Karlsruhe Institute of Technology

Community-driven & Work-integrated Creation, Use and Evolution of Ontological Knowledge Structures

Zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaften (Dr. rer. pol)
von der Fakultät für Wirtschaftswissenschaften
des Karlsruher Institut für Technologie (KIT) genehmigte Dissertation

von

Dipl.-Mediensystemwiss. Simone Braun

Tag der mündlichen Prüfung:	30. November 2011
Referent:	Prof. Dr. Rudi Studer
Korreferent:	Prof. Dr. Johannes Magenheim
Prüfer:	Prof. Dr. Andreas Oberweis
Vorsitzender:	Prof. Dr. Thomas Lützkendorf

To my grandmothers

Acknowledgements

First, I would like to thank my advisor Prof. Dr. Rudi Studer for his advice, for providing this great research environment and giving me the freedom to pursue my studies. I also like to thank my co-advisor Prof. Dr. Johannes Magenheim for his encouraging feedback and Prof. Dr. Andreas Oberweis and Prof. Dr. Thomas Lützkendorf for taking the time to review my thesis and attend my doctoral oral examination.

Among all, I would like to give my deepest appreciation to Andreas Abecker who supervised my work, encouraged me to finish it, guided me and helped me to improve it. This thesis would not have been possible without him and his support.

I thank the whole Semantic Karlsruhe research group, my colleagues and students at FZI Forschungszentrum Informatik and all the people I worked with in many research projects for their feedback, their help and collaboration in realizing this work.

I'm particularly grateful to Andreas Schmidt and Valentin Zacharias for many constructive discussions, their valuable feedback and significant help.

My colleagues at FZI not only provided me an environment for research and work but full of life and friendship. I would like to thank them for their patience, their support of any kind and for making all these years such an enjoyable experience. In particular, I want to thank Mark Hefke, Hans-Jörg Happel, Asarnusch Rashid, HeikoMax (HallerVölkel), Andreas Walter, Mercè Müller-Gorchs and Heike Döhmer. My heartfelt thanks to Peter Wolf and Athanasios Mazarakis for being my friends and emotional support especially in the last year and for convincing me to carry on.

Finally, I want to thank René Herrmann for his love, patience and encouragement at each second of this work, my grandmothers inspiring me on earth and from heaven, and my wonderful parents who have been the greatest support in all steps of my life.

It took years to write this thesis. Without all these people it would have taken forever.

Abstract

The thesis aims to support the collaborative development of ontological knowledge structures by communities of knowledge workers in order to facilitate the organization and sharing of information within their domain.

One big challenge for today's organizations and knowledge workers is the focused discovery of new information that is likely to be interesting and useful in order to generate new knowledge. But it is also the organization thus information that had once be found and identified as such can be rediscovered and shared. This is not only about the information itself but also about the people behind who hold the knowledge and e. g., may quickly provide assistance if there are questions. Knowledge about competencies and capabilities of its employees also is an essential need for an organization and its development. This encompasses activities like team staffing or identifying training needs.

Research on ontology-based semantic (web) applications has shown that ontologies are well-suited for organizing and retrieving relevant resources being it people or documents because they connect information resources with machine processable background knowledge. However in practice, ontology-based applications still haven't made their breakthrough. This might be traced back to the high effort and complexity of ontology development. On the other hand, folksonomy-based systems recently have proven to be agile and user-driven approaches for the same application area. They enable their users to collect, manage and share information resources in an easy and lightweight way. However, their lack of semantics also causes a number of problems plaguing tagging and hampering tag-based retrieval.

To that end, this thesis explores how we can combine folksonomy- and ontology-based approaches so that we keep their particular advantages and avoid their disadvantages thus supporting communities of knowledge workers in organizing and maintaining a shared information repository. This is investigated in the application of Social Semantic Bookmarking and Semantic People Tagging.

We present Ontology Maturing as a new perspective and conceptual model for the collaborative development of ontological knowledge structures. It supports (1) the development of a shared understanding, (2) the translation of Web 2.0 approaches to ontology engineering for more active participation, (3) the incremental formalization, (4) application-orientation & work-integration and (5) usable evolving models. To that end, we analyze the advantages and challenges of ontologies and ontology-based knowledge organization systems and make a comparison and consolidation of ontology spectra in literature as well as of ontology development methodologies and tools.

A conceptual design framework complements the ontology maturing model. It supports software developers in deriving and realizing socio-technical systems that scaffold and guide ontology maturing in the application of Social Semantic Tagging for a given organizational setting. It considers technical as well as non-technical aspects. It organizes and provides methods and tools with that end users without modeling expertise can collaboratively organize their information with ontologies and develop the latter one in

a work-integrated way. To that end, we analyze the advantages and challenges of folksonomies and folksonomy-based systems and classify tagging motivations and categories in literature. On this basis, we develop a general definition and model of social semantic tagging and its specializations of social semantic bookmarking and semantic people tagging.

The SOBOLEO framework presents a flexible culture-system-fit framework and reference implementation of the conceptual model and conceptual design framework. It provides a configurable and extensible architecture as well as reusable reference data models for social semantic bookmarking and semantic people tagging and competence ontology maturing. The review of related work shows that SOBOLEO is a pioneer for SKOS editors and the first implementation for semantic people tagging ever.

Following the methodology of design-based research, the model, conceptual design framework and technical framework have been validated and iteratively improved in nine case studies with more than 250 participants involved.

Zusammenfassung

Eine Herausforderung im Alltag von Unternehmen und Wissensarbeitern ist es nicht nur relevante Informationen zu entdecken um daraus neues Wissen zu generieren, sondern auch diese einmal gefundenen und als interessant und nützlich befundenen Informationen so zu organisieren, so dass sie später wiedergefunden und mit anderen ausgetauscht werden können. Hierbei geht es häufig nicht nur um die Informationen selbst, sondern auch um die dahinter stehenden Personen, die über das entsprechende Wissen verfügen und z.B. bei Fragen schnell weiterhelfen können. Das Wissen über Kompetenzen und Fähigkeiten der Mitarbeiter ist auch wesentlich für ein Unternehmen und dessen organisationale Entwicklung, zum Beispiel bei der Zusammenstellung von Teams oder auch der Identifikation von Weiterbildungsbedarfen.

Die Forschung zu ontologiebasierten semantische Anwendungen hat gezeigt, dass Ontologien durch die Verknüpfung von Informationsressourcen mit Hintergrundwissen, das Organisieren und (Wieder-)Auffinden relevanter Ressourcen (wie Dokumente und Personen) sehr gut unterstützen können. In der Praxis jedoch haben ontologiebasierte Anwendungen den Durchbruch auf breiter Ebene noch immer nicht geschafft – was besonders auf den hohen Aufwand und Komplexität der Ontologieentwicklung zurückgeführt werden kann. Folksonomiebasierte Anwendungen hingegen, haben sich in letzter Zeit als agile und nutzergetriebene Ansätze für das gleiche Anwendungsgebiet bewiesen. Sie befähigen ihre Nutzer auf einfache und leichtgewichtige Weise relevante Informationen und Wissen über andere zu sammeln, zu organisieren und auszutauschen. Jedoch führt bei diesen Ansätzen das Fehlen von Semantik auf verschiedenen Wegen u.a. zu einer verringerten Wiederauffindbarkeit.

Diese Dissertation behandelt die Frage wie folksonomiebasierte und ontologiebasierte Ansätze kombiniert werden können, so dass die jeweiligen Nachteile vermieden und die Vorteile erhalten bleiben und auf diese Art und Weise eine Gemeinschaft von Wissensarbeitern in der Organisation und Pflege eines gemeinsamen Informationsbestands (Webressourcen und Personen) unterstützt werden kann.

Mit der Ontologiereifung wird eine neue Sichtweise und konzeptuelles Modell zur kollaborativen Entwicklung ontologischer Wissensstrukturen vorgestellt. Es unterstützt (1) die Entwicklung eines gemeinsamen Verständnisses, (2) die Übertragung von Web 2.0 Ansätzen auf die Ontologieentwicklung für mehr aktive Teilnahme, (3) die inkrementelle Formalisierung, (4) die Anwendungsorientierung und Arbeitsintegration, (5) in der Nutzung sich weiterentwickelnde Modelle. Hierzu werden die Vorteile und Herausforderungen von Ontologien und ontologiebasierten Systemen analysiert sowie Ontologiespektren in der Literatur konsolidiert und Ontologieentwicklungsmethoden und -werkzeuge verglichen.

Ein konzeptuelles Design Framework erweitert das Ontologiereifungsmodell. Es unterstützt Entwickler in der Ableitung und Realisierung eines solchen sozio-technischen Systems für einen bestimmten Organisationskontext unter Berücksichtigung sowohl technischer als auch nicht-technischer Aspekte. Es organisiert und präsentiert Methoden

und Werkzeuge, mit denen Gruppen von Endnutzern ohne Modellierungswissen gemeinsam Informationsbestände mit Hilfe von Ontologien organisieren und dabei arbeitsintegriert weiterentwickeln können. Hierzu werden die Vorteile und Herausforderungen von Folksonomien und folksonomiebasierten Systemen analysiert sowie Tagging-Motivationen und Tagging-Kategorien in der Literatur klassifiziert. Auf dieser Basis wird eine Definition und ein Model für Social Semantic Tagging vorgestellt.

Das technische Framework SOBOLEO schließlich ist eine Referenzimplementierung des Design Framework. Es bietet eine konfigurier- und erweiterbare Architektur sowie wiederverwendbare Referenzdatenmodelle für Social Semantic Bookmarking, Semantic People Tagging und Ontologiereifung. Der Vergleich mit verwandten Arbeiten zeigt, dass SOBOLEO ein Vorreiter für SKOS-Editoren und die erste Implementierung für Semantic People Tagging überhaupt ist.

Das Modell, das konzeptuelle Design Framework sowie das technische Framework wurden der Design Research Methodologie folgend iterativ in neun Fallstudien mit insgesamt über 250 Teilnehmern entwickelt und evaluiert.

Contents

1. Introduction 1

1.1. Motivation and Problem Statement . . . 1

1.1.1. How It's Dealt With . . . 2

1.1.2. Research Question . . . 3

1.2. Approach . . . 4

1.3. Methodology . . . 5

1.4. Contributions . . . 6

1.5. Thesis Overview . . . 7

1.6. Publications . . . 7

I. Foundations 15

Overview . . . 16

2. Folksonomies and Social Tagging Systems 17

2.1. Web 2.0 . . . 17

2.2. Social Software . . . 17

2.3. Enterprise 2.0 . . . 19

2.4. From Tags to Social Tagging . . . 19

2.5. Folksonomies . . . 20

2.5.1. Formal Model of Folksonomy . . . 21

2.5.2. Different Types of Folksonomies . . . 21

2.6. Folksonomies and their Applications . . . 22

2.6.1. System Design and Implications . . . 22

2.6.2. Social Bookmarking . . . 23

2.7. Characteristics of Folksonomy-based Systems . . . 24

2.7.1. Tagging Behavior and Motivations . . . 24

2.7.1.1. User Types . . . 25

2.7.1.2. Categorizing Tagging Motivations . . . 26

2.7.1.3. Impact of Social Presence . . . 30

2.7.2. Semantics of Tags and Tag Assignments . . . 30

2.7.3. Folksonomy and Emerging Vocabulary . . . 35

2.8. Advantages and Challenges of Folksonomies and Social Tagging Systems . 36

2.8.1. Advantages of Folksonomies and Social Tagging Systems . . . 38

2.8.2. Challenges of Folksonomies and Social Tagging Systems . . . 39

2.9. Conclusions . . . 40

3. Ontologies and Ontology-based Knowledge Organization 41

3.1. Notion of Ontology . . . 41

3.1.1. Ontology – Origin . . . 41

3.1.2. Ontology in Computer Science . . . 41

3.1.2.1. Ontology Community and Commitment . . . 42

3.1.3. Dimensions of Ontologies . . . 43

3.1.3.1. Varying Scopes of Ontologies . . . 43

3.1.3.2. Varying Levels of Formality 44
3.1.3.3. Ontology Languages 50
3.1.3.4. Uses of Ontologies 53
3.2. Ontology Engineering 54
3.2.1. Ontology Design Criteria 55
3.2.2. Knowledge Engineer-driven Ontology Development 56
3.2.2.1. Generic Methodology 56
3.2.2.2. Ontology Engineering Methodologies 58
3.2.2.3. Ontology Engineering Tools 63
3.2.2.4. Discussion 66
3.2.3. Collaborative Ontology Engineering 68
3.2.3.1. Collaborative Ontology Engineering Methodologies 68
3.2.3.2. Collaborative Ontology Engineering Tools 75
3.2.3.3. Discussion 77
3.2.4. (Semi-) Automatic Ontology Learning Techniques 77
3.2.4.1. From Folksonomies to Ontologies – Extracting Semantics of Folksonomies 79
3.2.4.2. Discussion 82
3.3. Semantic Annotation 82
3.3.1. Manual Semantic Annotation Approaches 83
3.3.2. Automatic Semantic Annotation Approaches 86
3.3.3. Discussion 87
3.4. Advantages and Challenges of Ontologies & Ontology-based Knowledge Organization 87
3.4.1. Advantages of Ontologies and Ontology-based Knowledge Organization 87
3.4.2. Challenges of Ontologies and Ontology-based Knowledge Organization 89
3.5. Conclusions 90

II. Solution **91**
Overview 92

4. Ontology Maturing **94**
4.1. The Model of Ontology Maturing 95
4.2. Different Contributors in Ontology Maturing 96
4.3. The Artifact, Knowledge and Social Dimensions of Ontology Maturing . . . 97
4.4. Ontology Maturing for Competence Ontologies 98
4.4.1. Phases of Competence Ontology Maturing 99
4.4.2. Variable Level of Formality in Competence Ontology Maturing . . . 101

