design of IP Components

AKMs have been deployed to collaborative design of an IP (Intellectual Property) component in the frame of the MAPPER project in an experiment that integrates two IP components development companies from Germany and Poland that aim at designing an advanced USB IP component. An IP component is required for hardware implementation of standard serial communication protocols. Obviously, this design needs to fulfil constraints on “time to market”, security of design data, and the customer-specific final product functionality. This IP design is being performed in a truly distributed collaborative engineering network that is being built upon the MAPPER infrastructure. Each node in this network constitutes an engineer’s workspace equipped with design tools or just a remote tool that processes automatically design data.

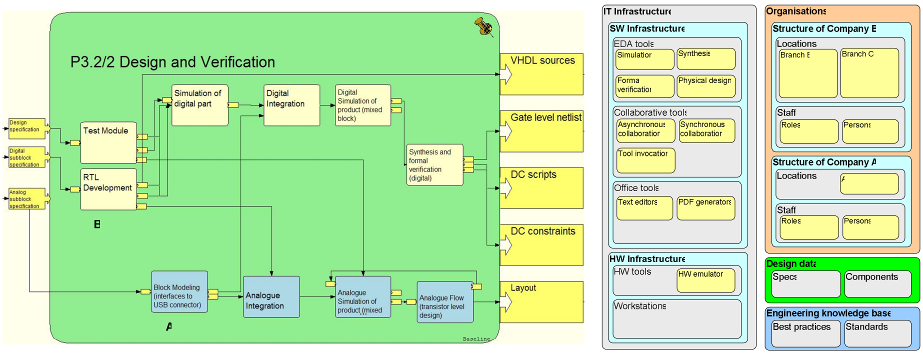


Fig. 7. A part of an IP component design represented as a visual active knowledge model

The presented model comprises a wide spectrum of information related to the current joint product, namely information on: the internal organizations of both involved companies A and E (e.g., a company structure, geographical locations, human resources, competence skills of staff), the available IT infrastructure (e.g., design automation, administration, and office tools), the current project organization (e.g., project responsibilities), the detailed structure of the joint product, and the project plan (e.g., management and design workflows). The model can comprise a large amount of elements, i.e., objects and relations, not easy to grasp, therefore there is a need for information scoping. It is possible to control visibility of model elements by creating different views with selected model components only. More details on this distributed collaborative engineering IP component design can be found in [11].

5 Conclusions and Further Work

A collaboration platform of model-configured services has been introduced in the paper. The multi-layered MAPPER architecture of services offers a powerful means for configuration of services for particular application areas. Experiments conducted in the use cases in automotive and electronics design domains prove utility of the presented approach and the platform for collaborative knowledge-intensive engineering processes conducted in dispersed groups. Selected conclusions from these experiments are enumerated below:

- Integration into one homogeneous knowledge model of various aspects of business processes, design processes, workflows, design environments with human and tool resources, is an important step towards knowledge-based engineering;

- Visual AKM representation of design processes, workflows and task patterns is natural to designers. Still work needs to be invested in elaboration of more electronic engineer specific profiles and views at the METIS web portal;
- Asynchronous communication and collaboration services provided by CURE were well adopted by engineers in their design flows.

Deployments of the presented collaborative infrastructure and the profound validation process resulted also in refined and new requirements for enhanced versions of the infrastructure, services and tools. Further applications will focus on concrete business services supported by the enhanced participative engineering methodology. Information on additional experiments and lessons-learned is available on *mapper.eu.org*.

References

1. Active Knowledge Modeling, `http://activeknowledgemodeling.com/`
2. CURE open source project, `http://sourceforge.net/projects/cure/`
3. Concert Chat open source project, `http://sourceforge.net/projects/concertchat`
4. Metis Enterprise 5.2, Troux Technologies, `http://www.troux.com/`
5. TRMS – Tool Registration and Management, `http://www.ecolleg.org/trms`
6. Haake, J., et al.: Supporting flexible collaborative distance learning in the CURE platform. In: Proc. of the 37th Hawaii Int. Conf. on System Sciences (HICSS-37). IEEE Press, Los Alamitos (2004)
7. Johnsen, S., et al.: Model-based Adaptive Product and Process Engineering. In: Rabe, M., Mihók, P. (eds.) New Technologies for the Intelligent Design and Operation of Manufacturing Networks, pp. 7–27. Fraunhofer IRB Verlag, Stuttgart (2007) ISBN 978-3-8167-7520-1
8. Jorgensen, H.: Interactive process models. PhD thesis, Norwegian Univ. of Science and Technology (2004), `http://www.diva-portal.org/ntnu/theses/abstract.xsql?dbid=4&lang=en`
9. Lillehagen, F., Krogstie, J.: Active Knowledge Modeling of Enterprises. Springer, Heidelberg (2008) ISBN 978-3-540-79415-8
10. Pawlak, A., Fraś, P., Penkala, P.: Web services-based collaborative system for distributed engineering. In: Camarinha-Matos, L., Picard, W. (eds.) PRO-VE'08, Pervasive Collaborative Networks, Poznań, Poland, September 8-10, pp. 463–472. Springer, Heidelberg (2008)
11. Pawlak, A., et al.: Distributed Collaborative Design of a Mixed-Signal IP Component. In: Proc. IEEE DSD2009 Conference, Patras, Greece, August 27-29 (2009)
12. Salminen, V., et al.: Global engineering networking – Turning engineering knowledge into an accessible corporate asset. In: Proc. of the 11th Conf. on Engineering Design (ICED'97), Tampere, Finland, vol. 25, pp. 165–172 (1997)
13. Ranta, J.: Globalisation of manufacturing – A new paradigm emerging from electronics industry. In: IST'99, European Conf. on Information Society Technologies, Helsinki (1999)
14. Schümmer, T., Lukosch, S.: Patterns for Computer-Mediated Interaction. Wiley, Chichester (2007) ISBN 978-0-470-02561

Using Context to Improve Information Supply in the Medical Sector

Oliver Koch and Elisei Rotaru

Fraunhofer Institute for Software and Systems Engineering ISST,
Emil-Figge-Str. 91, 44227 Dortmund, Germany
{oliver.koch,elisei.rotaru}@isst.fraunhofer.de

Abstract. The consideration of context information in physicians search for medical knowledge and information offers good prospects to improve the quality of the delivered search results from heterogeneous eResources. Within the scope of this contribution the authors introduce their information logistic approach towards a context-based information supply for physicians. Basically, this approach encloses a comprehensive context model, which is divided into the four sub models “process”, “clinical case”, “physician” and “environment”. The prototypical implementation of a context application uses an excerpt from the model. The application is based on a context middleware and was used as a basis for an empiric validation of the approach during practical tests and two laboratory experiments. The contribution concludes with an outlook on future needs for research regarding context-based information supply and information logistics in general.

Keywords: information logistics, context-based information supply, context modeling, information need, context-awareness.

1 Introduction

Physicians are knowledge workers. During their medical activities (anamnesis, diagnostics, therapy etc.) they do not only rely on patient- or case-related medical data, but also on medical knowledge [1], [2], [3], [4], [5]. Physicians need this information to come to profound diagnostic and therapeutic decisions. Lack of adequate information leads to delays in medical treatment processes or even to erroneous decisions chargeable to patients.

Considering the information supply of physicians in practice, it is obvious that hospital information systems (HIS), practice management software and electronic medical records can give an efficient access to patient-related data. The supply with medical knowledge from web-based electronic information sources (eResources) is neither meeting the physicians' information need nor is it efficient [6]. In particular, physicians complain about the known phenomenon of information overload, i.e. the information processing requirements regularly exceeding the information processing abilities of physicians [7], [8], [9], [10].

Among the reasons causing the information overload problem are the following three aspects:

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 192–203, 2010.
© Springer-Verlag Berlin Heidelberg 2010

- The **quantity** (i) of available medical knowledge duplicates every 19 years [11]. The number of medical journals has centupled in the last 100 years [11]. More than 130.000 journal articles have been published in the Medline database of the National Library of Medicine (NLM) in the last ten years on the topic “cancer” [12]. These figures impressively document the meaning of “information capacity” as a factor causing the physicians' information overload problem.
- The **time** (ii) physicians' are able to spend for searching and processing medical information - relevant e. g. for treatment decisions - is strongly limited. The Eurocommunication Study that was carried out in six European countries[1] from 1996 to 1999 shows that general practitioners spend 10.7 minutes per patients visit [13], [14]. This includes examination, conversation and if necessary, searching for medical knowledge. Physicians work in hospitals under comparable conditions of time pressure. Searching for information should be limited to 30 seconds in order to be acceptable to physicians [15].
- The third factor of influence is the **characteristics** (iii) of medical information which excels in its high degree of complexity, ambiguity and novelty. As a result, profound data processing abilities of physicians are needed.

In order to improve the information supply of physicians it is necessary to weaken the negative effects of the mentioned factors of influence. Within the scope of this work, the authors have compiled an information logistic approach from the area of context-based information supply, have implemented it in a context application and have evaluated its effect on the reduction of physician's information overload.

2 Related Work

The inclusion of context information, in particular the surroundings of a (mobile) device, has been discussed since the early 1990s in the mobile and ubiquitous computing area [16], [17]. Since 2003 there is a growing awareness in computer and information sciences that a context model goes beyond the scope of simply describing location and time information. It also has a useful effect on information retrieval and supply. In 2004 a workshop on the subject “Information Retrieval in Context” was carried out in Sheffield on the annual congress of the ACM Special Interest Group (SIG) Information Retrieval for the first time.[2] These research activities are based on the insight, that “[…] *IR research is now conducted in multi-media, multi-lingual, and multi-modal environments, but largely out of context. However, such contextual data can be used effectively to constrain retrieval of information thereby reducing the complexity of the retrieval process. To achieve this, context models for different modalities will need to be developed so that they can be deployed effectively to enhance retrieval performance.*”[18]

---

[1] 190 physicians and 3.674 patients from Belgium, Germany, UK, Netherland, Switzerland and Spain participated in the study.

[2] This workshop was followed by other workshops, e.g.: IRiX-Workshop (ACM SIGIR Conference 2005), CIR-Workshops (CONTEXT Conference 2005 and 2007). Fundamental publications: [31], [32], [33].

In parallel with these research activities different practical attempts were undertaken to improve the quality of physicians' information supply. The COWSPOTS-Project, the SmartQuery-Solution and the KnowledgeLink-Approach are examples of practical approaches [19], [20], [15], [21]. In these projects context information has been used only to a small extend for improving the information supply. With his Infobutton concept Cimino [22], [23] suggests another approach, which incorporates a small set of context information for the first time. An Infobutton is "[...] *a point-of-care information retrieval application that generates and sends queries to electronic health information resources (e-resources) using patient data extracted from the electronic medical record and context information* […]" [24]. Our approach of context-based information supply considers the results of these four approaches in terms of related work and tries to refine the basic information logistics principle.

Substantial advancements exist thereby in the development of a differentiated and comprehensive context model, a close linking to the underlying medical processes (e.g. clinical pathways or guidelines) and the physicians' means of query and retrieval modifications (see figure 1).

	HL7-Infobutton	SmartQuery	COWSPOTS	Our Approach
Facilities for Query Alteration by Physician		X		X
Automatic Query Formulation	X	X	X	X
Consideration of User Preferences and User Profiles	X			X
Based on sophisticated ad comprehensive Context Model				X
Empiric Evaluation of Approach	X			X
Based on comprehensive Process Model				(X)
Flexible Binding of different eRessources	X	X		X

Fig. 1. Comparison of approaches

3 Concept of Context-Based Information Supply

The development and validation of our information logistic concept towards a context-based information supply included four basic steps (see figure 2).

In the first step a comprehensive analysis of the requirements of efficient information supply was accomplished [6]. A meta analysis of 14 international studies on physicians' information needs and information behaviour was conducted. Beyond that, we carried out our own survey among 2.500 physicians in Germany concerning their information need and need for IT-technical assistance [4]. Finally, related work regarding the improvement of the physicians' information supply described in the preceding chapter was analyzed. In the further steps the development of a comprehensive context model (Chapter 4), the implementation of an information logistic context middleware and application (Chapter 5) as well as the empirical examination of the overall approach (Chapter 6) took place.

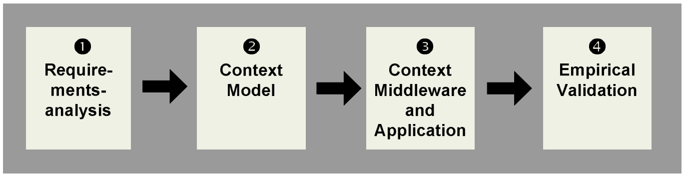


Fig. 2. Procedure Model

During the development of the overall concept it turned out that the "classic" information logistic approach of an straight to the point information supply with the right information (single data or simple information structures), at the right point of time and at the location where the information is needed, has to be modified to meet the requirements of an adequate information supply of physicians.

Therefore we implemented the **concept of selective information spaces** (see figure 3) [6]. The physician isn't supplied with single data, but provided with a space of information and/or documents based on the actual context in consideration of its potential relevance. The more precisely the context aspects of the physician's information need can be covered and mapped on the information supply by a search request, the more relevant information will be contained in the information space. Then the physician can select these pieces of information from the selective information space, which are actually relevant for him.

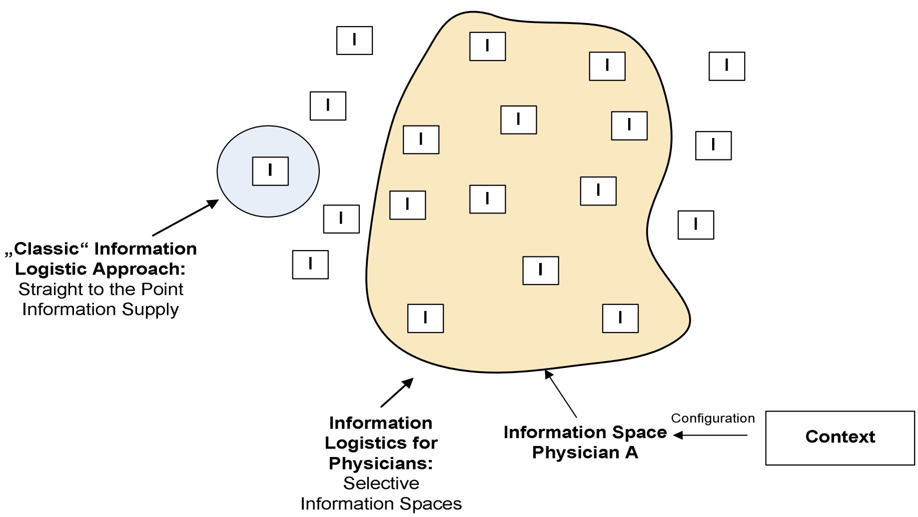


Fig. 3. Concept of Selective Information Spaces

Because of the physician's active role in selecting information, the fuzziness occurring during the semantic indexing of medical specialist literature and the capturing of his information need can be reduced. Thereby a binary relevance ranking becomes achievable.

4 Context Model of Physicians' Information Supply

Most context models are designed for use within ubiquitous computing and focus on location, identity and time [25]. Unlike the context within ubiquitous computing [17], the context in information retrieval refers to information which is directly linked to the information search and helps characterizing the user's information needs.

Koch proposes a process-oriented context model that supports the information retrieval for physicians [26]. This generic context model can be used to improve the usability and effectiveness of information retrieval systems. Since it is not designed with respect to a certain application or an application class, many of them can benefit from this model. Therefore, the proposed model does not claim to be complete. Moreover it is extensible and adaptable. It builds a basis for improving the precision of the retrieved information and reducing the fall-out.

The proposed model is divided into four areas for information push respectively five areas for information pull access. These areas are presented in figure 4 and are as follows: **clinical case context, physician context, process context, environment context** and optional search terms. Each one of these areas is modelled as a UML class diagram describing context attributes and relations [6]. Some examples of context attributes are: primary diagnosis, secondary diagnosis, therapy, patient's basic claim data (age, gender, etc.) actual activity, next activity, physician's preferences (preferred language, type of information, etc.) specialisation and experience and main environment data (terminal, location etc.) [26].

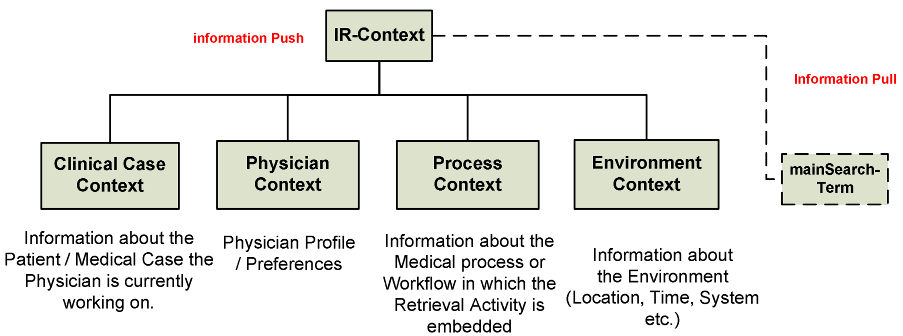


Fig. 4. Context Model of Information Retrieval at Medical Workplaces

Not all identified context attributes were included while implementing the middleware for context-based information supply. Two criteria have to be met for successfully using context attributes to improve the efficiency of information retrieval: (i) the acquisition of context information has to be feasible and (ii) the used electronic resources have to be sufficiently semantically tagged and ideally support semantic queries.

While many bibliographic databases, like Medline [27], are tagged using ontologies and taxonomies, other knowledge sources, like the clinical guideline database from the German Association of the Scientific Medical Societies (AWMF), are not tagged at all. A clear structure of clinical guidelines is missing and the access to such documents is sometimes difficult. In summary, we identified the following challenges in binding electronic resources in context-based information supply: (i) the semantic tagging of medical information, (ii) the access and availability of information and (iii) the structure of supplied information.

5 A Middleware for Context-Based Information Supply

We developed a middleware to provide context-based information supply in the medical sector. Based on a service-oriented architecture (SOA), the middleware serves as an intermediate layer between the care personnel client applications and the available electronic resources. It can be integrated in any web-capable system, like HIS. An information request can be sent through web services or through an elementary HTTP request.

To create a functional context application we connected the middleware to an open source HIS, the care2x HIS. It acts as a test environment and provides the main context information (patient- and physician-related data).

5.1 Architecture

The main components of the context-based information supply middleware are shown in figure 5. For the complete specification of the middleware we refer to [6].

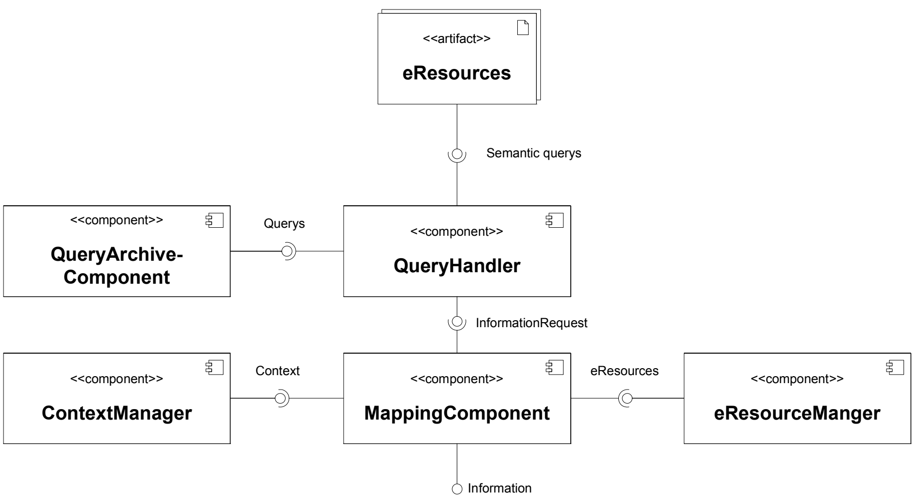


Fig. 5. Main middleware components

In ubiquitous computing, the acquisition of context information is usually being carried out by sensing hardware. This is not feasible for the healthcare domain because the context information is mostly available and stored in different locations. The acquisition has to be performed by a middleware which connects to different data sources (e.g. HIS, workflow engines, profiles). This is being realized by the **ContextComponent** in Figure 5.

The **eResourceManager** stores and manages information (e.g. URI, interfaces, ports, query structure etc.) about available electronic resources. Then, users can choose and configure preferred resources.

A central component in the overall architecture is the **MappingComponent**. It performs the semantic mapping between the acquired context information and the available electronic resources. Thereby, it analyses the structure of the chosen electronic resource and the acquired context information. Then it applies complex mapping techniques (e.g. computing and transformation algorithms) to create search terms for building queries.

The **QueryHandler** builds up queries from the mapped terms and sends them to the electronic resource by using available context information. Then, it processes the retrieved information to be sent as an answer to the user.

An optional component is the **QueryArchiveComponent**. It can be used to improve information filtering techniques and to adapt the retrieval to the users' behaviour.

5.2 Challenges and Achievements

While implementing the middleware, we identified challenges that have to be addressed in order to create efficient context-based health information supply systems. One of them is the acquisition and processing of context information. The middleware needs access to patient-related health information. In Germany, such information is categorized as specific type person-related information that is covered by data protection acts. Appropriate authentication and authorisation techniques as well as encryption and decryption techniques for acquiring personal data have to be implemented by the ContextManager.

A second challenge is closing the gap between the context information, the modelled information needs and the available electronic resources. Ideally, all of them are semantically tagged, but they often use different ontologies. For example, the patient's diagnosis is usually coded in ICD-10 while bibliographic databases, like Pubmed, tag diagnosis related papers and articles using MeSH descriptors. In order to create semantic queries a transformation has to be performed. A study [28] shows that this is not trivial. A new approach presented in [6] can improve the mapping ratio from ICD-10 to MeSH, identifying for almost every ICD-10 code one ore more appropriate MeSH terms. While about half of the mappings conducted use UMLS, the other half also use the UMLS SPECIALIST Lexicon loosing the guarantee for semantic correctness. Tests (presented in Chapter 6) show that information retrieved in this way is more relevant than without the mapping. Moreover, the semantic correctness is not crucial for the middleware which only generates documents of lower relevance.

A third and essential challenge is the formulation and configuration of semantic queries. In order to fulfil the defined information need, complex queries have to be composed. A study [29] shows that the degree of specificity of an IR request corresponds to the length of a search query. Our middleware does not only compose long and complex queries, but also uses an iterative process to retrieve information spaces that match the user's information need. The user can then select the required space.

6 Empirical Concept Validation

The validation of the information logistic concept of context-based information supply represents an important step within the research work as well as for identifying further challenges [6]. The choice of appropriate measurement parameters for improving the information supply was based on the factors causing the information overload. The improvement will be measured in the following terms: can the expenditure of time (factor "time") and the amount of information (factor "quantity") be reduced in such way that it can be handled by physicians? Did it succeed to offer only the information that is relevant for the physician with regard to a specific case within the selective information space (factor "information characteristics")?

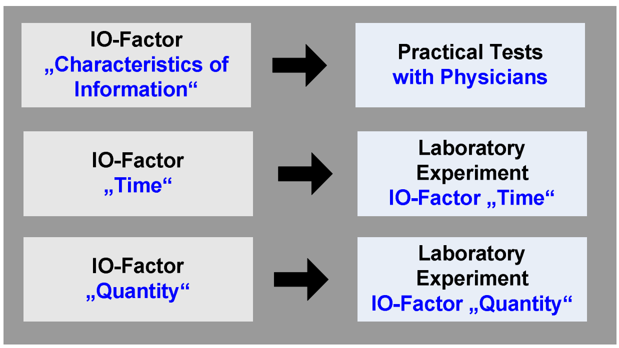


Fig. 6. Mapping of empiric instruments to information overload factors

In order to validate the improvement reached by our context-application and the therein implemented section of the context model, laboratory experiments and practice tests were carried out (see figure 6).

6.1 Practical Tests with Physicians

Within a sensible process of selection, 7 physicians (physicians in private practice and hospital) were chosen to test the context application. In order to respect the timetable of the actual practice test, the physicians received a form in advance, in which they could enter the test data (context information on 5 patients). The test data was entered into the context application in form of patient records. In this way, the data of a total of 35 patients was available for the laboratory experiments and practice tests.

Within the practice test, the physicians evaluated how exact the search results from the Medline – Database matched the basic diagnosis of a patient and the concrete treatment case. Each physician tested two randomly chosen search results per patient (breadthways test) and afterwards the first ten search results for a randomly chosen patient (in-depth test). In this way, 126 search results were evaluated.

92.86 % of the search results from the Medline – Database were classified as potentially relevant to the underlying disease of a patient. A total of 37.1 % (breadthways test) and 40 % (in-depth test) of the search hits were categorized as relevant for the concrete case and the basic diagnosis. This means that almost half of the hits within the information space were classified by the physicians as helpful for the concrete patient case. Despite the complex information characteristics of medical technical literature, choosing relevant pieces of information becomes considerably easier.

6.2 Laboratory Experiment IO-Factor "Time"

Objective of the laboratory experiment "factor time" was to asses and compare the time expenditure for the search - with and without context application - in the Medline – Database for ten randomly chosen patients. In the first part of the experiment, the time the system needed to find ten hits for ten randomly chosen patients, based on context information was recorded automatically. This time span had an arithmetic average of 4.80 sec with a standard deviation of 1.40 sec. In the second part of the test the single steps of the automatic search of the context application were simulated manually. This part of the experiment was performed by the authors and a student assistant with comprehensive research experience. The average search time took 15 min 21 sec with a standard deviation of 4 min 41 sec. The context application reduced the time for search for information for both considered patient cases by more than 15 minutes, which shortened the IO-factor "time" significantly.

6.3 Laboratory Experiment IO-Factor "Quantity"

In order to measure the extent to which the context application reduces the IO-factor "quantity", the search within the Medline – Database started with an initial search term, which considered only the basic diagnosis as context feature, and search results were counted. Gradually, further context features were added. Each added iteration reduced the number of hits and consequently the amount of information the physician potentially has to process. For the test two patients of each participating physician were randomly chosen (14 in total) and the number of search hits for each iteration was recorded. The experiment showed that the amount of information was reduced significantly by employing context-based information search. While the initial search for a basic diagnosis showed over 8.000 hits, in average only 67 remained after the 8th iteration cycle.

The empirical validation proves that the information logistic concept of context-based information supply for physicians has a positive impact on all three parameters which cause the information overload.

7 Conclusion

In this paper, we presented a concept for using context to improve information supply in the medical sector by applying context-based information retrieval approaches to provide physicians with information tailored to suit their needs. For this purpose we developed a comprehensive process-oriented context model which consists of four areas. The new approach of selective information spaces provides the user with an initial selection of information, which can be refined.

Furthermore we developed a middleware in order to implement the presented context. A service-oriented architecture allows the provision of medical workplaces with context-based information by a context application. We developed techniques and methods to acquire context information from different sources, algorithms that map different ontologies to each other and query formulation methods to create and send semantic queries to electronic resources.

An empirical test with physicians validated the concept. The time costs have been reduced from an average of 15 min and 5 sec to less than 30 sec for a case-related information search. In this way, retrieved data can be used in conversations with patients. Also the amount of documents has been reduced in relation to a manual search by a factor of over 100.

8 Future Research

In our future work we are planning to further improve the concept of using context for efficient medical information supply. We will extend the implemented middleware and refine the architecture. The presented solution consists of one intermediate layer. Facing the growing number of available resources, of the acquired context information and of the bound sensors, a multi-layered architecture could be necessary.

We will also investigate the possibility of binding other electronic resources to the middleware. Many hospitals and health insurance companies maintain their own health-related databases. In order to integrate such data in our middleware semantic tagging methods have to be developed, as well as the investigation and classification of available resources has to be performed. In addition to electronic resources paper-based documents can be integrated (e.g. info-brochures, flyers or care offers).

The embedding of the process related context into information retrieval is of high importance in order to efficiently predict next steps in the workflow and the potential information need. The next step is to analyze the influence and impact of information supply in the current workflow. Workflow control options could be integrated in the retrieved information to e.g. help physicians in taking decisions. Also, workflow changes could affect the user's information need. An information push could inform and provide him with relevant information.

Another future research direction is the user's information need. In our concept application this is realized through the notion of profiles. They can store preferences, settings and options set by the user, thereby defining what information is relevant (e.g. weight of context attributes) and how this information should be handled. Such profiles have the disadvantage to be static and unrelated to the context of information retrieval. Context-aware user profiles [30] though, associate the user profile to potential situations. This is highly relevant for the medical domain. Methods have to be developed to improve the linkage from the user profile to possible context situations.

An unimplemented component of the middleware is the QueryArchiveComponent. Information filtering techniques adapt the user's behaviour by learning from the past. The QueryArchiveComponent should store and process queries and react to the users feedback. This could be realized through relevance feedback or observing user's behaviour.

References

1. Andrews, J.E., Pearce, K.A., Ireson, C., Love, M.M.: Information-seeking behaviors of practitioners in a primary care practice-based research network (PBRN). Journal of the Medical Library Association 93(2), 206–212 (2005)
2. Bryant, S.L.: The information needs and information seeking behaviour of family doctors. Health Information and Libraries Journal 21(2), 84–93 (2004)
3. Gonzalez-Gonzalez, A.I., Dawes, M., Sanchez-Mateos, J., Fuertes-Riesgo, R., Escortell-Mayor, E., Sanz-Cuesta, T., Hernandez-Fernandez, T.: Information Needs and Information-Seeking-Behavior of Primary Care Physicians. Annals of Family Medicine 5(4), 345–352 (2007)
4. Koch, O., Kaltenborn, R.: Wissensmanagement am Arbeitsplatz - Mehr Zeit für Patienten durch bessere Information. Deutsches Ärzteblatt online (2005), `http://www.aerzteblatt.de/auf-saetze/0506`
5. Seol, Y.-H., Kaufman, D.R., Mendonça, E.A., Cimino, J.J., Johnson, S.B.: Scenario-based Assessment of Physicians' Information Needs. In: Fieschi, M., Coiera, E., Li, Y.-C.J. (eds.) Medinfo 2004 - Proceedings of the 11th World Congress on Medical Informatics, Amsterdam, pp. 306–310 (2004)
6. Koch, O.: Informationslogistische Konzeption einer kontextorientierten Informationsversorgung in medizinischen Behandlungsprozessen. Dissertation (PhD.), Philipps-Universität Marburg, Marburg (2009)
7. Carlson, C.N.: Information Overload, Retrieval Strategies and Internet User Empowerment. In: Haddon, L., Mante-Meijer, E., Sapio, B., Kommenon, K.-H., Fortunati, L., Kant, A. (eds.) The Good, the Bad and the Irrelevant: The User and the Future of Information and Communication Technologies, Conference Proceedings (Cost 269), Helsinki, pp. 169–173 (2003)
8. Eppler, M.J., Mengis, J.: The Concept of Information Overload. MCM Research Paper, St. Gallen (2002)
9. Schneider, S.C.: Information Overload: Causes and Consequences. Human Systems Management 7(2), 143–153 (1987)
10. Speier, C., Valacich, J.S., Vessey, I.: The influence of task interruption on individual decision making: An information overload perspective. Decision Sciences 30(2), 337–359 (1999)
11. Wyatt, J.C.: Uses and sources of medical knowledge. Lancet 338, 1368–1372 (1991)
12. Hunt, R.E., Newman, R.G.: Medical knowledge overload: a disturbing trend for physicians. Health Care Management Review 22(1), 70–75 (1997)
13. van den Brink-Muinen, A., Verhaak, P., Bensing, J., Bahrs, O., Deveugele, M., Gask, L., Mead, N., Leiva-Fernandez, F., Messerli, V., Oppizzi, L., Peltenburg, M., Perez, A.: Doctor-patient communication in different European health care systems: relevance and performance from the patients' perspective. Family Practice 20(4), 478–485 (2003)
14. Deveugele, M., Derese, A., van den Brink-Muinen, A., Bensing, J., De Maeseneer, J.: Consultation length in general practice: cross sectional study in six European countries. BMJ 325, 472–474 (2002)
15. Maviglia, S.M., Yoon, C.S., Bates, D.W., Kuperman, G.: KnowledgeLink: Impact of Context-Sensitive Information Retrieval on Clinicians' Information Needs. JAMIA 13(1), 67–73 (2006)
16. Dey, A.K.: Providing Architectural Support for Building Context-Aware Applications. Dissertation (PhD.), Georgia Institute of Technology, Atlanta (2000)
17. Weiser, M.: The Computer for the 21st Century. Scientific American 265(2), 94–104 (1991)

18. Ingwersen, P., van Rijsbergen, K., Belkin, N., Larsen, B. (eds.): Information Retrieval in Context. Proceedings of the ACM-SIGIR Workshop 2004, Sheffield (2004)
19. Dahlem, D., Bychkov, Y., Kawasme, L., Jahnke, J.H.: Towards Context Oriented Web Services of Smart Personal Object Technologies (COWSPOTS). In: OOPSLA Workshop on Reference Architectures and Patterns for Pervasive Computing, Anaheim (2003)
20. Jahnke, J., Bychkov, Y., Dahlem, D., Kawasme, L.: Context-Aware Information Services for Health Care. In: Roth-Berghofer, T., Schulz, S. (eds.) Proceedings of the KI-04 Workshop on Modelling and Retrieval of Context, Ulm (2004)
21. Price, S.L., Hersh, W.R., Olson, D.D., Embi, P.J.: SmartQuery: Context-Sensitive Links to Medical Knowledge Sources from the Electronic Patient Record. In: Kohane, I.S. (ed.) Proceedings of the AMIA Annual Symposium 2002, pp. 627–631 (2002)
22. Baorto, D.M., Cimino, J.J.: An 'Infobutton' for enabling patients to interpret on-line Pap smear reports. JAMIA 7(supplement), 47–50 (2000)
23. Cimino, J.J.: Use, Usability, Usefulness, and Impact of an Infobutton Manager. In: Hersh, W. (ed.) AMIA Annual Symposium Proceedings (AMIA 2006), pp. 151–155 (2006)
24. Del Fiol, G., Cimino, J.J., Rocha, R., Maviglia, S., Strasberg, H., Kuhn, T.: Infobutton Communication Standard Proposal. Draft Standard for Trial Use (DSTU) HL7, Salt Lake City (2006)
25. Strang, T., Linnhoff-Popien, C.: A Context Modelling Survey. In: Workshop on Advanced Context Modelling, Reasoning and Management, UbiComp 2004, Nottingham (2004)
26. Koch, O.: Process-based and context-sensitive information supply in medical care. In: Doan, B.-L., Jose, J., Melucci, M. (eds.) Context-Based Information Retrieval, Proceedings of the CIR'07 Workshop on Context-Based Information Retrieval in Conjunction with CONTEXT-07, Computer Science Research Report, Roskilde, vol. (114), pp. 93–104 (2007)
27. Giglia, E.: Medline/pubmed revisited: new, semantic tools to explore the biomedical literature. European journal of physical and rehabilitation medicine 45(2), 293–297 (2009)
28. Bodenreider, O., Nelson, S.J., Hole, W.T., Chang, H.F.: Beyond Synonymy: Exploiting the UMLS Semantics in Mapping Vocabularies. In: Lorenzi, N.M. (ed.) Proceedings of the AMIA Annual Symposium 1999, pp. 815–819 (1999)
29. Phan, N., Bailey, P., Wilkinson, R.: Understanding the relationship of information need specificity to search query length. In: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '07, Amsterdam, pp. 709–710 (2007)
30. Kirsch Pinheiro, M., Villanova-Oliver, M., Gensel, J., Berbers, Y., Martin, H.: Personalizing Web-Based Information Systems through Context-Aware User Profiles. In: The Second International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, pp. 231–238 (2008)
31. Belkin, N.J., Muresan, G., Zhang, X.-M.: Using User's Context for IR Personalization. In: Ingwersen, P., van Rijsbergen, K., Belkin, N., Larsen, B. (eds.) Information Retrieval in Context. Proceedings of the ACM-SIGIR Workshop 2004, pp. 23–25. Sheffield University (2004)
32. Ingwersen, P., Järvelin, K.: The Turn: Integration of Information Seeking and Retrieval in Context. Springer, Dordrecht (2005)
33. Järvelin, K., Ingwersen, P.: Extending Information Seeking and Retrieval Research Toward Context. In: Ingwersen, P., van Rijsbergen, K., Belkin, N., Larsen, B. (eds.) Information Retrieval in Context. Proceedings of the ACM-SIGIR Workshop 2004, Sheffield, pp. 11–15 (2004)

Validation and Use of Information Demand Patterns in Higher Education

Kurt Sandkuhl

Jönköping University, School of Engineering
Box 1026, 55 111 Jönköping, Sweden
Kurt.Sandkuhl@jth.hj.se

Abstract. The context for work presented in this paper is the field of information logistics and addressing the reduction of information overload. Among the approaches for reducing information overload, the concept of information demand patterns has been developed for capturing organizational knowledge on how to improve information supply. This paper investigates the use of information demand patterns (IDP) in higher education by performing an exercise with 22 students divided into 12 groups in a master course on information logistics. The intention is to explore whether the IDP structure reached a level of maturity to transfer it to actors outside the development team and whether it also can be applied for capturing information demand in general. The main result of the work is that the respondents participating in the exercise managed to produce information demand descriptions using the pattern structure, which indicates that they understood the structure as such and were able to apply it.

Keywords: Information demand, information demand pattern, demand modeling, validation.

1 Introduction

The context for work presented in this paper is the field of information logistics, which aims at demand-oriented information supply [1]. The intention of information logistics is to contribute to reducing information overload, which more and more is perceived as problem in enterprises (see [2] for a survey in this context). Among the many concepts and approaches for reducing information overload, modeling and analysis of information demand have been proposed [3] and the concept of information demand patterns has been developed. Information demand patterns are considered as a way of capturing organizational knowledge about how to establish demand-oriented information supply.