5. Conceptual Design Framework **103**
5.1. Scaffolding Ontology Maturing 103
5.1.1. Seeding, Evolutionary Growth and Reseeding 103
5.1.2. Ontology Maturing and SER 104
5.1.3. Scaffolding Methods and Functions 105
5.1.3.1. Nudge, Agree, and Aggregate 105
5.1.3.2. Categorizing Methods and Functions 105
5.2. Social Semantic Tagging 109
5.2.1. Social Semantic Tagging Model 110

5.3. Design Framework . . . 110
5.3.1. General Design Options . . . 110
5.3.1.1. Resources . . . 111
5.3.1.2. Ontology . . . 114
5.3.1.3. Tag Assignments . . . 117
5.3.2. Design Options for Semantic People Tagging Systems . . . 120
5.3.3. Design Options for Social Semantic Bookmarking . . . 121
5.3.4. Search Heuristics Using Social Semantic Tags . . . 122
5.3.4.1. Semantic People Tagging Specific Aspects . . . 122
5.4. Concluding Remarks . . . 123

6. Technical Framework **124**
6.1. Overview of the Technical Framework . . . 124
6.2. Architecture . . . 125
6.2.1. Implementation Frameworks . . . 126
6.2.2. Event Bus Communication . . . 128
6.2.3. Configuration . . . 129
6.3. Data Models . . . 129
6.4. Ontology Editor and Management . . . 136
6.4.1. Variable Level of Formality in Competence Ontology Maturing . . . 137
6.4.1.1. Implicit Annotations . . . 138
6.4.2. Ontology Gardening . . . 141
6.4.2.1. SKOS Analysis . . . 142
6.4.2.2. Concept Relationship Analysis . . . 143
6.4.2.3. Activity Analysis . . . 143
6.5. Dialog and Discussion . . . 144
6.6. Social Semantic Bookmarking . . . 147
6.6.1. Bookmarking and Annotating Documents . . . 147
6.6.2. Searching for Documents . . . 149
6.6.2.1. Query Refinements & Relaxations . . . 151
6.6.3. Ontology-based Navigation through the Shared Document & Dialog Space . . . 151
6.7. Semantic People Tagging . . . 153
6.7.1. Bookmarking and Annotating People . . . 153
6.7.2. People Directory and Profiles . . . 154
6.7.2.1. People Directory . . . 154
6.7.2.2. People Profiles . . . 154
6.7.3. Searching for People . . . 157
6.7.3.1. People Search and Ranking Algorithm . . . 157
6.7.4. Ontology-based Navigation through the Annotated People Space . . . 158
6.7.5. Expertise Analytics . . . 158
6.8. Supporting Ontology Maturing with SOBOLEO . . . 159
6.8.1. In the Application of Social Semantic Bookmarking . . . 160
6.8.2. In the Application of Semantic People Tagging . . . 161
6.9. Summary . . . 165

III. Usage and Assessment **169**
Overview . . . 170

7. Social Semantic Bookmarking 172
7.1. Case Study I: The Collaborative Knowledge Construction Challenge 172
7.1.1. Procedure 172
7.1.2. Results 172
7.1.3. Discussion 173
7.2. Case Study II: The Rapid Prototyping Research Community 174
7.2.1. Procedure User Tests 1 175
7.2.2. Results User Tests 1 177
7.2.3. Procedure Participatory Design and User Tests 2 181
7.2.4. Results Participatory Design and User Tests 2 182
7.2.5. Discussion 184
7.3. Case Study III: Workshop Reflections on Knowledge Modeling 185
7.3.1. Procedure 186
7.3.2. Results 187
7.3.3. Discussion 189
7.4. Case Study IV: Course Experiments for the Coll. Dev. of Knowledge Structures 189
7.4.1. Procedure 189
7.4.2. Results 190
7.4.2.1. Results of Sub-Experiment 1 190
7.4.2.2. Results of Sub-Experiment 2 192
7.4.2.3. Results of Sub-Experiment 3 192
7.4.3. Discussion 193
7.5. Case Study V: Course Experimental Study in Coll. Tagging & Sensemaking 193
7.5.1. Procedure 194
7.5.2. Results 195
7.5.3. Discussion 196
7.6. Case Study VI: The Classic Roman Civil Engineering Community of Practice 196
7.6.1. Procedure Phase 1 Formative Evaluation and Participatory Design Activities 197
7.6.2. Results Phase 1 Formative Evaluation and Participatory Design Activities 198
7.6.3. Procedure Phase 2 Formative Evaluation 200
7.6.4. Results Phase 2 Formative Evaluation 201
7.6.4.1. Results of the Translation and Case-testing 201
7.6.4.2. Results of the User Experience 203
7.6.4.3. Results from the Facilitators 206
7.6.4.4. Results of Interviews 206
7.6.5. Discussion 208
7.7. Conclusions 210

8. Semantic People Tagging 211
8.1. Case Study VII: People Tagging Field Experiments 211
8.1.1. Procedure 211
8.1.1.1. Materials and Instruments 212
8.1.2. Results 214
8.1.3. Discussion 215
8.2. Case Study VIII: HR Experts Focus Group 217
8.2.1. Procedure 217
8.2.2. Results 218
8.2.3. Discussion 219

8.3. Case Study IX: British Career Guidance 219
8.3.1. Procedure Phase 1 Formative Evaluation and Participatory Design Activities 221
8.3.2. Results Phase 1 Formative Evaluation and Participatory Design Activities 221
8.3.3. Procedure Phase 2 Formative Evaluation 222
8.3.3.1. Materials and Instruments 223
8.3.3.2. Introduction Workshop I 225
8.3.3.3. Introduction Workshop II 225
8.3.3.4. User Feedback Workshop III 227
8.3.4. Results Phase 2 Formative Evaluation 227
8.3.4.1. Results from Introduction Workshop I 227
8.3.4.2. Design and Developments between Introduction Workshop 1 and Workshop 2 231
8.3.4.3. Results from Introduction Workshop II 231
8.3.4.4. Results User Feedback Workshop III 234
8.3.5. Discussion 237
8.4. Conclusions 238

IV. Conclusions **241**

9. Related Work **242**
9.1. Social Semantic Software 242
9.1.1. Social Semantic Wikis 244
9.1.2. Social Semantic Bookmarking 247
9.1.2.1. Comparison 249
9.2. Competence Management 255
9.2.1. Competence Management in General 255
9.2.2. Getting Competency Profiles 256
9.2.3. Social Tagging for Gathering Competence Information 256
9.3. Methods and Tools Related to Ontology Maturing 258
9.3.1. Collaborative and Community-driven Ontology Construction 258
9.3.2. Gardening as Scaffolding Method 260
9.3.3. Collaborative SKOS-Editors 262

10. Summary and Outlook **264**
10.1. Contributions and Impact 264
10.2. Future Research 266

V. Appendix **269**

A. Technical Details **270**
A.1. Data Model Notation 270

B. Case Study II: The Rapid Prototyping Research Community **271**
B.1. User Tests 1 271

C. Case Study VI: The Classic Roman Civil Engineering Community of Practice **273**
C.1. User Experience 273
C.2. Participants Interviews 275

D. Case Study IX: British Career Guidance 279
D.1. Task-oriented Guide for Introduction Workshops . . . 279
D.2. Pre-Questionnaire . . . 282
D.3. Post-Questionnaire . . . 284
D.4. Results Introduction Workshop I . . . 290
D.4.1. Pre-Questionnaire Results . . . 290
D.4.2. Post-Questionnaire Results . . . 291
D.4.3. Technical Feedback . . . 293
D.5. Results Introduction Workshop II . . . 294
D.5.1. Pre-Questionnaire Results . . . 294
D.5.2. Post-Questionnaire Results . . . 295
D.5.3. General Feedback . . . 297
D.5.4. Technical Feedback . . . 299
D.6. Results User Feedback Workshop III . . . 299
D.6.1. Open Discussion . . . 299
D.6.2. Aspects of Design Framework . . . 300

Bibliography 303

List of Tables

1.1. Outputs of Design Research by Vaishnavi and Kuechler (2004) 6

2.1. Tagging motivations by Ames and Naaman (2007) and adapted by Panke and Gaiser (2009) 27
2.2. Tagging motivations by Marlow et al. (2006) 28
2.3. Tagging motivations by Zollers (2007) 28
2.4. Tagging motivations by Wash and Rader (2007) 28
2.5. Tagging motivations by Heckner et al. (2009) 29
2.6. Tagging motivations by Nov et al. (2009) 29
2.7. Tagging motivations by Mirzaee and Iverson (2009) 29
2.8. Tag Categories 1 33
2.9. Tag Categories 2 34

3.1. Comparing knowledge engineer driven ontology development approaches . 67
3.2. Comparing collaborative ontology development approaches 78

6.1. SOBOLEO framework configuration 130
6.2. SKOS Concept Model Details 132
6.3. Competences Model Details 133
6.4. User Model Details 134
6.5. Dialog Moves & Openers 147

7.1. Statistical data overview of the Collaborative Knowledge Construction challenge 173
7.2. Statistical data overview of the user test 178
7.3. Results from the interview on the users' overall impression 180
7.4. Statistical data overview of the workshop "Reflections on Knowledge Modeling as a Maturing and Learning Process" 188
7.5. Statistical data overview of sub-experiment 1 191

8.1. Statistical data overview of both field experiments in comparison 214
8.2. Survey data overview of both field studies in comparison 216
8.3. Workshop I: Results on user satisfaction 230
8.4. Workshop II: Results on user satisfaction 233
8.5. Personal statements by the workshop participants 235

9.1. Comparing Social Semantic Bookmarking Systems regarding Users 250
9.2. Comparing Social Semantic Bookmarking Systems regarding Resources . . 251
9.3. Comparing Social Semantic Bookmarking Systems regarding Ontology . . . 252
9.4. Comparing Social Semantic Bookmarking Systems regarding Tag Assignments 253
9.5. Comparing Social Semantic Bookmarking Systems regarding Scaffolding Methods 254

B.1. Results from observations and screen recording 272

C.1. General feedback on the experience 276
C.2. Specific feedback on the experience 277
C.3. Specific feedback on the experience (continued) 278

List of Figures

1.1. Reasoning on Design Cycle by Takeda et al. (1990) 5
1.2. Three pillars of contributions 8

2.1. Social Software Triangle by Koch and Richter (2008) 18
2.2. Web 2.0 & Social Software Classification by Peters (2009) 18
2.3. FLATNESSES framework by Hinchcliffe (2007) 20
2.4. Tag Category Model overview and Functional Category Model by Heckner et al. (2008a) 31
2.5. Relationship btw. community influence and personal tendency by Sen (2009) 37

3.1. Ontology Spectrum by Guarino et al. (2009) 47
3.2. Ontology Spectrum by McGuinness (2003) 47
3.3. Ontology Spectrum by Smith and Welty (2001) 47
3.4. Ontology Spectrum by Obrst (2003) 48
3.5. Ontology Spectrum by Weller (2007) 48
3.6. Mapping of criteria to the five main principles. 56
3.7. Ontology Engineering Activities adapted from Simperl and Tempich (2006) 57
3.8. Techniques and resources and their type of contributions to derive ontologies from folksonomies 80
3.9. Solution overview 93

4.1. The four phases of the ontology maturing model 95
4.2. Relating the involvement of different types of contributors to the four phases of the ontology maturing model 97
4.3. Ontology Maturing for Competence Ontologies 100
4.4. Core competence model adapted from Schmidt and Kunzmann (2006) . . . 102

5.1. Categorizing scaffolding methods and functions 106
5.2. Design Aspects for Users 111
5.3. Design Aspects for Resources 113
5.4. Design Aspects for the Ontology 116
5.5. Design Aspects for Tag Assignments 119

6.1. Overview of the technical framework 125
6.2. SOBOLEO architecture 127
6.3. Document tagging model 131
6.4. Person tagging model 131
6.5. Competence ontology model 132
6.6. Tagged Person Model Details 134
6.7. Person Annotation Model Details 134
6.8. Tagged Document Model Details 135
6.9. Document Annotation Model Details 135
6.10. Rating Model Details 135
6.11. SOBOLEO Ontology Editor 136
6.12. SOBOLEO competence ontology editor 139

6.13. SOBOLEO ontology editor with gardening recommendations 142
6.14. An ontology maturing dialog about the concept 'aqueduct'. 145
6.15. Detailed transcript of a maturing dialog about the concept 'aqueduct'. . . . 146
6.16. SOBOLEO annotation tool with different types of tag recommendations. . . 148
6.17. SOBOLEO page displaying results of the semantic search 150
6.18. Browsing web documents. 152
6.19. Annotate People interface. 154
6.20. People Directory interface. 155
6.21. Profile of an individual person. 156
6.22. Guidance overview 159
6.23. Bookmarking and annotating a resource in SOBOLEO 160
6.24. Starting an ontology maturing dialog about a concept 161
6.25. From a tag idea to the formalized concept 162
6.26. Adding a new person to the collaboration space. 163
6.27. Timeline of case studies 170

7.1. Questionnaire results on task appropriateness. 177
7.1. Questionnaire results on learnability. 179
7.1. Questionnaire results on satisfaction. 179
7.2. Questionnaire results of remote user tests. 183
7.3. Statistical Results. 184
7.4. Road map Phase 2 Formative Evaluation 202
7.5. Case Study VI: Key topics and web documents created during the experience 204
7.6. Case Study VI: Concepts created during the experience with the SOBOLEO editor 205

8.1. Case Study VII: Paper-based poster used in field experiments 212
8.2. Case Study VII: People Tagging Setting of Field Experiment I 213
8.3. County Northumberland 220
8.4. Introduction Workshop II with personal career advisers. 226
8.5. Workshop I: Results on motivation and barriers 229
8.6. Workshop II: Results on motivation and barriers 236

9.1. Integrated Model of Competence Management 255

A.1. Data model notation. 270

C.1. Example Dialog in Detail 274

Third party graphics have been redrawn for the purpose of a common thesis design

The First approached the Elephant,
And happening to fall
Against his broad and sturdy side,
At once began to bawl:
"God bless me! but the Elephant
Is very like a wall!"

The Second, feeling of the tusk
Cried, "Ho! what have we here,
So very round and smooth and sharp?
To me 'tis mighty clear
This wonder of an Elephant
Is very like a spear!"

The Third approached the animal,
And happening to take
The squirming trunk within his hands,
Thus boldly up he spake:
"I see," quoth he, "the Elephant
Is very like a snake!"

(John Godfrey Saxe 1873[1])

1. Introduction

1.1. Motivation and Problem Statement

Living in the transition towards a knowledge-based society as it is postulated by the European Union (European Council, 2000), it is not surprising that organizations and individuals rely on their knowledge to successfully compete and persist in a knowledge-driven business environment. With the Lisbon Process the European Commission aimed to become "*the most dynamic and competitive knowledge-based economy in the world*" in the last decade (European Council, 2000). An ambitious goal that has not been achieved yet and is continued in Europe 2020[2]. This also shows that there is still a high potential to exploit for knowledge management, recognized as an important method to improve the capabilities of organizations.

One big challenge for today's organizations and knowledge workers is how to "*deal with the information deluge [and] pluck the diamond from the waste*" to create new knowledge (Economist, 2010). Therefore it is not only a question on the focused discovery of new information that is likely to be interesting and useful. It is also the organization of the information so that information that had once be found and identified can be rediscovered and shared.

Similarly challenging but essential is the knowing on who knows what (Groth, 2004). Knowledge about competencies and capabilities of its employees is an essential need for an organization and its development. This encompasses activities like finding the right person to contact, team staffing or identifying training needs.

We would like to illustrate both of these challenges with two exemplary application scenarios that will accompany us through this thesis.

[2] http://ec.europa.eu/europe2020

Scenario I: The Rapid Prototyping Research Community Let's take the case of a research community of applied science in the area of rapid prototyping bringing together experts from various disciplines like plastics, ceramics, and mechanical engineering (see also Section 7.2). Even though research, there are similar information and knowledge deficiencies like in industry and market globalization leading to increased competitive pressure, speed of innovation and shortened product cycles. In the area of plastics and their market these high dynamics are particularly obvious. New materials or new forms of existing ones frequently enter the market; brand names and manufacturers are permanently changing and hardly trackable; attributes of a chemical substance retrievable using its brand name today, are very hard to find once it is sold under a different label. Because there is no general up to date database listing manufacturer and brand names of currently available forms of plastics the users rely on the Internet and search engines like Google. However conventional tools lack the focus on the users' domain. At this point, using annotation and retrieval tools might help; e. g., when a colleague already found the new brand name of a product and tagged it with the old one you are looking for. Semantically enriched, it is possible to extend or refine the search in order to reduce irrelevant results and to guide the user.

Scenario II: Career Guidance in Northern England Looking at the case of a British career guidance organization, it is highly content dependent and rapidly changing due to the dependence on Labour Market Information (see also Section 8.3). Because of the geographical distribution, the career advisers' knowledge about the specialties and expertise of their colleagues across the offices is very limited and finding the right colleague to talk to is difficult. Typically, employee directories, which simply list staff and their areas of expertise, are insufficient. One reason is that information contained in the directories is outdated; or it is not described in an appropriate manner; or it focuses too much on 'experts'; and they often do not include external contacts (cf. Biesalski and Abecker, 2005; Schmidt and Kunzmann, 2007). Also Human Resource development needs to have sufficient information about the needs and current capabilities of current employees to make the right decisions. In service delivery contexts that must be responsive to the changing needs of clients, like career advising services, it is necessary to establish precisely what additional skills and competencies are required to keep up with new developments. Thus the HR development wishes to get a better overview on dynamics, especially new emerging topics.

1.1.1. How It's Dealt With

The first wave of ontology-based semantic (web) applications has shown that ontologies are well-suited for sophisticated ways of organizing and retrieving relevant resources being it people or documents because they connect information resources with machine processable background knowledge. Through this background knowledge, representing the implicit user context, applications may be able to better understand and fulfill the users' needs. However in practice, ontology-based applications still haven't made their breakthrough. This might be traced back to the high effort and complexity of ontology development. Additionally, traditional ontology engineering methodologies and tools suffer from the underlying assumption that a few modeling experts have to create an ontology for many users. Thus, development of an ontology and its usage in applications are separated: Knowledge engineers create and maintain the ontologies together with the domain experts in advance.

In order to keep the ontology in line with the intended usage, cumbersome procedures are introduced that lead to delayed and often error-prone updates to the ontology (cf. Barker et al., 2004; Hepp et al., 2006; Hepp, 2007): Knowledge engineers often have only a limited understanding of the domain and thereby the resulting ontologies are inaccurate, incomplete or incomprehensible for the community. Instead of having users making changes to the ontology by themselves, new requirements need to be collected from the users and understood by the knowledge engineers; in turn, changes to the ontology need to be explained to the users. This time consuming process often results in less frequent updates to the ontology, so that needed ontology elements, e. g., for annotating currently relevant resources, are not available in time; in particular for fast moving domains. This leads to questionable relevancy of the ontologies for the actual purpose and thus at best to unsustainable success. The users are frustrated and discouraged to further contribute to the ontology development process. Thus, it is hardly surprising that in practice ontology engineering principles are often ignored (Cardoso, 2007).

Additionally, ontologies are acknowledged as shared and common understanding of a domain (Decker et al., 2000). Likewise Gómez-Pérez et al. (2004) state that "*the model can only be considered an ontology if it is a shared and consensual knowledge model agreed by community*". However, it is not about *"what ontologies are but how they become shared formal specifications of a domain"* (Leenheer 2009). So far, it is neglected that such a shared understanding also has to be developed among the stakeholders and developing a shared understanding definitely is a learning process that requires active participation (O'Keeffe, 2002; Allert et al., 2006). Additionally, shared conceptual models are never static but constantly evolving especially in fast changing domains like rapid prototyping. Indeed, if we have a closer look, in particular at knowledge work processes, we discover that the users are almost constantly constructing and negotiating shared meaning in collaboration and social interaction with others by augmenting and evolving a community vocabulary.

For a while, folksonomy-based systems have enjoyed great popularity for organizing and retrieving relevant resources. Folksonomy-based systems are agile and user-driven approaches that enable their users to collect, manage and share information resources in an easy and lightweight fashion (Peters, 2009). These systems make use of tags to facilitate the organization, navigation and searching. Tags are arbitrary keywords that are used by the users to further describe the information resources in order to aid their retrieval. Additionally, these systems make collecting and organizing information resources a social experience by allowing the users to share their resources and tags with others. The users can use other users' tags, see which tags and annotated resources they have in common, or what is annotated with the same tags. In this way, they can find people with similar interests and discover new interesting resources. The popularity and high participation of such folksonomy-based systems have shown that this organizing principle with tags and folksonomies evolving from these is much easier accessible for users than structured and controlled vocabularies; in particular for collaborative applications. However, this missing structure is also the root cause for a number of problems plaguing tagging and hampering tag-based retrieval and accessibility of shared understanding: problems such as spelling variants, tags on different levels of abstraction, homonyms or synonyms (see Section 2.8.2 for a detailed discussion).

1.1.2. Research Question

Starting from the presented exemplary application scenarios and the shortcomings of existing ontology-based technologies, we have derived the following research question:

> How to support a community of knowledge workers in the collaborative development of ontological knowledge structures that are used for organizing and sharing information resources.

This broad research question entails several smaller sub-research questions:

- **Model:** What is a suitable conceptual model to describe and understand collaborative ontology development.
- **Formalism:** What are appropriate formalisms and knowledge structures to capture the information about documents as well as who knows what in this setting.
- **Tool and Organizational Environment:** What are the needed tools and organizational processes in order for the collaborative development of ontological knowledge structures to work.
- **Feasibility of End User-driven Ontology Engineering:** Differing from the traditional model of knowledge engineer-driven ontology development it has to be shown that groups of knowledge workers (non-modeling experts) can collaboratively create and maintain an ontology.

1.2. Approach

To tackle this problem we suggest a collaborative ontology development approach that considers the following aspects:

- **Shared Understanding:** We have to conceive ontology development not only as eliciting and formalizing a shared understanding of all stakeholders but rather as a social and collaborative learning process within a community of users where the shared model emerges and is formalized at the same time (a perspective motivated by constructivist views on learning, cf. Allert et al. 2006). The involved individuals incrementally deepen their understanding of the real world and of an (appropriate) vocabulary to describe it.
- **Active Participation:** This requires active participation. Active participation of all stakeholders not only distributes the effort of ontology development but also fosters the acceptance and up-to-dateness of the ontology. Folksonomy-based approaches have shown how to empower the individual to take part in community activities by lowering the barriers. Therefore, we want to have a look at how we can translate the Web 2.0 phenomenon into the area of ontology engineering.
- **Variable Level of Formality:** For more participation we need, however, ontology structures users can understand and deal with. High formality is counterproductive (cf. Hepp, 2007) and we need to start with lower formality levels. Therefore, we want to have a look at how we can support smooth and continuous transitions between the two worlds of folksonomies and ontologies.
- **Application-orientation & Work Integration:** Ontology building is not the primary activity of most users and vocabularies emerge in their daily work (often implicitly). Thus to foster motivation, we need a quick, simple and work-integrated way to engage in ontology development activities; i. e. , when using ontologies (e. g., for annotation or navigation). A work-integrated approach may also improve the acceptance and appropriateness of an ontology for the task at hand.

- **Usable Evolutionary Models:** Ontological structures are not only needed for the later use, i. e. knowledge and information organization and sharing, but also during the development of a shared understanding and the vocabulary itself because they facilitate their accessibility, e. g., in identifying and understanding different abstraction levels. At this point, folksonomies are not sufficient. This also requires to make the still evolving models usable.

1.3. Methodology

The approach follows the Design (Science) Research methodology (Takeda et al., 1990; Vaishnavi and Kuechler, 2004; Hevner et al., 2004) that *"involves the analysis of the use and performance of designed artifacts to understand, explain and very frequently to improve on the behavior of aspects of Information Systems"* (Vaishnavi and Kuechler 2004). Figure 1.1 illustrates the general process of a typical design science research approach (Takeda et al., 1990): it starts with the *Awareness of a Problem* that is followed by the *Suggestion* phase, in which possible problem solutions are found based on existing knowledge. In the *Development* phase we implement the designed artifact according to the problem solution. The implemented artifact is then *Evaluated* according to implicit or explicit criteria from the Awareness of Problem phase. Suggestions, Development and Evaluation are frequently iterative because results from the Evaluation and additional information from the Development generate new knowledge that initiates another Suggestion round. This is indicated by the *Circumscription* and *Operation of knowledge & Goal* arrow. Circumscription is of particular importance as *"it generates understanding that could only be gained from the specific act of construction"* (Vaishnavi and Kuechler 2004). The process terminates with the *Conclusion* when the results are decided to be "good enough" meeting the criteria for adequacy. The newly gained knowledge may be applied in a new design research effort.

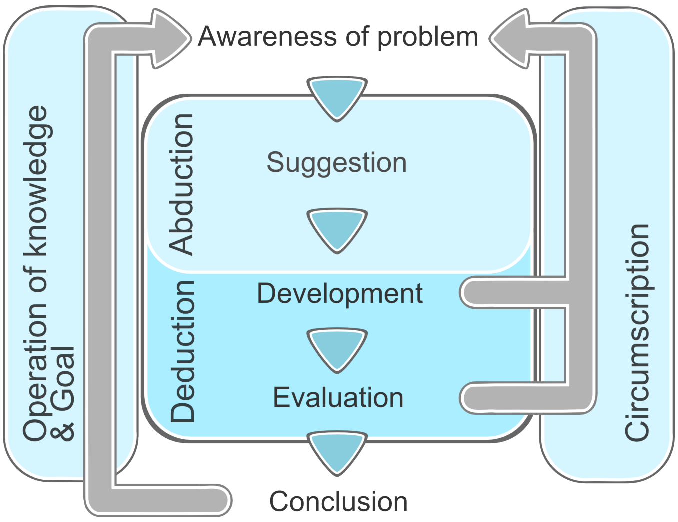


Figure 1.1.: Reasoning on Design Cycle by Takeda et al. (1990)

From a Design Research effort, we can obtain five outputs summarized in Table 1.1:

Constructs provide the vocabulary to define and communicate problems and solutions. They emerge when the problem is conceptualized and are refined during the design process. Models represent a real world situation by using design problem and solution

Table 1.1.: Outputs of Design Research by Vaishnavi and Kuechler (2004)

Output	Description
Constructs	The conceptual vocabulary of a domain
Models	A set of propositions or statements expressing relationships between constructs
Methods	A set of steps used to perform a task – how-to knowledge
Instantiations	The operationalization of constructs, models and methods
Better theories	Artifact construction as analogous to experimental natural science

statements. Methods provide guidance and define steps how to perform a task and solve problems. Instantiations operationalize constructs, models and methods and show that they are implementable in a real environment. Instantiations also help to learn about *"the real world, how the artifact affects it, and how users appropriate it"* (Hevner et al. 2004). The artifact construction, which can be an experimental proof and/or exploration of method, and the exposure of the relationships of the artifact elements during the construction and evaluation (in this way increasing the understanding of the elements) help to gain better theories.