Application and validation of information demand patterns so far focused on enterprise contexts and industrial applications, like in collaborative engineering [4]. This paper aims to extend work on information demand patterns by adding two new perspectives: the use of information demand patterns for expressing information demand and the use of such patterns for teaching the subject of information logistics in higher education. The main contributions of this paper are (1) results from validation the concept of information demand pattern in a master course and (2) experiences from using the concept of information demand pattern in higher education.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 204–213, 2010.
© Springer-Verlag Berlin Heidelberg 2010

The remaining part of this paper is structured as follows: section 2 describes the background for the work by introducing the concept of information demand and information demand patterns. Section 3 discusses the development and validation process of the concept of information demand pattern. Section 4 describes the set-up of the exercise used for validating information demand patterns, and summarizes the collected data. Section 5 discusses the results. Section 6 summarizes the work and presents an outlook on future work.

2 Background

Work on understanding the nature of information demand and on identifying and structuring patterns of information demand form the background for this paper. This background will briefly be summarized in this section.

2.1 Information Demand

The notion of information demand is closely related to work in the area information logistics, which considers understanding information demand as key aspect of information logistics solutions [4]. Information demand usually includes different dimensions, like the content required, the time of delivery, the location, the presentation and the quality of information. The research field information logistics explores, develops, and implements concepts, methods, technologies, and solutions for the above mentioned purpose.

Lundqvist explored the nature and characteristics of information demand in an enterprise context in an empirical investigation [5]. The conclusion from the study is information demand of employees in an organization is to a large extent based on the organizational role and the tasks an employee has. This role-centric perspective with task and responsibilities as primary characteristics has been the starting point for developing the a method for information demand analysis [3].

Lundqvist defines information demand as: "*Information demand is the constantly changing need for relevant, current, accurate, reliable, and integrated information to support (business) activities, when ever and where ever it is needed." [5, p. 61]*

2.2 Information Demand Patterns

The concept of information demand pattern originates from work in the research and development project Information Logistics for SME (infoFLOW). infoFLOW included seven partners from automotive supplier industries, IT industry and academia. The objectives were to develop a method for information demand analysis [6] and to identify recurring elements in information demand, i.e. patterns of information demand.

The general idea of information demand patterns (IDP) is similar to most pattern developments in computer science: to capture knowledge about proven solutions in order to facilitate reuse of this knowledge. In this paper, the term information demand pattern is defined as follows:

An information demand pattern addresses a recurring information flow problem that arises for specific roles and work situations in an enterprise, and presents a conceptual solution to it.

An information demand pattern consists of a number of essential parts used for describing the pattern: pattern name, organisational context, problems addressed, conceptual solution (consisting of information demand, quality criteria and timeline), and effects. These parts will be described in the following.

The *pattern name* usually is the name of the role the pattern addresses.

The *organisational context* explains where the pattern is useful. This context description identifies the application domain or the specific departments or functions in an organisation forming the context for pattern definition.

The *problems* of a role that the pattern addresses are identified. The tasks and responsibilities a certain role has are described in order to identify and discuss the challenges and problems, which this role usually faces in the defined organisational context.

The *conceptual solution* describes how to solve the addressed problem. This includes three parts:

- *Information demand* of the role, which is related to the tasks and responsibilities. The different elements of the information demand are identified in the pattern.
- *Quality criteria* for the different elements of the information demand. The criteria include the general importance of the information demand element, the importance of receiving the element completely and with high accuracy, and the importance of timely or real-time information supply.
- *Timeline* indicating the points in time when the different information demand elements should be available.

The *effects* that play in using the proposed solution are described. If the different elements of the information demand should arrive too late or are not available at all this might affect the possibility of the role to complete its task and responsibilities. Information demand patterns include a description of several kinds of effects:

- potential economic consequences;
- time/efficiency effects (i.e. whether the role will need more time for completing the task or will be less efficient);
- effects on increasing or reducing the quality of the work results;
- effects on the motivation of the role responsible;
- learning and experience effects;
- effects from a customer perspective.

The above parts of a pattern are described in much detail in the *textual description* of the pattern. Additionally, a pattern can also be represented as a *visual model*, e.g. a kind of enterprise model. This model representation is supposed to support communication with potential users of the pattern and solution development based on the pattern.

So far, four actual information demand patterns have been identified and developed following the above structure

- Proposal writing responsible
- Material specification responsible
- Responsible for quote preparation
- Responsible for branding

An example for an actual pattern for the role of "Material Specification Responsible" in a manufacturing enterprise can be found in [4].

3 Elaboration and Validation of Information Demand Patterns

The elaboration and validation of the concept of information demand pattern in general and of actual information demand patterns for specific organizational roles was performed in an iterative way consisting of various cycles, including both structure and content. Figure 1 illustrates the overall approach.

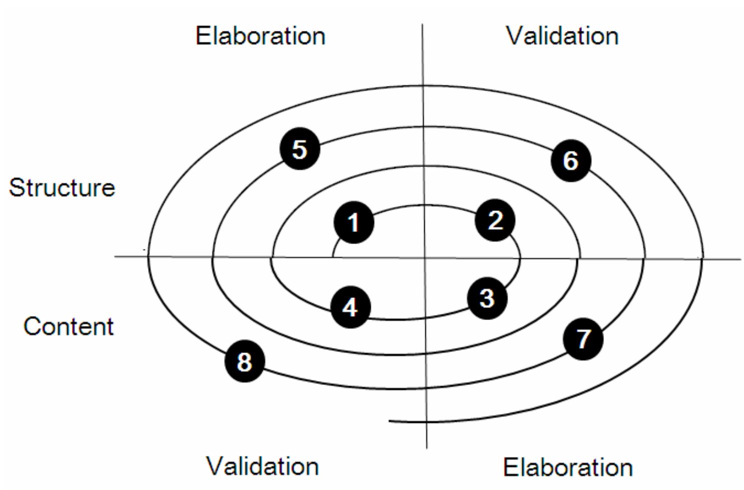


Fig. 1. Iterative process of elaboration and validation of information demand patterns

As indicated in section 2.2, the concept of information demand patterns is inspired by other work in computer science. Thus, the initial work on elaboration of the concept "information demand pattern" (marked as "1" in figure 1) focused primarily on the structure of such patterns based on the literature in the field. The validation of the initial structure ("2" in figure 1) consisted of checking this structure for internal consistency and soundness, and using it for capturing a very simple pattern example. The elaboration of the first actual pattern using the initial pattern structure consisted of selecting a simple but sufficiently complex organizational role and developing the content of the pattern for the selected role (step "3"). Afterwards, the initial content of the pattern again was validated by checking internal consistency and completeness, and by using own experiences (step "4"). Steps 1 to 4 form the first iteration, which was performed in a lab environment without involvement from actors outside the research team.

The second iteration started with the next elaboration phase for the pattern structure (step "5"), which was based on experiences from elaborating and validating the first actual pattern (steps 3 and 4). Improvement of the pattern structure and the following step of validation for the improved version (step "6") now also involved additional actors, namely the industrial partners of the infoFLOW project. Later iterations (i.e. steps beyond “8”) included elaboration and validation activities involving actors outside the project team and outside academia.

The use of the information demand pattern structure in a teaching context as described in section 4 can be considered a contribution to further validating the information demand pattern structure, even though the significance of the results will be limited (see section 5 for a discussion).

4 Use and Validation of IDP in Higher Education

The intention of using information demand patterns in higher education was twofold: we wanted to explore whether the concept and structure of information demand pattern is mature enough to be transferred to actors outside the team developing the concept of IDP and we wanted to explore whether the structure of an information demand pattern also can be used for describing information needed for acting in a certain role or performing a task. The context for exploring these two aspects was a course on information logistics as a part of a master program in the area of business informatics in a medium-sized European university. The instrument we used was to define an exercise in this course in order to collect data. This section describes the set-up for this data collection and selected results.

4.1 Set-Up for Data Collection

The data collection was performed in an exercise within the above mentioned master course. The 22 students in the course came from 10 different countries, 19 of them had a bachelors degree in computer science, 3 in business informatics. As a preparation for the exercise, the students were introduced to the area of information logistics by four lectures on basic principles of information logistics, demand modeling approaches, semantic matching and typical applications. This included scientific papers supporting the lectures. One of the lectures was dedicated to introducing the concept of information demand pattern including examples.

The exercise consisted of a practical task and a questionnaire to be filled in after the practical task. The students were allowed to work alone or in groups of two students, which resulted in 12 solutions handed in. The task was introduced by the teacher, but there was no further guidance provided during the work on the exercise.

The practical task was to describe the information demand for a task or a role, where the students felt that they are experts or have at least a lot of experience. They were encouraged to consider different areas when deciding on the task or role they want to describe, including their private or social background, the university context or any other field. The description of the information demand had to be structured like an information demand pattern (see 2.2); a template was provided for this purpose.

The questionnaire included 9 different questions, four of them with a five-point Likert scale, one with a nominal scale, one with an interval scale and three for free text answers.

4.2 Data Analysis

All 12 student groups, which in the following will be called respondents, submitted information demand descriptions following the pattern template and all groups filled in the questionnaire. The information demand descriptions addressed the following subjects:

- Apply for an exchange program at University of Technology of Troyes
- Organizing a cultural event
- Order delivery responsible
- Management of a small software project
- Planning public transport in Mexico City
- Event planning in hotels
- Supply Chain Management for commercial product
- Pizza delivery services responsible
- Campus party for 2000 people
- Inventory management responsible
- Data warehouse security manager
- Requirements manager for social web applications

In the questionnaire, the respondents were asked whether the elements in the IDP structure are needed in order to describe information demand or whether there are superfluous elements. Nine respondents answered that all elements are needed, three responded that one element was not needed: timeline, quality criteria or task/responsibilities, respectively.

When asked whether anything is missing in the IDP structure, 4 respondents answered that the source of the information demand should be explicitly included. The other 8 groups considered the structure as complete.

Table 1. Response distribution for "How difficult to *identify* were the different parts?"

Element of IDP structure	Very difficult	difficult	Neither difficult nor easy	easy	Very easy
Context	1	3	6	1	1
Problem	4	2	3	2	1
Tasks/responsib.	1	3	8		
Info demand	2	4	5	1	
Quality Criteria	2	5	4	1	
Timeline	1	2	4	5	
Effects		8	1	3	

When asked "How difficult to *identify* were the different parts of structure?", the following distribution of answers was observed (see table 1). This distribution follows in most rows the Gaussian distribution. An exception are the effects of not receiving information or receiving it too late. Here the 2/3 of respondents perceived it as "difficult" to identify these effects.

The question "How difficult to *describe* were the different parts of structure?" resulted in the following response distribution (table 2). Again, most of the responses follow Gaussian distribution with exception of 2 aspects. To describe information demand and to describe tasks/responsibilities of a role are considered at least difficult by 2/3 of the respondents.

Table 2. Response distribution for "How difficult to *describe* were the different parts?"

Element of IDP structure	Very difficult	difficult	Neither difficult nor easy	easy	Very easy
Context	3	1	6	2	
Problem	2	3	5	2	
Tasks/responsib.		8	3	1	
Info demand	2	7	1	2	
Quality Criteria	3	3	4	2	
Timeline	2	4	2	3	1
Effects		4	3	5	

Regarding the aspect "How much time was needed to identify and describe the different parts?", table 3 shows the distribution of responses. Here, describing the information demand was considered the most time consuming activity, as 75% of the respondents answered that "very much" or "much" time was needed. Describing context and problem statement also were considered time consuming by roughly 60% of the respondents.

Table 3. Response distribution for "How much time was needed for the different parts?"

Element of IDP structure	Very much	much	Neither much nor little	little	Very little
Context	2	6	3		1
Problem	2	5	3	1	1
Tasks/responsib.	1	2	9		
Info demand	5	4	3		
Quality Criteria	2	5	4	1	
Timeline	1	3	3	2	
Effects		4	3	5	

When it comes to the time needed to develop the information demand description, a substantial part of the respondents needed more than 10 hours and only 25% less than 6 hours (see table 4). Not surprisingly, the "better" demand descriptions (see section 5) needed more time to be elaborated. But the "better" descriptions were not necessarily the ones with the largest number of pages.

Table 4. Response distribution for the overall time needed for completing the exercise

Time needed	Number of responses
less than 2 hours	0
2 hours – 4 hours	2
4 hours - 6 hours	1
6 hours - 8 hours	4
8 hours - 10 hours	0
10 hours - 12 hours	3
More than 12 hours	2

5 Discussion

The main conclusion from the performed exercise is that there are clear indicators to believe that

- The respondents understood concept and structure of information demand patterns and were able to apply it on their own, and
- The structure proposed for information demand patterns is also suitable for describing information demand in general.

Both conclusions are supported by the fact that 12 complete information demand descriptions for different tasks or roles were developed by the participants in the exercise without guiding or supporting them in the actual development process of these demand descriptions. This indicates that the students learned how to use the information demand pattern structure for capturing information demand descriptions. The pattern structure was judged suitable and complete by the clear majority of the students. The weight of these observations is relatively low due to the limited experience level of the students, but nevertheless contributes to the validation of the IDP concept.

The list of developed information demand descriptions shows a wide bandwidth from more socially oriented tasks (like “organizing a party”) to strictly business-oriented (“delivery service responsible”) or IT-oriented roles (e.g. “data warehouse security manager”). The quality of these patterns was only evaluated by the teaching team in the course who checked the consistency between the different parts of the description, the understandability of the different textual elements, the completeness of the description, and whether the descriptions were sound and reasonable. This “perceived” quality of the patterns might be sufficient as initial check, but would have to be complemented with an “in-use” check of the description, i.e. applying them in a real-world situation for performing the task/role under consideration.

With respect to the perceived quality, the impression was that those patterns who addressed quite small and very specific roles or tasks were the better ones. Examples are “apply for an exchange program at University of Technology of Troyes” or “organizing a cultural event in the master program”. Furthermore, it was observed that some of the business related demand descriptions were very detailed and high-lighted interesting challenges. An example is the “data warehouse security manager”, which was reflecting a lot of experience, since the respondent developing it had several years of job experience on this position from his time before starting the master education. Thus, the results of the exercise also include some information demand descriptions, which can be considered as candidates for future information demand patterns.

The answers from the questionnaire regarding difficulty and time needed to identify and to describe the different elements of a demand description can be used as basis when improving the exercise and the course on information logistics. More emphasis in next year's teaching should be put on repeating organizational concepts like "role" and "task" and how to actually describe them in a proper way. This aspect is covered in a course on enterprise modeling, which is a mandatory prerequisite for the information logistics course. However, a repetition of the essential parts should contribute to developing sound information demand descriptions.

Furthermore, the quality of the information demand descriptions probably could be improved by several measures. One aspect could be to add more examples for information demand patterns and perform the stepwise development together with students in order to improve the understanding of the concept and the different elements of an IDP. Furthermore, the evaluation of the patterns could be made part of the course, by providing a guidance for the evaluation process and letting the students evaluate another group's demand description.

When evaluating the free form questions of the questionnaire, an interesting observation surfaced: the students had to learn that they really have to be experts in the domain in order to be able to describe the information demand. More than half of the groups indicated that the hardest task in the exercise was to scale down the role or task under consideration to a scope which really reflected their own area of expertise.

6 Summary and Future Work

This paper investigated the use of information demand patterns in higher education by performing an exercise with 22 students divided into 12 groups in a master course on information logistics. The intention was to explore whether the IDP structure has reached a level of maturity to transfer it to actors outside the development team and whether it also can be applied for capturing information demand in general. The main result of the work is that the respondents participating in the exercise managed to produce information demand descriptions using the pattern structure, which indicates that they understood the structure as such and were able to apply it. This is considered a small contribution to validate the concept of information demand patterns in the ongoing iterative elaboration and validation process. The work also resulted in some proposals for improving the IDP structure, e.g. by putting more weight on actual information sources, and in improving the way of teaching information demand modeling, e.g. by spending more efforts on enterprise modeling aspects such as role modeling.

The main shortcoming of the work is the limited significance of the results due to the small number of participants in the exercise, the university context of the exercise and the missing step of validating the developed information demand descriptions. These three aspects motivate continuous work into at least three directions:

- A similar validation effort should be made outside the academic context, i.e. to transfer the concept of information demand patterns to an industrially oriented community and evaluate the results of modeling information demand in such a setting

- In order to validate the IDP concept, the quality of the actual demand descriptions developed with this concept also has to be evaluated in a more systematic way than what was done in the exercise described
- A larger number of real-world information demand patterns should be developed, made publicly available and be continuously improved. This will form a more direct contribution to the elaboration and validation process of IDP than what was performed in the exercise in higher education.

All three aspects will be addressed in the continuation of the infoFLOW project, which has been started in May 2010.

Acknowledgements

Part of this work was funded by the R&D project "Information Logistics for SME (infoFLOW-2)" financed by the Swedish Knowledge Foundation. The author wishes to thank all infoFLOW co-workers involved in pattern development and validation.

References

1. Sandkuhl, K.: Information Logistics in Networked Organizations: Selected Concepts and Applications. In: 9th International Conference on Enterprise Information Systems, ICEIS 2008. LNBIP 12, pp. 43–54. Springer, Heidelberg (2008)
2. Öhgren, A., Sandkuhl, K.: Information Overload in Industrial Enterprises - Results of an Empirical Investigation. In: Proceedings ECIME 2008, London, UK (2008)
3. Lundqvist, M., Sandkuhl, K., Seigerroth, U., Stirna, J.: Method Requirements for Information Demand Analysis. In: 2nd International Conference on Adaptive Business Systems, ABS 2008 (2008)
4. Sandkuhl, K.: Supporting Collaborative Engineering with Information Supply Patterns. In: Euromicro PDP 2010. IEEE/CS, Pisa (2010)
5. Lundqvist, M.: Information Demand and Use: Improving Information Flow within Small-scale Business Contexts. Licentiate Thesis, Dept of Computer and Information Science, Linköping University, Linköping, Sweden (2007) ISSN 0280-7971
6. Lundqvist, M., Holmquist, E., Sandkuhl, K., Seigerroth, U., Strandesjö, J.: Information Demand Context Modelling for Improved Information Flow: Experiences and Practices. 2. In: Persson, A., Stirna, J. (eds.) PoEM 2009. LNBIP, vol. 39, pp. 8–22. Springer, Heidelberg (2009)

Context-Based Information and Knowledge Logistics for Self-organisation of Web-Services in Smart Environments

Alexander Smirnov, Tatiana Levashova, and Nikolay Shilov

Institution of the Russian Academy of Sciences
St.Petersburg Institute for Informatics and Automation of RAS
39, 14th line, St. Petersburg, 199178, Russia
{smir,oleg,nick}@iias.spb.su

Abstract. The paper offers an approach to self-organisation of resources of a smart environment for the purpose of taking joint resource actions to assist the people involved in a situation. The approach incorporates concepts of information and knowledge logistics and conforms to the idea of reducing information overflow by utilizing the paradigm of a demand-oriented information supply. For representation of the needs or demands of the people involved in the situation context model is used. The resources of the smart environment are represented by Web-services. Due to this representation self-organisation of the resources is replaced with that of the Web-services. The proposed approach, service-oriented architecture of the smart environment, and a protocol for Web-serves self-organisation are discussed in the paper.

Keywords: smart environment, Web-services, service-oriented architecture, self-organisation protocol.

1 Introduction

Information logistics is an innovative technology aiming at delivery of the right information product, in the right format, at the right place at the right time for the right people in a customer demand driven approach [1]. One of the information logistics goals is equipment of interested parties with information for the task at hand. Up-to-date information technologies are capable to provide with enormous amount of information. Such amount presents considerable difficulty in analyzing and processing the information. So, to achieve the information logistics goal it is necessary to reduce information overflow. A way to do this is utilizing the paradigm of a demand-oriented information supply [2].

The approach presented in this paper pursues the goal of information logistics extending it with providing the interested parties with knowledge representing the task at hand (the needs of the parties) and knowledge for solving this task. The approach conforms to the principal of the demand-oriented information supply. For representation of the interested parties' demands context model [3] is used.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 214–225, 2010.
© Springer-Verlag Berlin Heidelberg 2010

Issues investigated in the approach are intended for applying in smart environments. Actors of the smart environments expect the resources of these environments will sensible cooperate to assist them [4]. The approach considers the actors to be the interested parties mentioned above.

It is supposed that in the smart environment there are resources capable to perform same functions or fill same roles. Examples of different resources performing the same functions, e.g. of providing weather conditions, are sensors, Web-sites, satellite services. A simple example of different resources that can fill the same role is a group leader and a member of this group. Often, the leader is capable do fill all the roles the members of the group fill.

The idea behind the approach is self-organisation of the resources of the smart environment according to a situation that takes place in this environment. The purpose of the resources self-organisation is taking joint resource actions to assist the people involved in the situation. At that an efficient community of the resources is supposed to be organised. As criteria of the efficiency resources' functionalities, availabilities, reliabilities, access times, and costs are used.

In the approach the resources of a smart environment are represented by Web-services. This representation allows self-organisation of the resources to be replaced with that of the Web-services.

At present, in part of Web-services self-organisation research communities focus on automated composition of Web-services and Web-services orchestration and choreography. Techniques of the automated composition detect interested parties' needs or demands, find out appropriate Web-services in the Internet, compose them in the right order, and execute [5]. Techniques of Web-services orchestration and choreography concern Web-service interactions and messaging [6].

The approach presented here is a step toward an automated service composition through Web-services self-organisation in the sense of the service consumers do not have to select appropriate services and link them together. Though, a premise the approach is built upon is the desired Web-services already exist and known; there is no need to search them in the open environments.

The combination Information & Knowledge Logistics in the title and the header of the next Section is explained by the fact that the approach discussed in the paper deals with knowledge. In the knowledge management field it is recognised that knowledge is application of data and information. Therefore, having knowledge as a product of delivery we have information in this product as well.

The rest of the paper is structured as follows. Section 2 presents a brief description of the proposed approach and outlines the place of self-organisation of the Web-services in it. Section 3 introduces service-oriented architecture of the smart environment. A protocol of Web-services self-organisation is discussed in Section 4. Main results and drawings are summarised in Conclusion.

2 Context-Based Information and Knowledge Logistics

The smart environment is considered composed of a set of resources. The presented approach distinguishes three kinds of resources: *information*, *problem-solving*, and *acting*. *Information resources* are resources providing information from such information sources as sensors, Web-sites, databases, radio-frequency identification tags, *etc*. *Problem-solving resources* are computational modules, applications, services, *etc*. that can be used to solve problems requiring solutions in different situations. *Acting* resources are organisations and persons acting in the situations according to their roles.

Web-services represent the properties and implement the functions of the resources. Some examples of the resource properties are resource location, resource availability, resource reliability, *etc*. Functions of the information or problem-solving resources are information acquisition, data measurement, computations, *etc*. Functions of the acting resources are functions carried out by people or organisations according to the roles these people or organisations play in a certain situation.

Knowledge of the application domain is represented by an application ontology (AO). AO represents the knowledge using sets of classes, non-instantiated class attributes, attribute domains (ranges for the attribute values), and constraints. AO specifies conceptual and problem-solving knowledge of the domain [7].

Surjective mappings are specified between names of the classes or attributes contained in AO and attributes contained in the WSDL-descriptions of the Web-services. The mappings specification is based on discovering attributes occurred in the WSDL-descriptions, values of which can serve as values of the AO's class attributes. It is supposed that an AO-attribute can take on a value provided by a WSDL-attribute if the name of the WSDL-attribute is semantically close to the name of the AO-attribute [8].

The mappings enable the Web-services to gain the semantics of the application domain and become capable to exchange information and interact. With reference to AO the mappings specify (i) which values of the class attributes serve as input data for the functions implemented in Web-services; (ii) which class attributes take their values as results of executions of the functions implemented in Web-services; (iii) what types of acting resources Web-services represent.

A situation in the smart environment is modelled at two levels: abstract and operational. At the abstract level the situation is represented by the *abstract context* that is an ontology-based model of this situation. A requirement to the procedure of the abstract context building agrees with the aims of the information logistics technology – the abstract context has to represent knowledge in a demand-oriented way. The abstract context eventuated following this procedure comprises (i) knowledge from AO relevant to the situation and (ii) references to Web-services WSDL-descriptions of which have mappings to the abstract context knowledge. The Web-services the abstract context refers to form a set of contextual Web-services – WS_c.

The knowledge comprising the abstract context specifies information to be used to characterise the situation and problems to be solved in this situation. These problems represent needs or demands of the parties involved in the situation. The references to the contextual Web-services indicate that only those resources will be considered as candidates for joint actions which provide the information relevant to the situation or solve the specified tasks. Introducing the subset of the contextual Web-services enables to reduce the search space of information and solutions.

Instantiation of the class attributes occurred in the abstract context and solving the problems specified in it are primary goals of Web-services self-organisation. An instantiated abstract context is the *operational context*. The operational context corresponds to the operational level of situation modelling. It represents the situation scene. This context is the basis for achievement of secondary goals of Web-services self-organisation; namely, self-organisation of the acting resources to participate in joint actions expected in the situation.

3 Service-Oriented Architecture of Smart Environment

For the implementation of the approach to context-based self-organisation of the Web-services in the smart environments service-oriented architecture of such an environment is proposed (Fig. 1). It distinguishes two sorts of Web-services: *core* Web-services and *operational* Web-Services.

The core Web-services are intended for creating, storing, maintaining the abstract context, registering the Web-services, and monitoring the smart environment. The set of core Web-services comprises:

- AO access service that provides access to AO for the Web-services;
- abstract context service that creates, stores, maintains and reuses the abstract contexts;
- monitoring service that receives information provided by the environmental smart sensors and makes this information available to the other Web-services. The information concerns the type of the current situation and other situation characteristics that the sensors can provide (*e.g.,* the weather conditions);
- registration service that registers the Web-services in the service register.

The operational Web-services are intended for representing properties and functions of the resources of the smart environment and participating in the resource self-organisation instead of the resources. The set of operational Web-services comprises:

- sets of Information Resource and Problem-Solving Resource (IPSR) services that are responsible for (i) sending / receiving messages by the information and

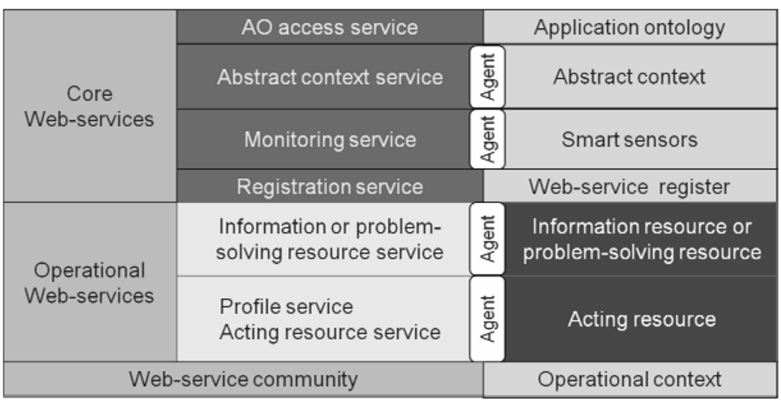


Fig. 1. Service-oriented architecture of smart environment

problem-solving resources, (ii) processing information provided by these resources (format conversions and information integration), (iii) problem solving, and (iv) instantiating the abstract context;

– a set of profile services that are responsible for (i) storing and updating personal information about the acting resources (name, phone number, address, *etc.*), (ii) ascertaining the role that an acting resource performs in the current situation, (iii) learning the resource competence, and revealing the resources' preferences;
– a set of acting resources' services that are responsible for (i) interacting the acting resources; (ii) providing context-sensitive help to the acting resources; (iii) displaying the picture of the current situation at the screens of devices (notebooks, mobile phones, PDAs, *etc.*) the acting resources work with; (iv) providing the acting resources with feasible action plans in the current situation; (v) accessing the operational context by the acting resources in case these resources have to enter or change some values represented in this context.

Below, the protocol for self-organisation of the services comprising the architecture is discussed. AO access service denoted in the architecture does not participate in the process of self-organisation since it interrupts its functions at the moment when the abstract context has been created and resumes only after the moment of finishing the join actions.

4 Protocol of Web-Services Self-organisation

The process of the Web-services self-organisation starts as soon as the abstract context has been created. As it is said above, one of the inputs for the abstract context creation is the type of the situation taking place in the smart environment. This type is captured by the smart sensors. The monitoring service sends the type of the situation to the abstract context service. Based on this type the abstract context service creates an abstract context.

The protocol for self-organisation of the Web-services supports 5 operations. An operation is implemented as a set of requests, replies, and notifications sent / received by the Web-services to execute the operation. The following operations are supported:

– **Function** – selects from the set WS_c a set WS_f of IPSR services based on the services' functionalities;
– **Role** – selects from the set WS_c a set WS_f of acting resource services based on the services' roles;
– **Availability** – selects from the set WS_f a set WS_a of Web-services representing any kinds of resources based on the services' availabilities over a predicted interval $[t_0, T]$, where t_0 – the time instant when the Web-services start initiating the abstract context, T – the time instant when the Web-services come to the decision about starting the joint actions;
– **Weight** – selects from a set of the alternative services containing in the set WS_a a set of Web-services based on the services' weights, so that the resulting set of Web-services would not contain the alternative ones.

The operations listed above are executed one after another starting from the operations **Function** and **Role** and finishing by the operation **Weight**. The result of the operation executions is a community of Web-services united to take joint actions according to the current situation. Besides the described operations, the protocol supports one more operation

– **Activity** – sustains activities of the Web-services within the community.

The requests, replies, and notifications implementing the protocol's operations are listed in Table 1. They have the following meanings:

– *GetAttribute* is request of IPSR service to the abstract context service for a list of class attributes occurred in the abstract context;
– *ReturnAttribute* is reply of the abstract context service to IPSR service in the form of the list of class attributes requiring to be instantiated;
– *Function* is notification the abstract context service sent by an IPSR service in the form of the list of functions which results (values of the functions' output arguments) can be used for instantiation of the class attributes contained in the abstract context;
– *GetRole* is request of an acting service to the abstract context service for a list of class names that represent roles performed by the acting resources in the current situation;
– *ReturnRole* is reply of the abstract context service to the acting service in the form of list of the requested class names (RR);
– *Role* is notification the abstract context service sent by an acting service in the form of the list of roles $(WSRR)$ the acting resource can perform $(WSRR \subseteq RR)$;
– *Include_FS* is reply of the abstract context service notifying an IPSR service or acting service about inclusion in (1) / exclusion from (0) the further negotiations this service based on a principal of *maximum functionality*;
– *GetInterval* is request of IPSR service or an acting service to the abstract context service for time values t_0 and T ;
– *ReturnInterval* is reply of the abstract context service with the requested values;
– *Intervals* is notification the abstract context service sent by an IPSR service or an acting service in the form of time intervals when the IPSR or acting resource is available;
– *Include_AS* is reply of the abstract context service notifying IPSR service or acting service about inclusion in (1) / exclusion from (0) the further negotiations this service based on a principal of *maximum access interval*;
– *GetCompetency* is request of the abstract context service to IPSR service or an acting service for the resource competence (in case of the acting resource) or the resource reliability (in case of IPSR);
– *ReturnCompetency* is reply of IPSR service or the acting service to the abstract context service with the resource competence;
– *GetTime* is request of the abstract context service to IPSR service for the function execution time (service waiting time);

Table 1. Requests, replies, and notifications implementing protocol operations

Operation	Request	Request parameters	Reply	Reply parameters
Function	GetAttribute	Abstract context identifier	Return Attribute	A list of class attributes occurred in abstract context
	Function	URI of Web-service, a set of names of functions implemented in Web-service, a set of names of output functions' arguments	Include_FS	1 or 0
Role	GetRole	Abstract context identifier	ReturnRole	A list of names of classes (occurred in abstract context) representing roles
	Role	URI of Web-service, a set of roles	Include_FS	1 or 0
Availability	GetInterval	Abstract context identifier	Return Interval	Time instant of abstract context creation, expected moment of taking joint actions
	Intervals	URI of Web-service, a list of resource availability intervals	Include_AS	1 or 0
Weight	Get Competency	URI of Web-service	Return Competency	URI of Web-service, competence
	GetTime	URI of Web-service, function name	ReturnTime	URI of Web-service, function name, function execution time
	GetAccess	URI of Web-service	Return Access	URI of Web-service, Web-service access time
	GetCost	URI of Web-service	ReturnCost	URI of Web-service, resource cost
	Include_WS	1 or 0	-	-
	Ready	-	-	-
Activity	LookInput	Function name, a list of input arguments	SendResults	Function name, value of output argument

- *ReturnTime* is reply of IPSR service with the function execution time;
- *GetAccess* is request of the abstract context service to IPSR service for the resource access time;
- *ReturnAccess* is reply of IPSR service with the resource access time;
- *GetCost* is request of the abstract context service to IPSR service or an acting service for the cost of the provided services;

- *ReturnCost* is reply of IPSR service or the acting service to the abstract context service with the cost of the services they provide for;
- *Include_WS* is notification of the abstract context service to IPSR service or acting service of inclusion in (1) / exclusion from (0) the further negotiations this service based on a principal of *minimum weight*;
- *Ready* is notification of the abstract context service to the registration service that the service community is organised;
- *LookInput* is request to the registration service of (i) IPSR service for input arguments of the functions implemented in IPSR service or (ii) an acting service for identifiers of the acting resources;
- *SendResults* is reply of IPSR service or the monitoring service to the registration service as the result of the executed function.

The decisions about inclusion in / exclusion from the further negotiations the services are made based on principals of *maximum functionality* (*Include_FS* reply)*, maximum access interval* (*Include_AS* reply)*, and minimum weight* (*Include_WS* notification).

Principal of maximum functionality. For the resources that support several functions or roles involvement in the community of one resource providing for several functions or performing several roles is considered more expedient than involvement of several resources providing the same functions or roles separately.

Principle of maximum access interval. If there exist several resources that are accessible over the interval $[t_0, T]$ at different time intervals $\{\Delta t\}$, then the use of less number of resources whose overall access intervals are closest to the interval $[t_0, T]$ is considered to be more efficient. At that, the interval $[t_0, T]$ must be fully covered by the intervals Δt_i^n (Δt_i^n – the i^{th} access interval for the n^{th} resource, $\Delta t_i^n \in \{\Delta t\}$), *i.e.* $[t_0, T] \subseteq \bigcup_{n=1}^{NR} \bigcup_{i=1}^{nt} \Delta t_i^n$ (nt – number of intervals Δt_i^n over the interval $[t_0, T]$ for the n^{th} resource, NR – number of resources in the set WS_a).

Principle of minimum weight. This principle is used to evaluate the efficient use of alternative resources. Alternative resources are represented by Web-services implementing the same functions or roles, and differing in their locations, costs, *etc.* The weight of a resource is represented by the weight of its Web-service. Weight of a Web-service is calculated as $W = \alpha(1-N) + \beta T_r + \gamma C$, where N – the Web-service's competence / reliability, $N = (0, 1)$; T_r – the average access time to the Web-service relatively to the access time that is maximum among the access times for the alternative Web-services; C – the cost of information acquisition from the Web-service relatively to the cost of information acquisition that is maximum among the costs for the alternative Web-services; α, β, γ – relative importance of the parameters for the particular Web-service ($\alpha + \beta + \gamma = 1$).

An example of messaging between the abstract context service and three IPSR services for the purpose of making decision about participation of these IPSR services in the further negotiations is given in Fig. 2. As it can be seen, IPSR service 2 has been excluded from the further negotiations since it is sufficient to take the results of the functions implemented in IPSR service 1 and IPSR service 3 to instantiate all the class attributes containing in the abstract context.

As soon as the service community is organised, the Web-services belonging to it start initiating the abstract context, *i.e.* producing an operational context, whereupon joint actions can be taken.

Services' messaging within the community is demonstrated by an example of services' negotiations to produce an operational context of an accident situation and then take joint actions in the accident response operation (Fig. 3). For the services' interactions the mechanism of blackboard is used. The registration service stands as the blackboard.

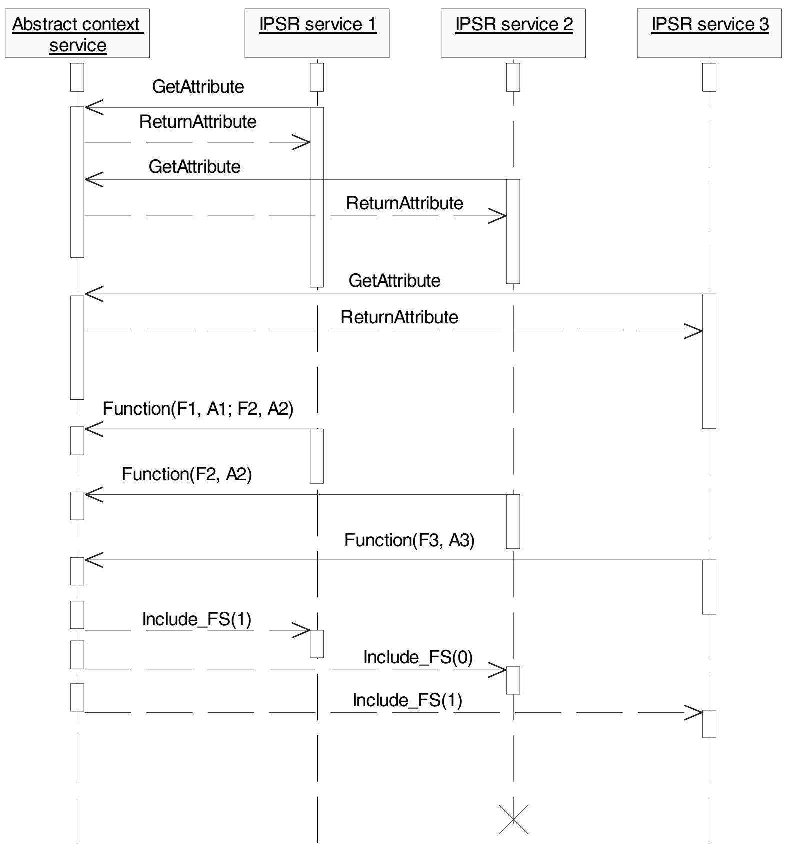


Fig. 2. An example of Web-services messaging when forming a set of IPSR services based on their functionality

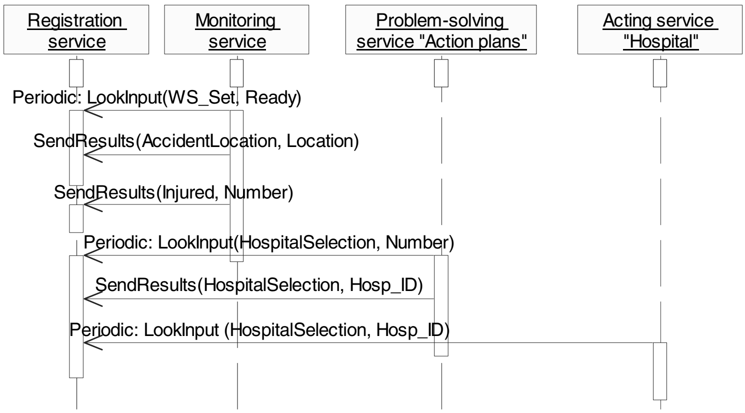


Fig. 3. Web-services' messaging within the community

In Fig. 3 the monitoring service periodically requests the registration service (request *LookInput*). As soon as the monitoring service found out that the service community has been organised (notification *Ready*), it executes the functions it implements. Some of the results of these functions are values for the attributes used to characterize the accident, namely accident location (*Location*) and number of people injured in the accident (*Number*). The functions of the monitoring service as well as functions of some other IPSR services can be executed immediately after the service community has been organised. The matter concerns services implementing functions that do not need values for their input arguments. The function results are sent (*SendResults*) to the registration service to be published.