Our approach puts emphasis on the iterative cycle of design, implementation, evaluation and redesign that is grounded in real-world contexts with social interaction and collaboration with practitioners in the style of design-based research[3]. Therefore, we made essential use of participatory design methods to evaluate the suitability of the technical framework for its purpose and to validate the conceptual models. These were based on software and paper-based prototypes.

Participatory design methods aim to actively involve users in the design process, i. e. not only in interviews and questionnaires, and in this way *"give the end users a voice [...] thus enhancing the quality of the resulting system"* (Bødker et al. 2000) and so *"the ultimate users of the software make effective contributions that reflect their own perspectives and needs, somewhere in the design and development lifecycle of the software"* (Muller et al. 1997) . This not only improves the quality of the software design and its development but also the acceptance by the end users.

1.4. Contributions

Our approach is based on three pillars (see fig. 1.2):

- The first pillar provides a methodological framework with a sound and comprehensive conceptual model that describes the process of ontology development with the metaphor of "*maturing*". It structures this maturing into four characteristic phases, ranging from emergence of ideas, consolidation in communities via formalization up to axiomatization. This ontology maturing model is complemented by a conceptual socio-technical design framework for social semantic bookmarking and

[3] Design-based research is the further development of design science paradigms in education, especially for designing learning environments, which Wang and Hannafin (2005) define as a *"methodology aimed to improve educational practices through systematic, flexible, and iterative review, analysis, design, development, and implementation, based upon collaboration among researchers and practitioners in real-world settings, and leading to contextually-sensitive design principles and theories."*

semantic people tagging; i. e. the application and use of the ontology for organizing and sharing web resources respectively the knowing on who knows what.

- The second pillar provides a technical framework with its instantiations based on the conceptual process model and design frameworks that enable domain experts to construct and maintain ontologies in a collaborative and continuous way, integrated into the usage of these ontologies within their daily work that is organizing and sharing web resources and knowledge about people.
- The third pillar provides a framework of empirical application-oriented insights gained from ten case studies; six of them instantiating the social semantic bookmarking case respectively three of them the semantic people tagging case.

1.5. Thesis Overview

This thesis is structured in four parts. The first part provides the theoretical foundations for this thesis and aims to raise the awareness of the problem. Chapter 2 starts with an introduction and discussion of advantages and problems of folksonomies and social tagging systems that helps us in how we can translate the Web 2.0 phenomenon into the area of ontology engineering. In Chapter 3 we will introduce ontologies and ontology-based knowledge organization systems and discuss in detail ontology engineering methods and methodologies before we span the problem space in a concluding discussion.

The second part of the thesis presents the solution. We will present the methodological framework as the suggestion of problem solution according to design science research. We illustrate the ontology maturing theory in Chapter 4. Chapter 5 first describes methods to scaffold and guide ontology maturing and then the socio-technical design framework for the application of social semantic bookmarking and semantic people tagging. We detail their implementation by the SOBOLEO technical framework in Chapter 6.

In the third part of the thesis we describe the usage and assessment. We will give an overview on nine case studies conducted during the design and development process in Chapter III and present gained empirical and application-oriented insights in Chapter 7 & 8.

The last part concludes the thesis with discussing related work in Chapter 9 and with a summary and outlook on future work and research directions in Chapter 10.

1.6. Publications

Parts of this thesis' contents have been published previously and have been updated to adhere to the thesis' common framework. The the German national funded research project "Im Wissensnetz – Vernetzte Informationsprozesse in Forschungsverbünden"[4] and the European Integrated Project MATURE[5] have been serving as background and case studies of this thesis and core contents have been published within their context. Especially, results of the implementation (Chapter 6) and evaluation (Chapter 7 and 8 have also been detailed in Braun et al. (2008); Hefke et al. (2009); Braun et al. (2009a);

[4] http://www.im-wissensnetz.de/

[5] http://mature-ip.eu

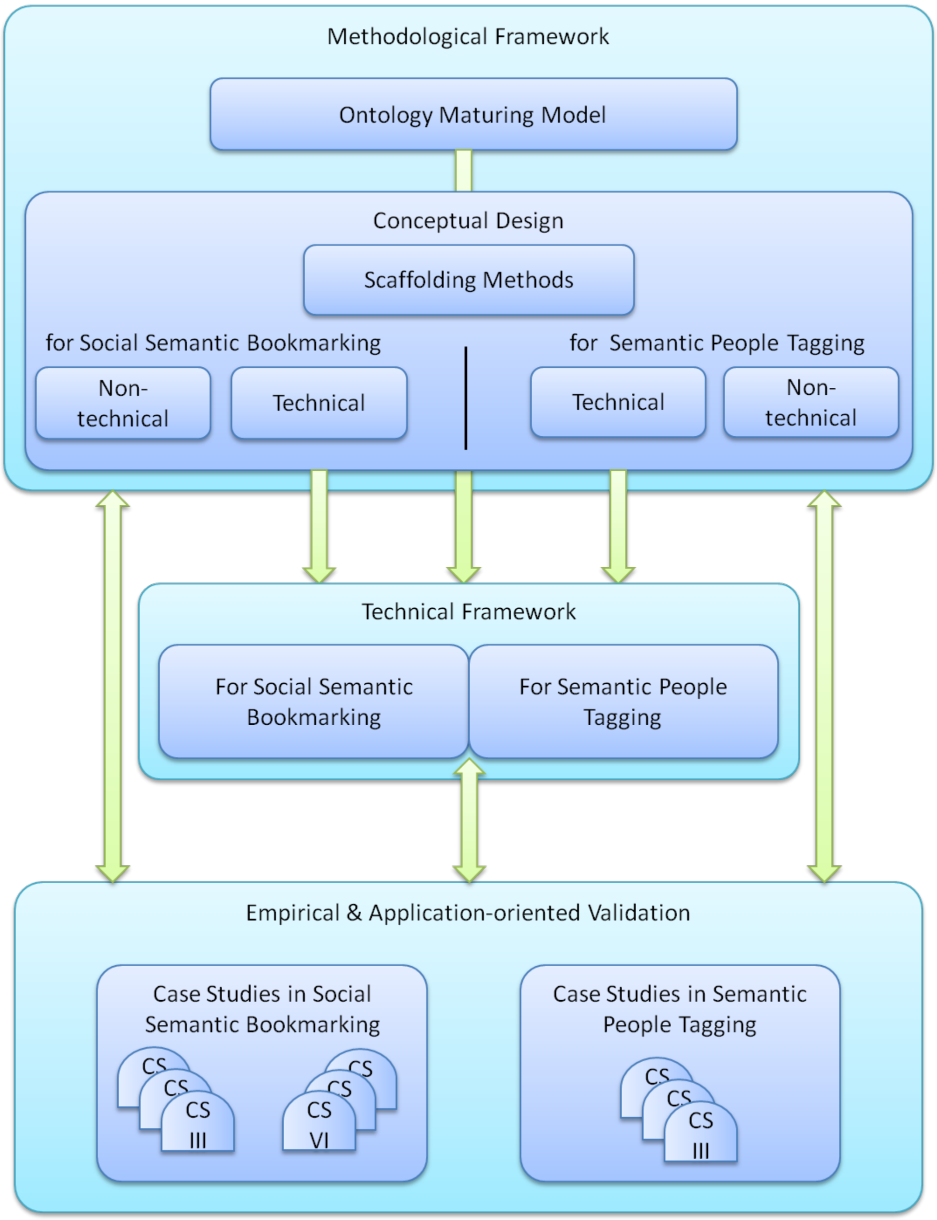


Figure 1.2.: Three pillars of contributions

Witschel et al. (2009); Bradley et al. (2010); Nelkner et al. (2011); Kump et al. (2011); Ravenscroft et al. (2010a).

In the following we detail the relation to further publications of the author that have already been peer-reviewed.

The approach of ontology maturing (Chapter 4) as a novel concept for collaborative ontology development together with a first tool support with SOBOLEO and first evaluation results have been presented in:

- Braun, S., Schmidt, A., Walter, A., Zacharias, V. (2008): Von Tags zu semantischen Beziehungen: kollaborative Ontologiereifung. In: *"Good Tags and Bad Tags – "Social Tagging in der Wissensorganisation"*, Waxmann, 2008, pp. 163-173
- Braun, S., Schmidt, A., Walter, A., Zacharias, V. (2008): Using the Ontology Maturing Process Model for Searching, Managing and Retrieving Resources with Semantic Technologies. In: *OnTheMove Federated Conferences 2008 (DAO, COOP, GADA, ODBASE)*, Monterrey, Mexico, Lecture Notes in Computer Science vol. 5332, Springer, 2008, pp. 1568-1578
- Braun, S., Schmidt, A., Walter, A., Zacharias, V. (2007): The Ontology Maturing Approach to Collaborative and Work-Integrated Ontology Development: Evaluation Results and Future Directions. In: *Proceedings of the International Workshop on Emergent Semantics and Ontology Evolution (ESOE) at the 6th International Semantic Web Conference (ISWC 2007)*, Busan, Korea, CEUR Workshop Proceedings vol. 292, 2007, pp. 5-18
- Zacharias, V., Abecker, A., Vrandecic, D., Borgi, I., Braun, S., Schmidt, A. (2007): Mind the Web, In: *New Forms of Reasoning for the Semantic Web: Scalable, Tolerant and Dynamic 2007. Proceedings of the First International Workshop, ISWC 2007*, Busan, Korea, November 11, 2007, CEUR Workshop Proceedings vol. 291, 2007
- Braun, S., Schmidt, A., Walter, A., Nagypal, G., Zacharias, V. (2007): Ontology Maturing: a Collaborative Web 2.0 Approach to Ontology Engineering. In: *Proceedings of the Workshop on Social and Collaborative Construction of Structured Knowledge at the 16th International World Wide Web Conference (WWW 2007)*, Banff, Canada.
- Braun, S., Schmidt, A., Zacharias, V. (2007): Ontology Maturing with Lightweight Collaborative Ontology Editing Tools, In: *Workshop on Productive Knowledge Work: Management and Technological Challenges (ProKW), 4th Conference on Professional Knowledge Management – Experiences and Visions (WM 2007)*, Potsdam, Germany, March 28-30 2007, Vol. 2, pp. 217-226

We have presented the notion of Social Semantic Bookmarking together with a first analysis of tools that enhance social bookmarking with semantics in four publications. This analysis has informed the development of our conceptual design framework for social semantic tagging and the scaffolding methods for ontology maturing in Chapter 5. Section 9.1.2 presents an updated version of the analysis.

- Braun, S., Schmidt, A., Zacharias, V. (2009): Mit Social Semantic Bookmarking zur nützlichen Ontologie. In: *i-com: Zeitschrift für interaktive und kooperative Medien, Themenheft "Nutzerinteraktion im Social Semantic Web"*. Oldenbourg Wissenschaftsverlag, 2009, pp. 12-19

- Braun, S., Schora, C., Zacharias, V. (2009): Semantics to the Bookmarks: A Review of Social Semantic Bookmarking Systems. In: *5th International Conference on Semantic Systems - I-SEMANTICS'09*, Verlag der TU Graz, 2009, pp. 445-454
- Zacharias, V., Braun, S., Schmidt, A. (2009): Social Semantic Bookmarking with SOBOLEO. In: *Handbook of Research on Web 2.0, 3.0 and X.0: Technologies, Business, and Social Applications*. IGI Global, 2009.
- Braun, S., Zacharias, V., Happel, H.-J. (2008): Social Semantic Bookmarking. In: *International Conference on Practical Aspects of Knoweldge Management, PAKM'08*, LNCS, vol. 5345, Spring, 2008, pp. 62-73

We have published Semantic People Tagging as an novel approach for collaborative competence management (Section 4.4) together with SOBOLEO for tool support (Section 6.7) and first evaluation results (Chapter 8). These experiences has guided us in the development of our conceptual design framework (Chapter 5).

- Braun, S., Kunzmann, C., Schmidt, A.: Semantic People Tagging & Ontology Maturing (2012): An Enterprise Social Media Approach to Competence Management. In: *International Journal of Knowledge and Learning*, Inderscience, 2012
- Abecker, A., Biesalski, E., Braun, S., Hefke, M., Zacharias, V. (2011): Semantics in Knowledge Management. In: *Rudi Studer – A Review on Semantic Web Research*, Springer Verlag, 2011
- Braun, S., Schmidt, A., Zacharias, V. (2010): People Tagging - Aspekte und Möglichkeiten zur Gestaltung. In: *Proceedings of "Mensch und Computer" 2010*, Oldenbourg Verlag.
- Braun, S., Kunzmann, C., Schmidt, A. (2010): People Tagging & Ontology Maturing: Towards Collaborative Competence Management. In: *From CSCW to Web 2.0: European Developments in Collaborative Design*, CSCW Series, Springer, 2010, pp. 133-154
- Braun, S., Schmidt, A. (2009): Mit "People Tagging" zum Kollaborativen Kompetenzmanagement, In: *SoSoft 09 - Social Software @ Work. Collaborative Work, Communication and Knowledge Management in Theory and Practice. Proceedings of the 1st Workshop of the Heinrich-Heine-University Düsseldorf*, Germany, September 28th and 29th, 2009, vol. 591, CEUR Workshop Proceedings, 2009, p. 65-71
- Braun, S., Schmidt, A. (2008): People Tagging & Ontology Maturing: Towards Collaborative Competence Management. In: *Proceedings of the 8th International Conference on the Design of Cooperative Systems (COOP '08)*, Institut d'Etudes Politiques d'Aix-en-Provence, pp. 231-241
- Braun, S., Schmidt, A., Graf, U. (2008): Partizipative Entwicklung von Kompetenzontologien, In: *Nutzerinteraktion im Social Semantic Web. Workshop Proceedings der Tagungen Mensch & Computer 2008, DeLFI 2008 und Cognitive Design 2008, M&C2008*, Lübeck, Germany, Sept. 8, 2008, Logos Verlag Berlin, pp. 168-173

Parts of the SOBOLEO reference architecture and implementation (Chapter 6) have been described in:

- Braun, S., Zacharias, V. (2010): SOBOLEO - Editor and Repository for Living Ontologies. In: *Proceedings of the 1st Int. Workshop on Ontology Repository and Editors for the Semantic Work ORES 2010 at the Extended Semantic Web Conference 2010*, 2010

- Braun, S., Hefke, M. (2009): Im WISSENSNETZ - Vernetze Informationsprozesse in Forschungsverbünden. In: *community of knowledge*, 2009

- Braun, S., Schmidt, A., Zacharias, V. (2007): SOBOLEO: vom kollaborativen Tagging zur leichtgewichtigen Ontologie. In: *Mensch & Computer - 7. Fachübergreifende Konferenz (M&C 2007)*, Munich, Germany, Oldenbourg Verlag„ pp. 209-218.

- Braun, S., Hefke, M., Schmidt, A., Sevilmis, N., (2007 Im Wissensnetz: Linked Information Processes in Research Networks, In: *Proceedings of the German e-Science Conference 2007 (GES 2007)*, Baden-Baden, Germany, May 2-4 2007

- Braun, S., Hefke, M., Schmidt, A., Sevilmis, N., (2007): Im Wissensnetz: Vernetzte Informationsprozesse in Forschungsverbünden, In: *4th Conference on Professional Knowledge Management - Experiences and Visions (WM 2007)*, Potsdam, Germany, March 28-30 2007, Vol. 1, pp. 434-435

- Braun, S., Zacharias, V. (2007): SOBOLEO - Social Bookmarking and Lightweight Ontology Engineering, In: *Proceedings of the Workshop on Social and Collaborative Construction of Structured Knowledge at the 16th International World Wide Web Conference (WWW 2007)*, Banff, Canada.

We presented our ideas of providing specific dialog support for the development of a shared understanding and knowledge structures (Section 6.5 and 6.8.1) in:

- Ravenscroft, A., Braun, S., Nelkner, T. (2010): Combining Dialogue and Semantics for Learning and Knowledge Maturing: Developing Collaborative Understanding in the 'Web 2.0 Workplace'. In: *Proceedings of the International Conference on Advanced Learning Technologies (ICALT'10)*, IEEE Computer Society, pp. 176-180

- Ravenscroft, A., Braun, S., Cook, J., Schmidt, A., Bimrose, J., Brown, A., Bradley, C. (2009): Ontologies, Dialogue, and Knowledge Maturing: A Design Study and Mashup. Technical Report, 2009

- Ravenscroft, A., Braun, S., Cook, J., Schmidt, A., Bimrose, J., Brown, A., Bradley, C. (2008): Ontologies, Dialogue and Knowledge Maturing: Towards a Mashup and Design Study. In: *Proceedings of 1st International Workshop on Learning in Enterprise 2.0 and Beyond, European Conference on Technology-Enhanced Learning (ECTEL '08)*, Maastricht, The Netherlands

Further ideas and experiences, especially of the conceptual design framework (Chapter 5), that are related to the issue variable levels of formality, gardening, and motivational and cultural aspects have been introduced in:

- Cook, J., Schmidt, A., Kunzmann, C., Braun, S. (2011): The Challenge of Integrating Motivational and Affective Aspects into the Design of Networks of Practice. In: *2nd International Workshop on Motivational and Affective Aspects in Technology Enhanced Learning (MATEL '11), ECTEL 2011*, Palermo, Italy, CEUR Workshop Proceedings, 2011

- Mazarakis, A., Kunzmann, C., Schmidt, A., Braun, S. (2011): Culture Awareness for Supporting Knowledge Maturing in Organizations. In: *Proceedings of the Workshop Motivation und kulturelle Barrieren bei der Wissensteilung im Enterprise 2.0 (MKBE '11)*, M&C 2011, Chemnitz, Germany, 2011

- Mazarakis, A., Braun, S., Zacharias, V. (2011): Feedback in Social Semantic Applications. In: *Int. Journal of Knowledge Engineering and Data Mining*, 1(4), Inderscience, 2011, pp. 291-302
- Ramezani, M., Witschel, H.F., Braun, S., Zacharias, V. (2010): Using Machine Learning to Support Continuous Ontology Development, In: *Proceedings of International Conference on Knowledge Engineering and Knowledge Management (EKAW 2010)*, Springer, 2010, pp. 381-390
- Zacharias, V., Braun, S. (2008): Tackling the Curse of Prepayment - Collaborative Knowledge Formalization Beyond Lightweight, In: *1st Workshop on Incentives for the Semantic Web, 7th International Semantic Web Conference ISWC2008*, October 27th, 2008, Karlsruhe, Germany, CEUR Workshop Proceedings 2008.
- Braun, S., Schmidt, A., (2007): Wikis as a Technology Fostering Knowledge Maturing: What we can learn from Wikipedia, In: *Proceedings of the 7th International Conference on Knowledge Management (I-KNOW 2007), Special Track on Integrating Working and Learning* (5.-7. Sept. 2007, Graz, Austria), Springer Verlag, 2007, pp. 321-329
- Braun, S., Schmidt, A., (2007): Expert Finding as Informal Learning Support: Competency-Aware and Socially-Aware, In: *First International ExpertFinder Workshop (EFW'07)*, Berlin, Germany, January 16 2007

We also have organized several workshops on the topic of collaborative knowledge modeling; i. e. ontology maturing, collaborative competence development and motivational, affective and cultural aspects on knowledge management and informal learning, that have contributed heavily to a better understanding of the topic of the thesis:

- Mazarakis, M., Richter, A., Stocker, A., Braun, S., Kunzmann, C., Schmidt, A., Koch, M. (2011): Workshop Motivation und kulturelle Barrieren bei der Wissensteilung im Enterprise 2.0 (MKBE '11), Mensch und Computer 2011, September 2011, Chemnitz, Germany
- Braun, S., Cress, U., Holocher-Ertl, T., Kunzmann, C., Mazarakis, A., Müller, L., Rivera-Pelayo, V., Schmidt, A. (2011): 2nd International Workshop on Motivational and Affective Aspects of Technology Enhanced Learning (MATEL '11), European Conference on Technology-Enhanced Learning (ECTEL '11), September 2011, Palermo, Italy
- Schmidt, A., Braun, S., Cress, U., Holocher-Ertl, T., Kunzmann, C., Mazarakis, A. (2010): 1st International Workshop on Motivational and Affective Aspects of Technology Enhanced Learning and Web 2.0 (MATEL '10), European Conference on Technology-Enhanced Learning (ECTEL '10), September 2010, Barcelona, Spain
- Braun, S., Kunzmann, C., Schmidt, A. (2010): Continuous Competence Development and Knowledge Maturing, Professional Training Facts 2009, November 2009, Stuttgart Germany
- Schmidt, A., Jarrar, M., Ceusters, W., Braun, S. (2008): 3rd International Workshop on Ontology content and evaluation in Enterprise (OntoContent '08), OnTheMove Federated Conferences, November 2008, Monterrey, Mexico
- Schmidt, A., Attwell, G., Braun, S., Lindstaedt, S., Maier, R., Ras, E., Wolpers, M. (2008): 1st International Workshop on Learning in Enterprise 2.0 and Beyond (LEB'08), European Conference on Technology-Enhanced Learning (ECTEL 08), September 2008, Maastricht, The Netherlands

- Braun, S., Ley, T., Schmidt, A. (2008): Reflections on Knowledge Modelling as a Maturing and Learning Process, Summer School on Technology Enhanced Learning and Knowledge Management 2008, Juni 2008, Ohrid, Macedonia

Part I.

Foundations

Overview

This part aims to raise the awareness of the problem. We provide an introduction to folksonomies and social tagging systems in Chapter 2. We present an analysis and discussion of their advantages and challenges that helps us in how we can translate the Web 2.0 phenomenon into the area of ontology engineering. We also analyze and classify tagging motivations in the literature. This classification of literature serves us as a guideline for the analysis of the community and organization where to implement our solution. We will show that there is furthermore a great variety of tag categories related to the implicit semantics of tags and their assignment in literature. Therefore, we provide a mapping and classification of tag categories in the literature. This classification supports (1) making folksonomy literature more accessible, (2) further cross-platform investigations regarding what tag types and categories are used, and foremost (3) guiding the specification of "semantics of tagging" when instantiating our solution.

In Chapter 3 we introduce ontologies and ontology-based knowledge organization systems. We make a comparison and consolidation of the ontology spectra in literature. This comparison helps us understanding the different perspectives in the community. Its consolidation shows the usability or functionality of ontologies. Whereas semantically richer representations with formal axiomatizations provide powerful reasoning capabilities, lesser formal ones are much more easier to develop and maintain and provide lesser computational costs. It also helps for the continuous transition between folksonomies and ontologies. Further, we present a review and comparison of knowledge engineer-driven and collaborative ontology development methodologies and tools. We analyze the advantages and challenges of ontologies and ontology-based knowledge organization systems before we before we span the problem space in a concluding discussion.

2. Folksonomies and Social Tagging Systems

2.1. Web 2.0

With the "Web 2.0", since about the year 2004, a new form of web applications have emerged that focus on an "architecture of participation", lightweight and easy-to-use user interfaces, and the provision, consumption and reuse/-mixing of information from multiple services and users to harness collective intelligence (O'Reilly, 2005). Indeed, this shows that Web 2.0 is not about certain new technologies but general principles that new technologies enable and facilitate whilst the voluntary and active participation makes up the largest difference to the Web 1.0.

2.2. Social Software

Applications that especially emphasis the social aspects are also known as "Social Software". The term "Social Software" was coined in the year 2002 by Clay Shirky's event called "Social Software Summit", even though its origins might go back much earlier and some people like Allen (2004) even locate them in Vannevar Bush's Memex in 1945 (Bush, 1945). We'd like to follow the perspective of Peters (2009) considering Social Software not synonymously to Web 2.0 but as part of.

Coates (2005) describes Social Software as: *"Software that supports, extends, or derives added value from human social behavior"*. (Bächle, 2006, p. 121) defines it as *"software systems that support human communication and collaboration"*[1]. Social Software can be grouped according to its basic functions. Schmidt (2006) proposed the "Social Software Triangle" based on the 3-K model for groupware by Teufel et al. (1995). It structures Social Software in supporting online-based identity management (self-representing in the web), social relationship management (networking and socializing), and information management (selecting, consuming and managing information). Koch and Richter (2008) adapted this triangle and augmented "identity management" by "identity and social network management" and replaced "social relationship management" by "communication" as corners (see Figure 2.1).

Peters (2009) divides social software according to its main functions into a group for communication and socializing (e. g., social networks and microblogging), for building up a knowledge base (e. g., wikis and podcasts), and for resource management (e. g., social bookmarking and video or photo sharing); in this way differentiating with the latter two groups the creation of new content and knowledge from its management (see Figure 2.2).

[1]translated by the author

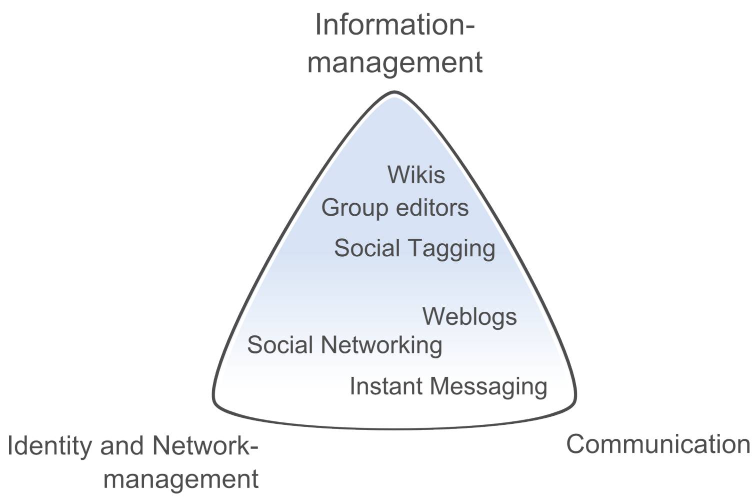


Figure 2.1.: Social Software Triangle by Koch and Richter (2008)