The services that have not had values for the input arguments of the implemented in them functions, periodically request the registration service. As soon as the registration service has received the needed values, these services execute the implemented functions and send the results of these functions to the registration service.

The acting resource services periodically requesting the registration service wait when identifiers of the resources these services represent will be published. The resource identifiers are the result of the functions implemented in a problem-solving service (the service "Action plans" in Fig. 3) that generate a set of feasible plans of actions [9] and offers one based on some predefined criteria. The criteria can be the minimal time of the response operation, the minimal cost of the operation or a compound criterion. The published resources' identifiers mean that these resources have been selected by the problem-solving service as the possible participants of the joint actions.

The acting resource services in Fig. 3 are represented by the service "Hospital". This service represents characteristics of a hospital as the hospital location, its capacity, specialisation, *etc*. As soon as the service "Hospital" as well as the other acting resource services have found the identifiers of the resources they represent, the resources represented by these services may start the joint actions according to the proposed plan or decline this plan. The latter case is out of consideration in this paper.

Generally, it may relate to further negotiations, regeneration of the plan of actions, and even to a service community reorganisation.

The resources (the human representatives of them) access the operational context using any Internet browser (a browser supported by a notebook, PDA, mobile phone, *etc*.). Fig. 4 presents a piece of the operational context that depicts the route of ambulance transportation from its current location to the location of the accident (the big dot in the screen) and then to a hospital according to the proposed plan.

By now, the presented approach has been implemented as a centralized system. The second step in the approach development is implementation of the protocol of the Web-services self-organisation discussed in the paper and incorporation it in the existing system. Likely application scenarios will be emergency response management and enterprise resource management.

Fig. 1. A piece of operational context

For a preliminary estimation of the protocol AO representing knowledge of the disaster management domain was used. AO contained more than 160 class attributes and around 20 classes representing roles. The abstract context built to characterise a traffic accident situation contained 40 class attributes to be instantiated and 6 classes representing roles. Taking an assumption that one Web-service implements one resource function or plays one role a preliminary rough conclusion can be made that the protocol of context-based Web-services' self-organisation will enable to reduce information search space 3 – 4 times.

5 Conclusion

An approach to self-organisation of resources of a smart environment for the purpose of taking joint resource actions to assist the people involved in a situation is proposed. In the approach ontology-based context is used to represent the needs of the people involved in the situation and to reduce information overflow.

For the implementation of the approach service-oriented architecture of smart environment is developed. It is shown that self-organisation of the resources can be substituted for self-organisation of the Web-services representing these resources. In the process of the self-organisation Web-services produce the picture of the current situation and propose feasible plans of actions for the acting resources. This prevents the decision makers as well as other acting resources from information overload.

To make the Web-services capable of self-organising, an agent-based service model is used. Agents enable Web-services to be the "active" components. Web-services' negotiations are supported by the formal interface agreement defined by the technology of Web-services enriched with the ontology semantics.

The protocol of Web-services self-organisation described in the paper enables efficient organisation of the community of the Web-services depending on the Web-services' functionalities their availabilities and presence of Web-services representing the alternative resources.

The approach is intended for application in domains facing problems of dynamic resource management like logistics, disaster management, enterprise resource planning, *etc*.

Acknowledgments. The paper is due to the research carried out as part of projects funded by grants 08-07-00264 and 09-07-00436 of the Russian Foundation for Basic Research, and project 213 of the research program "Intelligent information technologies, mathematical modeling, system analysis and automation" of the Russian Academy of Sciences.

References

1. Willems, A., Willems, J., Hajdasinski, A.: Information Logistics Research Report: Framework in the healthcare industry, Technical report, no. 09-04, The Nyenrode Research and Innovation Institute (NRI), The Netherlands (2009), `http://www.nyenrode.nl/facultyandresearch/nri/publications/Documents/digitale%20versie%2009-04.pdf`
2. Meissen, U., Pfennigschmidt, S., Wahnfried, T., Sandkuhl, K.: Situation-based Message Rating in Information Logistics and its Applicability in Collaboration Scenarios. In: EUROMICRO 2004, pp. 484–491 (2004)
3. Dey, A.K.: Understanding and Using Context. Personal and Ubiquitous Computing J. 5(1), 4–7 (2001)
4. Reisse, C., Kirste, T.: A Distributed Mechanism for Device Cooperation in Smart Environments. In: Proc. Sixth Intl. Conf. on Pervasive Computing – Pervasive 2008, Sidney, Australia, pp. 53–56 (2008)
5. Feng, Y., Liu, X., Kerridge, J.: An Aspect-Oriented Component-Based Approach to Seamless Web Service Composition. Intl. Transactions on System Science and Applications 5(2), 116–125 (2009)
6. Peltz, C.: Web Services Orchestration and Choreography. IEEE Computer 36(10), 46–52 (2003)
7. Smirnov, A., Pashkin, M., Chilov, N., Levashova, T.: Constraint-Driven Methodology for Context-Based Decision Support. J. Decision Systems 14(3), 279–301 (2005)
8. Smirnov, A., Levashova, T., Krizhanovsky, A., Shilov, N., Kashevnik, A.: Self-Organizing Resource Network for Traffic Accident Response. In: Landgren, J., Jul, S. (eds.) 6th Int. Conf. on Information Systems for Crisis Response and Management (ISCRAM 2009), Gothenburg, Sweden, CD-ROM (2009)
9. Smirnov, A., Pashkin, M., Chilov, N., Levashova, T., Krizhanovsky, A.: Agent-Based Intelligent Support to Coalition Operations: A Case Study of Health Service Logistics Support. In: Shalamanov, V., Johnson, G., Fay, J. (eds.) Information & Security. An Intl J., vol. 16, pp. 41–61. ProCon Ltd., Sofia (2005)

Use Cases for Business Metadata – A Viewpoint-Based Approach to Structuring and Prioritizing Business Needs

Daniel Stock, Felix Wortmann, and Jörg H. Mayer

Institute of Information Management, University of St.Gallen,
Müller-Friedberg-Straße 8, 9000 St.Gallen, Switzerland
{daniel.stock,felix.wortmann,joerg.mayer}@unisg.ch

Abstract. Business metadata plays a crucial role in increasing the data quality of information systems. Despite its importance, business metadata is primarily discussed from a technical perspective, while its business value is scarcely addressed. Therefore, this article aims at contributing to the further development of existing design approaches by explicitly accounting for the use cases of business metadata. In the course of this article a classification of use cases will be developed, which can be utilized when identifying the respective informational requirements a business metadata system needs to satisfy. A banking case is presented that demonstrates how this conceptual model has successfully been applied in practice.

Keywords: Business metadata, requirements engineering, user acceptance.

1 Introduction

1.1 Motivation and Objectives

In recent years, the challenge of "making better use of information" has gained importance and now ranks among the top five priorities of IT executives [1]. This trend is linked to the prevailing significance of Business Intelligence (BI), where data quality is a crucial factor for the perceived net benefits to the end-user [1-3]. In this context the scope of data quality is not limited to factual dimensions like data accuracy, completeness, and objectivity, but also includes individual-related dimensions such as data believability, ease of understanding, and accessibility [4-6]. Especially concerning the individual-related dimensions, business metadata plays a vital role in increasing data quality and therefore the acceptance of BI systems [7-9].

Despite its increasing relevance to practitioners [10], explicit discussions of the benefits and the challenges of implementing respective solutions remain rare [11, 12]. Especially the discussion of benefits of business metadata is generally led on a rather abstract level [8, 13-16]. Foshay et al. [8], for example, base the positive effect of business metadata on the overall usage of BI systems.

Furthermore, business metadata is, in many cases, treated solely from a technical solution perspective [12, 17, 18]. Shankaranarayanan and Even [12], for instance, explore the drawbacks of commercial off-the-shelf products for managing metadata

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 226–237, 2010.
© Springer-Verlag Berlin Heidelberg 2010

and designing and implementing metadata solutions. A comprehensive approach, which combines business requirements on the one hand with the respective design and implementation scenarios on the other, is missing so far. Existing approaches to metadata systems design lack specificity since they provide only insufficient guidance to the identification of (business) metadata requirements and the respective implications for their implementation [19-21].

This article thus contributes to a structured analysis of business metadata requirements by proposing a classification of potential use cases for business metadata. This classification can be utilized applying all the advantages inherent in a viewpoint-oriented requirements engineering approach [22-24] to the design of business metadata systems. By singling out the business metadata requirements of the respective stakeholders, each use case can be examined individually for its business benefits and associated costs of provisioning. Nevertheless, this approach helps maintaining a business perspective by focusing on the application of business metadata rather than developing technical solutions detached from actual business needs.

1.2 Research Methodology and Article Structure

This article applies the design research paradigm in order to accomplish utility. According to Hevner et al. [25] and March and Smith [26], the outcomes of a construction process under the design research paradigm can be classified as constructs, models, methods, and instantiations. Several reference models for the construction process have been proposed. The most recent, consolidated process model by Peffers et al. [27] specifies six phases (ignoring the phase 'communication'):

Identify the problem and motivate: We used a full-text keyword literature research [28] to find the relevant literature on business metadata [6-8, 11-17, 19, 29-39] and business metadata systems design [12, 17-21] to substantiate the missing perspective on the actual business value in business metadata systems design.

Define objectives of solution: The solution was to be designed to match the guidelines for model design in March and Smith [26].

Design and development: We used aggregation in order to develop a classification of business metadata use cases. We consulted a literature review on wide-spread concepts in data processing in a BI environment [40-42], as well as data consumption [43-45], data management [46, 47], and data production processes [40-42] in order to identify the single use cases. It should be noted that the documentation of the respective use cases is out of scope of this contribution.

Demonstration: We applied the conceptual model based on the viewpoint oriented requirements engineering approach to a specific case so that the applicability of the model could be demonstrated.

Evaluation: Finally, the design and the utility of the proposed classification were evaluated, first by us through analytical argumentation and then by focus groups in which interviews were led.

The article is structured as follows. Section 2 provides the conceptual foundations. Section 3 will then proceed with the development of a classification for use cases of business metadata by exploring the distinct stages in data processing. Section 4 demonstrates how this artifact was applied in a banking case and offers an evaluation of the artifact's design and utility. Section 5 concludes with final remarks.

2 Conceptual Foundations

2.1 Business Metadata

Even though the literature offers several definitions that are more precise [e.g., 14, 21, 36], 'data about data' has evolved as the most widespread definition of metadata. In order to be more specific we will adapt the following definition of Dempsey and Heery [36]: "Metadata is data associated with objects which relieves their potential users of having full advance knowledge of their existence or characteristics" (e.g. a definition, an index, or a manual). This definition highlights two crucial aspects of metadata. First, metadata is associated with a certain object, which means it has no meaning by its own. Therefore derived or aggregated data like descriptive statistics (e.g. mean and variance) or KPIs (e.g. turnover and earnings) cannot be regarded as metadata. Second, metadata does not necessarily have to be associated with a single data element, but can also be associated with a whole set of data (e.g. a report) or a physical object (e.g. an IT system).

While business metadata is a sub-category of metadata that is used by the business side, technical metadata in contrast is used by IT [8, 10, 11, 14, 21]. It should be noted that the two sets are not disjoint. This means there is metadata that can be regarded business as well as technical relevant (e.g. descriptions). Seven distinct categories of business metadata can be gleaned from the literature:

1. Definitional metadata [7, 8, 11, 12, 31, 33, 35, 37] – addressing the questions: What do I have and what does it mean?
2. Data quality [7, 8, 11, 12, 31] – addressing the question: Is the quality of data appropriate for its intended use?
3. Navigational metadata [7, 8, 11, 12, 33, 35, 37] – addressing the question: Where can I find the data I need?
4. Process metadata [7, 8, 11, 12] – addressing the questions: Where did this data originate from and what has been done to it?
5. Audit metadata [7, 11, 12, 31] - Who owns the data and who is granted access?
6. Usage metadata [11, 12, 31, 33] – addressing the questions: How frequently is a specific data set/report requested and what user profiles can be derived?
7. Annotations (semi-structured comments) – Which additional circumstances or information do I need to consider when reading this data?

Even though the category 'Annotations' is not mentioned in the examined literature, many BI and metadata tools provide a functionality to add comments or upload additional information, for example in the context of a deviation analysis.

2.2 Viewpoint-Oriented Requirements Analysis

A requirement is a "condition or capability needed by a user to solve a problem or achieve an objective" [48]. Requirements engineering involves capturing, analyzing and decomposing of many ideas, perspectives and relationships at varying levels of detail [24]. Methods based on rigid global schemes do not adequately deal with the diversity of issues presented by requirements engineering problems. To remedy this problem, methods based on the notion of viewpoints have evolved (e.g. Controlled Requirements Expression 'CORE' and Viewpoint-Oriented Requirements Definition 'VORD') [22-24]. The general steps in a VORD approach are illustrated in **Fig. 1**.

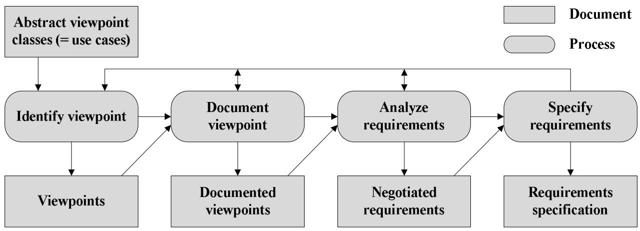


Fig. 1. Viewpoint-oriented requirements definition approach [24]

A viewpoint is a stakeholder and his or her task-related concerns (requirements) [22-24]. The definition of a viewpoint is thus intricately linked to the term 'use case', which describes functional requirements from a specific scenario perspective [49]. The biggest difference is the level of granularity. While a viewpoint is associated to one stakeholder only, a use case can describe the interaction of several stakeholders. Therefore, we will use the term 'use case' to refer to abstract viewpoint classes (see **Fig. 1**), which subsume stakeholders and their respective concerns regarding a specific operational scenario. So, the proposed classification can be conceived of as an overview of abstract viewpoint classes on a business metadata system, which can be used to identify a company's relevant viewpoints.

3 Derivation of Use Cases for Business Metadata

Regardless of the specific architecture, three phases in data processing within a BI context can be distinguished: data consumption, data staging, and data production [40-42]. Consequently, on the basis of this threefold categorization, three groups of business stakeholders can be made out: Data consumers, data managers (e.g. data stewards, business analysts), which ensure that data processing (incl. data staging) runs efficiently, and data producers. In the following, the three business stakeholder groups will serve as the fundamental use case categories. In sections 3.1-3.3, the relevant use cases in each of these three domains will be derived. It should be noted that the naming of the herein listed use cases does not adhere to common UML practices (e.g. containing an active verb) in favor of consistency with the functional areas of a company, where they typically originate from.

3.1 Use Cases in the Data Consumer Domain

Business organizations can be conceptualized as consisting of three different levels: senior management, middle management, and operational management [44, 45, 50]. While the senior management is responsible for corporate management (strategic/ operational planning and forecasting), the middle management controls corporate performance management (increasing organizational performance in order to achieve the goals set by the senior management), and the operational management monitors the daily activities. These three fields of duty strongly depend on analytical information and therefore constitute use cases, which can benefit from business metadata.

Each of these groups of data consumers satisfies their data requirements through two general types of reports: standard and ad-hoc reports [43]. Whereas standard reports anticipate data requirements in advance and are tailored to a specific role, ad-hoc reports meet spontaneous data requirements. However, data consumers may not be as familiar with data provided by ad-hoc reports and might thus need more information to assess its information value. We are therefore adding two extra use cases. First, 'ad-hoc information retrieval', which comprises activities by all data consumers to satisfy their spontaneous data needs. Second, in case of more sophisticated ad-hoc analysis, which requires the assistance of a 'knowledge worker' [46, 47], we will speak of 'advanced analytics'.

In total, five data consumer use cases can be distinguished: corporate management [44, 51], corporate performance management [44, 52], operational management [44], ad-hoc information retrieval [43], and advanced analytics [46, 47].

3.2 Use Cases in the Data Manager Domain

The Data Management Body of Knowledge (DMBOK) of the Data Management Association (DAMA) names 10 functions of data management in its functional reference model [46, 47]. Since this model provides a comprehensive list of data management activities within any kind of organization, not all of them fall under the responsibility of the business side. Therefore, we only include all the relevant activities, which state a business role (data steward, knowledge worker, data quality analyst, and business process analyst) as operationally responsible. **Table 1** lists the respective activities per relevant data management function.

Table 1. DAMA DMBOK business responsibilities [46]

Functions	DMBOK business responsibilities
Data governance	• Manage and resolve data related issues • Monitor and ensure regulatory compliance • Communicate, monitor, and enforce conformance with data policies, standards, procedures, and architecture
Data development	• Validate information requirements • Prepare for data deployment
Data security management	• Understand data security needs & regulatory requirements • Define data security policy • Define data security standards
Reference & master data management	• Indentify reference data sources and contributors • Define and maintain matching rules • Establish golden records • Define and maintain hierarchies and affiliations • Manage changes to reference and master data
Data quality management	• *All activities concerning data quality management fall under the responsibility of the business side*

3.3 Use Cases in the Data Producer Domain

Independently of the specific function, a data producer is responsible for data entry and maintenance [40-42]. This includes all activities that lead to the creation, update, or deletion of a data set. Since the data production process is the source of several data quality issues (e.g. inconsistencies), business metadata can be a means of proactive data quality management by providing support and increasing transparency. The use case 'data entry and maintenance' is the only one in the data producer domain. In total, eleven use cases could be identified, which are listed in **Fig. 2**.

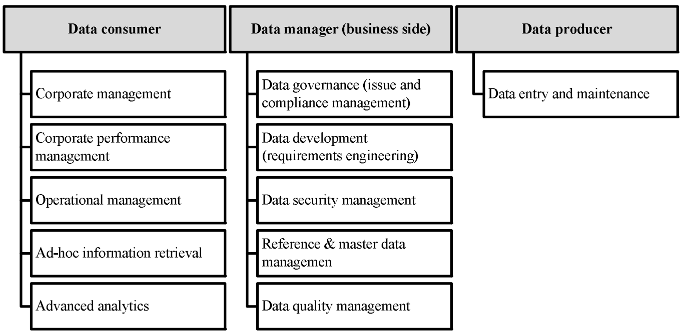


Fig. 2. Classification of use cases for business metadata

4 Demonstration and Explorative Evaluation

In this section, we will demonstrate the applicability of our conceptual model by means of a case study at "European Universal Bank" (EUB). Section 4.1 will introduce the initial situation at EUB and section 4.2 will give details of how the proposed classification was applied. Eventually, we will evaluate the design and utility of our conceptual model according to the criteria of March and Smith [26].

4.1 Initial Situation

EUB offers its 15 million customers all the services of a universal bank and considers Central and Eastern Europe as their home market. EUB started its initiative on business metadata in early 2008. Primarily driven by the IT department, the focus was laid on introducing a corporate wiki as a means of encouraging a common and unambiguous terminology. This effort was confronted with much resistance from the business side as no obvious business need could be seen.

In order to make the next advance to business metadata a success, EUB was seeking for a more business-driven approach. Rather than starting again from a certain implementation scenario, this time the overarching concept of business metadata and its benefits for the business should be explored.

4.2 Use Case Approach to Business Metadata Systems Design

As described in the previous section, the IT department of EUB decided to revise their current approach to business metadata. Therefore, we applied our proposed use case classification following the presented VORD approach presented above (see **Fig. 1**). As a first step, EUB prioritized the use cases in order to focus the analysis on the most relevant ones (section 4.2.1). We then documented in an exemplary fashion the requirements for the top priority use case in a collaborative workshop (section 4.2.2).

4.2.1 Use Case Prioritization

The EUB chose two employees of business and two of IT to identify the use cases with the biggest relevance for EUB. All of them were asked to study a comprehensive set of use case assessment templates and to provide their personal perspectives on the business potential and associated costs/burdens of their implementation per use case.

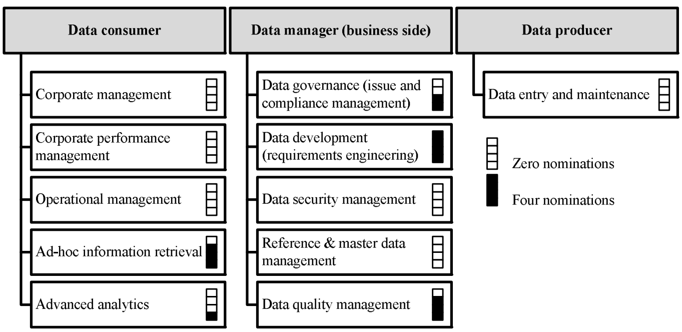


Fig. 3. EUB's prioritization of use cases for business metadata

Each template included a high level description of the use case, typical challenges, and likely business metadata needs to address these challenges. Subsequently, the employees were asked to name their top three to five use cases and explain the underlying rationales. **Fig. 3** illustrates the collective results. From the given use cases, the EUB employees identified 'data development' as the most relevant for EUB and therefore the prime candidate for exemplary viewpoint documentation.

4.2.2 Viewpoint Documentation

The use case 'data development' was further assessed in a workshop with 13 employees from EUB's IT and risk department. Within this group the use case was broken down into viewpoints. The group identified five relevant stakeholders and seven activities which benefit from business metadata.

Table 2. Business metadata requirements per activity in use case 'data development'

Activities per stakeholder	Definitional	Data quality	Navigational	Process	Audit	Usage	Annotations
Business units							
• Specifying change request	√		√		√		√
• Monitoring implementation		√					
Legal authority							
• Monitoring compliance	√	√			√		
Project manager							
• Estimating effort	√		√	√		√	
IT							
• Analyzing impact to the IS landscape	√	√	√	√	√	√	
Sponsor							
• Approving change request	√		√		√		√
• Monitoring change request		√					

For each activity, the workshop participants were asked to assess the usefulness of the seven business metadata categories listed in section 2.1 (see **Table 2**). Further discussions revealed that the workshop participants would value 'definitional', 'process', and 'usage' business metadata the most across all activities. 'Definitional' since almost everybody benefits from it, 'lineage' due to unsolved challenges in effort estimation, and 'usage' in order to install a report life cycle management that identifies unused reports that can be abandoned.

4.3 Evaluation of Design and Utility

Two aspects of the presented classification of use cases for business metadata is up for evaluation: its design and its utility. With respect to its design, the classification needs to fulfill the following requirements: 'completeness', 'fidelity with real world phenomena', 'internal consistency', 'level of detail', and 'robustness' [26].

Regarding its utility it needs to demonstrate that it contributes to either effectiveness and/or efficiency in the design of business metadata systems. Starting with its design, we argue that our conceptual model is likely to satisfy the requirements 'completeness', 'fidelity with real world phenomena', 'internal consistency' and 'robustness' due to our analytical-deductive approach of aggregating existent and widespread models. Only for the dimension 'level of detail' we cannot argue with our analytical-deductive approach, since granularity is a subjective measure and therefore has to be assessed in a specific context.

Table 3. Statements of focus group on design and utility

Statements regarding design (focus on the level of detail)
• Very clearly structured framework • Missing a use case 'risk management' • Complete list of topics to address/cover • Use case 'advanced analytics' too abstract
Statements regarding utility
• Helps in identifying the business value of business metadata • Facilitates step by step approach instead of rushing too early into technical discussions • Thrives the identification of business value • Better understanding of benefits by linking it to roles and activities • Good to maintain a business perspective • It will be hard to come up with a quantifiable business case (just not feasible)

Furthermore, we evaluated the classification's design and utility by means of a focus group with seven employees of EUB, which had participated in the use case prioritization and viewpoint documentation workshop discussed before (see **Table 3**).

Regarding the design, the participants challenged the features 'completeness' and 'level of detail'. In order to complement the classification they suggested including a use case 'risk management'. Since risk management is a function which holds special significance in financial institutions, including it as an extra use case would compromise the classification's generality. Concerning the second point, further investigation is necessary as to whether the use case 'advanced analytics' is too abstract because it subsumes too many distinct functions or whether the notation is not clear enough.

Concerning the utility, the participants were generally very satisfied with the utility of the conceptual model and the associated viewpoint-oriented approach. They particularly appreciated that it focused the discussions are concentrated on a business perspective rather than exploring technical details. Even though this is all in all a very positive evaluation result, it does lack representativeness. The participants of the focus group were all from the same company and thus this conceptual model needs further evaluation in different contexts to reliably prove its applicability and utility.

5 Concluding Remarks

This article has derived a classification for business metadata use cases. The classification comprises use cases in the data consumer, data management, and data producer domain. Its applicability was demonstrated in a banking case and subsequently evaluated within a focus group. The evaluation is not representative and must be subjected to further research. Especially the use case 'advanced analytics' needs further investigation in order to comply with the design requirement 'level of detail'.

In addition, there is potential for future research that builds on the newly developed classification presented above. First, this conceptual model could be used to derive information requirement patterns for business metadata to apply situational method engineering to business metadata systems design. Second, the typical business metadata requirements per use case could be extended via a representative evaluation which would pave the way to a stronger alliance between business value and technical implications.

References

1. Luftman, J.N., Kempaiah, R., Rigoni, E.H.: Key Issues for IT Executives 2008. MISQ Executive 8, 151–159 (2009)
2. Sabherwal, R., Jeyaraj, A., Chowa, C.: Information System Success: Individual and Organizational Determinants. Management Science 52, 1849–1864 (2006)
3. Wixom, B.H., Watson, H.J.: An Empirical Investigation of the Factors Affecting Data Warehousing Success. MIS Quarterly 25, 17–41 (2001)
4. Wang, R.Y.: A Product Perspective on Total Data Quality Management. Communications of the ACM 41, 58–65 (1998)
5. Wang, R.Y., Strong, D.M.: Beyond Accuracy: What Data Quality Means to Data Consumers. Journal of Management Information Systems 12, 5–34 (1996)
6. Jarke, M., Lenzerini, M., Vassiliou, Y., Vassiliadis, P.: Fundamentals of Data Warehouses. Springer, Heidelberg (2000)
7. Foshay, N.: The Influence of End-user Metadata on User Attitudes toward, and Use of, a Data Warehouse. IBM, Somers (2005)
8. Foshay, N., Mukherjee, A., Taylor, A.: Does Data Warehouse End-User Metadata Add Value? Communications of the ACM 50, 70–77 (2007)
9. Bucher, T., Wlk, U.: Bereitstellung und Nutzung von Endbenutzungs-Metadaten im Kontext der Informationslogistik - Empirische Erkenntnisse und Gestaltungsempfehlungen. In: DW 2008 - Synergien durch Integration und Informationslogistik, pp. 75–93. Gesellschaft für Informatik, Bonn (2008)
10. Schulze, K.-D., Besbak, U., Dinter, B., Overmeyer, A., Schulz-Sacharow, C., Stenzel, E.: Business Intelligence-Studie 2009. Steria Mummert Consulting AG, Hamburg (2009)
11. Shankaranarayanan, G., Even, A.: Managing Metadata in Data Warehouses: Pitfalls and Possibilities. Communications of the Association for Information Systems 14, 247–274 (2004)
12. Shankaranarayanan, G., Even, A.: The metadata enigma. Communications of the ACM 49, 88–94 (2006)
13. Chengalur-Smith, I., Ballou, D.P., Pazer, H.L.: The Impact of Data Quality Information on Decision Making. IEEE Transactions on Knowledge and Data Engineering 11, 853–864 (1999)
14. Even, A., Shankaranarayanan, G., Watts, S.: Enhancing Decision Making with Process Metadata: Theoretical Framework, Research Tool, and Exploratory Examination. In: Proceedings of the 39th Hawaii International Conference on System Sciences, pp. 1–10. IEEE Computer Society, Los Alamitos (2006)
15. Fisher, C.W., Chengalur-Smith, I., Ballou, D.P.: The Impact of Experience and Time on the Use of Data Quality Information in Decision Making. Information System Research 14, 170–188 (2003)
16. Shankaranarayanan, G., Watts-Sussman, S.: A Relevant Believable Approach for Data Quality Assessment. In: Proceedings of the MIT International Conference on Information Quality, pp. 178–189 (2003)

17. Hüner, K.M., Otto, B.: The Effect of Using a Semantic Wiki for Metadata Management: A Controlled Experiment. In: Proceedings of the 42nd Hawaii International Conference on System Sciences, pp. 1–9. IEEE Computer Society, Los Alamitos (2009)
18. Nevo, D., Wand, Y.: Organizational Memory Information Systems: a Transactive Memory Approach. Decision Support Systems 39, 549–562 (2005)
19. Marco, D.: Building and Managing the Meta Data Repository. John Wiley & Sons, Inc., New York (2000)
20. Melchert, F.: Integriertes Metadatenmanagement: Methode zur Konzeption von Metadatenmanagementsystemen für das Data Warehousing. Logos, Berlin (2006)
21. Tannenbaum, A.: Metadata Solutions: Using Metamodels, Repositories, XML, and Enterprise Portals to Generate Information on Demand. Addison-Wesley, Boston (2002)
22. Darke, P., Shanks, G.: Stakeholder viewpoints in requirements definition: A framework for understanding viewpoint development approaches. Requirements Engineering 1, 88–105 (1996)
23. Leité, J.C.S.P., Freeman, P.A.: Requirements Validation Through Viewpoint Resolution. IEEE Transactions on Software Engineering 17, 1253–1269 (1991)
24. Kotonya, G., Sommerville, I.: Requirements Engineering: Processes and Techniques. John Wiley & Sons Ltd., New York (1998)
25. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly 28, 75–105 (2004)
26. March, S.T., Smith, G.F.: Design and Natural Science Research on Information Technology. Decision Support Systems 15, 251–266 (1995)
27. Peffers, K., Tuunanen, T., Gengler, C.E., Rossi, M., Hui, W., Virtanen, V., Bragge, J.: The Design Science Research Process: A Model for Producing and Presenting Information Systems Research. In: Proceedings of the First International Conference on Design Science Research in Information Systems and Technology, pp. 83–106 (2006)
28. vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In: Proceedings of the ECIS 2009, pp. 2206–2217 (2009)
29. Benander, A., Benander, B., Fadlalla, A., James, G.: Data Warehouse Administration and Management. Information Systems Management 17, 71–80 (2000)
30. Levine, M., Siegel, J.: What the Accountant Must Know About Data Warehousing. The CPA Journal 71, 37–42 (2001)
31. Niland, J.C., Rouse, L., Stahl, D.C.: An Informatics Blueprint for Healthcare Quality Information Systems. Journal of the American Medical Informatics Association 13, 402–417 (2006)
32. Pipe, P.: The Data Mart: A New Approach to Data Warehousing. International Review of Law Computers 11, 251–261 (1997)
33. Vaduva, A., Vetterli, T.: Metadata Management for Data Warehousing: An Overview. International Journal of Cooperative Information Systems 10, 273–298 (2001)
34. West, L.A., Hess, T.J.: Metadata as a knowledge management tool: supporting intelligent agent and end user access to spatial data. Decision Support Systems 32, 247–264 (2002)
35. Xie, G., Yang, Y., Liu, S., Qiu, Z., Pan, Y., Zhou, X.: EIAW: Towards a Business-friendly Data Warehouse Using Semantic Web Technologies. In: Aberer, K., Choi, K.-S., Noy, N., Allemang, D., Lee, K.-I., Nixon, L.J.B., Golbeck, J., Mika, P., Maynard, D., Mizoguchi, R., Schreiber, G., Cudré-Mauroux, P. (eds.) ASWC 2007 and ISWC 2007. LNCS, vol. 4825, pp. 857–870. Springer, Heidelberg (2007)
36. Dempsey, L., Heery, R.: Metadata: A Current View of Practice and Issues. Journal of Documentation 54, 145–172 (1998)

37. Müller, R., Stöhr, T., Rahm, E.: An Integrative and Uniform Model for Metadata Management in Data Warehousing Environments. In: Proceedings of the International Workshop on Design and Management of Data Warehouses, pp. 12–28. ACM Press, New York (1999)
38. O'Neil, B.: `http://www.tdan.com/view-articles/4934`
39. Shankaranarayanan, G., Ziad, M., Wang, R.Y.: Managing Data Quality in Dynamic Decision Environments: an Information Product Approach. Journal of Database Management 14, 14–32 (2003)
40. Ariyachandra, T.R., Watson, H.J.: Data Warehouse Architectures: Factors in the Selection Decision and the Success of the Architectures (2005)
41. Inmon, W.H., Strauss, D., Neushloss, G.: DW 2.0: The Architecture for the Next Generation of Data Warehousing. Elsevier Science, Amsterdam (2008)
42. Kimball, R., Ross, M.: The Data Warehouse Toolkit. John Wiley & Sons, Inc., New York (2002)
43. Horváth, P.: Controlling. Franz Vahlen, München (2006)
44. Laudon, K.C., Laudon, J.P.: Management Information Systems - Managing the Digital Firm. Prentice-Hall, Upper Saddle River (2006)
45. Mintzberg, H.: The Rise and Fall of Strategic Planning. Prentice-Hall, Englewood Cliffs (1994)
46. DAMA: The DAMA Guide to the Data Management Body of Knowledge. Technics Publications, New Jersey (2009)
47. Mosley, M.: DAMA-DMBOK Functional Framework Version 3. DAMA International (2008)
48. IEEE: IEEE Standard Glossary of Software Engineering Terminology. IEEE Computer Society, New York (1990)
49. Eriksson, H.-E., Penker, M.: Business Modelling with UML - Business Patterns at Work. John Wiley & Sons, New York (2000)
50. Anthony, R.N.: Planning and Control Systems: A Framework for Analysis. Harvard University Press, Boston (1965)
51. Vandenbosch, B., Huff, S.L.: Searching and Scanning: How Executives Obtain Information from Executive Information Systems. MIS Quarterly 21, 81–107 (1997)
52. Nord, J.H., Nord, G.D.N.: Executive Information Systems: A Study and Comparative Analysis. Information & Management 29, 95–106 (1995)

LIT 2010 Workshop Chairs' Message

Piotr Stolarski[1], Tadeusz Tomaszewski[1], and John Zeleznikow[2]

[1] Poznan University of Economics, Department of Information Systems,
al. Niepodleglosci 10, 60-967, Poznań, Poland
[2] Laboratory of Decision Support and Dispute Management,
School of Information Systems, Victoria University
Room G.03, Land Titles Office, 283 Queen St., Melbourne, Victoria, Australia
{P.Stolarski,T.Tomaszewski}@kie.ue.poznan.pl,
John.Zeleznikow@vu.edu.au

The 3rd Legal Informatics and Legal Information Technology Workshop will be remembered as a successful event – both to the organizers as well as the workshop community. This time the workshop was held in Berlin on 3rd May 2010. The workshop was organized in conjunction with the 13th International Conference on Business Information Systems. We are now pleased to provide to readers the post-proceeding containing the accepted papers.

Our pleasure arises from three distinct facts: 1. the event was well received by all participants including quite a few number who are not Legal Information community members; 2. the high quality of the accepted papers; 3. the event supports the integration and cooperation among the community members, which in our view is an extremely positive aspect.

LIT 2010 had far more papers submitted to it, than the previous Legal Informatics and Legal Information Technology Workshops. In this edition, 16 papers were submitted, out of which 7 works were accepted for publication. In comparison to the 2nd LIT, two important factors should be raised here: 1. the number of overall submissions has doubled while at the same time 2. the organizers managed to lower the acceptance rate (43.7%).

All this, combined with fact that despite the result we managed to stick to the rigorous procedures in terms of accepting papers, similar to the ones applied to the previous edition[1] we are certain that all this factors inevitably lead to better quality of the workshop and specifically the works contained in this publication.