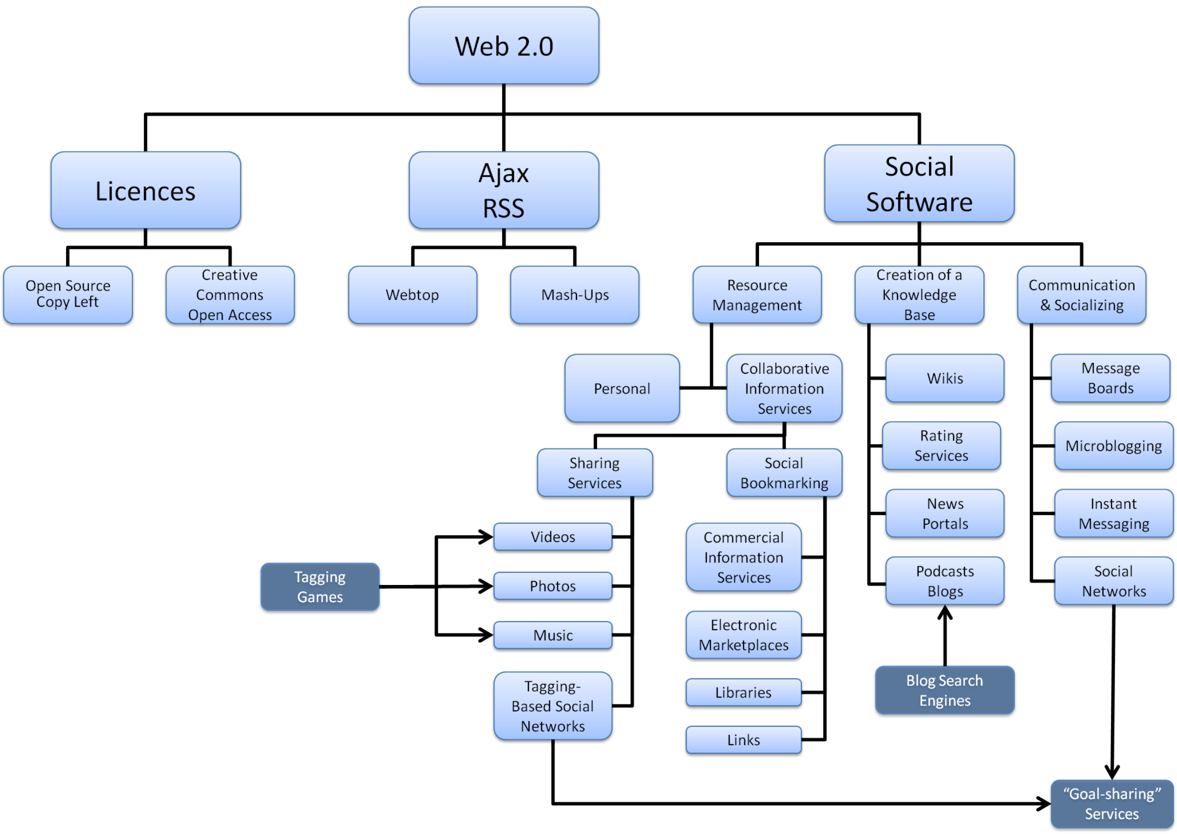


Figure 2.2.: Web 2.0 & Social Software Classification by Peters (2009)

2.3. Enterprise 2.0

"Enterprise 2.0" describes the use of social software in an organizational context to support the collaboration of the employees in pursuit of the organizational goals. McAfee (2006a,b, 2009) focuses especially on *"emergent social software platforms"* that are used by organizations (within organizations or between these and their partners or customers) *"in order to make visible the practices and outputs of their knowledge workers"*. Here, *emergent* means to be "freeform", i. e. optional, free of imposed structure, egalitarian and free in data types, and letting patterns and structures arise from the people's interactions. The author summarizes the technical features of Enterprise 2.0 social software with the acronym *SLATES*, i. e.

- **Search**: information must be easily discoverable
- **Links**: deeply connecting information into an interactive and interdependent community
- **Authoring**: everyone must have easy access to contribute and edit contents
- **Tags**: organic on-the-fly structuring and organization of data with own terms from every point of view thus meaningful for the users
- **Extensions**: providing suggestions and recommendations by mining patterns and user activity
- **Signals**: notification of new available content of interest and changes.

Not being defining enough, Hinchcliffe (2007) extended the six features to capture the social, emergent and freeform aspects as well as the new network-oriented aspect, i. e. making also small junks of content addressable and reusable, forming the mnemonic FLATNESSES (see also Fig. 2.3).

McAfee (2006a) identifies four requirements in order to make Enterprise 2.0 to work:

- Building an open and receptive organizational culture
- A common platform that allows emerging collaborations
- A viral, informal rollout guided by the employees' need instead of sticking to formal processes
- Managerial commitment and support that encourages usage.

2.4. From Tags to Social Tagging

One commonality especially of information respectively resource management applications is that they enable their users to annotate resources with arbitrary keywords or labels, so called "tags", for personal or shared organization and retrieval of the information. In principle, "tagging" as the process of adding keywords to information resources is nothing new. It's done by librarians, indexers or also machines. What's new regarding humans adding tags is that not only dedicated experts like librarians or the resource creators provide the annotations but also laypersons, who consume the information resources, can do the subject indexing without being bound to any predefined rules or controlled vocabulary.

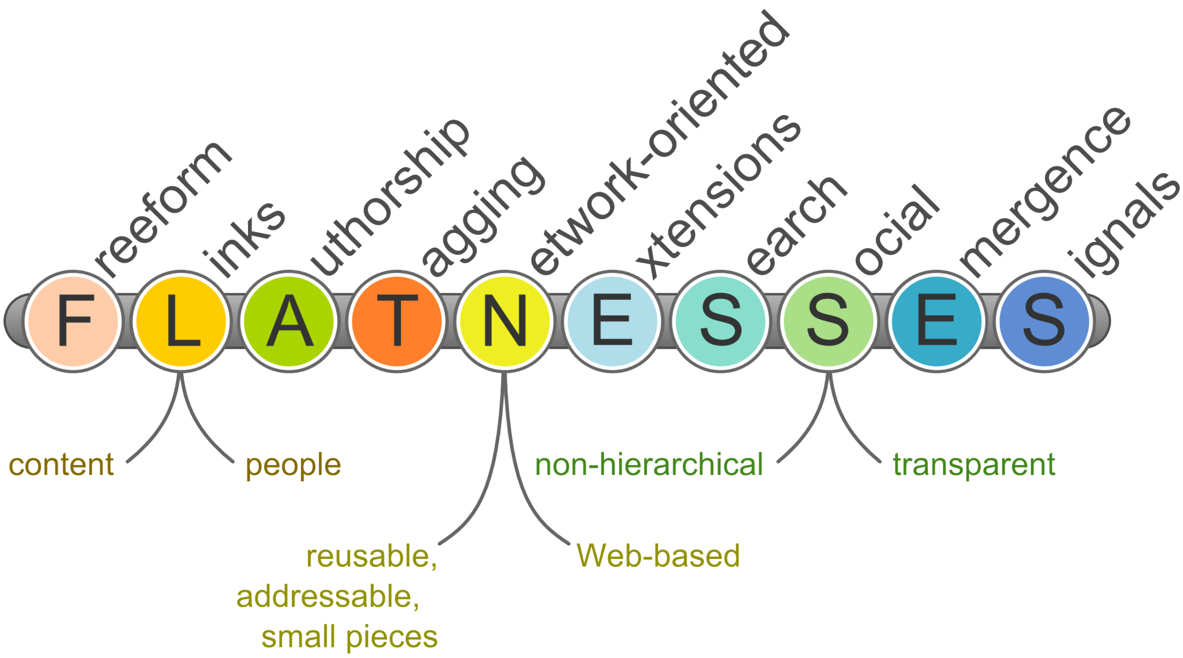


Figure 2.3.: FLATNESSES framework showing the key aspects of Enterprise 2.0 social software by Hinchcliffe (2007)

If tagging takes place in a social-technical environment, in which tags, resources and users are publicly visible and shared, we speak of "social tagging" (Trant, 2009). Thus, Golder and Huberman (2006) define social tagging as *"the process by which many users add metadata in the form of keywords to shared content"*. Social tagging might also be referred to as "user tagging" (Hayman, 2007) or "collaborative tagging" (Golder and Huberman, 2006)[2].

2.5. Folksonomies

Resulting from social tagging is the interconnectedness of tags through the aggregation of all tags assigned by the users in a system, which is called "folksonomy". Thomas Vander Wal coinded the term folksonomy in 2004 and has defined it as follows:

> *"Folksonomy is the result of personal free tagging of information and objects (anything with a URL) for one's own retrieval. The tagging is done in a social environment (usually shared and open to others). Folksonomy is created from the act of tagging by the person consuming the information."* (Wal 2007)

The term Folksonomy is composed of 'folk' and 'taxonomy', whereas the 'taxonomy' part is discussed very controversially and might be misleading because folksonomies aren't a classification with clearly defined relations but more a shared vocabulary to categorize the information resources (Mathes, 2004). Folksonomy is also referred to as "social classification" (Smith, 2004; Hammond et al., 2005), "ethnoclassification" (danah boyd, 2005; Merholz, 2004), "grassroots taxonomy" (Weinberger, 2005), "tagsonomy" (Hayes

[2] Strictly speaking, collaborative tagging is a subclass of social tagging because, despite giving the same definition, the authors exclude from collaborative tagging those tagging activities where the users can only tag resources they own, e. g., photographs on flickr

et al., 2007), "tagosphere" (Gruber, 2007) and many other (cf. Peters, 2009). This variety of terms describing the same phenomenon also shows that there's no consistent definition of "folksonomy" itself. Additionally, in literature it is often synonymously used with 'tagging' or 'social tagging' whereas that's properly speaking the activity from which folksonomies emerge. Confronted with the same difficulty, we'd like to stick in this thesis to what's considered as a folksonomy by Schmitz (2007) going along with the formal definition in the following: *"a number of users who attach arbitrary tags to a set of resources, thus creating a set of tag assignments"*

2.5.1. Formal Model of Folksonomy

A folksonomy can also be viewed as a tripartite (undirected) hypergraph consisting of disjoint finite vertex sets of users $U = \{u_1, ..., u_k\}$, tags $T = \{t_1, ..., t_l\}$ and resources $R = \{r_1, ..., r_m\}$ and a set of ternary relations between them, the tag assignments[3] $A \subseteq U \times T \times R$ (Mika, 2005; Marlow et al., 2006; Lambiotte and Ausloos, 2006). Thus a folksonomy is a quadruple $F := (U, T, R, A)$ or represented as hypergraph $G(A) = (V, E)$, where the set of vertices is $V = U \cup T \cup R$, and the set of hyperedges is $E = \{\{u, t, r\} | (u, t, r) \in A\}$. Some researchers extend this definition by time, i. e. the moment of the tag assignment (Maass et al., 2007; John and Seligmann, 2006), or subtag/supertag relations (Hotho et al., 2006c).

Based on the co-occurences of its elements, we can reduce the tripartite hypergraph into three bipartite graphs (Mika, 2005): the UT graph aggregating users and tags associations, the UR graph aggregating users and resources associations, and the TR graph aggregating tags and resources associations. Out of these graphs, we can gain additional information on implicit social networks, networks of resources and tags. For instance, we can fold UT respectively UR into one-mode networks[4]. We gain on the one hand a social network of people connected by the same used tags/resources and weighted by the number of tags/resources they have both used. This shows us people with overlapping tag sets (shared vocabulary) and resource sets. On the other hand we gain a network of tags/resources that have overlapping sets of people with the weight between pairs of tags/resources given by the number of people who have used both. Regarding the gained network of tags, Mika (2005) also refers to lightweight ontology of tags. Similarly we gain a lightweight ontology of tags based on overlapping resources from the TR graph.

2.5.2. Different Types of Folksonomies

We can further distinguish between two different types of folksonomies – "*broad*" and "*narrow*" folksonomies (Wal, 2005) – according to what Sen et al. (2006) call the "*tag scope*" and which describes whether a tag might be assigned multiple times to the same resource or only once. Marlow et al. (2006) also speak of "*bag model*" for a broad folksonomy and of "*set model*" for a narrow folksonomy.

Thus, in broad folksonomies any number of users may assign their set of tags from their personal vocabulary to a resource and retrieve the resource based on these tags. That means individual users own the tag assignments. A broad folksonomy-based system is, for example, the Delicious[5] social bookmarking system.

[3] In literature tag assignments are also called tag applications or tag posts.
[4] Networks with relations between a single set of nodes.
[5] http://delicious.com

In narrow folksonomies, users collectively tag a resource and tags are only assigned once to a resource. This means the community owns the tag assignment. This is often done only by a small group of people or by a single user – the user who added the resource to the system and, in the course of this, applied the first tags. Only a few or no other users add additional tags. Thus, it's not possible to analyze tag frequencies or distributions on the level of resources. An example of such a narrow folksonomy-based system is Flickr[6].

In literature, narrow folksonomies are also often seen in a more limited way (e. g., by Damme et al. 2008 or Bogers 2009) and described as folksonomies that emerge when only the resource creator (the one who added the resource to the system) can assign tags to the resource, but others can use these tags e. g., for retrieval, like in the case of YouTube. Indeed, it's to arguable, if that's still a folksonomy because the social cohesion is missing.

2.6. Folksonomies and their Applications

2.6.1. System Design and Implications

As previously seen, the different types of emerging folksonomies are inherently determined by the social tagging systems' design; for instance who is allowed to tag and if there are individual or community tag assignments. Marlow et al. (2006) discusses additional design dimensions in which folksonomy-based system can be classified together with their implications:

- **Tagging Rights:** Systems can specify tagging rights at different levels of permissions; ranging from any user tagging any resource (called "*free-for-all tagging*"), over restricted permission to specific groups (e. g., friends, family and contacts in Flick), to only the resource owner being allowed to tag the resource ("*self-tagging*"). Likewise, the right to remove tag assignments may vary between no one, anyone, the tag assignment creator or resource owner. This might influence the nature and type of assigned tags as well as the role of tags in the system. For instance, possibly many tags per resource, synonymous or divergent in meaning, in case everyone can tag versus few and subjective tags per resource in case of the resource creator as the only tagger.

- **Tagging Support:** Systems may provide tag recommendations to support the tagging activity. Marlow et al. (2006) differentiate between three types: *blind tagging* with no recommendation, *viewable tagging* showing tags already assigned to the resource, and *suggestive tagging* recommending tags based on, e. g., previously assigned tags by the user or others or extracted context metadata. Thus, suggestive tagging might support quicker convergence and consolidation of the tagging vocabulary on the one hand. On the other hand, showing tags can create a bias thus overweighted tags, because the users only confirm the displayed tags instead of adding new ones.

- **Aggregation:** Systems may allow that users can assign tags multiple times to the same resource or only once. Whereas only the first option allows to exploit aggregate statistics.

[6] http://flickr.com

- **Type of Object:** Systems may support different kind of contents, e. g., bookmarks, photos, people, location etc., as resources being tagged. This might influence the nature and type of assigned tags. For instance, textual resources might be differently tagged than multimedia resources.
- **Source of Material:** The content being tagged may be user-supplied (e. g., videos on YouTube), system-supplied (e. g., audio on last.fm!) or taken from another web source. This might have an impact on the users motivation to tag as well as on the nature and type of assigned tags. For instance, users might tag their own video differently to bookmarks (see also Section 2.7.1).
- **Resource Connectivity:** Resources may be connected beyond the users' tags via direct links or groups. This might lead to connected resources being similarly tagged.
- **Social Connectivity:** Systems may allow their users to link each other or to form groups. This might lead to convergence on their vocabulary.

Sen et al. (2006) mention an additional design dimension "*tag sharing*" that can also be seen as an addition to Marlow et al.'s Tagging Rights dimension. Thus, systems can specify to which extent tag assignments are visible to other users. At one end, there are systems which display all tag assignments to all users and at the other end, tag assignments are only visible to the tagger him-/herself. Whereas we leave out the latter one because that does not follow anymore our definition of social tagging.

Additionally, there are systems in the middle enabling their users to control the visibility of their tag assignments. This might be further influenced by default settings (e. g., always private or always public). In this way, social tagging systems represent social translucent systems, i. e. *"systems that support coherent behavior by making participants and their activities visible to one another"* (Erickson and Kellogg 2000). In concrete, tags in social tagging systems make their users aware of their individual tags as well as of the tags and resources other users contribute and make use of.

We will further detail the impact of these design dimensions on the emergence of common understanding and vocabulary in Section 2.7.3 and take up these design dimensions for designing our framework in Section 5.3.

2.6.2. Social Bookmarking

There is a vast range of social tagging applications. *Social bookmarking* was one of the very first ones that gained widespread popularity. The resources that are stored, shared, and organized and managed with tags are references[7], so called bookmarks, to online web documents (Hammond et al., 2005). Instead of collecting links in the browser on the desktop, users can store their bookmarks in a centralized repository, usually a social bookmarking service in the web. This makes the users independent of the desktop, as they can access their bookmarks from any computer with an Internet connection. Usually bookmark collections are personally created and maintained but typically visible to others. One of the earliest and most popular service is Delicious that was founded in 2003 by Joshua Schachter and in 2005 acquired by Yahoo!. Delicious is a broad folksonomy-based system as every user owns his/her own set of tag assignments for the bookmarks stored in the system. For storing bookmarks, users have to be registered. Browsing through public bookmarks and tags is possible for anyone. When saving a bookmark,

[7] not the content

the URL is used as identifier of the resource. The user can additionally specify a title, a description or note and a set of tags (separated with spaces) and if it should be kept private or not (the default is public). The date and time are stored as well when the user submits the data. All submitted information together is also called a post. Delicious provides further features to support the users like suggestive tagging or tag bundles that allow the users to group tags together under a common topic.

As with any social tagging application, social bookmarking systems rely on browsing instead of searching for bookmarks, especially by dint of tags. Thus, social bookmarking support pivot browsing usually for all of the three folksonomy dimensions, i. e. users, tags and bookmarks, and in reverse chronological order. For instance, clicking on a particular user name shows the user's bookmark collection together with the tags he/she has assigned. The bookmarks and tags are again links to pages for these bookmarks and tags, listing in the first case all users that have tagged a resource together with the assigned tags and in the latter case all bookmarks together with the users this tag is assigned to. Combinations of the dimensions are also possible. This tight linkage allows a quick navigation through the information space. Additional visual navigation support provide tag clouds that show e. g., all tags of a user and where size and color of the tags indicate their assignment frequency. Further popular features are feeds, e. g., to subscribe to a specific tag, or user, or tag of a user etc., and social networking that allows to directly share bookmarks with specific users, e. g., via special tags of particular format like "for:<username>".

Social bookmarking is not limited to web documents only. It has also been applied to enterprise contexts where the employees collect and share references to non-public intranet documents, e. g., with Dogear by and within IBM (Millen et al., 2006) or to manage literature references, e. g., with BibSonomy[8], Connotea[9] or CiteULike[10].

2.7. Characteristics of Folksonomy-based Systems

2.7.1. Tagging Behavior and Motivations

It's not only the system design choices that affect the emerging folksonomy but also the users' tagging behavior and motivation that play a significant role. Thus, several researchers started to investigate the dynamics of social tagging systems' use. These studies are primarily of descriptive statistics nature. Several studies had a look at tag (assignment) distributions on the level of the whole database, the users and the resources. They could show that, on the database as well as on the resource level, there is either a power law or inverse-logistic distribution. Both of them show the characteristic of a "long tail", i. e. the majority of tags are only seldom used. These tags represent personal or specialized tags of only few people (see also the different tag types in Section 2.7.2) whereas the highly frequently used tags of the beginning are generally meaningful for the community (Guy and Tonkin, 2006).

Both distributions vary in their beginnings: in a power law distribution, there are only a few tags with a high frequency, the curve harshly drops off and quickly ends in the long tail, whereas the inverse-logistic distribution shows a long trunk of similar high frequent tags, the curve smoothly drops off to the inflection point and fades to the long tail. This

[8] http://www.bibsonomy.org
[9] http://www.connotea.org
[10] http://citeulike.org

long trunk might be interpreted in a way that the resource / data base comprises different meanings that need to be described with several tags (Kipp and Campbell, 2006; Peters, 2009).

Additionally, we can observe variations of spellings and notations, acronyms, neologisms and synonyms and relatedness among tags. Users typically tag with nouns that consist of single words. This may also be traced back to the fact that many tagging systems don't support multi word keywords and use the space character as delimiter for tags. This leads to the separation of multi word keywords into multiple tags and several strategies of compound words forming in order to overcome these systems' inability, e. g., by using punctuation or CamelCase[11] (Kipp, 2009; Heckner et al., 2008a,b; Spiteri, 2007).

What's interesting is that the tag distributions stabilize once a resource reaches a certain number of tag assignments respectively the database reaches a critical mass of sufficiently tagged resources, i. e. the set of top tags remains unchanged and only their ranking varies (Golder and Huberman, 2006; Maier and Thalmann, 2008; Robu et al., 2009). In general, systems with narrower folksonomies show fewer tags per resource than in systems with broader folksonomies.

Similarly for the users, most tag assignments are generated by relatively few users with up to three or four tag assignments per resource in median (Farooq et al., 2007; Sen et al., 2006; Kipp and Campbell, 2006; Xu et al., 2006; Lee, 2006). At the same time, the median seem to vary between different social tagging systems as shown by Heckner et al. (2009). Additionally, we can observe a correlation between the users' resource collection size and their tag vocabulary size (Farooq et al., 2007; Santos-Neto et al., 2007). When tagging a resource, users seem to tend to assign first more general (community-agreed) tags, representing a kind of basic level, that are followed by more specific and personal ones (Golder and Huberman, 2006). Over time, we can also observe a refinement in the tags a user is using. The authors explain this phenonmenon on the basis of seeing tagging as a retrospective process in the sense of sensemaking; i. e. in the way of re-discovery, users find a hitherto unnoticed distinction in meaning (see also Section 2.7.3).

2.7.1.1. User Types

Further examinations by Körner (2009) and Körner et al. (2010) reveal that the users' vocabulary size and number of tag assignments per resource also depend on the users' type and tagging motivation. They differentiate between *categorizers* and *describers*. Categorizers have a more limited but discriminative and rather stable tag vocabulary of high reuse that they use for navigation aid for later browsing and where two tags per resource might already be sufficient. On the other hand, describers are more focused on later searching and therefore precisely describe their resources with many tags. They have an open and likely to be larger tag vocabulary containing redundant (e. g., synonyms) and rarely used tags.

Thom-Santelli et al. (2008) investigated in a qualitative study the tagging behavior in enterprise social tagging systems and could identify five user types:

- **Community-seeker:** search and tag with tags in order to show or create their social connection with a community

[11]CamelCase, aka. medial capital, is the practice of forming compound words by omitting the spaces between and capitalizing the initial letter of each single word within the compound

- **Community-builder:** assigns tags used and known within the community thus current and potential members can easily find the resources
- **Evangelist:** assigns tags, known within the community, consistently cross several tagging systems in order to guide others to related information and to his/her resources and in this way enhancing his/her and the community's visibility and reputation within the organization
- **Publisher:** are broadcasters and in order to disseminate resources to various target groups, they use the tags their intended recipients search with. As opposed to the community-builder or evangelist, the publisher's interest is not in the own reputation or social linking.
- **Small Team Leader:** uses tags and thus introduces a specific terminology that is only understandable by a small group in order to well-directed distribute resources within this specific group.

As another interesting point, the authors mention the users' general awareness of having an audience and with this the attempt for tag consistency with previously used own tags but also with other users' tags within and across systems.

As a mixture of the previous one, Panke and Gaiser (2009) identified four user types:

- **Ego Taggers:** see themselves as the information elite and seek for recognition and publicity through tagging
- **Everyday Archivers:** organize their web activities through tagging
- **Broadcasters:** tag for public sharing of their resources
- **Team Players:** use tags for information exchange information with personal networks

2.7.1.2. Categorizing Tagging Motivations

It also interesting to see that Körner et al.'s user types have a focus on personal organization strategies, whereas these seem to take a back seat in the enterprise context of Thom-Santelli et al.'s user types that focus more on the aid in information sharing and discovery. Even though there's much debate on how to categorize user motivations, we can identify two general groups: information organization or management and information sharing (Angus and Thelwall, 2010; Heckner et al., 2009).

In this way, Marlow et al. (2006) categorize tagging motivations into two high-level practices *organizational* and *social*. Organizational motivations arise from the attempt of developing a personal digital structured filing and social motivations from the attempt to express themselves and communicate with other users. Similarly, Hammond et al. (2005) categorize into *selfish* and *altruistic*: "*There is a range from a 'selfish' tagging discipline, where the users are primarily tagging their own content for their own retrieval purposes, right through to a more 'altruistic' tagging discipline, where the user is tagging others' content for yet others to retrieve*".

Ames and Naaman (2007) extended the notions of *organizational* and *social* and provide a taxonomy along the two dimensions *function* and *sociality*. The sociality dimension relates to intended target group of the tags, i. e. oneself or others (called social) comprising friends, family and public. The function dimension relates to the tag's purpose, i. e. to aid organization and retrieval or to aid in communication. Whilst tagging motivation

Table 2.1.: Tagging motivations by Ames and Naaman (2007) and adapted by Panke and Gaiser (2009)

Amnes 2007 / Panke & Gaiser 2009:		Function	
		Organization	Communication
Sociality	Individual	searching, categorizing and retrieving resources	documenting one's own ideas, views, memories and context information
	Group	Sharing resources with others	providing contextual descriptions and reviews about resources
		Attract attention	expressing social cohesion

tends to be driven by either self or social motives, organization and communication can play a dual role (Angus and Thelwall, 2010). Table 2.1 shows the resulting 2x2-matrix in which Ames and Naaman (2007) locate different tagging motivations. These were more generalized by Panke and Gaiser (2009).

- **individual – organization:** users tag for individual knowledge organization purposes, i. e. searching, categorizing and retrieving resources
- **social – organization:** users tag for group-related knowledge organization purposes, i. e. sharing resources with others and/or drawing others' attention to one's own resources
- **individual – communication:** users tag for self-communication purposes creating knowledge cues *"to revoke a previously experienced mental state, when used"* (Völkel 2010), i. e. documenting one's own ideas, views, memories and context information
- **social – communication:** users tag for group communication purposes, i. e. providing contextual descriptions and reviews about resources to others and expressing social cohesion to provide context information for others.

In the following, we will classify the tagging motivations in the literature into the adapted matrix by Panke and Gaiser (2009).

Marlow et al. (2006) investigated a range of tagging motivations of which they see "future retrieval" as the foremost reason (see Table 2.2):

- **Future Retrieval:** to mark resources for personal retrieval also as an activity incentive or reminder to oneself or others
- **Contribution and Sharing:** to increase access to resources and make resources available to others
- **Attract Attention:** to draw others attention to own resources
- **Play and Competition:** to produce tags based on certain internal or external rules or competitions
- **Self Presentation:** to present and show the own identity and interests
- **Opinion Expression:** to articulate and share the own opinion and judgment.

Table 2.2.: Tagging motivations by Marlow et al. (2006)

Marlow et al. 2006:		Function	
		Organization	Communication
Sociality	Individual	Future retrieval	
	Group	Contribution & Sharing	
		Attract Attention	
		Self-Presentation	
			Opinion Expression

Table 2.3.: Tagging motivations by Zollers (2007)

Zollers 2007:		Function	
		Organization	Communication
Sociality	Individual		
	Group		Opinion-expression
		Self-Presentation (to guide others to own resources)	Performance (inc. Self-Presentation)
		Activism (tagging according to community-rules & to show social cohesion)	

Zollers (2007) deduced three tagging motivations (see Figure 2.3): (1) opinion expressing; (2) performance inc. self-presentation, using tags as a show-off technique to others; and (3) activism, to broadcast a group view on a certain topic to a certain audience.

Wash and Rader (2007) categorize the identified motivations into three groups of (1) later retrieval; (2) resource sharing; (3) social recognition (see Table 2.4).