Dr. Thomas F. Gordon from Frauenhofer Fokus gave an inspiring opening address. He discussed the lessons learned from the LKIF framework construction. In the paper of Boer et al. the problem of mediation towards knowledge representation for reconstruction of traces to sources of law as well as to implementation knowledge

[1] Review notes were independently granted by at least 3 reviewers assigned randomly from the Programme Committee. The cases of possible conflicts of interest were avoided in advance. The topic bidding procedure have been adapted in order to take into account reviewers' topical preferences. In some cases of largely diversified reviews we took advice from additional reviewers.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 238–239, 2010.

resources is being scrutinized. This research is especially vital for knowledge maintenance issue in administrative organizations together with legal information analytic tasks. Kubiak et al. propose an alternative for framework of qualified electronic signatures. The expected improvement is to offer much better properties in terms of usability, costs, risk level and security. Peruginelli and Francesconi in their study investigate feasibility of development of a collaborative Web platform with automatic, language-independent procedures for the support of Multicultural Legal Domain thesauri. Such a platform is the key for ensuring future multilingual legal information access at semantically interoperable level. Another text of Boonchom and Soonthornphisaj is the continuation of the quest of authors for a marvel algorithm for legal ontologies learning. The suggested solution uses some known ant colony techniques for enriching ontology seeds. Bueno et al. presents the developments in the Ontojuris Project. This large system for legal information retrieval is now extended with the multilingual ontology editor. The performance biasing is achieved with introduction of new concept of Universal Words. Another work switches form the topic of interoperability and ontologies towards Natural Language Processing. This is the continuation of effort announced last year in which German court decisions are being automatically analyzed and the content structure is labeled. Last but not least is the paper of Bellucci et al. This research focuses on fairness within information technology tools to support negotiation and mediation. The key issue at this stage of is the use of providing BATNAs and integrative bargaining in this class of systems.

This year's event covered a huge area of themes with many challenges appearing on the horizon.

We would like to direct our cordial thanks to all the authors as well as participants whose work and attendance enabled us to organize the whole workshop in this successful way. We are aware that there would be no workshop in this shape without the participating community. We also would like to express our deepest thankfulness towards the Programme Committee members who supported us with their knowledge and judgments. Without the silent work of those people the selection of the works would not be feasible. Finally, as always we would like to express our profound gratefulness to the organizers of the hosting conference (BIS2010). Without their help, the organization of the 3rd Legal Informatics and Legal Information Technology Workshop would not have been possible.

This year the workshop took place at exceptional place for the European information technology. The Zuse Center was built to honor the first computing machinery in Europe[2]. We hope that the next pages will be no less inspiring for the Reader than the visit at the place. We also look forward to next year's Legal Information Technology Workshop.

---

[2] Konrad Zuse (1910-1995) is accepted to be the constructor of the first programmable, binary computer. His early works did not endure the WW II. Zuse has been the winner of many international awards and *doctor honoris-causa* title around the world.

An Overview of the Legal Knowledge Interchange Format

Thomas F. Gordon

Fraunhofer FOKUS
Berlin, Germany
thomas.gordon@fokus.fraunhofer.de

The Legal Knowledge Interchange Format (LKIF) is an XML Schema for representing theories and arguments (proofs) constructed from theories ESTRELLA Project (2008). A theory in LKIF consists of a set of axioms and defeasible inference rules. The language of individuals, predicate and function symbols used by the theory can be imported from an ontology represented in the Web Ontology Language (OWL). Importing an ontology also imports the axioms of the ontology. All symbols are represented using Universal Resource Identifiers (URIs). Other LKIF files may also be imported, enabling complex theories to modularized.

Axioms are named formulas of full first-order logic. The heads and bodies of inference rules are sequences of first-order formulas. All the usual logical operators are supported and may be arbitrarily embedded: disjunction ($\wedge$), conjunction ($\vee$), negation ($\neg$), material implication ($\rightarrow$) and the biconditional ($\leftrightarrow$). Both existential ($\exists$) and universal ($\forall$) quantifiers are supported. Free variables in inference rules represent schema variables.

Terms in formulas may be atomic values or compound expressions. Values are represented using XML Scheme Definition (XSD) datatypes. Atomic formulas are reified and can be used as terms, allowing some meta-level propositions to be expressed.

The schema for atomic formulas has been designed to allow theories to be displayed and printed in plain, natural language, using Cascaded Style Sheets (CSS). An atomic proposition may first be represented in propositional logic, using natural language, and later enriched to become a first-order model, by marking up the variables and constants of the proposition and specifying its predicate using an XML attribute. This feature of LKIF is essential for enabling domain experts, not just computer specialists, to write and validate theories.

Support for allocating the burden of proof when constructing arguments from theories in dialogues is provided. An *assumable* attribute is provided for atomic formulas, to indicate they may be assumed true until they are challenged or questioned. An *exception* attribute is provided for negated formulas, to indicate that $\neg P$ may be presumed true *unless* P has been proven. Making a negated formula, $\neg P$, an exception shifts the burden of proof for P to the opponent of P in the dialogue.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 240–242, 2010.
© Springer-Verlag Berlin Heidelberg 2010

Arguments in LKIF link a sequence of premises to a conclusion, where both the premises and the conclusion are atomic formulas. Attributes are provided for stating the direction of the argument (pro or con), the argumentation scheme applied and the role of each premise in an argument. Arguments can be linked together to form argument graphs. The legal proof standard each proposition at issue must satisfy, such as "preponderance of the evidence" or "beyond reasonable doubt" may be specified. Attributes are provided for recording the relative weight assigned to each argument by the finder of fact, such as the jury, or some other audience, as well as the status of each issue in the proceeding. Arguments represented in LKIF can be evaluated and analyzed using a variety of computational models of argument, including both ASPIC Prakken (2010) and Carneades Gordon et al. (2007); Gordon and Walton (2009)

All of the main elements of an LKIF file may be assigned Universal Resource Identifiers, allowing them to be referenced in other documents, anywhere on the World Wide Web. Cross references between elements of legal source documents and the elements of the LKIF document which model these sources may be included within the LKIF file, using a sequence of *source* elements. The scheme allows m to n relationships between legal sources and elements of the LKIF model to be represented.

LKIF builds on and uses many existing World Wide Web standards, including the XML, Universal Resource Identifiers, XML Namespaces, the Resource Description Framework (RDF) and the Web Ontology Language (OWL). However, for a variety of reasons it does not use other XML schemas for modeling legal rules, such as Common Logic, RuleML, the Semantic Web Rule Language (SWRL), or the Rule Interchange Format. Common Logic is an ISO standard for representing formulas of first-order classical logic. While LKIF includes a sublanguage for first-order logic, LKIF has been designed to allow formulas of first-order logic to be represented in human readable form in natural language, to ease development, maintenance and validation by domain experts. Moreover, the ISO Common Logic standard does not look like it will be widely adopted within the World Wide Web community, which has its own standards body, the World Wide Web Consortium. RuleML, SWRL and RIF, among other efforts, are competing to become the Web standard for rules. At the beginning of the ESTRELLA project, SWRL was the leading candidate. In the meantime, during the development of LKIF in Estrella, RIF has become the leading contender. But neither SWRL nor RIF are currently expressive enough for the legal domain. Legal rules can best be understood as domain-dependent defeasible inference rules. They cannot be adequately modeled as material implications in first-order logic. However, an LKIF theory can in principal import a first-order theory represented in any XML format, to be used as part of the axioms of the theory. This feature of LKIF enables a part of the legal theory to be represented in first-order logic and imported into an LKIF theory, similar to the way LKIF currently can import ontologies represented in OWL, using whatever format eventually becomes the World Wide Web standard for rules.

Acknowledgements

The work reported here was supported by a grant from the European Commission, in the ESTRELLA project (IST-4-027655).

References

ESTRELLA Project. The legal knowledge interchange format (LKIF). Deliverable 4.1, European Commission (2008)

Gordon, T.F., Walton, D.: Proof burdens and standards. In: Rahwan, I., Simari, G. (eds.) Argumentation in Artificial Intelligence, pp. 239–260. Springer, Berlin (2009)

Gordon, T.F., Prakken, H., Walton, D.: The Carneades model of argument and burden of proof. Artificial Intelligence 171(10-11), 875–896 (2007)

Prakken, H.: An abstract framework for argumentation with structured arguments. Argument & Computation (in press, 2010)

How Information Technology Can Support Family Law and Mediation

Emilia Bellucci, Deborah Macfarlane, and John Zeleznikow

School of Information Systems, Victoria University, Melbourne, Australia
{emilia.bellucci,deborah.macfarlane,john.zeleznikow}@vu.edu.au

Abstract. In Australia, before a divorcing couple can have their case heard by the Family Court, they must undertake mediation. Thus it is useful to develop information technology tools to support negotiation and mediation in family law. Most negotiation support systems focus upon integrative bargaining. In doing so, they tend to ignore issues of fairness. In Australian Family Law, the interests of the children, as opposed to those of their parents/guardians, are paramount. We investigate the use of providing BATNAs and integrative bargaining in providing family mediation decision support. The discussion is highlighted with examples taken from the domain of Australian Family Law.

Keywords: Negotiation Support Systems, Principled Negotiation, bargaining and trade-offs, Australian family law.

1 Introduction

The idea that Information Technology (IT) can be gainfully used by mediators to provide advice and support for disputing couples is not one that is intuitively obvious. Its impersonality does not appear to lend itself well to the skilful balancing act required to settle such disputes. We argue, however, that computers can help negotiation by providing quick and easily accessible decision support. In this article, we will demonstrate several computer systems capable of complementing existing methods of interaction by providing support for assessment and information management.

[15] reflects on the introduction of IT to negotiation as a move from resolving disputes outside the courtroom (ADR) to meetings taking place somewhere in cyberspace or “directly to each individual’s personal computer.’

Many commentators argue that the most important aspect of mediation is face-to-face communication [10]. There are many circumstances, however, where face-to-face communication is either not feasible or undesirable. Examples include, but are not limited to:

- Parties who have a history of violent conflict, or
- Parties for whom the costs of being in the same room are exorbitant, or
- Parties who are in different time zones, or
- Parties who cannot agree upon a joint meeting time.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 243–255, 2010.
© Springer-Verlag Berlin Heidelberg 2010

The lack of in-person interaction can actually be an advantage for disputes in which the emotional involvement of the parties is so high that it is preferable that they do not see each other.

The development of family mediation centres, resourced with qualified mediators, using software that can help provide fast, efficient, accurate and cheap advice, can be very useful. In section 3, we will describe a computer software tool which provides advice on asset division in family law, currently being trialed at the Brisbane Office of Relationships Australia (RAQ).

In this article, we will outline our approaches for using Information Technology to provide decision support for family mediators. However, we wish to emphasise we are providing support for mediators whose goal is among other things to help divorcing couples resolve any disputes with minimal conflict. At no stage do we wish to automate decision making. To have durable just agreements, it is important for the parties, rather than a computer, to have control of the negotiation process. However, useful negotiation decision advice can help the parties reach a quicker and more congenial resolution of their dispute.

2 Negotiation Support Systems in Australian Family Mediation

2.1 Australian Family Mediation

In Australia, mediation – generally facilitative mediation – has been used to handle disputes in the family arena for about twenty years[1]. The 'family arena' generally comprises property or child-related disputes arising between parents, whether married or not, and whether they have lived together or not. Such disputes may also involve other people, related to the children or who have cared for the children; however, the overwhelming majority of disputes mediated are couple-related.

The typical approach to mediation expected the couple enter mediation voluntarily. However, this is not the situation Australia today: at least one meeting with a specialist "family dispute resolution practitioner" is mandatory before lodging an application for a parenting order in the Family Court. The reasons for the change are complex. Certainly, a major factor in the past two decades has been the marked increase in the breakdown of family relationships, resulting in excessive workloads for courts dealing with family matters.

Arguably, economic drivers have also played a significant role for at least two reasons, first, because of the limitations in the number of available judges and funding to the courts; and, second, because of the cost to the public purse of court action for matters which can be dealt with more expeditiously and cheaply by other means. Also influential is evidence that points to a major emotional cost to the family of court action; a situation, at odds with the 'best interests of the child' principle which theoretically governs family court actions. Finally, the success of voluntary mediation

[1] Although there was a focus on alternative dispute resolution in the Family Court of Australia from its inception in 1975, conciliation and counselling were initially employed and it was not until the late 1980s that mediation began to be used.

has convinced decision-makers that mediation is a better alternative than adjudication for the majority of family disputes, whether undertaken voluntarily or not.

Australian family dispute resolution has undergone rapid transformation in the space of 20 years. Beginning in the 1980s, mediation was considered an alternative form of dispute resolution. Over the next 15 years, mediation became a primary approach to dispute resolution and from 1 July 2008 has been a compulsory first step if parents cannot reach agreement on arrangements for their children. That mediation has proven an effective way of dealing with family relationships, together with the existence of a substantial body of research demonstrating the detrimental effects of conflict between parents on their children, provides logical justification for Government enacting compulsory dispute resolution.

2.2 The DEUS and Family Negotiator Systems

DEUS [39] represented our earliest attempt at building negotiation support systems in Australian Family Law. It is a template-based system[2]. [39] developed a model of family law property negotiation, which relies upon building a goal for each of the litigants, with the goals being supported by their beliefs. Goals can only take real number values, because in simplifying the model it is assumed that the goal of each party is a monetary figure. Beliefs, which support the goals, are expressed in natural language. In the system, which has been implemented using this model, goals are used to indicate the differences between the parties at a given time. The model calculates the agreement and disagreement between the litigants' beliefs at any given time. The agreement and disagreement are only in relation to the beliefs and hence do not resolve the negotiation. In order to reach a negotiated settlement, it is essential to reduce the difference between the goals to nil. DEUS is a useful tool for gaining an understanding of what issues are in dispute and the extent of the dispute over these issues.

Our earliest Negotiation Support System was Family_Negotiator [5]. It utilises a hybrid rule-based and case-based system to provides disputants with advice on how to best resolve the issues in an Australian Family Law dispute. While evaluating the Family_Negotiator system, we discovered that Family Law negotiation was not an appropriate domain in which to apply either Case-based or Rule-based Reasoning, due principally to the open textured nature[3], of the domain. Nor did the overall framework of Family_Negotiator provide in-depth solutions expected from real-life negotiations.

---

[2] Traditionally, negotiation support systems have been template based, with little attention given to the role the system itself should play in negotiations and decision-making support. The primary role of these systems has been to demonstrate to users how close (or far) they are from a negotiated settlement. The systems do not specifically suggest solutions to users. However, by informing users of the issues in dispute and a measure of the level of the disagreement, they provide some decision support.

[3] Open textured legal predicates contain questions that cannot be structured in the form of production rules or logical propositions and which require some legal knowledge on the part of the user in order to answer.

2.3 Family_Winner and Asset_Divider

In this section we will detail our most recent contributions to supporting family mediation through IT. Both computer systems facilitate asset division in family law. Family_Winner was developed foremost, and proved to be a successful tool in generating media interest in our work[4].

Family_Winner takes a common pool of items and distributes them between two parties based on the value of associated ratings. Each item is listed with two ratings (a rating is posted by each party), which signify the item's importance to the party. A rating in Family_Winner is a number in value from 0-100 (0 being of no importance; 100 to signify absolute importance). The algorithm to determine which items are allocated to whom works on the premise that each parties' ratings sum to 100; thereby forcing parties to set priorities. The program always checks this is the case, and if not, it scales ratings to ensure all sum to 100. The basic premise of the system is that it allocates items based on whoever values them more. Once an item has been allocated to a party, the ratings of the remaining items are modified (by firing trade-off equations) to ensure the items (and their associated ratings) are ready for the next round of allocation [3].

Family_Winner allocates items to one of two parties in the dispute. Family_Winner's method of decision support involves a complex number of techniques, including the incorporation of an Issue Decomposition Hierarchy, a Compensation and Trade-off strategy, and an Allocation strategy. The trade-offs pertaining to a disputant are graphically displayed through a series of trade-off maps, while an Issue Decomposition Hierarchy enables disputants to decompose issues to any required level of specification.

Family_Winner was evaluated by a number of family solicitors at Victoria Legal Aid (VLA) and was presented to mediators from RAQ. While both the solicitors and mediators were very impressed with the way Family_Winner suggested trade-offs and compromises, they had one major concern – that in focusing upon negotiation, the system had ignored the issues of justice by not acknowledging the items in dispute have a monetary value [6]. Family_Winner simply allocates property to parties based on their interest in the item. It does not allow for monetary values to influence the allocation process. The dollar value of items is important to the dispute because each party wants to be allocated the right or 'just' amount of money. This concept contrasts with assigning an interest value to an item, which is intrinsically different. An interest is an evaluation based on the significance of the item to a person. For example, party A may be very fond of a lamp that has been passed down throughout the generations, and consequently they give it a rating of 50. The remaining items are not as important to party A, and so are given much lower ratings. While using interests to negotiate is a

[4] In late 2005, the Family_Winner system achieved much media attention, including over a dozen radio interviews in all Australian states and on BBC Radio 5 and separately the BBC World Service, articles in the Sydney Morning Herald, the Times of London, the Australian Financial Review and the Economist. The inventors were asked to compete on ABC (Australian Broadcasting Commission) New Inventors television show on November 16 2005. They won their heat – see http://www.abc.net.au/tv/newinventors/txt/s1504763.htm last accessed February 3 2008.

very interesting exercise, it does not in any way reflect the dollar value of the item. This is where Family_Winner fails to support the mediation process realistically.

The AssetDivider system [4] incorporates the basis of Family_Winner's allocation and trade-off strategy to decide upon the allocation of assets based on interests and an item's monetary value.

2.4 Asset_Divider

Family_Winner and AssetDivider's input and output

AssetDivider accepts a list of items together with ratings (two per item) to indicate the item's importance to a party. In addition it also accepts the current monetary value of each item in dispute. We assume this dollar value has been negotiated (if necessary) before AssetDivider is used [5]. Hence, only one dollar value is entered per item. The proposed percentage split is also entered; this reflects what percentage of the common pool each party is likely to receive in the settlement. The system is not capable of determining the percentage split; this figure has to be derived from the mediator's knowledge in past cases or from computer systems such as SplitUp [26], which can provide a percentage split given certain characteristics and features of divorce cases.

AssetDivider's output consists of a list of items allocated to each party. All of the items (except one) on the allocation lists were provided in the intake screen by the disputants. The additional item is a "payout" item, which reflects the amount of money a disputant would need to pay the other party for the items they have been allocated and collectively are valued greater than the percentage split offers them. For example, party A has been allocated a total value of $100,000 in assets, and party B $115, 000. Under a 50/50 split, party B will need to pay $15,000 to party A to satisfy the percentage split.

AssetDivider's Allocation Strategy

The order by which issues are allocated is of paramount importance in a negotiation. Through interviews with mediators from the (then) Family Mediation Centre in Mitcham, we discovered issues attracting little disagreement should be presented foremost for allocation, so as to help foster a positive environment in which to negotiate. By summing the ratings of issues to 100, the level of discourse surrounding an issue can be measured by calculating the numerical distance between the ratings of an issue assigned by each of the parties. For example, if two parties assign the same high rating to an item, then it is expected the level of disagreement surrounding the issue to be substantial (because both parties want the item), whereas large differences between the ratings of parties indicate the issue will be resolved much more quickly. AssetDivider use this strategy in deciding the order by which items are presented for allocation.

---

[5] Sometimes the parties cannot agree on the monetary value of the item. In this case, mediators would reference standard objective tables and the like to reach a consensus. For example, if parties are arguing over the value of a car, then mediators may access websites that gave independent valuations, such as redbook.com.au.

AssetDivider's allocation strategy works by allocating an item to the party whose rating is the highest ie to parties according to whoever values them the most. It then checks the dollar value of items it has been allocated previously (that is, their current list of items), the dollar value of the item presently allocated and the dollar amount permitted under the percentage split given by mediators. If by allocating the item in question the party exceeds its permitted amount, the item is removed from its allocation list and placed back into negotiation. In this case, the item has not been allocated to a party. If the dollar value of the item was within the limits of the amount permitted under the percentage split rule, then the allocation proceeds.

Once an item has been allocated to a party, the remaining ratings (of items still in dispute) are modified by trade-off equations. These modifications try to mimic the effect losing or gaining an item will have on the rest of the items still in dispute. The equations directly modify ratings by comparing each one against that of the item recently lost or won (each party's set of ratings are modified as a result of an allocation). The equations update ratings based on a number of variables - whether the item allocated was lost or gained, the value of the allocated item in relation to items still in dispute and the value of the item whose rating will change as a resultit is Only the 'losing party' in AssetDivider is compensated by the trade-off equations modifying ratings (whereas in Family_Winner both winning and losing parties were affected). The extent to which ratings were modified was determined through an analysis of data we collected from mediation cases provided by the Australian Institute of Family Studies. These are detailed in [3]. For more technical details on how the ratings are modified in AssetDivider, ie applicable equations and functions, see [4].

An example using AssetDivider

This section will review the process and outcome of a Family Law case on AssetDivider. The aim of this exercise is to demonstrate AssetDivider's operation in practice.

The case description of this real-life divorce scenario and the relative point allocations have been extracted from [7][6]. The case Jolis v Jolis, began on December 5th, 1980, and concluded on October 30th, 1981. The case was heard in New York City, at a time when a new law subjecting all martial property to a 50–50 split was being introduced. The couple had been married for 41 years, and had spent 33 of these years together. The Wife had given up her early and successful career to care for the couple's four sons. The couple had lived together in substantial wealth, primarily due to the expansion of the Husband's diamond business.

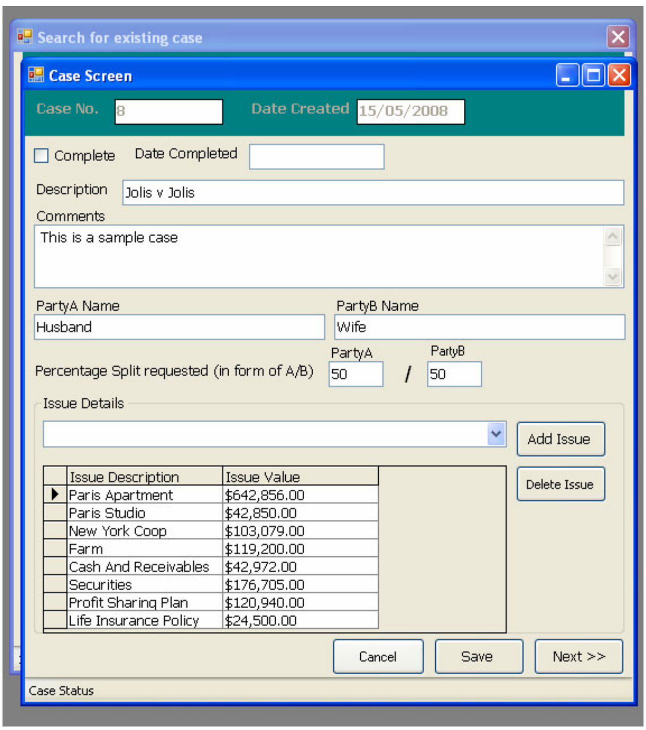
There are both real estate and liquid assets to be divided. The Husband's diamond business is not treated as marital property as its growth was primarily due to market forces, especially the diamond boom of the 1970s. The children's welfare is not included as an issue as they were no longer considered minors at the time of separation.

[6] At page 105.

Table 1. Point allocations and dollar valuations

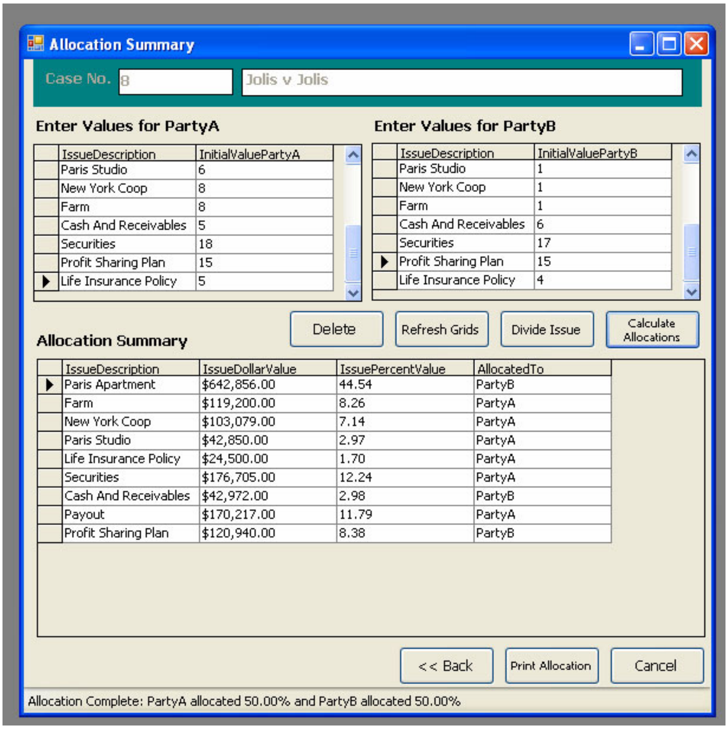
Issues	H's ratings	W's ratings	Dollar value of asset
Paris Apartment	35	55	$642,856
Paris Studio	6	1	$42,850
New York Coop	8	1	$103,079
Farm	8	1	$119,200
Cash And Receivables	5	6	$42,972
Securities	18	17	$176,705
Profit Sharing Plan	15	15	$120,940
Life Insurance Policy	5	4	$24,500
Total	100	100	$1,273,102

First we will discuss the case how the case is presented to the AssetDivider system. The relevant case information is entered in screen 1.



Screen 1. Intake screen for negotiation

The next screen (screen 2) that appears lists the issues in dispute, their ratings and the allocation summary, which is filled in when the user clicks button "Calculate allocations".



Screen 2. Final output screen from the AssetDivider system. It gives the user the allocation list for each party; which includes a payout figure allocated accordingly.

According to AssetDivider, the preferred outcome, taking into account each party's' priorities (ratings) and percentage split is:

Table 2. Allocation list for Husband (party A) and Wife (party B) using AssetDivider

Husband (Party A)	**Value of**	**Wife (Party B)**	**Value of**
Farm	$119,200	*Paris Apartment*	$642,856
New York Coop	$103,079	*Cash and receivables*	$42,972
Paris Studio	$42,850	*Profit Sharing Plan*	$120,940
Life Insurance Policy	$24,500		
Securities	$176,705		
Payout	$170,217		-$170,217
Total:	$636,551	*Total:*	$636551

In analysing the case, we can see both parties wanted the Paris apartment above all else; though Wife (party B) valued it more than the husband (Party A). As a consequence, both parties gave the rest of the items relatively low values. On the whole, both parties received the items they valued considerably (except for Party A's loss of Paris apartment to Party B – since she valued it more highly). The only item valued equally by the parties was the profit-sharing plan [15]. It was given to Party B. Party B also needed to pay out Party A the amount of $170,217 to ensure the split was exactly 50%.

The current version of Asset Divider allows for negative contributions from the partners (such as credit card and gambling debts). It also allows for the parties to make cash payments to each other – for example if the wife is given the house, which is the only major asset arising from the marriage, then she may need to make a cash payment to the husband.

3 Current and Future Research

AssetDivider has not been evaluated extensively by industry. While RAQ has installed the software on their network, they are yet to commence its use or evaluation. To aid evaluation, we have developed two questionnaires to be conducted with mediators and an interview schedule with managers, based on the CCCF System Operational Context [13].

The questionnaires ask uses to comment on the operation and use of the system, and include:

1) Mediator usage census, and
2) Mediator survey

The mediator's usage census (1), assesses how mediators use the program to assist in mediations. Ideally it should be filled in after a mediation has concluded (whether Asset Divider was used or not), and allows users to enter their thoughts on why Asset Divider was used (or why not), how it was used and their comments on its usefulness.

The survey for mediators (2) should be filled out at the end of the trial (evaluation) period. It allows mediators to reflect on the appropriateness of the program, its user friendliness, quality and overall feel of its usefulness to family mediation.

It is expected mediators at RAQ will test and evaluate the system in the near future. We are also in negotiations with other organisations for the purposes of evaluation and commercial use of the software. We expect an extensive evaluation campaign to reveal further improvements to the decision making module and user interface. Our research has revealed a lack of negotiation support systems for family law. We hope our collaboration with RAQ (or other organisations) will enable AssetDivider to be used in practice, being the first negotiation support systems to do so.

In March 2008, we commenced a large research grant from the Australian Research Council[7] to work with the Queensland Branch of Relationships Australia, to develop a negotiation decision support system using both interested-based and justice-based negotiation.

Online dispute resolution makes it possible to provide family law services to parties who are geographically dispersed. In such circumstances, the process of

[7] Linkage Project LP0882329 Developing negotiation decision support systems that promote constructive relationships following disputes.

separation and divorce could historically feature a large amount of correspondence and expensive litigation. Australia is both a country and a continent, the population is geographically spread and a part of the population lives in remote regions. The provision of government services to such communities is both difficult and expensive.

To meet such needs the Australian Government initiated the Family Relationship Advice Line (FRAL)[8]. The Advice Line is a national telephone service established to provide information to people affected by family relationship or separation issues. FRAL can also refer callers to local services that can provide assistance.

In a major addition to traditional family dispute resolution services, FRAL will organise telephone dispute resolution for people assessed as more appropriate for the telephone medium than for face-to-face services. The Telephone Dispute Resolution Service (TDRS) was established through funding from the Federal Attorney General of the Australian Government in 2007 [30] and is based in Queensland, operated by Relationships Australia (Qld) in partnership with Relationships Australia (NSW). Potential clients cannot directly contact or self refer to the TDRS; rather they need to be referred either through the Family Relationship Advice Line, a Family Relationships Centre, or any other government-funded Family Dispute Resolution provider.

Online Family Dispute Resolution services potentially broaden the options and capacity of existing services by offering another method of supporting families to reach agreements. Given that most people in disputes about the care of children are under 45, the great majority of the disputants will have familiarity with and access to the appropriate technology. For those who do not have access, the technology could be made available at locations such as local libraries, Community Legal Centres or Family Relationship Centres. Furthermore, such an approach could include provision of mobile resources in centres or hire facilities. Compared to the costs of litigation or even prolonged alternative dispute resolution, the investment in online technologies is potentially both value adding and cost saving.

In conjunction with Relationships Australia Queensland, we are developing an Online Family Dispute Resolution Service. We are also investigating principles for the successful negotiation of Information Technology Outsourcing Agreements [8].

Currently we are about to commence two projects that investigate the fairness and consistency of negotiation support systems in the legal domain. In a project with the title 'Developing negotiation support systems in law which encourage more consistent and principled outcomes' we argue that unless negotiation support systems are seen to advocate outcomes which arise from consistent and principled advice, disputants will be reluctant to use them. Thus we propose conducting research that will develop measures for assessing the outcomes of online negotiation in the legal domains of sentencing, plea bargaining and family mediation. Such measures will form the basis of a new model for evaluating justice and consistency within online dispute resolution systems. The model will inform the construction of fairer and more consistent systems of IT-based negotiation support in the future.

---

[8] See http://www.ag.gov.au/www/agd/agd.nsf/Page/Families_FamilyRelationshipServicesOverviewofPrograms_ForFamilyRelationshipServicesPractitioners_FamilyRelationshipAdviceLineResources last accessed 16 August 2009.

To meet this goal we will:

1. Develop models of consistency and justice based on two very distinct legal domains: sentencing and family law. Further, the knowledge about these domains will be shared from three distinct Common Law jurisdictions: Australia, Israel and USA.
2. Develop information retrieval techniques to extract knowledge from textual legal and negotiation data.
3. Use KDD techniques (such as association rules, Bayesian belief networks and neural networks) to compare litigated and negotiated family law cases.
4. Develop models of disputation and negotiation in both family law and sentencing. These models will then be tested to examine how closely they align with the notion of Bargaining in the Shadow of the Law (as compared to 'pure' interest-based negotiation).
5. Use Lodder and Zeleznikow's three step model for an Online Dispute Resolution Environment and Toulmin's theory of argumentation to construct a generic online dispute resolution environment. The development of such an environment on which to place various negotiation support systems will increase users' access to justice.
6. Develop and evaluate specific sentencing and negotiation support systems using our newly developed Online Dispute Resolution Environment.

While producing significant information technology and socio-legal advances, the project also develops innovative artificial intelligence techniques to help develop 'consistent' and 'just' negotiation support systems and an environment for online dispute resolution.

[12] argue that rather than develop negotiation support systems that are useful in resolving legal disputes, it is more important to develop systems that provide advice about how to avoid conflicts. In domains such as family law and body corporate disputes, it is important for the disputants to have ongoing relationships once the disputes have been resolved. Hence it is vital to develop software that focuses upon developing ongoing relationships, rather than focusing upon interests, justice (using BATNAs) or power. Negotiation support systems that incorporate machine learning and planning can help in this regard.

References

1. Abrahams, B., Bellucci, E., Zeleznikow, J.: Incorporating Fairness into Development of an Integrated Multi-agent Online Dispute Resolution. Submitted to Group Decision and Negotiation Journal (2010)
2. Astor, H., Chinkin, C.M.: Dispute resolution in Australia, 2nd edn. LexisNexis Butterworths, Sydney (2002)
3. Bellucci, E.: Developing Compensation Strategies for the construction of Negotiation Decision Support Systems. PHD thesis, La Trobe University, Bundoora 3086, Victoria, Australia (2004)

4. Bellucci, E.: AssetDivider: A New Mediation Tool in Australian Family Law. In: Hindriks, K.V., Brinkman, W.-P. (eds.) Proceedings of the First International Working Conference on Human Factors and Computational Models in Negotiation, pp. 11–18. Delft University of Technology, Delft (2008)
5. Bellucci, E., Zeleznikow, J.: Family-Negotiator: an intelligent decision support system for negotiation in Australian Family Law. In: Proceedings of the Fourth Conference of the International Society for Decision Support Systems, Lausanne, pp. 359–373 (1997)
6. Bellucci, E., Zeleznikow, J.: Developing Negotiation Decision Support Systems that support mediators: a case study of the Family_Winner system. Journal of Artificial Intelligence and Law 13(2), 233–271 (2006)
7. Brams, S.J., Taylor, A.D.: Fair Division, from cake cutting to dispute resolution. Cambridge University Press, Cambridge (1996)
8. Chandar, S., Zeleznikow, J.: Risks To Consider When Negotiating IT Outsourcing Agreements. In: Proceedings of GDN 2009, the 10th Annual Meeting INFORMS Section on Group Decision and Negotiation, Toronto, Ontario, June 14-17 (2009), `http://info.wlu.ca/~wwwmath/faculty/kilgour/gdn/papers.htm` (last accessed August 26, 2009)
9. Conley Tyler, M., McPherson, M.: Online Dispute Resolution and Family Disputes. Journal of Family Studies 12(2), 165–184 (2006)
10. Eisen, J.: Are We Ready for Mediation in Cyberspace? Byu L. Rev. 1305, 1322–1325 (1998)
11. Fisher, R., Ury, W.: Getting to YES: Negotiating Agreement Without Giving In. Haughton Mifflin, Boston (1981)
12. Gray, P., Gray, X., Zeleznikow, J.: Decision Negotiating Logic: For richer or poorer. In: Winkels, R. (ed.) Proceedings of Eleventh International Conference on Artificial Intelligence and Law, pp. 247–251. ACM Press, Palo Alto (2007)
13. Hall, M.J.J., Zeleznikow, J.: A method for evaluating legal knowledge-based systems based upon the Context Criteria Contingency guidelines framework. In: Proceedings of the Ninth International Conference on Artificial Intelligence and Law, pp. 274–283. ACM Press, Edinburgh (2003)
14. Jennings, N.R., Faratin, P., Lomuscio, A.R., Parsons, S., Wooldridge, M., Sierra, C.: Automated negotiation: Prospects, methods and challenges. Group Decision and Negotiation 10, 199–215 (2001)
15. Karamon, M.: ADR on the Internet, 11 OHIO ST. J. on Disp. Resol., 537–548 (1996)
16. Kersten, G.E.: Support for Group Decisions and Negotiations. In: Climaco, J. (ed.) An Overview, in Multiple Criteria Decision Making and Support. Springer, Heidelberg (1997)
17. Kersten, G.E.: Modeling Distributive and Integrative Negotiations. Review and Revised Characterization. Group Decision and Negotiation 10(6), 493–514 (2001)
18. Kolodner, J.L., Simpson, R.L.: The Mediator: Analysis of an Early Case-Based Problem Solver. Cognitive Science 13, 507–549 (1989)
19. Kraus, S.: Strategic Negotiation in Multi-Agent Environments. MIT Press, Cambridge (2001)
20. Lax, D.A., Sebenius, J.: The Manager as Negotiator. The Free Press, New York (1986)
21. Lodder, A., Zeleznikow, J.: Developing an Online Dispute Resolution Environment: Dialogue Tools and Negotiation Systems in a Three Step Model. The Harvard Negotiation Law Review 10, 287–338 (2005)
22. Matwin, S., Szpakowicz, S., Koperczak, Z., Kersten, G.E., Michalowski, G.: NEGOPLAN: An Expert System Shell for Negotiation Support. IEEE Expert 4(4), 50–62 (1989)

23. Peterson, M., Waterman, D.A.: Evaluating civil claims: An expert systems approach to evaluating liability cases. In: Walter, C. (ed.) Computer Power and Legal Reasoning, pp. 627–659. West Publishing Company, St. Paul (1985)
24. Ross, H.: Settled Out of Court, Aldine (1980)
25. Rubinstein, A.: Perfect equilibrium in a bargaining model. Econometrica 50, 97–109 (1982)
26. Stranieri, A., Zeleznikow, J., Gawler, M., Lewis, B.: A hybrid—neural approach to the automation of legal reasoning in the discretionary domain of family law in Australia. Artificial Intelligence and Law 7(2-3), 153–183 (1999)
27. Sycara, K.: Negotiation planning: An AI approach. European Journal of Operations Research 46, 216–234 (1990)
28. Sycara, K.: Machine Learning for Intelligent Support of Conflict Resolution. Decision Support Systems 10, 121–136 (1993)
29. Thiessen, E.M., McMahon, J.P.: Beyond Win-Win in Cyberspace. Ohio State Journal on Dispute Resolution 15, 643 (2000)
30. Thomson, M.: Annual Report of the Telephone Dispute Resolution Service to the Attorney General. Relationships Australia, Queensland (2009)
31. Van Velthoven, B.C.J., Ter Voert, M.J.: Geschilbeslechtingsdelta 2003. WODC/Boom, The Hague (2004)
32. Walton, R.E., Mckersie, R.B.: A Behavioral Theory of Labor Negotiations. McGraw-Hill, New York (1965)
33. Waterman, D.A., Paul, J., Peterson, M.: Expert systems for legal decision making. Expert Systems 3(4), 212–226 (1986)
34. Wertheim, E., Love, A., Littlefield, L., Peck, C.I.: Win: You Win Penguin Books, pp. 130–135. Ringwood, Victoria (1992)
35. Wilkenfeld, J., Kraus, S., Holley, K.M., Harris, M.A.: GENIE: A decision support system for crisis negotiations. Decision Support Systems 14, 369–391 (1995)
36. Williams, G.: Legal Negotiation and Settlement. West Publishing Co., St. Paul (1983)
37. Wooldridge, M.: Agent-Based Software Engineering. IEE Proc. Software Engineering 144(1), 26–37 (1997)
38. Zeleznikow, J., Bellucci, E.: Family_Mediator - adding notions of fairness to those of interests. In: Proceedings of Nineteenth International Conference on Legal Knowledge Based System, pp. 121–130. IOS Publications, Amsterdam (2006)
39. Zeleznikow, J., Meersman, R., Hunter, D., van Helvoort, E.: Computer tools for aiding legal negotiation. In: ACIS'95 — Sixth Australasian Conference on Information Systems, pp. 231–251. Curtin University of Technology, Perth (1995)

Traceability of the Implementation of Legal Rules in Public Administration

Alexander Boer, Tom van Engers, and Radboud Winkels

Leibniz Center for Law, University of Amsterdam, The Netherlands
{A.W.F.Boer,T.M.vanEngers,R.G.F.Winkels}@uva.nl

Abstract. While *isomorphism* of knowledge representation has been recognized as important, particularly to maintenance in legal knowledge representation, the requirements of the maintenance process in general get less attention. Traceability from knowledge resources used in the organization to the sources of law used in their production is a central maintenance issue in administrative organizations. This paper explores a mediating knowledge representation for reconstruction of traces to sources of law and to implementation knowledge resources, that should be helpful for analysis of the impact of changing sources of law.

1 Introduction

While *isomorphism* has been recognized as important to maintenance in legal knowledge representation, the maintenance process itself gets little attention. Traceability from knowledge resources used in the organization to the sources of law used in their production is a central maintenance issue in administrative organizations. This paper explores an *Agile* knowledge representation for reconstruction of traces to sources of law and to implementation knowledge resources, that should be helpful for analysis of the impact of changing sources of law.

In the Agile project, we focus on robust traceability, and on dealing with the risks involved in agile network arrangements with other organizations. The increasing interest in public administration for traceability to sources of law and modeling network arrangements is spurred by:

1. increasing commercial traffic between jurisdictions, which leads to increasing interest in the complications of *legal pluralism*[1] in its broader sense,
2. the increasing dependence of service provision on negotiated network arrangements between organizations, also within the same jurisdiction,
3. the increased formalization of interactions through ICT [1] (XML standards, the service paradigm, etc), and generally
4. the perception that the pace of change in society increases.

Network arrangements are for instance service contracts, resource sharing, automated and standardized data exchange, etc, between independent agents that only enter into such arrangements if and as long as these are beneficial to the participants. No larger

[1] The EU and its free traffic of goods, services, and persons should for instance come to mind here.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 256–267, 2010.
© Springer-Verlag Berlin Heidelberg 2010

organizational framework creates and enforces the network arrangement. Loose network arrangements make it harder to organize compliance to the law: meeting obligations no longer necessarily translates to control objectives, to allocation of responsibilities to roles in fixed business processes and to specific groups of employees in departments. The organization that depends on service delivery of others becomes responsible for (intelligently monitoring and reacting to) what happens outside its organization.

Traceability is the solution to the legal provenance issue involved in legally justifying the structures and knowledge resources of the organization. Following implementation traces back to sources of law is helpful for diagnosis when undesired or unintended results are produced, and is the starting point of impact analysis when the law changes.

Our approach to traceability, explained in section 4, and demonstrated in section 5, is to distinguish between three different universes of discourse, layered on top of eachother. The entities in each layer are explicitly linked to corresponding entities in adjacent layers by a logical theory, in order to be able to trace the impact of changes in sources of law all the way to implementation in an organization, and back from experiences on the work floor to the legislator's intentions. The logical theory will account for in what sense the sources of law and implementation knowledge resources are related. Provenance information about the sources of law, based on MetaLex (section 3), and applicability rules (section 4.2) also play an important role.

We expect that the interjection of an abstract legal institutional layer between law and implementation in the legal requirements model leads to more accurate documentation of the impact of changes, and better reflects the degrees of freedom the organization has in implementation. The *Agile* project, and the pilot implementation projects that are part of it, will test this hypothesis.

2 Sources of Law in Process and Product Development

The issue of traceability and impact analysis is of relevance not mainly for the primary business process, but for internal product and process development activities within public administration. Product and process development translates the legal requirements into implementation knowledge resources, which guide primary business processes, but it also takes input from the work floor and gives feedback to the legislator. As such the activities can be conceived of as taking place within two feedback cycles – one with the primary individual case handling processes, and one with the legislative process. See Fig. 1 for a visualization of these cycles, using the decomposition of problem solving activities proposed in [2].

The purpose of the legal requirements model mediating between sources of law and implementation knowledge resources presented in the next sections should be seen in the context of these two cycles. A simple set of traceability edges is not enough. The objective is to develop a knowledge representation methodology for a requirements model that is explicit enough about the relation between implementation and law to be useful, cost-effective to implement, and conceptually simple enough to be understandable to people with the skill set typically found in administrative product and process development activities.

The representation explored in this paper is focused on representing the meaningful uses the sources of law afford (in the sense advocated by for instance [3]), i.e. the things the organization can or should do, and the things that can happen to it, becuase the sources of law exist, or are changed. Legal rules afford avoidance of violations, the assumption of certain social roles and execution of predictable interaction protocols, monitoring of the behaviour of others, misrepresentation of our intentions to others to our own advantage, etc.

This contextual meaning of legal rules leads us to reject the notion that a knowledge base for a decision support system, or an ontology used to validate case data in a database, can be considered as the primary ("executable") specification of the way in which legal rules are implemented in the organization. The issue is rather which distinct uses the rule affords, and what design patterns the organization's knowledge engineers should consider when interpreting them.

A simple example: The legal protocol that guides buying and selling, and the formal evidence produced in this protocol in the form of orders, invoices, and receipts, provides the material for various fraudulent disbursement, money laundering, and tax evasion schemes. Besides the primary use of legislation about sales transactions for the execution of sales transactions, a tax administration may for instance use the same rules to organize information about tax evasion. A fraudulent disbursement is for instance transaction that follows the sales pattern, but without an actual duality of stock flows. A sale far below market price between family members is for instance not an arm's length transaction and is classified as a gift. The tax administration *uses* the legal rules in a monitoring and assessment activity, and changes of these rules are relevant to it. The role these rules play in the activity is however not apparent from studying the content of the rules.

This kind of misrepresentation, and associated monitoring activities, are quite common and constitute a design pattern; An immigration administration for instance takes an interest in marriages of convenience, and develops interpretation rules for those.

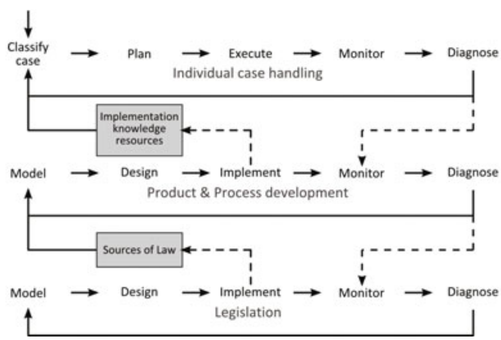


Fig. 1. The legislative and product and process development feedback cycles

3 Bibliographic Identity of Sources of Law

To implement traceability from knowledge representation to sources of law, the Agile project builds on the results of our work on MetaLex XML (cf. for instance [4,5,6]). MetaLex is a common document format, processing model, and metadata set for software development, standardized by a CEN/ISSS committee. MetaLex requires adherence to a URI[2] based, open, persistent, globally unique, memorizable, meaningful, and "guessable" naming convention for legislative resources based on provenance information. This provenance information can be directly used in OWL2 [7].

MetaLex and the MetaLex naming convention strictly distinguish the source of law as a published work from its set of expressions over time, and the expression from its various manifestations, and the various locatable items that exemplify these manifestations, as recommended by the Functional Requirements for Bibliographic Records (FRBR; cf. [8]).

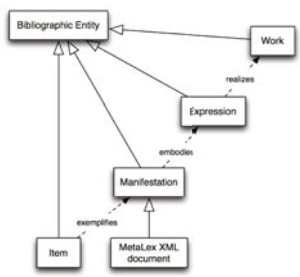


Fig. 2. Taxonomy of bibliographic entities in MetaLex, and their relata, based on FRBR

MetaLex extends the FRBR with a detailed but jurisdiction-independent model of the lifecycle of sources of law, that models the source of law as a succession of consolidated versions, and optionally *ex tunc* consolidations to capture the possibility of correction (errata corrige) or annulment after the fact of modifications by a constitutional court. In these cases the version timeline is changed retroactively: the conceptual time travel involved is an excellent example of the weird applications of constitutiveness. See for instance [9] for an explanation of the practical ramifications of annulment, and more generally an overview of the complexities involved in change of the law.

In current organizational practice links are more often than not made to locatable items, often even without formal agreements about the permanence of the used item identifiers *between different departments of the same organization*. Correct traceability to the right bibliographic abstraction (generally work or expression depending on the purpose of the reference) is – particularly at the levels below *formal* law – a notable weak point, and *ex tunc* change scenarios are not explicitly modeled, or even recognized. MetaLex makes this aspect of the traceability problem explicit, and provides tools to address it.

In the MetaLex metadata set, represented in an OWL ontology, the `realizes` property between expressions and works represents the connection between the two ontological levels at which documents exist that are of relevance to their real world use (see Fig. 2). The source of law on the expression level for instance *cites* other rules on the work level, while the legal rules we represent knowledge about are necessarily identified by their representation in expressions.

[2] Uniform resource identifier.

A citation *(text fragment) w applies to (concept) C* should for instance be read as *each legal rule that is represented by an expression-level text fragment that* `realizes` *work fragment w applies to C*. This representation technique, also an implementation of the idea of ontological stratification, also plays an important role in the Agile project.

The idea of the MetaLex standard is of course that provenance metadata will be supplied by the publisher of the used XML manifestation, and is extracted from it in OWL form by organizations that use it.

3.1 Actions Performed on Documents

In MetaLex provenance information is organized around actions performed on documents.

Provenance metadata more often than not exists of simple predicate-object statements about electronic documents, even though permitting different perspectives on the same action, because its identity was not made explicit, may yield incompatible metadata descriptions, results in unnecessary duplication of metadata, and separate occasions in which to make mistakes, therefore creates unnecessary maintenance, and, lastly, the loss of relevant references between documents [10].

An action for instance plays the mediating role between relevant entities and the resource the metadata description is about. The natural coherence between for instance between *author*, *publication date*, and *publication channel* information (e.g. state gazette bibliographic information) is apparent to all: all are participants in the publication (promulgation) event. There is for instance also a natural coherence between an old consolidation, the new consolidation, the modifying legislation, the modifying authority, and the modification date: the modification event links them together.

Because actions also play a central role on the other layers, we choose for a uniform representation of action inspired by MetaLex. Generally, we try to build on analogies between the legislative domain and the implementation domain, choosing the same representation solutions for both.

4 Agile Rules

Conceptually, Agile knowledge representation is based on a distinction between three universes of discourse (see Fig. 3): law, legal institutional reality, and implementation in brute reality. Between these three layers we find a simple and uniform interface, existing of a *representation*[3] relation, which dominates between first and second layer, and a constitutiveness or *counts as* relation, which dominates between third and second layer.

The Agile knowledge representation consists of logical rules describing these three layers, individually and as a whole. While there are logical rules bridging these layers, these rules are defeasible. Rules that describe the structure of one layer are not.

The first layer addresses the relevant provenance and efficacy information about the sources of law. The sources of law include not only legislation and case law, as already pointed out, but also policy statements and guidelines of lower level rule makers used in the organization.

[3] Or presentation if one prefers to take that ontological stance.

Fig. 3. Law, legal institutional reality, and implementation in brute reality

On the next layer we find abstract legal institutions, whose presence is produced by the sources of law [11,12,6]. We take the source of law to be a writing that may be used to back an argument concerning the presence of a certain institutional entity, often but not always a legal rule, created by the legislator in a certain legal institution [5]: The source of law is the result of a legislative act performed with the intent of creating that institutional entity, and functions as evidence of the legislative act. In this sense the sources of law are the media of legal institutional reality.

Finally, on the third layer, there is the implementation of the legal institutions in brute reality. The institutional reality as represented in the sources of law only comes to life through the brute reality that constitutes (or counts as) it. The raising of a hand for instance counts as a bid in an auction. The issuing of a document invented by some administrative agency tasked with issuing residence permits similarly counts as an official residence permit.

In the context of Agile the brute reality of interest is revolves around business processes and services. Related to these are *resources* such as documents and forms, fielded software, business process specifications, manuals, etc, but also the declarative, logical rules in fielded knowledge bases.

The *legal rules* presented by the sources of law constrain the structure of institutional reality and the superposition of institutional reality on brute reality. Any *logical* rule in the Agile model mapping brute reality (the raising of the hand) to institutional reality (the bid) may either be backed by the applicability of a legal rule presented by the sources of law or simply by knowledge of the domain, as for instance represented by business process and service specification resources. Both legal rules and the various rules one may find in software (production rules, integrity constraints) are from the perspective of the mediating model also *entities*, with the function of resource, in the domain being represented by the rules of the mediating requirements model.

The entities named in rules, including knowledge resources, are all URI-identifiable and may serve as anchors for reference from other knowledge resources. The ability to make some sort of URI-based hyperlinks from other knowledge resources is increasingly common, and doesn't impose a high implementation burden.

Agile knowledge representation is based on OWL2 axioms, augmented with MetaLex and a defeasible and more expressive rule formalism [13].

Agile rules are reified using the RDF reification mechanism, may be version managed on work and expression level, to reflect improved understanding over time, and may also carry provenance information. Provenance information on reified OWL2 rules create the possibility for a future belief base revision approach to defeasibility [14,6], and is also the preferred approach to extending the operational semantics of Agile Rules in other directions.

The intention in Agile is not to add to the state-of-the-art in defeasible logic, but rather to come to an account of which argument schemes are – in the proper context – a defensible basis for decision making in the organization.

4.1 Constitutiveness and Representation

Institutional events are constituted by events in brute reality. The main function of the constitutive rule is to define the interface through which the state of the institution can be changed through necessary and indicative conditions. Constitutive rules may be defeasible. Institutional rules map out a logical space of possible models of the institution: they form the institution's ontology, and can be interpreted as terminological axioms [11,6]. They are not to be considered defeasible as a matter of convention, because permitting ambiguity about the institutional ontology makes no sense from the administrative organization's perspective.

Constitutiveness is modeled through the *constitutes* (inverse *constitutedBy*) property. This property applies to *legal things*, and not legal propositions: we do not follow the custom of talking about *legal facts* arising from brute facts in representation. Newly created legal things *must* be constituted by another thing, which means in essence that one of the constitutive rules that can create a legal fact *must* be applicable [6]: a legal thing is constitutedBy some thing, and a legal rule is applicable to it.

Constitutiveness is often accompanied by (re)presentation, modeled through the *represents* (inverse *representedBy*) property. A formal act is an act of representation of a legal proposition, with the intent of being constitutive of that proposition. This pattern is typical for legislation itself, and is also often found in administrative transactions.

The use of formal acts creates confusion between proposition and the use of formal representations of a proposition as evidence. For immigration, the proposition that someone is married may for instance be legally relevant; In the implementation of this criterion this for instance becomes the proposition that someone has supplied a marriage certificate, but even if the marriage certificate must be renewed every year, a certificate may in fact be still valid at the moment of decision making while the marriage is not at that point in time in existence. Moreover, some issuers of certificates in foreign countries may be notoriously unreliable, etc. Behind a single proposition in the law are often complete forms, procedures, checks, constraints, etc, and many subtle differences in meaning in all those places where the proposition exists. The formal act of marriage remains of course a single, unique event.

4.2 Applicability

Applicability plays a central role for knowledge engineers as soon as the identity of the legal rule and the interpretation of its logical meaning are distinguished. The logical rule must assert explicitly that the legal rule is being applied. The law also frequently identifies rules: a special class of legal rules, *applicability rules* (e.g. [15,16,6], constrains the applicability of other rules, or make the application of one legal rule conditional on the application of another legal rule.

Applicability is modeled through the *applicable* (inverse *appliesTo*) property. This property applies to the *legal* thing the rule is about, regardless of its left hand or right hand side position in the axiom. Together *constitutedBy* and *applicable* explain how a legal thing arose.

Applicability is wherever possible attached to legal actions as a methodological choice. Since we do not determine confluence of applicability rules from subsumption between propositions, or sets of them (e.g. the *case description*), we have to be sure that rules about the same subject apply to the same thing. Actions are the focal objects; This immediately entails that we take a so-called ought-to-do perspective towards normative rules.

Generally one of the great challenges for us in understanding applicability is the distinction between its dispositional and categorical meaning (i.e. if the rule *were* applied to something, the result would have some quality that it wouldn't have if it were inapplicable, vs. the rule has been as a matter of fact applied). In its epistemological applications (of which [16] is an example) in defeasible reasoning, and in implementation resources, it is generally taken to be dispositional. Why make explicit the application of a rule at all, unless it is defeasible? We prefer to use it in the more clearcut categorical sense, aware of the fact that this doesn't does not address the inherent defeasibility of applicability.

5 Rules Example

The example source of law for this paper consists of two sentences presenting two simplistic rules:

t_1 The publication of a text presenting a rule counts as the creation of that rule.
t_2 Rule t_1 applies to text published by a rule maker.

The legislative example is *qua* design pattern representative of formal legal acts in bureaucratic environments.

Agile knowledge representation consists of making rigorous distinctions between entities in different domains.

Firstly, we distinguish the text, which is a MetaLex expression object, from the legal rules: t_1 represents legal rule r_1, and t_2 represents legal rule r_2. This distinction makes alternative interpretations of the same text, and the maintenance of provenance information about these different interpretations, possible. We moreover also distinguish the legal rules from the (logical) OWL2 rules a_1 and a_2 describing their meaning.

Rule r_1 is clearly constitutive. Rule r_1 (represented as an OWL axiom written in a compact Manchester syntax-like notation for purposes of readability) demonstrates an interesting pattern relating the *constitutes* and *represents* (inverse *representedBy*) relations:

```
if :Publication that
 (:resultsIn some (:Text that
  (agile:represents some :Rule)))
 then
  (agile:constitutes some (:Creation that
   (:resultsIn some :Rule) and
   (agile:applicable value :r1)))
```

This pattern is typical of implementation of legal acts as *formal* acts, and occurs often in public administration. Here we also see the major expressive limitations of OWL: it describes treelike patterns, and here we would like to add the constraint that the rule represented by the text is the rule being created (as in Fig. 4). Some workarounds are possible in OWL2, but these sacrifice readability. We are currently evaluating the extension to description graphs in [13].

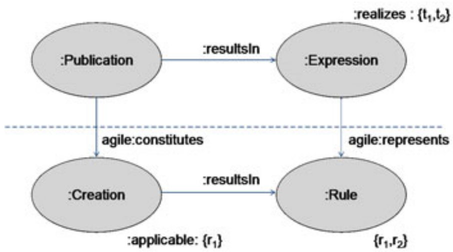


Fig. 4. The structure of interest in rule 1 of section 5

The second rule r_2 limits the applicability of r_1, but also of any other rules derived from any future or alternative version of t_1, as follows, showcasing the subtle bibliographic identity distinctions made by MetaLex:

```
if (agile:representedBy some
 (metalex:realizes value :t1))
 then
  (agile:appliesTo all ((:actor some :RuleMaker) and
   (agile:applicable value :r2)))
```

In the universe created by this small set of rules, the entities r_1, r_2, any instance of *Rule*, *Creation*, and *RuleMaker* are institutional (and this is enforceable through domain and range restrictions), and the rest are implementation entities. Texts t_1, t_2 represent information, while rules r_1, r_2 are institutional entities.

Note that, although these two rules are consistent, the applicability assertions may turn out to be in logical conflict with respect to some common types of cases. We intentionally do not resolve this defeasibility between t_2 and t_1 in the rules, because there are alternative, equally reasonable ways to resolve it depending on the knowledge representation language semantics used, and on the disposition one has towards these rules. The problem is to enumerate prototypical roles rules play in the organization and its environment.

We believe that this method of representing applicability will results in better and more durable isomorphism between sources of law and implementation resources [17] over time. More gains will come from the distinction between resource (here for instance t_1) and the thing it presents (r_1).

The example shows that the conceptual frameworks that relate the first and second layers and the second and third layers are not actually different: In both the constitutiveness and representation concepts play key roles. They are only different from the implementing organization's point of view, to the extent that it sees itself as addressee of formal legislation only, and perceives a disjointness between legislative and implementation domains that is non-existent ontologically in our interpretation of institutions.

We tend to focus on the representation issue when talking about the sources of law simply because this, and not the legislator's abstract legal act, is our point of access to the law as knowledge engineers. We focus on constitutiveness when talking about implementation, because implementation is to a large extent a matter of designing ways to perform legal acts. The centrality of constitutiveness is itself only a relatively recent realization in academic knowledge engineering and requirements engineeering, which hasn't been fully absorbed by the business community yet.

In implementation in administrative processes, representation however also plays a central role. It is the marriage certificate that represents the legally relevant marriage, the receipt that represents the financial transaction, the written administrative decision that represents the change of legal position, and an ODBC update in some database that represents the fact that the administrative organization officially recognized a legally relevant new fact in some procedure.

6 Conclusions and Discussion

In this paper we focused on ontological stratification as a modeling device, and on applicability reasoning using features of MetaLex. We believe these are areas were quick gains are to be made in large organizations, as opposed to for instance the normative reasoning aspect of legal knowledge representation which is often perceived as very academic on the work floor.

Generally, we distinguish between changes of the law that do and do not change institutional reality (for instance codification of case law or practices already developed within the organization), between changes of institutional reality (due to a new interpretation or new law) that do and do not necessitate change of implementation, and change of implementation for it's own sake, for instance due to work floor feedback.

If the requirement for a valid passport is for instance dropped from immigration law and replaced by the more general requirement that an immigrant should be identified, the implementing organization may still opt to retain the requirement for a valid passport in a fast track service implementation, and delegate the handling of cases where no valid passport can be shown to the fully manual appeal procedure. The business process remains the same, while its legal justification changes.

Common organizational problems in tracing resources and actions back to the sources of law are based on not understanding the distinction between abstract legal entities and their implementation. Organizations that are the unique in their institutional status are bound to suffer from this problem most.

In our approach, a deviation from mainstream legal knowledge representation is found in the rigorous ontological stratification (cf. [18,19]) of legal entities and implementation entities (as opposed to propositions) we have chosen for Agile.

Legal knowledge representation literature discusses the "counts-as" or constitutive rules (cf. for instance [20]), although it often considers them just one type of rule, among other (notably normative) ones [21]. On the other hand, in [22,6] we for instance find useful reconstructions of obligation and violation in terms of constitutiveness.

Good traceability tools are however not the only important aspect of impact analysis. Importantly, the organization also faces an increasing need to model its behaviour and the behaviour of agents it interacts with. To explain the services it provides, it needs to have an understanding of the effects its actions have on the agents that request those services, and the reasons why those services are requested. It, in other words, needs to engage in agent modeling to understand its own services and its performance of those services, and the impact of changes.

Network arrangements increase the stakes in agent modeling. When the organization itself plays the role of client requesting service of other organizations, it needs to understand why those services are provided and how those services are implemented. It needs to keep an eye on the legislation governing the performance of those services, and to resist the temptation to merely defend those arrangements against change.

The Agile project includes an agent simulation activity for simulation of candidate realizations of legal services and of the reaction of network partners and clients. This activity is not the focus of our work, and not of this paper. For the business process compliance aspect of this project, we can build on existing work in legal knowledge representation in for instance [23]. The business process compliance perspective does not, however, cover all relevant uses of the law, as indicated in section 2.

Acknowledgements

Agile is a Jacquard project funded by the Netherlands Organisation for Scientific Research (NWO). In the Agile project, The Leibniz Center for Law of the University of Amsterdam cooperates with the Technical University of Delft, and commercial and public administration partners.

References

1. Boer, A., van Engers, T.: Agile: From source of law to business process specification. In: Casellas, N., Francesconi, E., Hoekstra, R., Montemagni, S. (eds.) LOAIT 2009: 3rd Workshop on Legal Ontologies and Artificial Intelligence Techniques Joint with 2nd Workshop on Semantic Processing of Legal Text, Barcelona, Spain, IDT-UAB. IDT Series, vol. 2, pp. 27–44 (2009)
2. Breuker, J.: A suite of problem types. In: Breuker, J., de Velde, W.V. (eds.) Common KADS Library for Expertise Modelling, pp. 57–88. IOS-Press/Ohmsha, Amsterdam/Tokyo (1994)
3. Scarantino, A.: Affordances explained. Philosophy of Science 70(5), 949–961 (2003)
4. Boer, A., Winkels, R., Vitali, F.: Metalex XML and the Legal Knowledge Interchange Format. In: Casanovas, P., Sartor, G., Casellas, N., Rubino, R. (eds.) Computable Models of the Law. LNCS (LNAI), vol. 4884, pp. 21–41. Springer, Heidelberg (2008) ISBN: 978-3-540-85568-2
5. Boer, A., Vitali, F., de Maat, E.: CEN Workshop Agreement on MetaLex XML, an open XML Interchange Format for Legal and Legislative Resources (CWA 15710). Technical report, European Committee for Standardization, CEN (2006)
6. Boer, A.: Legal Theory, Sources of Law, & the Semantic Web. In: Frontiers in Artificial Intelligence and Applications, vol. 195. IOS Press, Amsterdam (2009)
7. Boer, A.: Metalex naming conventions and the semantic web. In: Governatori, G. (ed.) Legal Knowledge and Information Systems, pp. 31–36. IOS Press, Amsterdam (2009)

8. Saur, K.G.: Functional requirements for bibliographic records. UBCIM Publications - IFLA Section on Cataloguing 19 (1998)
9. Governatori, G., Rotolo, A.: Changing legal systems: legal abrogations and annulments in Defeasible Logic. Logic Jnl IGPL, jzp075 (2009)
10. Boer, A.: Using event descriptions for metadata about legal documents. In: Winkels, R., Francesconi, E. (eds.) Electronic Proceedings of the Workshop on Standards for Legislative XML, in Conjunction with Jurix (2007)
11. MacCormick, N.: Norms, institutions, and institutional facts. Law and Philosophy 17(3), 301–345 (1998)
12. Hindriks, F.A.: Rules & Institutions; essays in meaning, speech and social ontology. PhD thesis, Erasmus University Rotterdam (2005)
13. Glimm, B., Horridge, M., Parsia, B., Patel-Schneider, P.F.: A syntax for rules in owl 2. In: OWL Experiences and Directions (OWLED 2009), Washington D.C., USA (2009)
14. Halaschek-Wiener, C., Katz, Y., Parsia, B.: Belief base revision for expressive description logics. In: Online Proceedings of the Second OWL: Experiences and Directions Workshop (OWLED), Athens, Georgia, USA (2006), http://owl-workshop.man.ac.uk/accepted06.shtml
15. Prakken, H., Sartor, G.: On the relation between legal language and legal argument: assumptions, applicability and dynamic priorities. In: ICAIL '95: Proceedings of the 5th International Conference on Artificial Intelligence and Law, pp. 1–10. ACM, New York (1995)
16. Kowalski, R.A., Toni, F.: Abstract argumentation. Artificial Intelligence and Law 4, 275–296 (1996)
17. Bench-Capon, T., Coenen, F.: Exploiting isomorphism: development of a KBS to support British coal insurance claims. In: Sergot, M. (ed.) Proceedings of the Third International Conference on AI and Law, pp. 62–69. ACM, New York (1991)
18. Borgo, S., Guarino, N., Masolo, C.: Stratified ontologies: The case of physical objects. In: Proceedings of the ECAI-96 Workshop on Ontological Engineering (1996)
19. Smith, B.: An essay in formal ontology. Grazer Philosophische Studien 6, 39–62 (1978)
20. Boella, G., van der Torre, L.W.N.: Regulative and constitutive norms in normative multiagent systems. In: Proceedings of the 9th International Conference on the Principles of Knowledge Representation and Reasoning, Whistler, CA (2004)
21. Mazzarese, T.: Towards the semantics of "constitutive" in judicial reasoning. Ratio Iuris 12, 252–262 (1999)
22. Boella, G., van der Torre, L.W.N.: Obligations as social constructs. In: Cappelli, A., Turini, F. (eds.) AI*IA 2003. LNCS (LNAI), vol. 2829, pp. 27–38. Springer, Heidelberg (2003)
23. Governatori, G., Sadiq, S.: The journey to business process compliance. In: Handbook of Research on BPM, pp. 426–454. IGI Global (2009)

Legal Ontology Construction Using ATOB Algorithm

Vi-sit Boonchom and Nuanwan Soonthornphisaj*

Department of Computer Science, Faculty of Science
Kasetsart University, Bangkok, Thailand
g5184018@ku.ac.th, fscinws@ku.ac.th

Abstract. Ontology is a knowledge representation technique that can be applied in many areas such as information retrieval. Ontology construction is a tedious job and time consuming task for law experts. This paper proposes a system framework called ATOB that can automatically generate a seed ontology and extend the ontology using ant colony algorithm. Two ontologies are created by ATOB system which are succession law ontology and family law ontology. We compare the performance of these ontologies with the ontologies manually created by law experts in our previous system (TLOE) and found that ATOB provides satisfactory result.

Keywords: Automatic Ontology Building, Seed Ontology, Legal, Thai Law, Ant Colony Algorithm.

1 Introduction

Ontology can be used as common understanding explicit knowledge of specific domain [1]. Currently, there are many specific domain ontologies such as breast cancer ontology [2], agriculture-specific ontology [3], e-learning ontology [4], etc. Several legal ontologies have been constructed which are legal ontology for the Spanish e-government [5], documentary ontology of French law dedicated to information retrieval on the Web [6], etc.

The objectives of our research are to construct a Thai succession law and family law ontology using a set of Supreme Court sentences. The advantages of our ontologies are two-fold: 1) Legal students can easily conceive an overall concept of a domain specific law via the ontology. 2) The extended ontology can be applied to enhance the performance of the Supreme Court sentence retrieval.

We propose a system called ATOB that consists of three modules. The first module named automatic seed ontology building is responsible for building core ontology by randomly select a set of Supreme Court sentences. This module do word segmentation and find relationship among those words using ThaiLegalWordNet dictionary.

The second module employs an ant colony algorithm to do Supreme Court sentences retrieval. The function of the last module is the ontology expansion process.

The contributions of this research are the Thai succession law ontology and family law ontology that can be applied to the Supreme Court search engine. This paper is

* Corresponding author.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 268–279, 2010.
© Springer-Verlag Berlin Heidelberg 2010

organized as follows: section 2 reviews some related works. Section 3 presents our system framework. Experimental results are shown in section 4. The conclusion and future works for this paper are given in section 5.

2 Related Works

For more than a decade, legal ontologies have been developed. These ontologies concerned with the development of legal knowledge system and information management [7].

2.1 Legal Ontology Construction Methodology

A research done by Corcho, O. *et al.* [8] has shown how experts in the legal domain can develop their own ontologies. They proposed a methodology called METHONTOLOGY and WebODE as a tool for ontology construction.

Semi-automatic approach in ontology building using text mining was proposed by Parekh, V. *et al.* [9]. They used a Singular Value Decomposition (SVD) based clustering technique called Principal Direction Divisive Partitioning (PDDP). The proposed algorithm automatically extracted groupings of related terms/concepts from domain specific texts and glossaries/dictionaries. Text mining technique was used to create an E-commerce law ontology by Kayed, A. [10]. His corpus is law case related to E-commerce domain. He proposed an algorithm to reduce a number of links connected between nodes in order to reduce the complexity of the ontology.

Natural Language Processing techniques were applied to legal documents to generate an ontology [11]. Korean law ontology was constructed from law text books and research papers [12]. They employed Korean WordNet to facilitate the algorithm during the ontology construction process.

Walter, S. and Pinkal, M. [13] used computational linguistic analysis technique for information access. They presented a rule-based method for extracting and analyzing definitions from parsed text. The proposed method was evaluated on a corpus of 6000 German Court decisions within the field of environmental law.

Lame, G. [14] proposed a method to identify ontology components. His technique relied on NLP techniques to extract component and relations among these concepts. The method was applied in the legal documents related to Criminal Code, Civil Code and Intellectual Property Law.

Logic Programming technique was applied in the research done by Saias, J. and Quaresma, P. [15]. They used dynamic logic programming framework called EVOLP as a mechanism for web information retrieval system of the Portuguese Attorney General's Office.

Imsombut, A. [17] has developed a methodology for constructing a Thai ontology automatically from corpus, thesaurus and dictionary. She used NLP technique to extract ontological concepts and taxonomic relations using WordNet, corpus and thesaurus in the agricultural domain.

Automatic ontology construction has been done in several specific domains. In additions, each specific domain used same way for ontology building process.

2.2 TLOE System

The framework of Thai Law Ontology Expansion (TLOE) was proposed by Boonchom, V. and Soonthornphisaj, N. [16]. In TLOE system, the seed ontology was created by a legal expert. The ontology expansion is done using ant colony algorithm during the document retrieval process. The experiment was done on Thai Succession Law domain and the result showed that the Thai Succession Law ontology obtained from TLOE system provided a promising results.

3 The Framework of Automatic Thai Law Ontology Building (ATOB)

ATOB consists of 3 modules which are Automatic Seed Ontology Building module, Thai Legal Document Retrieval module and Thai Legal Ontology Expansion module (see Fig. 1)

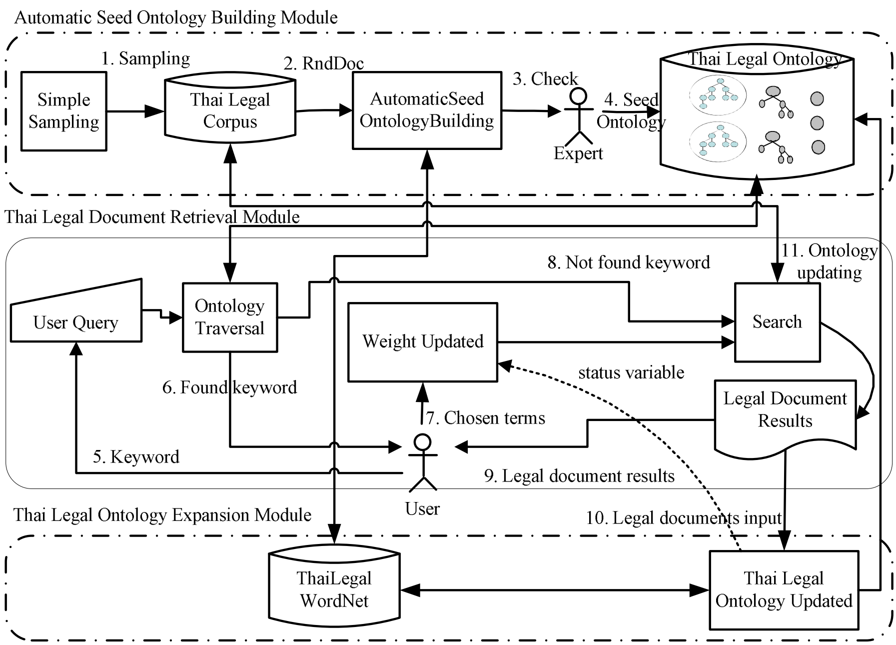


Fig. 1. The framework of ATOB

3.1 The Automatic Seed Ontology Building Module

Given a Supreme Court sentences corpus, five court sentences are randomly selected by the `RandomDoc` process. After that a `WordSegmentation` procedure is performed since Thai document has no word boundary. Then a `ThaiLegal-WordNet` is applied to find a superclass of each term in these documents. Note that the ThaiLegalWordNet is a dictionary that contains a set of legal terms. The structure of seed ontology is formed using the superclass legal terms. These superclass terms join a set of words obtained from the Supreme Court sentences (see Table 1).

Table 1. Automatic Seed Ontology Building Algorithm

```
Algorithm. AutomaticSeedOntologyBuilding
Input: ThaiLegalCorpus
Begin
          RndDoc=RandomDoc(ThaiLegalCorpus)
           For each Doc_i ∈ RndDoc
            Terms ← WordSegmentation(Doc_i)
            For each t_i ∈ Terms
             Superclass_i←Explore(ThaiLegalWordNet(t_i))
              If Superclass_i='Thing'or NodeInOntology
                Then JoinConcept (t_i,Ontology)
            End for
End
Output: SeedOntology
```

3.2 The Thai Legal Document Retrieval Module

There are three kinds of node in the ontology, a root node, branch nodes and leaf nodes respectively (see Fig.2). Each node contains a legal term, ID number and a pheromone level. The pheromone level is updated by each ant in the colony using equation (1) and (2).

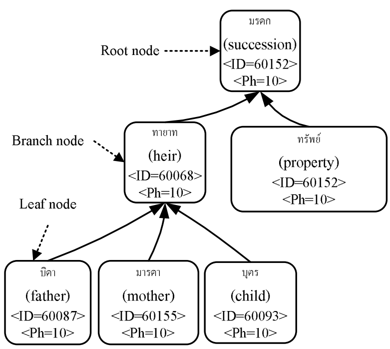


Fig. 2. Three kinds of nodes and three parameters in ontology

Our ant colony consists of 26 ants which are legal users. When user enter a keyword, an ant traverse through the ontology to find the target node. If the node is found by the ant, all child nodes in the first level will be provided to that user.

Given a set of proposed terms contain in those first level child nodes, the expert selects a subset of terms to be used in the document retrieval module. Then the pheromone updating process is started. The pheromone level Ph_{term_i} is accumulated for each chosen term. A parameter ξ is equal to 1 if the $term_i$ is chosen. On the other hand, the pheromone is evaporated for non-chosen term and the value of ξ is equal to -1. The amount of updated pheromone, $\left(\frac{1}{c}\right)$, depends on the number of proposed terms. More proposed terms mean less amount of pheromone level. If the algorithm found that the pheromone of the node is completely evaporated ($Ph_{term_i} \leq 0$) during the pheromone updating process, the node will be deleted from the ontology. It means that if the leaf node is not chosen by users for long period of time, it will be disappeared from the ontology.

$$Ph_{term_i} = Ph_{term_i} + \left(\xi \cdot \frac{1}{c}\right) \tag{1}$$

$$\mathrm{DocScore_j} = \sum_{i=1}^{n} \left(Ph_{term_i} \times \mathrm{tf_{ij}}\right) \tag{2}$$

The court sentences retrieval process is performed using an expanded query that consists of the original keyword and `n` chosen terms. Each court sentences' score is calculated using equation 2. The pheromone level Ph_{term_i} takes part in the document score calculation. The `Sort` function is used to sort the documents in the corpus.

Table 2. Thai Legal Document Retrieval Algorithm

```
Algorithm. ThaiLegalDocRetrieval
Input: User keyword, status
Begin
        If SearchConcept(keyword,Ontology) Then
          Terms←OntologyTraversal(keyword,Ontology)
          For each term_i ∈ Terms
            QExpand←ChosenTerm(term_i)
            If status<6 then
              Ph_term_i←WeightUpdated(QExpand,Ontology)
                If Ph_term_i<= 0 && term_i = LeafNode Then
                  DeleteNode(term_i, Ontology)
            End If
            For each Doc_j ∈ Corpus
              DocScore_j ← Ph_term_ij x tf_term_ij
         LegalDocRanked←Sort(DocScore,Corpus)
        End If
End
Output: LegalDocRanked
```

Considering Table 2, a `SearchConcept` function is started when user input a keyword to the system. If the algorithm finds the keyword in the ontology, all first level child nodes will be collected in `Terms`. In case that no matched term found in the ontology the `Terms` will contain the original keyword only. The `ChosenTerm` function displays a set of terms to user to filter out some terms that may not meet the user requirement. Then the `WeightUpdated` function performs pheromone updating.

The parameter that got from `ontology expansion algorithm` we called `status` and used for terminating the `WeightUpdated` and `DeleteNode` function. Finally, the retrieval process is done by calculating the `DocScore` using term frequency and weighted value from each keyword of each document in corpus using equation 2. The final results are shown in descending order (by `DocScore`) to user using `LegalDocRanked` function.

Table 3. Ontology expansion algorithm

```
Algorithm. OntologyExpansion
Input: LegalDocRanked ←{Doc_1, Doc_2,…, Doc_n}, Ontology,
       TotalDoc, status
Begin
If status<6 then
          For each Doc_j ∈ RankedDoc
            Term ← WordSegmentation(Doc_k)
            For each term_i ∈ Term
              OntologyTraversal(term_i)
              If notFound(term_i)
            Superclass_i←Explore(ThaiLegalWordNet(term_i))
                OntologyTraversal(Superclass_i)
                If Found(Superclass_i)
                  JoinConcept(term_i,Ontology)
                    If CountNewNode(JoinConcept)=1
                      status=status+1
            CountDoc++
            If CountDoc > TotalDoc
              Return Ontology
End.
Output: OntologyUpdated
```

3.3 Thai Legal Ontology Expansion Module

The `OntologyExpansion` module gets a set of `LegalDocRanked` from the retrieval module. `Max` variable is the total number of document sentences from the corpus. Given the first ranked document, the `WordSegmentation` function is performed to get a set of terms, `DocTerms`. For each $term_i$, the algorithm traverses through the ontology to find $term_i$. If `notFound` function is true, the algorithm will explore the `ThaiLegalWordNet` to get the synonym or the superclass of the $term_i$. In case that, the algorithm find a superclass of $term_i$, it will be used to join its concept to its superclass node by `JointConcept` function. In case that only one new node is inserted into the ontology for five times, the ontology expansion will be terminated (see Table 3).

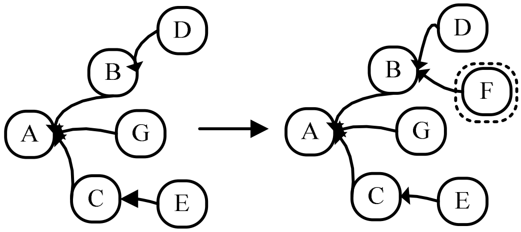


Fig. 3. An example ontology

For example, the term ‘F’ cannot be found in the ontology. Therefore, the algorithm explores the `ThaiLegalWordNet` and found that the superclass of ‘F’ is ‘B’. Then, node ‘F’ is connected to node ‘B’ (see Fig. 3).

In case that, the superclass of the $term_i$ is not found, the $term_i$ is discarded. The ontology expansion process stops when the number of nodes is stabilized.

We store the ontology structure in XML format which supports the `OntologyTraversal` process (see Fig.4).

```
<?xml version="1.0" encoding="windows-874"?>
<chunks>      <chunk>
      <class><term>มรดก (succession)</term>
                  <id>60152</id>
                  <ph>10</ph></class>
      <subclass><term>ทายาท (heir)</term>
                  <id>60068</id>
                  <ph>10</ph></subclass>
      <subclass><term>ทรัพย์ (property)</term>
                  <id>60068</id>
                  <ph>10</ph></subclass>
</chunk>      </chunks>
```

Fig. 4. Illustrates an example of ontology in the XML file

3.4 The Structure of ThaiLegalWordNet Dictionary

ThaiLegalWordNet a dictionary covers 2 Thai civil law domain which are succession law and family law. There are 282 concepts provided by law experts.

Considering table 4, SuperClassId is the concept identifier of superclass node, SuperClassName is the name of super class node, NodeId is the identifier of each current term, NodeName is the name of current node and Ph is the pheromone value of each term. For example, the succession (มรดก) node has 60152 as NodeId and its SuperClassName is 'Thing'. Thing node is the root node of the ontology, its superclass identifier is 60000. We found that, succession (มรดก) node is the first level child node of the root node. The heir (ทายาท) node has 60068 as the node identifier and its SuperClassName is the succession (มรดก) node. The pheromone levels of all nodes are initialized to be 10.

Table 4. Nodes and meanings of the seed ontology.

SuperClassId	SuperClassName	NodeId	NodeName	Ph
60000	Thing	60152	มรดก (succession)	10
60152	มรดก (succession)	60068	ทายาท (heir)	10
60152	มรดก (succession)	60065	ทรัพย์ (property)	10
60068	ทายาท (heir)	60087	บิดา (father)	10
60068	ทายาท (heir)	60093	บุตร (child)	10
60068	ทายาท (heir)	60155	มารดา (mother)	10

3.5 Data Pre-precessing

ATOB algorithm used word segmentation procedure for each court sentence. This function used a set of law terminologies that created by law expert to search and count the frequency of each term in each document using equation 2.

4 Experimental Results

4.1 Data Set

The Supreme Court sentences repository was collected from the Supreme Court search engine web site (www.supremecourt.or.th). There are 1,305 sentences related to the Thai succession law and family law (article No. 1435-1755).

A word segmentation procedure is applied to each court sentence. We obtained 126,996 terms from the corpus.

4.2 Results

4.2.1 The Result of Automatic Seed Ontology Building on Succession Law

The algorithm used Simple Sampling function for 5 sentences from our corpus and used word segmentation process for extraction the Thai law terms. After that, the ATOB algorithm used Automatic Seed Ontology Building module to create a seed ontology. There are 29 nodes in the seed ontology for succession law (see Table. 5 and Fig 5).

Table 5. Seed ontology obtained from 5 court sentences on succession law data set

SentencesNo.	Terms from document	Terms for ontology building	Terms
3146/2533	15	2	บุคคลตาย (dead person) ผู้มีส่วนได้เสีย (interested person)
6225/2539	24	6	เจ้ามรดก (owner) ทายาท (heir) มรดก (succession) ลูกหนี้ (debtor) บิดา (father) หนี้ (debt)
194/2535	36	14	ทรัพย์ (property) อา (uncle) สิทธิ (right) ผู้จัดการมรดก (administrator of an estate) หน้าที่ (duties) ตา (grandfather) น้า (uncle) บุตร (child) มารดา (mother) ผู้เยาว์ (a minor) ศาล (court) ย่า (grandmother) อำนาจ (authorize) เสียงข้างมาก (mojority of votes)
7719/2542	24	4	คำสั่ง (ruling) จัดการมรดก (property management) ที่ดิน (land) ส่วนแบ่ง (portion)
2078/2537	23	3	ครอบครอง (posses) โฉนด (land certificate) ชั้น (rank)

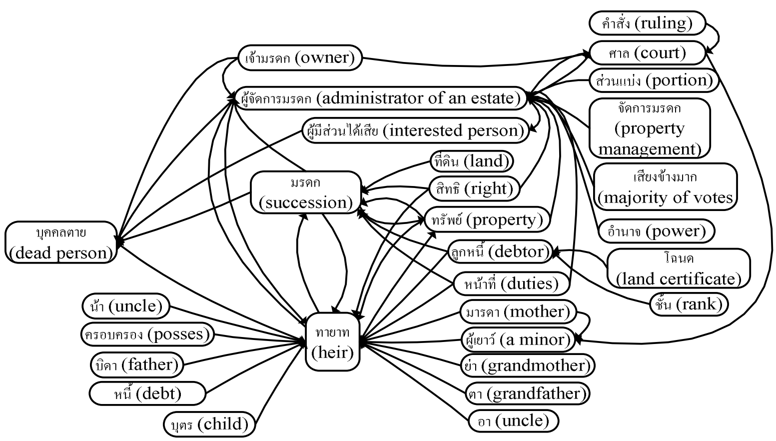


Fig. 5. The tree structure of the seed ontology for succession law

4.2.2 The Result of Automatic Seed Ontology Building on Family Law

Considering the family law domain, ATOB algorithm has an ability to create a seed ontology as shown in Table 6 and Fig.6.

In order to evaluate the performance of automatic seed ontology building using ant colony algorithm, a set of keywords are given to 26 law experts. Each expert randomly picks up the keywords and retrieves the set of court sentences using ATOB

Table 6. Seed ontology obtained from 5 court sentences on family law data set

SentencesNo.	Terms from document	Terms for ontology building	Terms
679/2532	9	3	ครอบครัว (family) สมรส (marriage) หนี้ (debt)
6695/2540	10	3	ทรัพย์ (property) ยินยอม (consent) สินสมรส (marriage portion)
8264/2538	13	4	จดทะเบียน (entry in the register) ถอน (withdraw) บิดา (father) หย่า (divorce)
5351/2545	8	1	โมฆะ (void)
6657/2539	23	0	-

algorithm and TLOE algorithm [7]. TLOE algorithm uses a seed ontology manually created by a law expert, whereas ATOB constructs a seed ontology automatically. We set up 4 experiments for each data set as follows: 1) ATOB using ant colony, 2) ATOB without ant colony means that the weight ontology updating function is removed from the algorithm, 3) TLOE using ant colony, and 4) TLOE without ant colony.

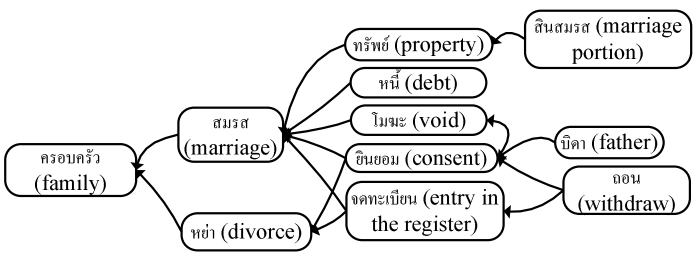


Fig. 6. The tree structure of the seed ontology for family law

The performances of extended ontologies obtained from those experiments are measured using precision value in equation (3)

$$\text{precision} = \frac{\text{number of relevant documents retrieved by algorithm}}{\text{number of retrieved documents}} \quad (3)$$

Considering the number of nodes in ontologies shown in Table 7, we found that there is no significant different in both data sets. That means our proposed algorithm (ATOB) has enough potential to reduce the burden of law expert in creating the seed ontology. The retrieval performances measured in term of precision value (see equation 4) are shown in Table 8. We found that ant colony plays an important role on the retrieval performance of both data sets. ATOB algorithm outperforms TLOE on family law data set. However the precision value of TLOE is a bit higher in Succession law data set.

Table 7. The characteristic of ontologies obtained from ATOB and TLOE using Ant Colony

Number of Nodes	Succession Law		Family Law	
	ATOB	TLOE	ATOB	TLOE
Seed ontology	29	19	11	9
Extended ontology	119	121	55	54
Root and branch node	45	42	20	18
Leaf node	74	79	35	36

Table 8. The precision values on ATOB and TLOE algorithm

Law	ATOB		TLOE	
	With Ant colony	Without Ant colony	With Ant colony	Without Ant colony
Succession Law	0.87	0.75	0.89	0.72
Family Law	0.88	0.71	0.86	0.69

5 Conclusions

This research proposes a new algorithm for automatic seed ontology building, the Supreme Court sentences retrieval algorithm and the ontology expansion algorithm. Our algorithm, ATOB, uses simple sampling method to construct a seed ontology using ant colony algorithm as a mechanism in order to increase or decrease pheromone level in each legal term proposed to users.

The result of experiments showed that, ATOB algorithm could create the seed ontology which has efficiency similar to the seed ontology from TLOE algorithm. ATOB algorithm used ThaiLegalWordNet created by law expert and our algorithm guarantee for the user's satisfaction by using ant colony on retrieval process. In the near future, we plan to extend our work to other areas of Thai law.

Acknowledgments

This work has been supported by Office of the Higher Education Commission, Thailand.

References

1. Thomas, R.G.: Toward Principles for the Design of Ontologies used for Knowledge Sharing. International Journal of Human-Computer Studies 43, 907–928 (1995)
2. Dridi, O., Ben Ahmed, M.: Building an Ontology-based Framework for Semantic Information Retrieval. Application to Breast Cancer. In: Information and Communication Technologies: From Theory to Applications (ICTTA), pp. 1–6 (2008)
3. Xie, N., Wang, W., Yang, Y.: Ontology-based Argricultural Knowledge Acquisition and Application. Computer and Computing Technologies in Agriculture 1, 349–357 (2008)

4. Henze, N., Dolog, P., Nejdl, W.: Reasoning and Ontologies for Personalized E-learning in the Semantic Web. Educational Technology & Society 7(4), 82–98 (2004)
5. Perez, A.G., Rodriguez, F.O., Terrazas, B.V.: Legal Ontologies for the Spanish e-Government. In: Marín, R., Onaindía, E., Bugarín, A., Santos, J. (eds.) CAEPIA 2005. LNCS (LNAI), vol. 4177, pp. 301–310. Springer, Heidelberg (2006)
6. Bourigault, D., Lame, G.: Analyse Distributionnelle et Structuration de Terminologies. Traitement automatique des, 129–150 (2001)
7. Despres, S., Szulman, S.: Construction of a Legal Ontology from a European Community Legislative Text. In: The Seventeenth Annual Conference, Jurix 2004, pp. 79–88. IOS Press, Amsterdam (2004)
8. Corcho, O., Lopez, M.F., Perez, A.G., Cima, A.L.: Building Legal Ontologies with Methontology and WebODE. In: Benjamins, V.R., Casanovas, P., Breuker, J., Gangemi, A. (eds.) Law and the Semantic Web. LNCS (LNAI), vol. 3369, pp. 142–157. Springer, Heidelberg (2005)
9. Parekh, V., Gwo, J., Finin, T.: Mining Domain Specific Texts and Glossaries to Evaluate and Enrich Domain Ontologies. In: Proc. Int Conf. Information and Knowledge Engineering (2004)
10. Kayed, A.: The Grid: Building e-Laws Ontology: New Approach. In: Meersman, R., Tari, Z., Herrero, P. (eds.) OTM-WS 2005. LNCS, vol. 3762, pp. 826–835. Springer, Heidelberg (2005)
11. Lenci, A., Montemagni, S., Pirrelli, V., Venturi, G.: NLP-based Ontology Learning from Legal Texts. A case study. In: Proc. LOAIT07-II Workshop on Legal Ontologies and Artificial Intelligence Technique, pp. 113–129 (2007)
12. Hwang, S., Jung, Y., Yoon, A., Kwon, H.C.: Building Korean Classifier Ontology Based on Korean WordNet. In: Sojka, P., Kopeček, I., Pala, K. (eds.) TSD 2006. LNCS (LNAI), vol. 4188, pp. 261–268. Springer, Heidelberg (2006)
13. Walter, S., Pinkal, M.: Automatic Extraction of Definitions from German Court Decisions. In: Proc. of COLING-2006, Workshop on Information Extraction Beyond The Document, pp. 20–28 (2006)
14. Lame, G.: Using NLP techniques to Identify Legal Ontology Components: Concepts and Relations. In: Benjamins, V.R., Casanovas, P., Breuker, J., Gangemi, A. (eds.) Law and the Semantic Web. LNCS (LNAI), vol. 3369, pp. 169–184. Springer, Heidelberg (2005)
15. Saias, J., Quaresma, P.: A Methodology to Create Logic Ontologies in a Logic Programming Information Retrieval System. In: Benjamins, V.R., Casanovas, P., Breuker, J., Gangemi, A. (eds.) Law and the Semantic Web. LNCS (LNAI), vol. 3369, pp. 185–200. Springer, Heidelberg (2005)
16. Boonchom, V., Soonthornphisaj, N.: Thai Succession Law Ontology Building Using Ant Colony Algorithm. In: Proceeding of the Third International Workshop on Juris-informatics (JURISIN), Campus Innovation Center, Tokyo, Japan, pp. 27–37 (2009)
17. Imsombut, A.: Automatic Thai Ontology Construction from Corpus, Thesaurus, and Dictionary. Ph.D. Thesis. Computer Engineering, Kasetsart University, Thailand (2007)

Semantic Interoperability among Thesauri: A Challenge in the Multicultural Legal Domain

Enrico Francesconi and Ginevra Peruginelli

Institute of Legal Information Theory and Techniques
Italian National Research Council
francesconi,peruginelli@ittig.cnr.it
http://www.ittig.cnr.it

Abstract. In the last few years crucial issues like cross-language legal information retrieval, document classification, legal knowledge discovery and extraction have been considered in theory and in practice. The availability of services allowing cross-language and cross-collection retrieval is a growing necessity. This paper focuses on the need to develop solutions for automatic, language-independent procedures to provide interoperability between mono/poly-lingual thesauri at national and European levels. This will guarantee sustainable and scalable services enabling to manage the multilingual complexity of the European Union legal context to be used for cross-language and cross-collection legal information retrieval. Wider use of the service can also be envisaged as support to legal translation services, as well as in general to promote integration and sharing of widespread and heterogeneous legal resources, providing new market opportunities for stakeholders to exploit the economic potential of public sector information in a multilanguage environment.

Keywords: Thesauri Semantic Interoperability, Legal Information Indexing, Cross-collection Information Retrieval, SKOS.

1 The Context

Nowadays the proliferation of transnational exchanges as part of a more general convergence of models of development inspired by principles of peaceful coexistence, respect for human rights, environmental protection, international solidarity, economic growth greatly increases interdependence among countries.

In the legal domain integration is an essential requirement, but serious difficulties arise in realising it. The principle of multilingualism in the law domain ensures not only transparency and fairness of rights among individuals, but guarantees legal certainty. The objective is the integration of different legal cultures while improving their legal identities that reflect the wealth of each legal order.

The multicultural reality of the European Union represents a privileged observatory of the phenomenon of multilingualism. In fact, the Union is based on the principle of unity in diversity: diversity of cultures, customs, beliefs and languages. One of the most direct expression of this cultural heterogeneity is undoubtedly its linguistic richness. The European Union currently consists

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 280–291, 2010.
© Springer-Verlag Berlin Heidelberg 2010

of twenty-seven Member States, with their twenty-three official languages, but there are more than sixty indigenous languages and scores of non-indigenous languages spoken by communities of different Member States. Some of these are prestigious languages of culture, others are dialects and only recently have they been recognised as official languages of history. All these languages represent cultural heritage and the way in which they were promoted, protected, but also persecuted vary from State to State. Any attempt aimed at exchanging legal knowledge among different communities to reach a common understanding must therefore take account of linguistic and conceptual problems that fall in the disciplines of comparative law, theory of law, linguistics and translation theory.

2 Enhancing Quality in Legal Information Indexing

At operational level the need to develop services allowing users who do not necessarily know the language of origin of legal material of different countries through multilingual search modalities is of paramount importance.

The difficult task to effectively access multilingual legal material is definitively matching and weighting legal terms across languages. This generally implies translating from the language of the query to that of the material to be found or vice-versa, while addressing the problem of word disambiguation which is greatly increased when mapping diverse legal languages. In fact, crossing the language barrier between search requests and documents requires that the problems of the system-bound nature of legal terminology have to be adequately faced, devising methods to map concepts between different legal systems.

In a multilingual access environment information is searched, retrieved and presented effectively without constraints due to the different languages and scripts used in the material to be searched and in the metadata, that is descriptive and semantic information allowing the retrieval of indexed documents. This implies that in creating multilingual access services, both users native language and the multiplicity and richness of world-wide languages are to be accommodated, and methods have to be developed to allow users to put queries expressed in any language and retrieve information resources independently of the language of documents and indexing. Therefore, the availability of services allowing cross-language and cross-collection retrieval is a growing necessity.

In this context semantic tools can greatly contribute to cross-language localization through the structure and functions of thesauri and ontologies. In the domain of law efforts are starting to be made in this direction. Example of these are: the Lexical Ontologies for Legal Information Sharing (LOIS) project [20], Jurwordnet and a number of linguistic tools like the Legal Taxonomy Syllabus (LTS), Eurovoc Thesaurus and Jurivoc, the legal thesaurus of the Swiss Federal Tribunal.

As concerns thesauri several approaches are used, based on recommendations made at international level for the creation of multilingual thesauri [2]. These concern the translation of a multilingual thesaurus already existing in one or more languages, the fusion of several monolingual thesauri or the creation of a multilingual thesaurus ex novo. Experts recommend to adopt a truly multilingual approach in the construction of thesauri in order not to privilege the original source language and to ensure that the description of concepts is equally detailed and analytical in all languages treated. In fact, a multilingual thesaurus is more than the incorporation of several monolingual thesauri, it must necessarily adopt the principle of equality between languages, providing the conceptual apparatus and terminology of each language represented. It is an extremely complex work requiring substantial resources. The guidelines of the International Organization for Standardization[1] for the development of multilingual thesauri recommend appropriate procedures for their construction, addressing issues relating to the existence of various levels of equivalence among linguistic terms that describe concepts and providing possible solutions and suggestions. These regard the use of more than one descriptors in the target language, the creation of neologisms, the preservation of descriptors in the source language with a link to the preferred term. Obviously, to implement a multilingual query system using this approach requires the translation of each term of the thesaurus for each new language considered [3,4].

There is no doubt that multilingual thesauri are complex instruments, requiring not only an onerous task for development and maintenance, but involving specific skills for indexing documents especially for a large collection of documents which are different in subject and type.

In practice, aligning vocabularies of two or more languages, especially in the legal domain, is a hard process. Ideally a multilingual legal thesaurus should include all concepts needed in searching by any user in any of the source languages, but difficulties arise in making the systems of legal concepts the same for all languages. In fact a different language often suggests a different way of classifying law material and a system needs to be hospitable to all of these. In such a context what is needed is mapping query terms from the source language to their possible multiple equivalents in the target language. However, each of these equivalents may have other meanings in the target language or may not have a precise equivalent, requiring a mapping to broader or narrower terms. This can lead to distorting the meaning of the original query. Multiple meanings can be disambiguated through users interaction, but the success of this approach depends on the quality of the hierarchy of concepts, the provision of well-structured cross-references, and the interface of the system.

After all, whether searching is by controlled vocabulary or by free text, it is definitely helpful to the user to browse a well-structured and well-displayed hierarchy of concepts in his or her language. This can be achieved by ensuring an adequate correspondence among thesauri.

Considering the need, but also the complexity, of aligning multilingual thesauri, automatic facilities able to map concepts of different thesauri to support the intellectual activities of thesaurus alignment are desirable.

[1] International Organisation for Standardisation. ISO 5964: 1985: Guidelines for the establishment and development of multilingual thesauri.

3 A Formal Characterization of the Schema-Based Thesaurus Mapping Problem

Thesaurus mapping can be seen as the process of identifying terms, concepts and hierarchical relationships that are approximately equivalent [7]. The problem therefore is to define the meaning of "equivalence" between concepts.

In literature "concept equivalence" is defined in terms of set theory: according to this vision two concepts are deemed to be equivalent if they are associated with, or classify the same set of objects [8] (*Instance-based mapping* [9]). This approach is characterized by the availability of data instances giving important insight into the content and the meaning of schema elements. On the other hand concepts may also be associated with semantic features and mappings can be based on equivalences between feature sets (*Schema-based mapping* [9]). This approach is the only possible when only schema information is available.

The most complete classification of state-of-the-art schema-based matching approaches can be found in [10], where schema-based methods are organized in two taxonomies with respect to the: *Granularity/Input interpretation* (classification based on the granularity of match (element or structure level)), *Kind of Input* (classification based on the kind of input (terminological, structural, semantic)). A similar mapping approaches classification, along with an overview of the main experiences in this field, can be found in [11]. Taking into account this classification and elementary techniques described in literature, in this paper a framework for schema-based thesaurus mapping is proposed along with a methodology to implement schema-based thesaurus mapping for the case study.

Thesaurus mapping for the case-study is a *Schema-based mapping* problem aimed at thesaurus terms alignment. The problem to be addressed for an automatic mapping is to identify the conceptual/semantic similarity between a term (simple or complex)[2] in the source thesaurus and candidate terms in a target thesaurus. This can be done providing similarity measures according to specific terms semantic representations and metrics.

In this perspective we propose to characterize the schema-based Thesaurus Mapping ($\mathcal{TM}$) problem as a problem of Information Retrieval ($\mathcal{IR}$). As in $\mathcal{IR}$ the aim is to find the documents, in a document collection, better matching the semantics of a query, similarly in $\mathcal{TM}$ the aim is to find the terms, in a term collection (target thesaurus), better matching the semantics of a term in a source thesaurus.

As the $\mathcal{IR}$ problem [12], the $\mathcal{TM}$ problem can be formalized as $\mathcal{TM} = [D, Q, F, R(q_i, d_j)]$ where:

1. D is the set of the possible *logical views* of a concept in a target thesaurus (a document in $\mathcal{IR}$);
2. Q is the set of the possible *logical views* of a term in a source thesaurus (a query in $\mathcal{IR}$);

[2] Hereinafter when the expression "term" is used for addressing the elements of a thesaurus, we refer to a "simple or complex term": for example *Parliament* is a simple term, *President of the Republic* is a complex term.

3. F is the framework of concept representations;
4. $R(q_i, d_j)$ is a ranking function, which associates a real number with (q_i, d_j) where $q_i \in Q$, $d_j \in D$, giving an order of relevance to the concepts in a target thesaurus with respect to a concept of a source thesaurus.

Having identified an isomorphism between $\mathcal{IR}$ and $\mathcal{TM}$, the implementation of a specific methodology for $\mathcal{TM}$ is connected to the instantiation of $\mathcal{TM}$ using different approaches for: (i) the identification of suitable frameworks F for $\mathcal{TM}$, namely the representations of concepts in source and target thesauri (Q and D respectively) for better representing the "semantics" of thesaurus concepts, and in a way amenable for computation; (ii) the identification of the ranking function $R(q_i, d_j)$ between the semantics of such concepts.

This framework can be implemented following recent directives for using open-standards coming from the W3C community and for the need to propose semantic Web oriented solutions. In particular for representing mapping relationships among thesauri the standards SKOS (Simple Knowledge Organisation System) can be used [13]. It provides a standard way to represent knowledge organisation systems using RDF to describe concept schemes such as thesauri, classification schemes, subject heading lists, taxonomies and other types of controlled vocabulary, thus guaranteeing interoperability among applications.

4 A Thesaurus Mapping Case Study

The methodology framework previously described has been implemented in a case study of interest for the Publications Office of the European Union; such case study is aimed at testing the interoperability among five thesauri: EUROVOC, ECLAS, GEMET, UNESCO Thesaurus and ETT. EUROVOC is the main EU thesaurus containing a hierarchical structure with inter-lingual relations. It helps a coherent and effective managing, indexing and searching information of EU documentary collections, covering 21 fields. ECLAS is the European Commission Central Libraries thesaurus, covering 19 domains. GEMET, the GEneral Multilingual Environmental Thesaurus, is utilised by the European Environment Agency (EEA). UNESCO Thesaurus is a controlled vocabulary developed by the United Nations Educational, Scientific and Cultural Organisation which includes subject terms for several areas of knowledge. ETT is the European Training Thesaurus providing support to indexing and retrieval vocational education and training documentation in the European Union.

According to the project specifications, a mapping between EUROVOC and the other thesauri of interest are expected. The basic mapping methodologies are applied to *descriptors* within corresponding microthesauri in their *English version* as a pivot language. The system workflow is here below described (Fig. 1 can be considered as reference).

4.1 SKOS Core Transformation and Term Pre-processing

Thesauri XML proprietary formats are transformed into an RDF SKOS Core representation using XSLT techniques. Moreover, to reduce the computational

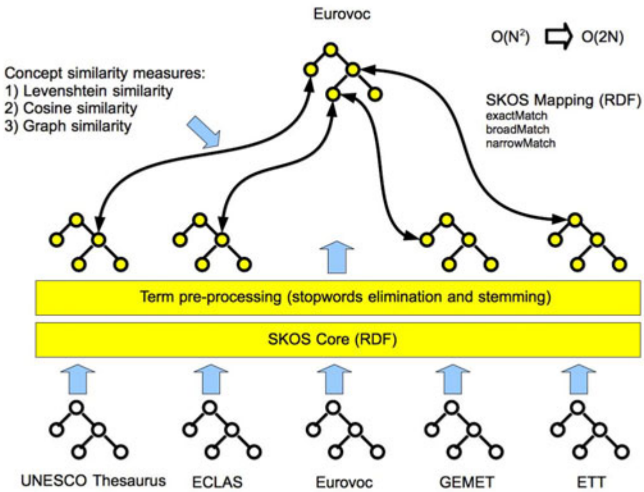


Fig. 1. Thesaurus mapping workflow

complexity of the problem, thesaurus terms are normalized so that digit characters and non-alphabetic characters (if any) are represented by a special character; then other operations as *stemming* or the use of word stoplists (*stopwords elimination*) are performed. Once basic terms have been identified, a vocabulary of terms for each thesaurus is created, containing all the terms which occur at least once in the set. The following steps consist in the instantiation of the components of our $\mathcal{TM}$ formal characterization.

4.2 Logical Views of Concepts in Source (Q) and Target (D) Thesauri

As discussed, term mapping between thesauri is a process which aims at matching term semantics (*concepts*) rather than finding their lexical equivalences. In traditional thesauri each meaning is expressed by one or more terms in the same language (for instance ‘pollution’, ‘contamination’, ‘discharge of pollutants’), as well as in different languages (for instance, the English term ‘water’ and the Italian term ‘acqua’, etc.). Moreover each word can have more than one sense, i.e. it can express more than one concept. In this view in order to effectively solve, automatically or semi-automatically, the $\mathcal{TM}$ problem, term (simple or complex) semantics has to be captured.

The semantics of a term is conveyed not only by its morphological characteristics, but also by the context in which the term is used as well as by the relations with other terms. In the $\mathcal{TM}$ problem F is exactly aimed at identifying the framework for term representations able to better capture term semantics in source and target thesauri. We propose to represent the semantics of a thesaurus term, a concept, according to an ascending degree of expressiveness, by: its *Lexical Manifestation*, its *Lexical Context*, its *Lexical Network*. A *Lexical Manifestation* is a thesaurus concept represented as a string of characters, normalized according to the pre-processing steps discussed in Section 4.1. A *Lexical Context* is a thesaurus concept represented by a vector d of term binary entries (statistics on terms are not possible since document collections are not available) composed by the term itself, relevant terms in its definition and linked terms of a T-dimensional vocabulary (F is T-dimensional vectorial space and linear algebra operations on vectors). A *Lexical Network* is a thesaurus concept represented as a *direct graph* where nodes are terms along with related ones, and the labeled edges are semantically characterized relations between terms. Lexical Contexts and Lexical Networks are built considering a strict adjacency relations between terms.

4.3 The Proposed Framework ($\boldsymbol{F}$)

Having identified thesaurus terms logical views, the *frameworks* in which the $\mathcal{TM}$ problem can be modeled are also identified. For *Lexical Manifestations*, the framework is composed by strings representing terms and the standard operations on strings. For *Lexical Contexts*, the framework is composed by T-dimensional vectorial space and linear algebra operations on vectors. For *Lexical Networks*, the framework is composed by graphs and the algebra operations on graphs. The identified frameworks can also provide the intuition for constructing a ranking function R, for the chosen representation of the space elements (terms).

4.4 The Proposed Ranking Functions ($\boldsymbol{R}$)

A ranking function R is able to provide a similarity measure between concepts in source and target thesauri; when extended to a set of target concepts such a function may provide a matching order among them.

For Lexical Manifestations a similarity string-based measure techniques are used, in particular the *Edit distance* (or Levenshtein distance, $dist_{lev}(s_i, s_j)$) applied on pre-processed strings (s_i,s_j) through language-based techniques (as tokenization, stemming and stopword elimination) normalized with respect to the longest string. In particular a similarity measure is used, defined as

$$sim_{lev}(s_i, s_j) = 1 - dist_{lev}(s_i, s_j), \quad sim_{lev} \in [0, 1]$$

Similarity measures for Lexical Contexts, being binary vectors, can be measured as the correlation between such vectors, quantified, for instance, as the cosine of the angle between these two vectors.

As regards Lexical Networks, basically direct graphs, a vast literature exists to measure the distance between them [14] [15] [16]. In literature the more frequently addressed graph similarity measure is the *Graph Edit Distance*, namely the minimum number of nodes and edges deletions/insertions/substitutions to transform a graph g_1 into a graph g_2. Because of computational complexity we have considered three variants of the Graph Edit Distance: *Conceptual similarity*, *Relational similarity* and *Graph similarity*.

The *Conceptual similarity* sim_c expresses how many concepts two graphs g_1 and g_2 have in common by an expression analogous to the Dice coefficient [17]:

$$sim_c = \frac{2n(g_c)}{n(g_1) + n(g_2)}$$

where g_c is the *maximum common subgraph* of g_1 and g_2 (it defines the parts of both graphs which are identical to one another, and, intuitively, it describes the intersection between the two graphs) and $n(g)$ is the number of concept nodes of a graph g. This expression varies from 0 when the two graphs have no concepts in common to 1 when the two graphs consist of the same set of concepts.

The *Relational similarity* sim_r indicates how similar the relations between the same concepts in both graphs are. In a way, it shows how similar the contexts of the common concepts in both graphs are. The Relational similarity sim_r is targeted to measure the proportion between the degree of connection of the concept nodes in g_c, on the one hand, and the degree of connection of the same concept nodes in the original graphs g_1 and g_2, on the other hand. With this idea, a relation between two concept nodes conveys less information about the context of these concepts if they are highly connected in the original graphs, and conveys more information when they are weakly connected in the original graphs. Using a modified formula for the Dice coefficient, sim_r can be obtained as:

$$sim_r = \frac{2m(g_c)}{m_{g_c}(g_1) + m_{g_c}(g_2)}$$

where $m(g_c)$ is the number of the arcs in the graph g_c, and $m_{g_c}(g_i)$ is the number of the arcs in the immediate neighborhood of the graph g_c in the graph g_i. The immediate neighborhood of $g_c \subseteq g_i$ in g_i consists of the arcs of g_i with at least one end belonging to g_c.

Considering a graph g to be matched with a graph g_T as reference, a possible similarity measure able to sum-up the previous two is the *Graph similarity* [18]:

$$sim_g = \frac{N_c(g, g_T) + E_c(g, g_T)}{N(g_T) + E(g_T)}$$

where $N_c(g, g_T)$ is the number of nodes shared by graph g and g_T; $E_c(g, g_T)$ is the number of edges common to g and g_T; $N(g_T)$ is the number of nodes in graph g_T; $E(g_T)$ the number is of edges in g_T.

4.5 Ranking among Candidate Concepts

Concepts of the target thesaurus, represented according to one of the models here presented, are matched with the chosen concept in a source thesaurus using a proper similarity measure. Candidate concepts of the target thesaurus are ranked according to the similarity measure values $\in [0, 1]$ and a semantics is assigned to the mapping relation using proper heuristic threshold values (T_1, T_2 $\in [0, 1]$)

if $sim < T_1$ $\Longrightarrow$ No Match
if $T_1 < sim < T_2$ $\Longrightarrow$ Partial Match (broadMatch or narrowMatch)
if $T_2 < sim$ $\Longrightarrow$ Exact Match
The established relations are expressed using RDF SKOS Mapping standard.

4.6 Mapping Implementation

The automatic mapping system is aimed at supporting the activities of an editorial staff of experts using an editorial environment and cooperating in the identification of matching concepts. In this framework:

1. the expert will be able to choose a concept in the source thesaurus; with this action the expert will implicitly submit a query to the target thesaurus;
2. the system will be able to retrieve a number of possible matching concepts as well as ranking them according to a similarity measure;
3. the expert will be able to: (i) confirm the system suggestions (matching concepts and related relations); (ii) select an alternative among those suggested by the system; (iii) browse the target thesaurus to search for matching concepts not identified by the system.

In this framework an effective hint is the system ability to retrieve the most part of matching concepts (irrespective to the matching type (untypedMatch)) in the minimum number of the system predictions, so to reduce the number of global browsing of a target thesaurus to search for matching concepts. This way the experts will be mostly required to select a proposed matching concept and, in case, to confirm or specify the type of relation proposed by the system.

5 Interoperability Assessment through a "Gold Standard"

Interoperability between thesauri has been assessed on a "gold standard" data set, namely an ideal collection of expected term mappings. SKOS Mapping relations, limited to the *exactMatch*, *broadMatch* and *narrowMatch* relations, are established according to the expert judgment [19]. In particular when a term in EUROVOC corresponds exactly to one or more terms in a target thesaurus according to the expert judgment, the relation is an exact match. A broad equivalence is one where the source term is more specific in some sense than the target term. Similarly a narrow equivalence is one where the source term is more general in some sense than the target term [19]. Following [7] a complete and optimal mapping is established. A *complete mapping* [19] is one where a source term, having no exact equivalence in the target, is matched to *at least* one target term that is semantically broader and *at least* one target term that is semantically narrower. An *optimal mapping* [19] is one where the aforementioned broader target term is the nearest broader term to the source one, and where the aforementioned narrower target term is the nearest narrower term to the source one.

In our case study the "gold standard" has been constructed by domain experts using an application called THALEN (THesauri ALigning ENvironment),

Fig. 2. A THALEN screenshot showing thesauri parallel views and a form to establish mapping relations

developed within the project on a relational database, able to provide a user-friendly access to thesauri and simple functionalities to establish term relations for mapping. THALEN provides functionalities of user authentication, thesaurus loading, parallel view of two thesauri, term browsing and searching, manual mapping implementation, summary of the established term mapping as well as exportation of mapping relations in RDF SKOS. Moreover this application can be used for human validation of the automatic mappings. In Fig. 2 a screenshot of THALEN as developed for the project is shown.

The work of legal experts, in their commitment to reach a flexible matching among thesauri, has undergone through several meetings which have resulted in a satisfactory cooperation. This has allowed to define uniform criteria to build the "gold standard".

6 Preliminary Test of the Proposed $\mathcal{TM}$ Approaches

The "gold standard" produced by experts has been used to assess the proposed methodologies for automatic thesaurus mapping: it includes 624 relations, of which 346 are exactMatch relations. The system performances are assessed with respect to the "gold standard" using, for each mapping relation type, the typical *Precision* and *Recall* measures. In particular for the case-study, the system *Recall* with respect to a less specified generic relation between concepts (untypedMatch) has been assessed, since system predictions will be specified and validated by humans.

Preliminary experiments showed satisfactory performances regarding the identification of untypedMatch relations between terms, and, as a consequence, the identification of "No Match" relations; on the contrary, good performances have been obtained regarding the selection of exactMatch relations, while the distinction between narrowMatch and broadMatch relations revealed a high degree of uncertainty. Global performances on specific types of relations are highly affected by this uncertainty, therefore meaningful results can be given with respect to untypedMatch and exactMatch relations with respect to the "gold standards".

The proposed logical views for thesaurus terms and the related ranking functions outperformed a simple string matching among thesaurus terms. In particular for (EUROVOC vs. {ETT, ECLAS, GEMET}) the Lexical Manifestation logical view and the Levenshtein similarity ranking function have given the best results (untypedMatch Recall = 66.2%, exactMatch Recall = 82.3%), while for EUROVOC vs. UNESCO Thesaurus the best results have been obtained using the Lexical Network logical view and the Conceptual similarity ranking function (untypedMatch Recall = 73.7%, exactMatch Recall = 80.8%).

7 Conclusion

Access to information regardless of language barriers is a key factor for effective global sharing of knowledge. This need has created a convergence of interests from various research communities around cross-language retrieval, which is becoming more and more popular between information specialists. In this context multilingual legal information retrieval systems do represent the necessary tools to encourage multilingualism in the law domain and have the chance to make it effective. Through linguistic-conceptual correspondence definition activities it can be possible to set up experimental applications, pilot projects and systems capable of fostering awareness and understanding of different countries legal concepts, thus facilitating worldwide access to law material.

This paper, presenting a methodological framework of schema-based thesaurus mapping on different five thesauri has offered preliminary results revealing that a simple Lexical Manifestation logical view and Levenshtein similarity ranking function produced the best results on most of the matches between the thesauri of interest. More complex descriptions of thesaurus concepts (Lexical Contexts and Lexical Networks logical views) suffer from problems of computational tractability (in particular as regards Lexical Networks). To improve the system performances using more expressive concept representations (as Lexical Contexts and Lexical Networks) different criteria of features selection can be tested to reduce the computational complexity, as well as the variability of the similarity measures.

References

1. Sadat, F., Yoshikawa, M., Uemura, S.: Exploiting thesauri and hierarchical categories in cross-language information retrieval. Springer, Berlin (2002)
2. Chan, L.M., Zeng, M.L.: Ensuring interoperability among subject vocabularies and knowledge (2002), http://www.ifla.org/IV/ifla68/papers/008-122e.pdf

3. Fluhr, C.: Multilingual information retrieval. In: Survey of the State of the Art in Human Language Technology (1996), http://cslu.cse.ogi.edu/HLTsurvey/ch8node7.html
4. Oard, D.W.: Alternative approaches for cross-language text retrieval (1997), http://www.glue.umd.edu/~dlrg/filter/sss/papers/oard/paper.html
5. Smith, T.F., Waterman, M.S.: Identification of Common Molecular Subsequences. J. Mol. Biol. 147, 195–197 (1981)
6. Curtin, D.M.: Citizens' fundamental right of access to eu information: an evolving digital Passepartout? Common Market Law Review 37(1), 7–41 (2000)
7. Doerr, M.: Semantic problems of thesaurus mapping. Journal of Digital Information 1(8) (2001)
8. Miles, A., Matthews, B.: Deliverable 8.4: Inter-thesaurus mapping (2004), http://www.w3c.rl.ac.uk/SWAD/deliverables/8.4.html
9. Rahm, E., Bernstein, P.: A survey of approaches to automatic schema matching. The International Journal on Very Large Data Bases 10(4), 334–350 (2001)
10. Euzenat, J., Shvaiko, P.: Ontology Matching. Springer, Heidelberg (2007)
11. Trojahn, C., Moraes, M., Quaresma, P., Vieira, R.: A Cooperative Approach for Composite Ontology Mapping. Journal on Data Semantics, 237–263 (2008)
12. Baeza-Yates, R., Ribeiro-Neto, B.: Modern Information Retrieval. Addison-Wesley, Reading (1999)
13. Miles, A., Brickley, D.: Skos simple knowledge organization system reference (2008), http://www.w3.org/TR/skos-reference
14. Sowa, J.: Conceptual Structures: Information Processing in Mind and Machine. Addison-Wesley, Reading (1984)
15. Zhong, J., Zhu, H., Li, J., Yu, Y.: Conceptual graph matching for semantic search. In: Priss, U., Corbett, D.R., Angelova, G. (eds.) ICCS 2002. LNCS (LNAI), vol. 2393, pp. 92–106. Springer, Heidelberg (2002)
16. Montes-y-Gómez, M., Gelbukh, A., López-López, A., Baeza-Yates, R.: Flexible Comparison of Conceptual Graphs. In: Mayr, H.C., Lazanský, J., Quirchmayr, G., Vogel, P. (eds.) DEXA 2001. LNCS, vol. 2113, pp. 102–111. Springer, Heidelberg (2001)
17. Rasmussen, E.: Clustering algorithms. In: Frakes, W.B., Baeza-Yates, R. (eds.) Information Retrieval: Data Structures & Algorithms (1992)
18. Cai, W.-T., Wang, S.-R., Jiang, Q.-S.: Address extraction: a graph matching and ontology-based approach to conceptual information retrieval. In: Proceedings of the Third International Conference on Machine Learning and Cybernetics, pp. 1571–1576 (2004)
19. Liang, A.C., Sini, M.: Mapping AGROVOC and the Chinese Agricultural Thesaurus: Definitions, tools, procedures. New Review of Hypermedia and Multimedia 12(1), 51–62 (2006)
20. Tiscornia, D.: The Lois Project: Lexical Ontologies for Legal Information Sharing. In: Proceedings of the V Legislative XML Workshop, pp. 189–204. European Press Academic Publishing (2007)

A Framework for Graph-Based Parsing of German Private Law Decisions

Florian Kuhn

Cognitive Science Center of Excellence / Dept. of Linguistics
University of Potsdam
Karl-Liebknecht-Str. 24-25, 14476 Golm, Germany
fkuhn@uni-potsdam.de

Abstract. We present a work in progress report of our research on automatically analyzing German court decisions. For accessibility to information retrieval and text summarization processing, we show concepts of a parsing framework dealing with linguistic features of this genre. To cover these features, we first developed a description language inferred from content structure analysis presented before. The main aspect of this paper is the presentation of a graph-based parsing framework concept. The parser will be able to label the content structure of German court decisions.

1 Introduction

Looking at German court decision documents, we find certain characteristics that permit structuring according to a particular content schema. Exploiting these characteristics for retrieval tasks is obvious since interested parties can make use of information compiled from document archives of related topics.

In this short paper, we will present a concept of a graph-based parsing framework that will label German court decision documents according to their content structure zones. To achieve flexibility of this framework, we first developed a lightweight description language capable of describing a range of linguistic features that are apparent in our document collection.

We restricted our view on private law decisions. While looking at a number of texts issued by different county courts we discovered frequent similarities in terms of overall structure of these documents. Respecting these facts, we collected a small corpus of 40 German private law decisions of 12 different county courts, trying to cover many legal domains and documents of different length between about four and 13 pages. Every document has been pre-segmented into *xdoc*-format to several divisions according to a set of rules, which is used in the MOTS-Framework of Potsdam-University.

2 Related Work

An early approach to legal content analysis is the SALOMON project [10] for structure analysis of Dutch criminal cases. The authors used a text grammar to

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 292–297, 2010.
© Springer-Verlag Berlin Heidelberg 2010

identify and summarize criminal cases for the purpose of guidance to lawyers. In two steps, relevant text units were extracted. First, a knowledge based SGML-syntax text grammar identified the case structure. Then, shallow statistical methods retrieved text segments of the alleged offenses and the opinion of the court. The system performed well with best precision values about 82% and 95% depending on the methods. Since a content structure approach was applied successfully, it seemed promising for related analysis on German legal documents.

Another project to summarize legal documents is the SUM project [7] summarizing an XML-corpus of judgments of the *House of Lords*. A rhetorical annotation scheme inspired by [17] was utilized here. Moreover, automated linguistic annotation was done by processing the documents with complex tokenizer (including part of speech tagging) and a linguistic analysis (including named entity recognition and clause detection). After annotation, a number of well-established machine learning techniques were used to train rhetorical role and relevance classifiers. Experiments based on cue phrase information, Support-Vector Machines and Maximum Entropy showed encouraging results, which motivated the authors to conclude that these steps may underlie the subsequent summarization algorithm as well as highlighting the utility of the rhetorical annotation scheme for legal discourse, thus its relevance for the research presented in this paper.

The only related work on German law decisions we are aware of is [18]. Here, the idea is to extract legal *definitions* by using a rule-based approach, working with a corpus of verdicts of German environmental law. The verdicts were first processed by a dependency parser to construct abstract semantic representations of the sentences. These were used to transform the definition's dependency patterns into a set of extraction rules based on connectors and signal words that consolidate the definitions found in the corpus. Different evaluation techniques were used and best precision values achieved were slightly above 70%.

Finally, aspects of the content zone structure concept for German court decisions as well as first frequency analysis results were shown by us in [15] and [16].

Research using graph-based algorithms for NLP tasks are not uncommon and have become popular in recent years while being used since decades for classification in general. Applications, for example, have been done in the field of number classification [11] and text summarization [9], where ranking methods are facilitated. Another example is [1], where relaxation labeling and edge pruning were applied.

3 Parsing Prerequisites: A Content Zone Description Language

By analyzing our corpus of German court decisions, we identified five macro-segments or *global zones* that can be further divided into more fine-grained content zones we call *local zones*. All global segments appear in a fixed sequence order while local zones have a less strict order of appearance. Sometimes they are optional. While the sequence and names of the global zones are motivated according to the German *Zivilprozessordnung*[1], the local zones are partially motivated on observations made in [4]. Local zones in particular are furthermore divisible into *functional* and *formal* ones. Functional zones in legal context cover argumentation and display of disputable views while formal zones contains information on the context of the case (e.g. names, court and file).

To appropriately describe the characteristics that involve the content zone concept, we developed a small XML markup language called *zone description language*(zdl) defining the analysis process of a given zone schema by feature-elements of certain linguistic manner. A content zone that has been defined as a valid element of a document of the corpus corresponds to a zone element of the description language. At the moment, a zone element is defined by three attributes.

- `zoneID` A unique integer.
- `name` String that should be self-explanatory
- `depID` optional reference to superordinate.

Next important part of zdl is the feature element. A feature is child-element of a local zone. Every feature is defined by a sequence of attributes that set conditions for this feature. Moreover, they refer to a specific text genre independent expert module that processes the active text segment accordingly. Return values of this module are evaluated by the parser. At the moment, we are using following attributes for `<feature>` elements:

- `featureID` A unique integer.
- `name` String that should be self-explanatory.
- `type` Classification of feature type.
- `sufficiency` Binary value for sufficient prerequisites.
- `weight` Frequency weight for this feature. Note this is a zone independent frequency.
- `parameter` A specific value passed on to the expert module.
- `moduleURI` Addresses the independent expert module.

An zone element example describing the features of the plaintiff's view, is shown in the listing below:[2]

```
<zone zoneID="22" name="Klaeger_Streitig" depID="18" sufficiency="true
    ">
<feature featureID="36" name="Klaeger_Streitig_keys0" type="keyword"
    sufficiency="true" weight="1.0" parameter="Meinung|Ansicht|
        behaupte(n|t)|tr.gt vor" moduleURI="www.ling.uni-potsdam.de
        "/>
<feature featureID="37" name="Klaeger_Streitig_mode" type="
    sentence_mode" sufficiency="false" weight="0.5" parameter="
    Konjunktiv I" moduleURI="wwww.ling.uni-potsdam.de"/>
</zone>
```

[1] Translation: *civil law legal proceeding prescription*.

[2] The attributes' values are basically placeholder at the moment and do not necessarily represent the final notation form.

4 Parsing Framework Concept

For purposes of parsing of German court decision documents, we decided to use a directed graph-based model for document model description. The segmentation of the documents into divisions, each of at least a header, in most cases a sequence of sentences, is suitable for a document graph of relatively small size. Moreover, the moderately strict sequence structure of court decision documents will generate graphs low on cycles and optionality. Regarding the xdoc divisions of the input data, we have either *header* or *paragraph* segments. Although we take this pre-processing information into account, we will also re-evaluate the state of each division to avoid adopting initial segmentation errors.

To build a graph-based document model that will be used for parsing court decisions, we first define some prerequisites for a Graph $G = \{V, E\}$:

- $v \in V$ is a local zone.
- Every edge $e \in E$ is a possible transition between two neighboring zones.

We extend this simple graph-model and add F, W, τ and ϕ, were F is a set of features, W a set of weights, ϕ a function mapping $v \in V$ to a subset of F and τ a function mapping $e \in E$ to $w \in W$. The coordinates of each w are constituted by the evaluation return values of a feature expert module assigned to every feature in that zone. A valid document passed to the parser will be represented by a possible path through the graph.

To extend the parsing framework by the patterns of linguistic features stated above, we make every transition probability feature value dependent: Every feature included in a local zone is evaluated by passing a specific input (an xdoc `<div>`-Element) on to a feature-expert module and then computing a final zone-probability based upon the modules' output value. The output value reflects a weight derived from the frequency of the feature in that zone. The final weight is mean of all process values completed.

Features that are sufficient for a zone to identify, do not necessarily bear a maximum weight for it might appear seldom, but nonetheless cause a positive zone identification. This is done by the attribute `sufficiency` which triggers the evaluation process to quit and allocates a maximum weight to the transition. Moreover, to keep the evaluation process cost-efficient, we order all feature-module-invocations according to their complexity, having low cost keyword-matching tasks on top and supporting syntactical analysis on lower levels. All steps of this weighting heuristic are summarized below:
At a vertex v

1. Choose the next most obvious *outdeg* e and invoke evaluation of features of corresponding v' with $e = outdeg(v) = indeg(v')$.
2. Start with feature expert module f_1 having lowest complexity and store return value.
3. Repeat with next lowest complexity module $f_2, ..., f_n$.
4. If a sufficient feature is identified, quit evaluation process and maximize transition probability.

5. Quit evaluation and allocate mean weight of all process values to transition, thus $\tau(e) \rightarrow w$.
6. Repeat 1.-4.
7. Compute probabilities and choose e with $\max(prob(e))$.
8. Repeat 1.-7. with e'.

The following segment, which forms the plaintiff's view, illustrates the concept of feature finding and weighting:

Der Kläger trägt vor, die Registrierung der Domain "friedrich.de" durch die Beklagte, die diesen Namen nicht führe, beinhalte eine Namensleugnung gemäß §12 BGB.[3]

Assuming that non-disputed facts in the facts-zone have already been identified, the local zone's node has two *outdegs*: either the local zone of plaintiff's view or plaintiff's proposal. At this moment, we have following zdl-entries for both zones: Starting with the evaluation of the latter which uses two features, the parser will first try to find a keyword phrase feature *Der/Die Kläger...(hat/haben) beantragt/beantragen* (the plaintiff(s) propose(s)). Since this sufficient feature cannot be found, the parser searches for the next feature, an *extended infinitive clause*[4], which is a specific German grammar feature. This clause also cannot be found, thus the parser cancels evaluation for plaintiff's proposal and switches to evaluate plaintiff's view zone features. This zone possesses two features. A sufficient keyword phrase feature (*(ist der) Meinung/ Ansicht/ behaupte(n/t)/ trägt vor*) is checked first. A positive match is found. Since the feature is sufficient, the parser can state a positive identification of the zone. In several cases, however, the supporting feature can contribute to this process. For the plaintiff's view, the second feature is a subjunctive sentence mode. Indeed in the subclauses of the example, the predicates found (*führe,beinhalte*) are in subjunctive mode.
Finally, to obtain the most probable document graph with the method sketched above, common search algorithms on weighted graphs like A^* can be used.

5 Summary

In this work in progress paper we presented a graph-based parsing architecture for pre-segmented court decision documents based on a linguistically motivated description language and content structure analysis shown before. Related work, first results and the fact of relatively strict rules of document writing in German court decisions seem promising enough to deepen research in this field of work.

References

[1] Angelova, R., Weikum, G.: Graph-based Text Classification: Learn from Your Neighbors. In: Proceedings of the 29th Annual International ACM SIGIR Conference, Seattle (2004)

[3] Translation: *The plaintiff states that the registration of the domain 'friedrich.de' by the defendant, who does not bear this name, is a denial of name according to §12 BGB.*

[4] In German: *Erweiterter Infinitiv mit "zu".*

[2] Bieler, H., Dipper, S., Stede, M.: Identifying formal and functional zones in film reviews. In: Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue, Antwerpen (2007)
[3] Dipper, S.: XML-based Stand-off Representation and Exploitation of Multi-Level Linguistic Annotation. In: Proc. of Berliner XML-Tage (BXML 2005), pp. 39–50 (2005)
[4] Engberg, J.: Signalfunktion und Kodierungsgrad von sprachlichen Merkmalen in Gerichtsurteilen. Hermes - Journal of Language and Communication Studies (1992)
[5] Freeman, J.: Dialectics and the macrostructure of arguments: A theory of argument structure. Foris, Berlin (1991)
[6] Hanneforth, T., Heintze, S., Stede, M.: Rhetorical Parsing with Underspecification and Forests. In: Proceedings of the HLT/NAACL Conference 2003, Companions Volume. Edmonton/AL (2003)
[7] Hachey, B., Grover, C.: Extractive summarization of legal texts. Artificial Intelligence and Law 14, 305–345 (2006)
[8] Marcu, D.: The Theory and Practice of Discourse Parsing and Summarization. MIT Press, Cambridge (2000)
[9] Michalcea, R.: Graph-based Ranking Algorithms for Sentence Extraction, Applied to Text Summarization. In: Proc. of ACL, Barcelona (2004)
[10] Moens, M.-F., Uyttendaele, C., Dumortier, J.: Abstracting of Legal Cases: The SALOMON Experience. In: Proc. of the 6th Int'l Conference on Artificial Intelligence and Law, Melbourne (1996)
[11] Radev, D.R.: Weakly-Supervised Graph-based Methods for Classification. U. Michigan Technical Report CSE-TR-500-04 (2004)
[12] Rave, D., et al.: Paraphrasen juristischer Texte, Darmstadt (1971)
[13] Stede, M., Bieler, H., Dipper, S., Suryiawongkul, A.: SUMMaR: Combining linguistics and statistics for text summarization. In: Proceedings of ECAI, Riva del Garda (2006)
[14] Stede, M., Sauermann, A.: Linearization of arguments in commentary texts. In: Proceedings of the Workshop on Multidisciplinary Approaches to Discourse, Oslo (2008)
[15] Stede, M., Kuhn, F.: Identifying the Content Zones of German Court Decisions. In: 2nd LIT 2009 Proceedings, Poznan (2009)
[16] Stede, M., Kuhn, F.: Document Structure and Argumentation in German Court Decisions. In: ICAIL 2009 NaLEA Workshop, Barcelona (2009)
[17] Teufel, S., Moens, M.: Summarizing Scientific Articles – Experiments with Relevance and Rhetorical Status. Computational Linguistics 28(4) (2002)
[18] Walter, S., Pinkal, M.: Linguistic support for legal ontology construction. In: Proceedings of ICAIL, pp. 242–243 (2005)

Mediated Signatures - Towards Undeniability of Digital Data in Technical and Legal Framework*

Przemysław Kubiak[1], Mirosław Kutyłowski[1],
Anna Lauks-Dutka[1], and Michał Tabor[2]

[1] Wrocław University of Technology
[2] Trusted Information Consulting, Warsaw
{przemyslaw.kubiak,miroslaw.kutylowski,anna.lauks}@pwr.wroc.pl,
michal.tabor@ticons.pl
http://kutylowski.im.pwr.wroc.pl

Abstract. We present an approach different from the classical framework of qualified electronic signatures that seem to offer much better properties regarding usability, costs, risk level as well as security in technical sense. Our approach is based on mediated signatures and chaining mechanisms that replace time stamping services.

Keywords: electronic signature, qualified signature, mediator scheme, chaining, Merkle tree, undeniability.

1 Introduction

Electronic documents play an increasing role in legal systems. Gradually, they replace paper documents; economic motivation as well as ease of use, erode the legal institution of written form. There have been an attempt to introduce an electronic equivalent version of written form - based on attaching qualified electronic signatures to electronic documents [1]. Despite of well designed legal framework these attempts have failed on a large scale. The reason for this phenomenon is not conservatism of business organizations or people, it is rather related to serious security problems and high risk for all participating parties. The high cost compared with low profit has been decisive for market failure so far.

After all problems related to qualified electronic signatures there is a tendency to replace them with solutions that are cheaper and easier to deploy. Unfortunately, quite often short time and long time consequences for legal safety are ignored and may lead to serious problems once the systems are really widespread.

Novelty of the paper. In this paper we discuss two techniques: mediated signatures and hash chains, that seem to be much better suited than qualified signatures and time stamping for implementing in a legal framework as standard

* Supported by Polish Ministry of Science and Higher Education fund for research & development in years 2009/2011, grant No. O R00 0015 07.

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 298–309, 2010.
© Springer-Verlag Berlin Heidelberg 2010

solutions. While the underlining cryptographic techniques are known for a long time, missing exchange of ideas between cryptographic community and social sciences prevented a deep discussion about their usage for solving practical challenges. As far as we know this is the first paper that opens a discussion about migration to a legal framework and usage scenarios of presented cryptographic mechanisms.

2 Concept of Advanced Digital Signatures

2.1 Written Form

Apart form the typical legal functions (such as finalizing the document by signing it), the institution of written form has a purely technological dimension. Namely, a signed paper document gives its verifier a proof about:

- the exact wording of a declaration,
- its integrity,
- and its originator.

Verification of a paper document has its strong and weak points:

- the results of signature verification are vague, coping techniques make it hard to distinguish between a copy and the original, when the inspection is made by a nonspecialist,
- verification is time consuming,
- serious verification requires other signatures of the same person, a pixel image of a signature is not really enough.

One of the serious disadvantages is that determining the signature creation time is impossible. On the other hand, over a long period of time paper documents offer quite good approximation of their age.

Risk issues. Relying document flow on paper documents has growing disadvantages due to changing economic and social conditions:

- as the signer and recipients of a document are dominantly not the members of the same local community, signature verification becomes unreliable,
- advance in coping techniques make it harder and harder to distinguish a copy and the original without a special equipment,
- paper documents can be modified after signing. This concerns in particular long documents that can be easily printed from electronic version just for signing.

2.2 Qualified Signatures

The concept of qualified electronic signatures is based on secret signature keys assigned to a physical person. In more detail, there are *signature creation data*, which are secret, and matching them public keys, called *signature verification data*. The signature verification data are linked to a person in a qualified certificate issued by a third party, a qualified certification services provider. The certificate is a digital document (in a standardized format) signed by the provider.

From technical point of view, verification of a qualified signature yields a strong proof that

– either the signature has been created with the secret signing keys or its equivalent, or
– the signature algorithm has been broken.

Additionally, the verification process should check the signature in the certificate, thus checking the link between a person and signature verification data. It has to be underlined that no check is possible that the declared signer has really used the secret signing keys for signature creation.

The legal concept of qualified electronic signatures [1] is to regard them as equivalent to handwritten signatures for fulfilling written form legal requirement. Since the owner of the keys may loose control over the secret signing keys, it is necessary to provide extra mechanism for dealing with this situation. Two mechanisms are implemented:

– each certificate has a limited validity period, after which validation process should yield a negative result,
– one can revoke a certificate, changing its status from valid to invalid.

As we see, while creating a qualified signature is an off-line process, signature verification requires checking status of the certificate, and therefore must be performed online. As each signature should be verified at least once (in many cases many times), the solution architecture is not an online one, as sometimes claimed. Depending on revoked certification lists (CRL) is risky for the verifier, since each CRL contains only data from the past. Hence the verifier may regard a signature as equivalent to handwritten one, while in fact it has lost its status.

Since determining status of certificates plays a crucial role in the trust system, verification depends very much on certification services provider. Therefore, legal systems formulate rules and control systems over qualified service providers. In this way security system's idea depending on cryptographic mechanisms reduces to organization security.

Standard implementation. In a standard implementation of qualified signatures the signing keys are implemented by a secure signature device - a cryptographic smart card. Access to the card is secured with a PIN only.

A person entrusting a signature has to check the status of a certificated with OCSP protocol or (less preferably) a recent CRL list.

Risk issues. In many countries qualified signatures are considered as the most advanced, secure and inter-operable solution. For this reason, quite often only advanced electronic signature with qualified signatures are considered to yield electronic documents with attribute of written form. This approach is reflected among others by the European directive [2] on services in the internal market. Unfortunately, one has to keep in mind many high level risks related to qualified signatures:

1. **Loosing control over signature creation device:** It is assumed optimistically that an owner can keep safely the cryptographic card holding

secret signature keys. This is not true in the real life, but the owner holds whole legal responsibility in case of malicious usage by a third party. Of course, to get access to the card it is necessary to provide a PIN. However, as the number of PIN numbers, passwords, logins is rapidly growing, one cannot expect that the owner does not keep a written record about his or her PIN. Even if this is not the case, the attacker can try to use the card of the owner exactly once, for instance when the owner is away from her office. Not all devices keep record in non-volatile memory about activation attempts.

2. **Poor randomness:** Many signature algorithms use random parameters for signature creation. This is the case for signatures based on the discrete logarithm problem – like DSA, ECDSA, but, notably, not for RSA. The problem is that in case of weak randomness the secret signing key may be leaked. While a lot attention is payed to the choice of cryptographic methods, verification of randomness quality is neglected. This is partially due to the fact that existing evaluation methods may provide negative results (poor randomness), while it is hardly possible to show that actually the randomness is good. Moreover, the problem is even harder due to black box principle. We are never sure that the inspected devices and the devices delivered to the end-users are built in the same way.
3. **Kleptography:** Apart from poor randomness, which reveals the keys to anybody inspecting signature-creation device, malicious kleptographic code can be inserted into the device (cf. a kleptographic channel in RSA signatures with *probabilistic* padding, i.e. in RSA-PSS [3]; note that kleptographic attacks are not applicable to RSA signature creation with *deterministic* padding like Full Domain Hash for example). We describe this problem in detail in Section 2.3.
4. **Retaining the key by the provider:** It is allowed that a provider of certification services not only issues qualified certificates, but also creates secret signing keys (signature creation data) for its clients. In this case it is strictly forbidden to retain a copy of the secret keys or the data that can be used for reconstructing the keys. However, enforcing this is left to criminal law only. There is no technical mechanism that would help to guarantee that the service provider is not retaining the keys, unless he stores the keys in a quite unprofessional way.
5. **Uncertainty about signing time:** As such, qualified signatures provide no verification mechanism for the time of their creation. The main idea for certificates is that signing is performed off-line, therefore no online mechanisms during signing time may be used. One the other hand, signature creation devices such as cryptographic smart cards have no internal clock and therefore can only sign time provided by a (untrustworthy) reader.
6. **Revoking certificates:** While certificate revocation is a mean of safety for signers, for a person entrusting an electronic document this is a legal risk. As the signer can demand to revoke his certificate at any time, a document with a qualified signature may not fulfill written form equivalence. In some legal systems, revoking the qualified certificate or its expiration means termination of equivalence with the written form.

7. **Non-qualified signatures based on qualified certificate:** Even if the European directive [1] defines advanced electronic signature as created with the means which are under sole control of the signer, it is unclear how to implement this requirement without a secure signature-creation device. On the other hand, there is a growing pressure on accepting advanced electronic signatures not created with secure signature-creation devices [2]. In particular, when accepting a signature based on qualified signature, one cannot be sure that the signature creation data is well protected. As a person entrusting the signature should be aware of this, in case of a fraud not necessarily the whole legal responsibility will be assigned to the signer.
8. **Decline of mathematical strength:** It turns out that progress in cryptanalysis and other attacks (among others, side channel attacks) against electronic signatures in unpredictable and result from new, striking ideas emerging suddenly. Good examples of this kind are lattice based methods for RSA cryptanalysis and Chinese attacks on SHA-1 function.
   On the other hand, in order to make ground for safe investment in electronic systems based on qualified signatures, in some countries the authorities declare by law which algorithms are secure (e.g. Austria, Poland), while the other ones (like Germany) only publish guidelines concerning these algorithms. Nevertheless, one has to be aware that these predictions are of little technical value. Moreover, in a case of serious failure, the state may be blamed for forcing to use wrong techniques (as for today, SHA-1 and RIPEMD-160 are the only hash functions allowed for qualified signatures in Poland [4]).
   Apart from the list of allowed algorithms, the legal system provides institution of revocation of certificates. In case of serious threats of this kind, all certificates concerning endangered public keys (signature verification data) must be revoked. This adjusts the legal situation to the real value of qualified signatures, but would mean a practical disaster in electronic information flow.
9. **Standards and obscure technical requirements:** There is a tendency to evaluate quality of products used for cryptographic purposes based on fulfilling technical standards. While technical standards help to keep many risk factors under control, especially when the designers have limited expertise, adhering to technical standards is not a guarantee of security. Each standard must concentrate on a set of issues that make the product inter-operable and comparable.
   As technical standards are created in a long discussion between interested parties, the outcome is a kind of compromise between different views. Also, it happened that already on the day of publication, security standard contained serious security flaws.

While various risks are inevitable in electronic systems, the concept of qualified signatures suffers from existence of many points of failure. Electronic documents become non trustworthy if any <u>single</u> event described below occurs:

- the device used to create the signature turns out not to be compliant with the security requirements,
- an unauthorized person gets access to the card used for signing,
- the qualified certification authority misbehaves,
- signing algorithm turns out to be insecure.

Potential signers aware about all these risks can be further discouraged to use qualified signatures due to limitations to deny the legal impact of a qualified signature (see for instance [5, Article 6]).

2.3 Malicious Cryptography and Qualified Signatures

One of the most serious threats for solutions based on devices holding secret cryptographic keys is technique called *kleptography* discovered by Adam Young and Moti Yung [6]. The idea is a kind of subliminal channel that is readable for holder of secret keys that are not inserted in the contaminated device.

The basic mechanism can be explained for the signatures such like ElGamal. Recall that one of the components of the signature is an element $r = g^k$ computed in an appropriate group, where k is a parameter chosen independently at random for each signature creation. One of the main points is that when k is revealed, then one can immediately compute the secret signing key. A kleptographic code computes one signature (r, s) according to the protocol, and uses the next signature $r', s')$ to leak the signing key. Namely, instead of creating k' at random, the device computes: $k' := \mathrm{HASH}(Y^k)$, where HASH is an appropriate hash function, and Y is a special parameter chosen by the author of the kleptographic code. Namely, the attacker knows x such that $Y = g^x$ in the group concerned. The rest of creation of (r', s') is as in the original procedure.

Now, when the signatures (r, s) and (r', s') are published, then the attacker can compute $\mathrm{HASH}(r^x)$. Since $r^x = (g^k)^x = (g^x)^k = Y^k$, the value computed is equal k' and, as mentioned before, the attacker can derive the signing key based on s' and k'.

Note that the method works even if the signing keys are implemented on secure cryptographic cards, that there is no extra communication with the device, and that the attacker remains fully anonymous, as all he needs is to get signatures. In particular, data retention performed allegedly for security reasons, is a perfect source of information for the attacker. Due to the properties of modular exponentiation and hash functions, it is infeasible to detect irregularities in the output of an infected device. If the side channel issues are treated properly (compare [7]), then the only way to detect a kleptographic code is invasive inspection during which the card gets destroyed.

3 Undeniability of Existence and Creation Time

One of the basic questions concerning electronic documents is not only their origin but also their existence at a certain time and place. According to administration law, and court procedures the most important is determining the date of creation of a document, or even the date of submitting it to a public body. In turn, in case of financial documents it might be necessary to determine the creation time with quite a big precision. On the other hand, the time horizon for keeping undeniability of electronic documents is much shorter for financial operations than for public administration. In the first case it might be something like two years, while in the second one long time security might be necessary.

3.1 Concept of Time Stamps

The simplest solution is a concept of time stamps issued by the third trusted parties. In this solution this service provider acts as a witness and creates a signed document stating that some document existed at a given time.

In order to be less dependent on honesty of time stamps provider it may be required by law to use secure time stamping devices and deploy a system guarding from manipulations of these devices. A secure time stamping device should provide mechanisms preventing changing the current time to a value earlier than already observed. Also it should coordinate time with a trusted source in a safe way.

The overall architecture requires quite a high investment in infrastructure to reach an acceptable assurance level. This in turn requires a large scale implementation in order to get return from the investment.

3.2 Hash Chains

There is a mechanism simpler than high-end time stamping devices and the system of propagating from trusted sources. The idea is that what really counts is a sequence of event and not their physical time. For instance, if we are talking about a date T of creating a document M, then it is sufficient to show that the event of starting a day T precedes the time of creation of M and the event of creation of M precedes the event of starting day $T + 1$.

For witnessing a sequence of events we may use hash chains. Namely, for each event E we compute a value $\mathcal{H}(\mathcal{E})$ in the way that:

- $\mathcal{H}(\mathcal{E})$ depends on some events prior to E,
- the event E is taken as input for calculating some values $\mathcal{H}(\mathcal{E}')$ for some events E' occurring after E.

The choice of the events mentioned might be random (then we are talking about random witness schemes), or arranged deterministically - typically we are composing Merkle trees of the events from, say a day, to a single hash value, or we mix deterministic and randomized approach.

Let us describe in detail some of the scenarios of this type:

1. **Sequence of events:** In this case the server composing the chain orders the event according to the request time. The stamps for event E_i is calculated as $H_i := H(E_i, H_{i-1})$. This guarantees that no values can be inserted in the past fragments of the chains, while appending the chain is very easy.

2. **Mixing sequences:** This is the same approach as the previous one, apart from some system events. Namely, from time to time the servers take events such as:
   – issuing random strings by public generators,
   – contact to a trusted time server,
   – linking messages received or sent to other server of this kind.

   Due to outside links that are obligatory, it becomes infeasible to slow down or to speed up the time flow apart from small intervals.
3. **Merkle trees:** A typical approach is to gather all events from a single day, to build a Merkle tree from them, and finally publish the root value together with a signature. Persons submitting the documents to the tree get a record (values adjacent to the path leading from the root to the leaf corresponding to this document) that may be used as a proof of inserting the document to the tree.

Note, however, that if some procedure utilizing a hash chain is used, then verification requires to look the chain or its authenticated copy up.

4 Mediated Signatures - Advanced Digital Signatures Revisited

In Section 2.2 we have seen that using traditional architecture of qualified signatures is connected to relatively high risk for the signer, for the recipient of electronic signatures as well as for the service providers. The risks have a direct influence on the costs for all participants and on very low popularity of qualified signatures in business. In fact, the only widespread usage of qualified signatures is due to legal obligations in contacts with public bodies.

Many of the risks can be easily eliminated, if we slightly reorganize the system architecture. A good example for such a solution are *mediated signatures*[8].

Mediated signature mechanism. In case of mediated signatures, there are multiple keys (at least two) necessary to create a signature. Say, for a public key K there is a pair of keys K_1 and K_2, such that the process of creating the signature looks as follows:

1. a hash value $H(M)$ of a document M and some meta-data are composed together to a structure D (details depend on the format used),
2. the key K_1 and value D are used to derive an intermediate value S', called presignature,
3. the key K_2, presignature S' and may be D are used to derive the signature S over M. This process will be called finalizing the signature of M.

The basic security properties are as follows:

– having K, K_1 and any number of signatures created with K_1 and K_2, it is infeasible to derive K_2,
– having K, K_2 and any number of signatures created with K_1 and K_2, and any number of presignatures created with K_1 it is infeasible to derive K_1,
– without K_1 or K_2 is is infeasible to create a valid signature S of M.

In particular, it is infeasible to derive any of the keys K_1, K_2 based on the public key K and valid signatures that can be verified with K. So we see that mediated signatures guarantee perfect separation of roles between the holder of K_1 and the holder of K_2.

There are many mediated signature schemes in the literature. Mediated signatures are particularly easy to implement in the RSA framework. Namely, given a public modulus n ($n = pq$), $\phi = \mathrm{GCD}(p-1, q-1)$, public exponent e and private exponent d we determine two secret keys as follows:

- we choose an arbitrary value $r < \phi$, r might be an output of another cryptographic algorithm, the only condition is that r must be co-prime with $p-1$ and $q-1$,
- we set $K_1 = d/r \bmod \phi$, $K_2 = r$ (however, in many protocols $d = K_1 + K_2$ is preferred due to potentiality to distributed generation of RSA keys.

Since

$$K_1 \cdot K_2 = d \bmod \phi \ ,$$

it follows that

$$m^d = (m^{K_1})^{K_2} \bmod n \ .$$

From this example we see that not only we get well defined mediated signature scheme, but the results a standard RSA signature, so nothing has to be changed on side of the verifier that works with RSA signatures.

One can also see that RSA admits splitting the private exponent into many parts so that more than two participants are necessary for creating a valid signature.

Another nice property is that the string S' cannot be used for any kind of substitute verification. The problem is that there is no public key that can be used for this purpose.

Risk analysis. Of course, using mediated signatures does not eliminate all threats. However, the major advantages are:

- elimination of a single point of trust (and failure), which is a secure signature creation device,
- possibility of controlling signing activity of the holder of K_1 by the holder of K_2.

How these two features can be used we discuss in the next section.

5 Application Scenarios of Mediated Signatures

5.1 Improving Safety of Signature Creation Devices

In a standard usage scenario it should be recommended to use a mediated scheme in store different keys on different, independent devices. For instance, a user might split his key between a cryptographic card (secure signature-creation device) and own PC (either laptop or a stationary computer), or even a cellular phone. The key advantages are:

– stealing a card or a laptop does not enable the thief to create valid signatures, stealing both of them is necessary,
– even when both devices are captured enabling them to work requires guessing the PIN and getting access to the computer. If the computer is well administered, this is harder than just guessing a 4 digit PIN. For instance, the computer might have a biometric verification of the user.

Last not least, one can also use devices coming form different manufacturers (and even form different countries). With careful implementation of mediated signature protocol it might happen that the system is immune against kleptographic attacks. The idea is that all random output of the first device are recomputed by the second device before outputting. In this way, the hidden channels contained in the output are destroyed.

5.2 Mediated Signatures for Public Administration

Mediated signatures can be used to solve many problems of authorization of documents submitted by the citizens to public bodies. In this scenario we are using a server called *Mediator* that holds the second key for each user. In this scenario signing a document M by a citizen A looks as follows:

1. the smart card of citizen A creates a presignature S' of M,
2. A submits S' to *Mediator*,
3. *Mediator* checks, if the the smart card of A is not on a black list of lost cards. If it is not on the list, then *Mediator* finalizes the signature and returns to A. (Optionally, *Mediator* can forward the signature to the final recipients declared by A.)

The formatted data D should contain also current time (or simply the current date). According to the signing policy, *Mediator* checks if the current time and the time contained in D do not differ too much before making decision to finalize the signature.

The keys used by *Mediator* can be generated in a way similar to creation of PIN numbers. Namely, the key assigned to A is $F(Z, A)$, where Z is the master secret of *Mediator* and F is a cryptographic one-way function. Therefore, *Mediator* can be implemented on a Hardware Secure Module (HSM) without holding a database with finalization keys.

Usability issues. Submitting an electronic document is performed online. So there is no substantial additional effort to contact *Mediator* for finalizing a signature. (At the time when the standard X.509 has been created the situation has been quite different - access to Internet was much more difficult.)

If signature creation data are implemented in an electronic citizen identity card, then we take advantage of a whole system of keeping the list of lost ID cards. If a card is declared as lost, then an appropriate record is inserted in a state registry. Moreover, *Mediator* can be immediately informed about revoking validity of the ID card in a fully automatic way.

Since time of signature creation is determined by *Mediator*, there is no need to use external services of time stamping. *Mediator* can implement chaining techniques to provide undeniability of existence of signed data at certain time.

Of course, *Mediator* HSM's may be cloned so that mediated signatures can be created in parallel at different locations. This helps to scale the system in a case of the massive use.

Risk analysis and advantages of mediated signatures. Thanks to *Mediator* it is easy to prohibit creating a signature with a stolen cryptographic smart card. Indeed, when a a smart card is declared to be stolen, then automatically it is placed on a *black list* and no action will be taken by *Mediator* to finalize a signature created by such a smart card. Since citizens are usually well educated to report about stolen citizen identity cards and not to borrow them or leave uncontrolled, the black list is maintained up to date.

The second advantage of mediator scheme is that *Mediator* can monitor finalization requests and react on suspicious circumstances (similar procedures are implemented for credit card activities and help to avoid frauds). For instance, if a request turns out to be suspicious it is possible to contact with the citizen and verify if he really started the signing process.

Mediator may offer additional security measure for the citizens. Namely, a citizen may temporarily suspend signing activities with his citizen identity card. This may be useful for instance in case of a serious surgery, when the citizen is not only unable to submit document electronically, but also control access to his identity card. Blocking the card for this period improves practical security.

From the point of view of recipient of a signature, mediated signatures offer much better level of confidence. Namely, positive verification result shows that at the time of the signature creation the citizen identity card was not on the list of the lost cards and nothing suspicious has been detected. Signature creation time can also be trusted.

Legal framework. Mediated signature schemes are compatible with the definition of advanced electronic signatures according to European Directive on electronic signatures [1]. The only difference is that apart from the secret key that remained solely under the control of the signer, there is additional mechanism based on the second secret that enables to prohibit uncontrolled usage of the signature-creation device.

In fact, the steps performed by *Mediator* is validation of the signature performed at the creation time. This is different than in the standard implementation of qualified signatures, where validation of a signature is performed later during signature verification. As we could see, the later approach has some deep practical problems.

One may ask how existence of *Mediator* is compatible with personal data protection regulations. In case of exchange documents with public bodies revealing signing pattern raises no problems, since finally the documents are known anyway to the authorities.

5.3 Mediated Signatures for Corporate Use

Of course, the mentioned mechanisms are quite attractive also for any organization, where signing activity need to be controlled. A local mediator installation has the following advantages:

– revoking signing authorization may be immediate and independent of returning or blocking the signature cards,
– there is a full accountability of signing activities, as the mediator can retain a record of signatures created,
– signing capability can be suspended temporarily, for instance preventing usage of signing card outside of the office hours.

In particular, mediated signatures is an attractive solution for controlling workflow of documents in the distributed environment with no well functioning other mechanism. Signatures created with mediator serve as a kind of the mailstones in electronic information flow.

Especially important might be applications of mediated signatures in the areas where on one hand high level of trust is necessary, while on the other hand the economic, legal or social consequences of decisions delivered on electronic way might be profound.

References

1. European Parliament and of the European Council: Directive 1999/93/ec of the european parliament and of the council of 13 december, on a community framework for electronic signatures. Official Journal of the European Communities L(13) (19/1/2000) (1999)
2. European Parliament and of the Council of Europe: Directive 2006/123/ec of the european parliament and of the council of 12 december, on services in the internal market. Official Journal of the European Communities L(376) (27/12/2006) (2006)
3. Young, A., Yung, M.: An elliptic curve backdoor algorithm for RSASSA. In: Camenisch, J., Collberg, C.S., Johnson, N.F., Sallee, P. (eds.) IH 2006. LNCS, vol. 4437, pp. 355–374. Springer, Heidelberg (2007)
4. Council of Ministers of Poland: Regulation on determining the technical conditions of organization for qualified certification service providers, certification policies for qualified certificates issued by these entities and the technical conditions for secure devices for the submission and verification of electronic signatures. Dziennik Ustaw 128(1094) (7/07/2002)
5. Polish Parliament: Electronic signature act. Dziennik Ustaw 130(1450) (18/09/2001)
6. Young, A., Yung, M.: Cryptovirology: Extortion-based security threats and countermeasures. In: IEEE Symposium on Security and Privacy, pp. 129–140. IEEE Comp. Soc., Los Alamitos (1996)
7. Kucner, D., Kutyłowski, M.: Stochastic kleptography detection. In: Alster, K., Urbanowicz, J., Williams, H.C. (eds.) Public-key cryptography and computational number theory (Warsaw 2000), pp. 137–149. Walter de Gruyter Inc. (2001)
8. Boneh, D., Ding, X., Tsudik, G., Wong, C.M.: Instantenous revocation of security capabilities. In: USENIX Security Symposium (2001)

Ontojuris Project: A Multilingual Legal Document Search System Based on a Graphical Ontology Editor

Tânia C.D. Bueno[1], Hugo C. Hoeschl[1,2], and César K. Stradiotto[1,2]

[1] I3G - Institute of E-government, Intelligence and Systems, Florianópolis -Brazil
+5548 30256609, CEP 88036-002
{cesar.stradiotto,tania.bueno,hch}@i3g.org.br
[2] Universidade Federal de Santa Catarina – UFSC, Florianópolis, Brasil

Abstract. The Ontojuris project consists of a multilingual system for legal information retrieval (IR) associated with an ontology editor. The system accepts queries in one language, and allows IR based on similarity of documents, written in one of several languages. The multilingual ontology editor works with the concept of Universal Words (UW´s). The UW´s are universal representations of things and objects, and it is developed inside the Universal Networking Language Project (UNL), allowing query expansion in a multilingual way. Query expansion can improve the searching process required by user, including additional terms which have similar meaning to the original query. In this study, we proposed a new expansion method which is based on domain ontology and UW´s, to achieve better performance for a multilingual IR. So, the Ontojuris project aims to facilitate access to information on legislation in the area of Intellectual Property Law, Consumer Rights and Electronic Law in the consortium formed by researches in Brazil, Chile, Spain and Argentina.

Keywords: Multilingual Ontology, Indexing, Query Expansion, Universal Words.

1 Introduction

Recently, there has been a growing interest in how institutions – together with its rules, norms, regulations and organizations – evolve, especially in settings beyond the nation state. In this context, IR in a multilingual legal domain is a topic of particular relevance because it is associated with high transaction and information costs.

Bridging the gap between legal text archives and legal structural knowledge is a principal task of studying the law [2], and the key challenge in multilingual legal IR. Term frequency / indexed document frequency methods (TF/IDF) do not help enough in law, as in other domains [10]. Together with that, in [12] a comparison research among TF/IDF and other two proposed indexing methods shows that the TF/IDF has inefficient parameter tuning and design choices. Currently, more and more organizations working in multilingual environments demand ontologies supporting different natural languages [5].

W. Abramowicz, R. Tolksdorf, K. Węcel (Eds.): BIS 2010 Workshops, LNBIP 57, pp. 310–321, 2010.
© Springer-Verlag Berlin Heidelberg 2010

The existing approaches use either translation of all documents into a common language, either automatic translation of the queries, or combination of both query or document translations ontology-based [7]. Unfortunately the problem of translating the knowledge expressed in natural languages is extremely complex. A feasible approach is to create an intermediate knowledge representation between a natural language and a computational one [4].

In this work, to support the process of multilingual information retrieval, the Ontojuris Project [9] has developed an environment for multilingual ontology construction based on the relationship between the conceptual classes that form the ontological model of knowledge representation (knowledge base) associated with Universal Words (UW´s). The UW´s are universal representations of things and objects developed by the project UNL [11]. This multilingual edition tool associates a tool-editing suite called Knowledge Engineering Suite (KES) [3] with the concept of UW´s [1].

Before to explain the ontology editor, we thought it was necessary to explain the document indexing and query expansion process, which will make the justification for the way the ontology editor shows the multilingual ontology.

This paper is organized as following: Section 2 introduces the background knowledge for this research – the Ontojuris project and its internal application of UW´s – and presents how its digital system for ontology-based information management is designed, to support the indexing process and multilingual query expansion. Section 3 presents the document indexing method that support the search process. Section 4 presents the query expansion method for the search process, using Universal Words. Section 5 explains the two proposed multiple document interface (MDI) ontology editors in its prototypical form: One in 2D space, and the other on the 3D space. In section 6 the paper is concluded. In section 7 we point out future research.

2 The Ontojuris Project

The Ontojuris project aims to facilitate access to information on law documents in the area of Intellectual Property Law, Consumer Rights and Electronic Law, stored on knowledge bases available under the management of the consortium formed by Brazil, Chile, Spain and Argentina[1]. For this account, since 2007 it has been formed a social network established between these countries, for the usage of Web-based technologies and methodologies, to help the construction of multilingual ontologies.

The motivations for the creation of legal ontologies are evident: knowledge base examination, common use and reuse of knowledge, knowledge acquisition and representation. The Ontojuris Project uses legal terms available in databases, to help the process for construction of multilingual ontologies (like is done in FAO – Food and Agriculture Organization Agrovoc Domain Ontologies[2], and IATE - Inter-Active

[1] Instituto de Governo Eletrônico, Inteligências e Sistemas - i3G (Brazil);Universidade Federal de Santa Catarina (Brazil); Universidade Politécnica de Madrid (Spain);Universidade de Fasta (Argentina);Canelo Abogados (Chile).

[2] http://aims.fao.org/website/Domain-Ontologies/sub

Terminology for Europe[3]). This "legal language", consisted by a complex structure of concepts, forms an abstraction derived from a text corpus, originally stored inside legal databases. Such legal structural knowledge does not only contain legal term interpretations, but also shows its (supposed) logical and conceptual structure.

2.1 System Description

2.1.1 The Ontology System

The Ontojuris system uses the ontologies created in the editor (Figure 1) to realize the information retrieval tasks for legal documents, collected from the web. The ontology representation for knowledge improves the search for legal documents by similarity, in the Ontojuris system. Particularly, the user can write some words, or even a complete text – copied integrally from some textual source – to be used as an entry for document searching.

It is possible to choose parameters like date intervals and knowledge-based sources by its title, where the searched documents are stored (the sources could be sites over the World Wide Web, or law document sets added by some semi-automatic way, inside a knowledge-base used as source).

Also inside this system, it is possible to filter legal documents by its source countries. At the present times, the system can retrieve documents stored in databases located on the following countries: Argentina, Brazil, Chile and Spain, which are the participants on the consortium.

Fig. 1. Ontojuris System interface

In the search by topics and sub-topics, instead of writing an entry text, the user can choose the fields of knowledge (topics) that are stored in the system database. This

[3] http://iate.europa.eu/iatediff/about_IATE.html

kind of interface, which groups documents by its contexts or similar subject, is quite similar to the interface presented in [14], and described in [15].

2.1.2 The Multilingual Editor for Legal Ontologies

In the development of this work to support the ontology engineering process, we use a methodology based on a description of semi-formal ontologies, called *Mind Engineering* [1]. In this methodology, instances of representation do not include description between objects, but only relations between them, within a domain.

In addition to methodologies, there are tools used to build ontologies. The KES is an independent computational structure for the development, creation and modification of ontologies, working partially on web and partially on desktop platforms. These editors are tools designed to assist the working team of engineers, to produce the knowledge and expertise in the construction of multilingual ontologies. The resulted ontologies consist in a network structure that connects words, considering their concepts in a defined field of knowledge, inside a specific application.

Such networks enable knowledge management applications to recognize the context of the searched documents, i.e., thanks to the help of these ontology networks, the applications based on this kind of knowledge representation can do contextualized searches for the stored documents.

The basic components for these ontology editors are classes (arranged in taxonomies), relations between terms (represent the type of interaction between the concepts of a domain), relations between a term and an UW, and subject domains (see Figure 2 below). In this project the ontology representation do not use axioms (used to model sentences always true) and do not use instances (used to represent specific elements, i.e. the data).

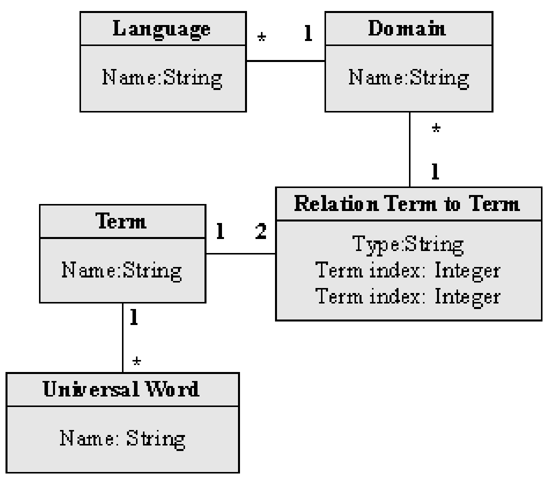


Fig. 2. UML Class Diagram for the Ontology Multilingual Editor

3 The Indexing Method

The Ontojuris system uses the ontologies created by the ontology editor for indexing and query information, contained in specific legal documents (i.e., norms and regulations). The terms created by ontologies construction method [1] are used for the indexing process.

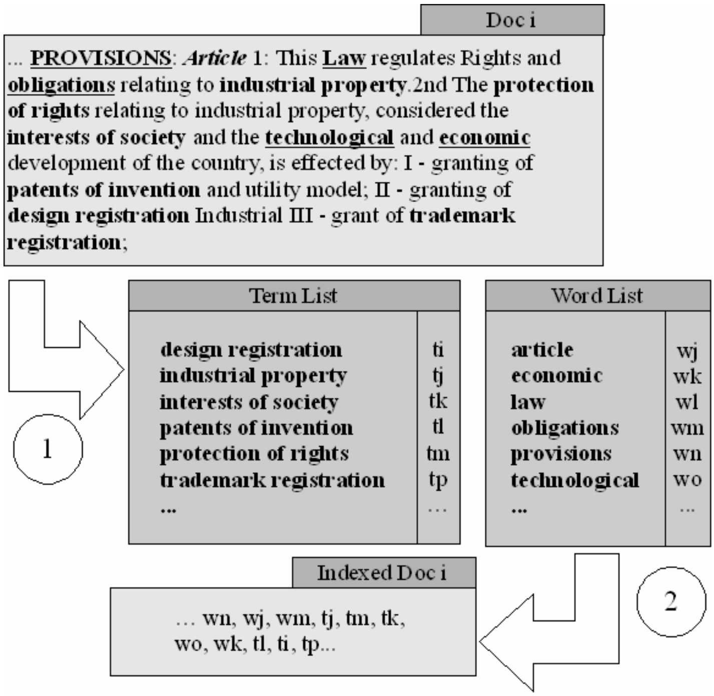


Fig. 3. Process for extracting words from the legal documents to a term and word lists, after preprocessing step, for document indexing

A semantic analysis tool (Mike Scott´s WordSmith)[4] is used to extract the terms from the document base, or some document samples. Some of these terms are considered relevant for future user to make their searching, and are accumulated on a term list (Fig. 3). The word list is generated from a dictionary in the language from where the document is originated. So, the documents are marked with the indexes of its contained terms and words (**Indexed Doc i** on Fig. 3). Observe that words **wi**, **wj**, … and terms **t0**, **t1**, **t2**, …, are present on the documents, and inside the word and term lists. The sample for Figure 3 was extracted from [16].

To explain how the system does a comparison between documents, a first diagram showing ontology with two domains is given on Figure 4, below. The relations are explained on Table 1. This ontology has the domain **Di** with the terms **t0** to **t7** with its corresponding relations, and the domain **Dj** containing almost the same terms: **t0**, **t2**, **t4**, **t5** and **t6**, but with different relations. For our model, the relations among terms depend on the domain it belongs.

[4] http://www.lexically.net/wordsmith/

Table 1. Symbols and weights of original terms relationships

Relationship	Symbol	Weight
Identical	**I**	1.00
Synonym	**S**	0.99
Part of	**P**	0.30
Type of	**T**	0.30

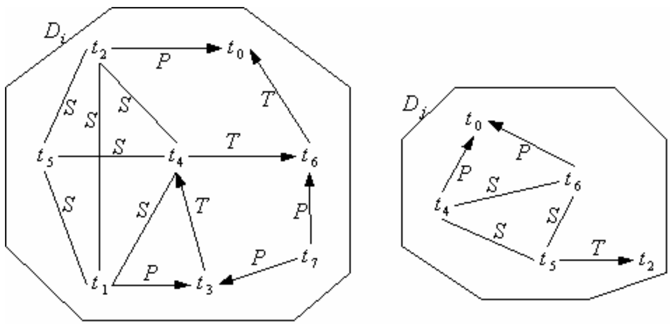


Fig. 4. An ontology sample, with two domains Di and Dj

Observe that synonyms are cliques: Given two terms **ti** and **tj** which are synonyms, if a third term **tk** is a synonym to **ti**, it is a synonym to **tj**, also. For the sample, we also have three example documents **Ui**, **DOCi** and **DOCj**, where **Ui** is the input text by the user, and **DOCi** and **DOCj** are documents on the knowledge base to be compared to **Ui**:

Table 2. Document content, to be compared using the ontology model, in the given sample

Document	Content (Terms and Words)
Ui (User Input)	t0, w0, t1, w1, w2, w3, t2, t3
DOCi	t1, t4, w3, w2, w1, t0, t5, t6, w1, t0, t1, t4
DOCj	t4, t3, t7, w0, w3, w4, t2, w0, w1

To compare **DOCi** and **DOCj** to **Ui**, given the domain **Di** in the ontology (the criteria to select the domain won´t be described) we have to construct one table for each comparison, as shown below. **ti** are the terms, **n** is the total number of terms in **Ui**, and **m** is the total number of words in **Ui**:

Table 3. Comparing Ui to DOCi, using domain Di

	t0 X 2	**t1 X 2**	**t4 X 2**	**t5**	**t6**	**n**	**m**
t0	2 X I				T	4	4
t1		2 X I	2 X S	S			
t2	2 X P	2 X S	2 X S	S			
t3		2 X P	2 X T				

Table 4. Comparing Ui to DOCj, using domain Di

	t2	t3	t4	t7	n	m
t0	P				4	4
t1	S	P	S			
t2	I		S			
t3		I	T	P		

The similarity values between documents **Ui** and **DOCi**, and **Ui** and **DOCj**, are given by this formula (defined in [1]):

$$S(U_i, DOC_i) = \frac{W_t \times \sum_{i=1}^{n} T_i + W_p \times \sum_{j=1}^{m} P_j}{W_t \times n + W_p \times m} \tag{1}$$

Where

S(U$_i$,DOC$_i$) is the similarity between the input text ***Ui*** and the law(document) ***DOCi*** in database;

n is the number of terms found in the input text;
m is the number of words found in the input text;
Ti represents the term ***i*** found in the input text;
Pj represents the key word ***j*** found in the input text;
Wt represents the weight given to the set of terms and
Wp represents the weight given to the set of key words.

The value of **Wt** and **Wp** are 2 and 1, respectively (which we have found experimentally as the difference about the influence of terms and isolated words on the results of the search process).

The similarity between **Ui** and **DOCj**, for the same domain **Di**, can be calculated by the same steps. For the domain **Dj**, the relation types inside the comparison tables may follow the connections inside domain **Dj**.

4 Query Expansion Using Universal Words

In this study, we propose a new expansion method which is based on domain ontology and UW´s to achieve initial results on multilingual information retrieval. With this method, we are looking for the ability to expand searches adding dictionaries on new languages, where these new dictionaries are mapped to the UW list.

After the user entered the document to be searched, and this document is converted to a word and term vector, as shown on Item 3 The Indexing Method, the system converts every **term** to its corresponding UW, and for each ontology in a different language associated to this UW list, the system converts the initial terms – inside the indexed document – to the terms on another languages.

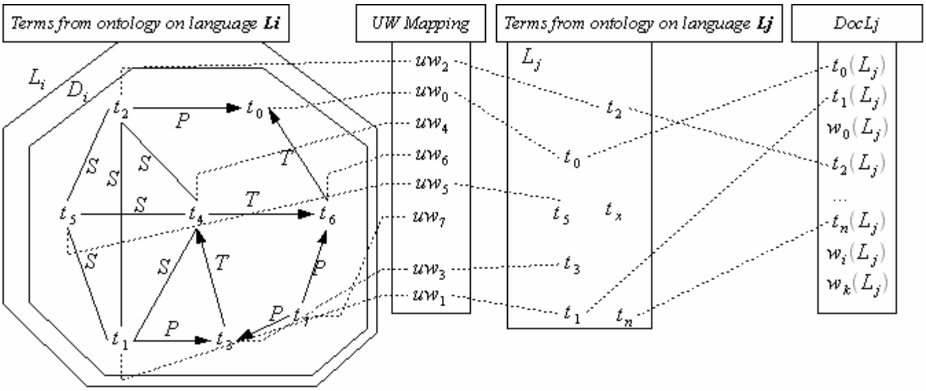


Fig. 5. UW Mapping for query expansion: Each term inside the given domain inside the starting language Li is associated to its corresponding UW, which is, equivalently, associated to another term on another target language

On the example on Figure 5 above, supposing that the terms found on the searched document are **t0** to **t7**, they have its corresponding UW´s from **uw0** to **uw7**. To make possible the comparison of this document and its corresponding terms to a document in another language **Lj**, the system has the UW equivalencies to terms on language **Lj** (**t2**, **t0**, **t5**, **t3** and **t1**). In this initial model, the system does not have the relations between terms in the ontology on the target language **Lj**. So, to calculate the similarity between the document entered by the user (**Ui**), and the document **DOCLj** **–** as made on the sample given from Tables 2 to 4, with **Ui**, **DOCi** and **DOCj** – we convert the terms from language **Li** to language **Lj** as shown on Table 5.

Observe that, for this initial model, for query expansion to another language, words are not taken on similarity calculation. We still don´t know how to treat them, for instance. Also, terms which are not mapped to the UW list are not compared (like **tn** on the sample).

Table 5. Comparison considering the UW equivalencies between two languages

	uw0≡t0(Lj)	**uw1≡t1(Lj)**	**uw2≡t2(Lj)**	**uwn≡tn(Lj)**
t0≡uw0	I		P	
t1≡uw1		I	S	
t2≡uw2	P	S	I	
t3≡uw3		P		

5 The Graphical Ontology Editor

Starting from the fact that ontologies based on the model described in Bueno [1] are composed by terms, relations and domains, and that such components can be described graphically – in this order – like vertexes, edges and envelopes, it was done the project and partial implementation of an experimental ontology network editor, which can offer to the final user a visual interface, allowing to that user to see, create and import another ontology networks from different sources, like a composition made by relationship networks and envelopes (Figure 6).

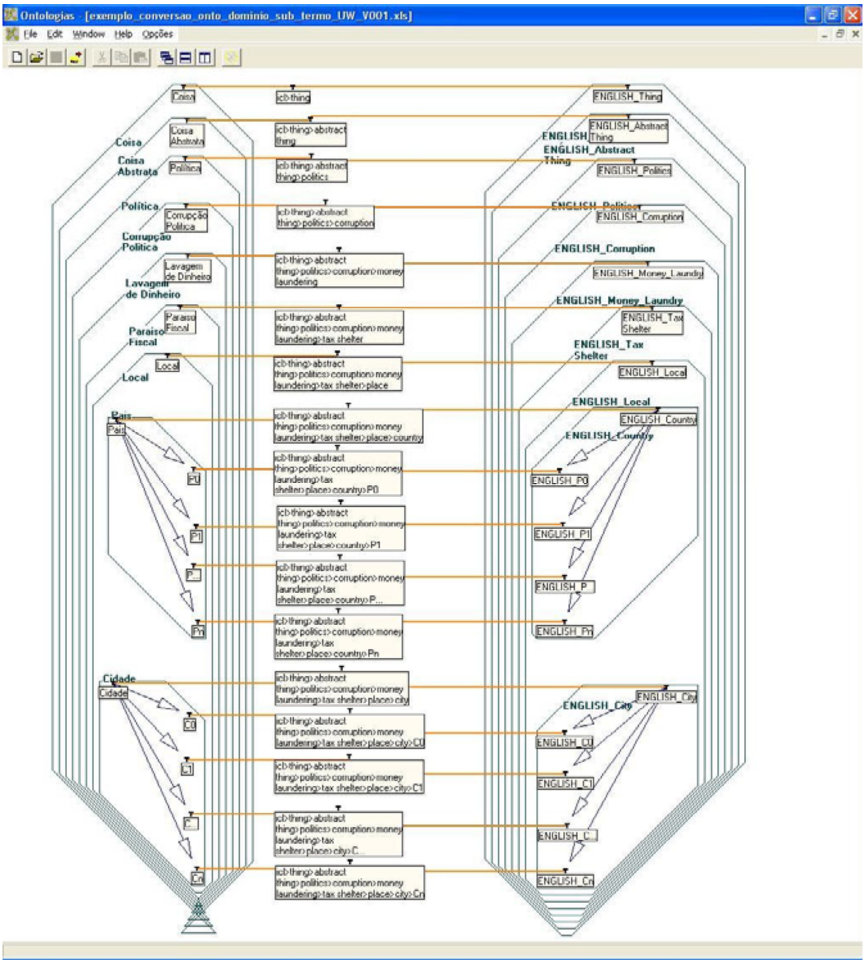


Fig. 6. Sample of two ontology instances, with its terms, relations, UW´s, and domains, represented as envelopes

The same figure shows a test to see how many chained domains (the envelopes around vertexes) can be shown on the graphical editor. Arrowed edges show the relations between terms inside the given domains. Clean edges show relations of equivalence between terms from each domain and UNL universal words. By this way, the knowledge engineer can see and edit these connections to allow the search system do make searches for multilingual documents.

The editor allows updating a centralized ontology database, but it still cannot show changes made by one user to another user in real time. So, it is not collaborative at real time. Due to the fact that vertexes and edges are one thread by element (different from the politics for envelopes, which are managed in limited number per thread), the graphical editor responds too slowly for a big amount of elements on screen, so, there are scale limitations. The last prototype (that on Figure 6) was made using Borland C++ Builder.

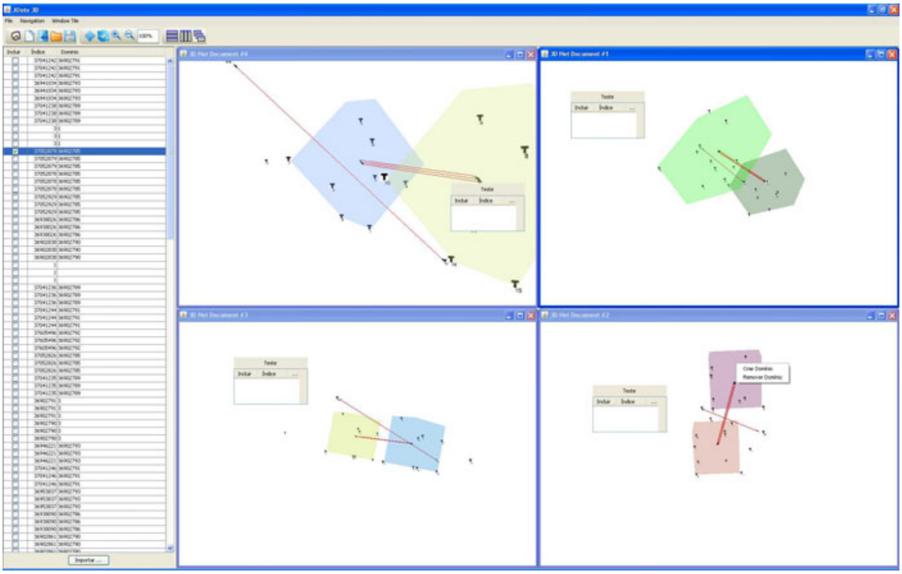


Fig. 7. Prototype of the new ontology editor, on its very early stage, showing terms and domain and volumetric envelops, using Java 3D technology

Actually, we are constructing a prototype of an ontology editor using Java 3D technology, which presented the result shown on Figure 7 above which, for instance, presents a good visualization about relations, terms and domains.

6 Conclusions

The prototype editor made in C++ Builder possess only functions for vision and creation of elements graphically (terms, UW´s, relations between terms, and relations between UW´s and terms, and creation of envelopes which represents domains, appearing like the results found applying the Graham´s Scan algorithm [6], [8], and vertex and edge decorators) and only two ways to importing data from filed sources: Microsoft Excel files and PostgreSQL databases.

The last prototype is made over Java 6 and Java 3D version 1.5.2, and is a MDI (Multiple Document Interface) application (since Java 3D architects resolved the problem caused by conflicts between lightweight and heavyweight Java components). For the database connection is used JDBC (for Postgres databases for the ontology model presented in [1]).

Comparing the two Development Environments, Eclipse with Java 5/6 and Java 3D is better suited for our needs because these have a huge amount of information available on internet, so it was chosen for future applications. The prototype application is a desktop due to the need for some dynamic link libraries to be copied to the Windows\System folder, for Java 3D to work properly, accessing the Graphic Processor Unit.

7 Future Work

The new ontology editor – shown on item 5 above – is being built having the final objective to allow a professional trained on ontology construction (Knowledge Engineer [1], [2]) to construct an ontology network from its elementary components, from any knowledge base.

The next goals of the multilingual and graphical ontology network editor are listed below:

- A lot of ontology data is stored on Microsoft Excel files. The new Java 3D prototype has to allow loading this kind of data from these formats;
- Load / save ontology networks from / to other formats beyond that for Microsoft Excel and PostGres databases;
- Implement some kind of semantic equivalence between UW´s ([4], [12]) and terms from the built ontology, to allow the implementation of inference rules, to make possible the search for constructed concepts contained in ontology networks made in different languages;

Acknowledgments

Our thanks to CNPq (Conselho Nacional de Pesquisa Científica) and the team of Ontojuris Project.: Sonali Bendin, Thiago Paulo Oliveira, Vinicius Mirapalhes, Roberto Lerena, Ana Di Iorio, Bibiana Luz Clara, Jesus Cardeñosa Lera, Carola Canelo and Álvaro Orellana.

References

[1] Bueno, T.C.D.: Engenharia da Mente: Uma Metodologia de Representação do Conhecimento para a Construção de Ontologias em Sistemas Baseados em Conhecimento. Florianópolis: UFSC. Tese (Doutorado em Engenharia de Produção) - Universidade Federal de Santa Catarina, Bueno (2005)

[2] Bueno, T.C.D., et al.: Juris Consulto: Retrieval in Jurisprudencial Text Bases using Juridical Terminology. In: Proceedings of the Seventh International Conference on Artificial Intelligence And Law. ACM, New York (1999), http://doi.acm.org/10.1145/323706.323789

[3] Bueno, T.C.D., Bedin, S.M., Cancellier, F., Hoeschl, H.C.: Evaluating the Effectiveness of a Knowledge Representation Based on Ontology in Ontoweb System. In: Larsen, H.L., Pasi, G., Ortiz-Arroyo, D., Andreasen, T., Christiansen, H. (eds.) FQAS 2006. LNCS (LNAI), vol. 4027, pp. 625–633. Springer, Heidelberg (2006)

[4] Cardeñosa, J., Gallardo, C., Iraola, L., de La Villa, M.A.: A new knowledge representation model to support multilingual ontologies. A case study. In: IKE08 – The 2008 International Conference on Information and Knowledge Engineering, Las Vegas, Nevada, USA, July 14-17 (2008)
[5] Espinoza, M., Perez, A.G., Mena, E.: Enriching an Ontology with Multilingual Information. In: Bechhofer, S., Hauswirth, M., Hoffmann, J., Koubarakis, M. (eds.) ESWC 2008. LNCS, vol. 5021, pp. 333–347. Springer, Heidelberg (2008)
[6] Graham, R.L.: An Efficient Algorithm for Determining the Convex Hull of a Finite Planar Set. Information Processing Letters 1(4), 132–133 (1972)
[7] Guyot, J., Radhouani, S., Falquet, G.: Ontology-Based Multilingual Information Retrieval. In: Working Notes for the CLEF 2005 Workshop, Vienna, Austria, September 21-23 (2005), http://clef-campaign.org/2005/working_notes/workingnotes2005/guyot05.pdf
[8] Heer, J., Card, S.K., Landay, J.A.: Prefuse: A Toolkit for Interactive Information Visualization. In: ACM Human Factors in Computing Systems (CHI), pp. 421–430 (2005), http://doi.acm.org/10.1145/1054972.1055031
[9] Ontojuris Project, http://www.i3g.org.br/ontojuris/sistema.html
[10] Schweighofer, E., Anton, G.: Legal Query Expansion using Ontologies and Relevance Feedback. In: Proceedings of the 2nd Workshop on Legal Ontologies and Artificial Intelligence
[11] Techniques (LOAIT-2007), Stanford University, Stanford, CA, USA (June 4, 2007), http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-321/paper9.pdf
[12] Universal Networking Digital Language Foundation, http://www.undl.org/ (accessed on September 2008)
[13] Joachims, T.: A Probabilistic Analysis of the Rocchio Algorithm with TFIDF for Text Categorization. In: Proceedings of the 14th International Conference on Machine Learning (ICML 1997), Nashville, Tennessee, USA, July 8-12, pp. 143–151. Morgan Kaufmann, San Francisco (1997) ISBN 1-55860-486-3, http://www.cs.cornell.edu/people/tj/publications/joachims_etal_97b.pdf
[14] Clusty (2008), http://clusty.com/ (Acessado em Novembro de)
[15] Valdes-Perez, R.: Introducing Clustering 2.0, http://searchdoneright.com/2008/01/introducing-clustering-2.0/ (accessed November 2008)
[16] Patents Act (1990) - Common Wealth of Australian Law - Attorney-Generalś Department - Australian Government, http://www.comlaw.gov.au/ComLaw/Legislation/ActCompilation1.nsf/0/FB66699B1BE7275BCA2576E300017F27/$file/Patents1990WD02.pdf (accessed April 2001)

Author Index