Heckner et al. (2009) only differentiate between (1) personal information management and (2) sharing which can be specific, e. g., targeted for a particular user, or unspecific (see Table 2.5).

Nov et al. (2009) investigated as individual tagging motivations: (1) enjoyment – users enjoy the act of sharing; (2) commitment – users feel loyal and obliged to the community, (3) self-development – users want to improve themselves through learning from others, and (4) reputation – users want to improve their reputation within the community (see Table 2.6).

Unfortunately, it has not yet been explored in a systematic way how tagging motivations relate to user behavior. The literature only gives anecdotal examples such as by Heckner

Table 2.4.: Tagging motivations by Wash and Rader (2007)

Wash & Rader 2007:		Function	
		Organization	Communication
Sociality	Individual	Later retrieval	
	Group	Sharing	
		Self-recognition	

Table 2.5.: Tagging motivations by Heckner et al. (2009)

<table>
<tr><th colspan="2" rowspan="2">Heckner et al. 2009:</th><th colspan="2">Function</th></tr>
<tr><th>Organization</th><th>Communication</th></tr>
<tr><td rowspan="3">Sociality</td><td>Individual</td><td colspan="2">Personal Information Management</td></tr>
<tr><td rowspan="2">Group</td><td colspan="2">Specific Sharing</td></tr>
<tr><td colspan="2">Unspecific Sharing</td></tr>
</table>

Table 2.6.: Tagging motivations by Nov et al. (2009)

<table>
<tr><th colspan="2" rowspan="2">Nov et al. 2009:</th><th colspan="2">Function</th></tr>
<tr><th>Organization</th><th>Communication</th></tr>
<tr><td rowspan="4">Sociality</td><td rowspan="2">Individual</td><td colspan="2">Enjoyment</td></tr>
<tr><td colspan="2">Self-development</td></tr>
<tr><td rowspan="2">Group</td><td colspan="2">Commitment</td></tr>
<tr><td colspan="2">Reputation</td></tr>
</table>

et al. (2009) where users 'overtagged' or tagged very intensively because they wanted to ensure to attract the attention of as many people as possible. It's only Mirzaee and Iverson (2009) who started to relate tagging motivations to behaviors. They identified seven tagging motivations and organized them into *organizing* and *socializing* similar to the high-level categories of Marlow et al. (2006) (see Table 2.7). Preliminary results show how the tagging motivations affect five specific tagging behaviors. For instance, users are motivated by creating social connection and therefore they adopt the tags in the community.

Table 2.7.: Tagging motivations by Mirzaee and Iverson (2009)

<table>
<tr><th colspan="2" rowspan="2">Mirzaee 2009:</th><th colspan="2">Function</th></tr>
<tr><th>Organization</th><th>Communication</th></tr>
<tr><td rowspan="6">Sociality</td><td rowspan="3">Individual</td><td>Describe content</td><td>Create context</td></tr>
<tr><td colspan="2">Improve search/retrieval</td></tr>
<tr><td colspan="2">Create social connections</td></tr>
<tr><td rowspan="3">Group</td><td colspan="2">Relate to like-minded people</td></tr>
<tr><td colspan="2">Increase/maintain social presence</td></tr>
<tr><td></td><td>Social signaling/notification</td></tr>
</table>

2.7.1.3. Impact of Social Presence

Even though not representative, the user interviews by Ames and Naaman (2007) indicate that users are especially motivated by social organization, i. e. that social motivations appear to be important for tagging. Other studies also support that social presence is affecting user tagging. Lee (2006) could show by analyzing Delicious data that users perceiving the presence of others are more likely to tag resources, i. e. more frequently, and to add more tag assignments to a resource. Similarly Nov and Ye (2010), who based their work on Ames and Naaman (2007), could confirm social presence as a prime motivator for tagging and that users tend to add more tag assignments to a resource when they perceive the presence of others, are self or publicly motivated or have a bigger resource collection. Arakji et al. (2009) investigated mechanisms of public resource contributions to social bookmarking sites. Their study reveal that publicly bookmarking is not a byproduct of organizing for personal purposes but intentional tagging for others; especially when the users believe that their resources are useful to others. Additionally, they find that the more users perceive that others contribute to the community, the more deliberately they tag resources for others.

2.7.2. Semantics of Tags and Tag Assignments

Besides tagging behavior and motivations, several researchers investigated the semantics of the tags and their assignment, i. e. their type and function. One of the most cited is the work by Golder and Huberman (2006) who identified seven types/functions of tags in their analysis of the Delicious system:

- **What or Who It Is About:** describing the topic of a resource with different levels of specificity and proper names (usually nouns)
- **What It Is:** identifying the type of the resource (e. g., "article", "book")
- **Who Owns It:** about the owner or creator of a resource or its content
- **Refining Categories:** tags that refine or qualify other tags
- **Qualities or Characteristics:** expressing the tagger's opinion (usually adjectives like "funny", "cool", "horrible")
- **Self Reference:** show the relation between the tagger and the resource (often beginning with "my" like "myown")
- **Task Organization:** describing future task-related use of a resource (e. g., "toread", "toemail", "courseplanning")

Others distinguish "location", "activity/event", "depiction" and "emotion/response" (Schmitz, 2006) or "content-based", "context-based", "attributes", "subjective" and "organizational" (Xu et al., 2006). Some reduce Golder & Huberman's categories to "personal", "subjective" and "factual" (Sen et al., 2006) or extend their categories by "location" and "time" (Bischoff et al., 2008). This shows that there is a great variety of tag categories with widely varying names.

The tag category model by Heckner et al. (2008a) gives an systematic overview on tag characteristics. It is sub-divided into three sub-models (cf. 2.4): (1) the *Linguistic Category Model* – relating to linguistic aspects of tags (i. e. word class, spelling, neologism and language); (2) the *Tag to Text Category Model* – relating to tag redundancy in terms of the tagged resource (i. e. identical to, variation from and not occurring in fulltext); (3) the

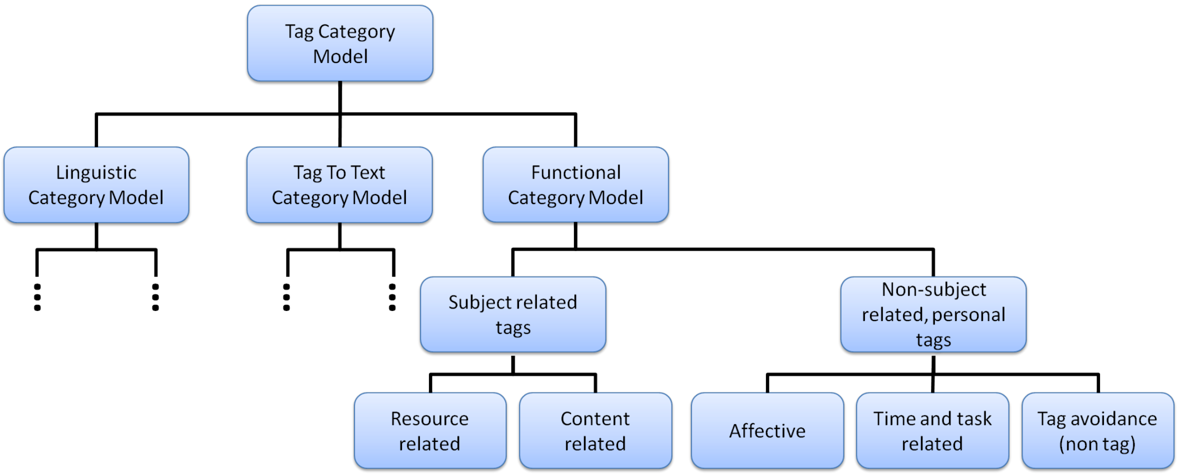


Figure 2.4.: Tag Category Model overview and Functional Category Model by Heckner et al. (2008a)

Functional Category Model – relating functional and semantic aspects of tags (i. e. "subject related tags" and "non-subject related, personal tags"). With Table 2.8 and Table 2.9 we provide a mapping of the different tag categories in the literature and classification into the Functional Category Model by Heckner et al. (2008a) that allows us to compare the different approaches.

Kipp (2008) uses the most similar categories to Heckner et al. (2008a). Whilst Kipp (2008) does not further differentiate subject related tags, she focused in her studies of Delicious, Connotea and CiteULike on time and task related tags such as 'toread' or 'todo' and on affective tags that describe an emotional state such as 'fun' and 'cool'.

Also similar are the categories of Sen et al. (2006). They name the group of subject related tags *factual* tags identifying "facts" of the resource or content (also including the content refining categories by Golder and Huberman (2006)) and the group of affective tags *subjective* tags. The third group the authors distinguish, are *personal* tags; i. e. tags with self-reference and organizational purpose. Tags related to time or location aspects are not named.

Besides Sen et al. (2006), it is only Dutta and Giunchiglia (2008) who also consider the content refining categories by Golder and Huberman (2006). Indeed, Dutta and Giunchiglia (2008) provide one of the most fine grained categorizations; especially for subject related tags. Interestingly, they include spatial and temporal tags in the class of content related and thus subject related tags whereas these are elsewhere viewed as non-subject related tags.

Being one of the few, Dutta and Giunchiglia (2008) also consider non-tags being unidentifiable (*unknown*) terms in Lin et al. (2006) or *junk* in Heymann et al. (2010). According to Heckner et al. (2008a) the group of non-tags or "tag avoidance" means that users deliberately don't use any tag or specific "avoidance" tags such as '-'. A cross system comparison of the distribution of non-subject related tags has shown that whilst this category never occurs in Connotea, it makes 30.77% in YouTube, 49,32% in Delicious and 98,69% in Flickr (Heckner et al., 2008b). That might be traced back to the fact that, in Flickr, tags are not critical and a link to the right album is engough for sharing and retrieval. Whilst YouTube users tag extensively to attract as many other users' attention as possible and for Connotea users tagging for information management seems to be an very important feature as such (Heckner et al., 2009).

Also interesting results revealed the study by Bischoff et al. (2008) in comparing tag type usage across different systems: whilst "Topic" tags are the most important tags in Flickr and Delicious, it's the "Type" category for last.fm. "Location" seems to be important only for photos, whilst last.fm shows a higher amount of subjective and opinion tags.

Lin et al. (2006) differ from the other categorizations in their detailed non-subject related tags. Whereas there is elsewhere only one group for affective/subjective/opinion tags, they distinguish humor, poetic, rating and emotion as individual categories. Similarly, they focus on detailed place, time and event descriptions.

Dubinko et al. (2006), who focus on the Flickr photo tagging analysis, stands out in describing only three quite narrow categories. These are personalities, events and social media tagging. The latter one depicts a random topic brought up by the Flickr community like "What's in your bag" applied to photos showing the contents of people's bags and that might be most likely seen as content related tags.

The six tag categories by Heymann et al. (2010) are a bit transversal. Objective and content-based tags do not refer to the annotator and/or describe the content. This category might overlap with the physical category when classifying the groups of Heymann et al. (2010) into the functional category model by Heckner et al. (2008a). Physical tags describe the object physically; this includes its type, format but also its location. The category personal, in contrast to Sen et al. (2006), do not only bear on self-reference and organizational purpose but also activities. Acronyms, elsewhere seen in the group of personal and task related tags, form a separate group. The last two groups are opinion and junk tags.

Table 2.8.: Tag Categories 1

Heckner et al. 2008		**Kipp 2008**	**Bischoff et al. 2008**	**Golder & Huberman 2006**	**Yanbe et al. 2007**	**Sen et al. 2006**	**Xu et al. 2006**
Subject related tags	**Resource related**	Subject related	Type	What It Is	n.n.	Factual	Attribute
			Author/ Owner	Who Owns It			
			n.n.	n.n			
	Content Related		Topic	What or Who It Is About	Content tags		Content-based
			n.n.	Refining Categories	n.n.		n.n.
Non-subject related tags, personal tags	**Time & task related**	Time & task related	Location	n.n.	n.n.	n.n.	Context-based
			Time				
			Usage Context	Task Organization		Personal	Organizational
			Self Reference	Self Reference			
	Affective	Affective / emotional	Opinions / Qualities	Qualities & Characteristics	Sentiment tags	Subjective	Subjective
	Tag avoidance	n.n.	n.n.	n.n.	n.n.	n.n.	n.n.

Table 2.9.: Tag Categories 2

Heckner et al. 2008		Schmitz 2006	Sigurbjörnsson & van Zwol 2008	Dutta & Giunchiglia 2008	Mathes 2004	Lin et al. 2006	Dubinko 2006	Heyman et al. 2010	
Subject related tags	**Resource related**	n.n.	n.n.	Resource type	Genre	n.n.	n.n.	Objective & Content-based	Physical
					Photographic tags	Photographic			
				Resource level	n.n.	n.n.			
				Identifier					
				Format				Physical	
				Author					
				Fequency					
				Language		Language			
				Date		n.n.			
	Content Related	Depiction	Artifact/ Object	Topical terms	Technical tags	n.n.	Social media tagging	Objective & Content-based	
					Colors				
				Corporate name	n.n.				
				Meeting name					
				n.n.		Living Thing			
			Person/ Group	Personal name		Person	Personalities		
		n.n.	n.n.	Spatial		n.n.	n.n.		
				Temporal					
				Content categorization					
Non-subject related tags, personal tags	**Time & task related**	Place	Location	n.n.	Place names	Place-name	n.n.	Physical	
						Place-general			
		Event/ Activity	Time		Year	Time		n.n.	
			Event/ Action	Action	n.n.	Event	Event	Personal	Acronym
		n.n.	n.n.		Self organization / Reminder	n.n.	n.n.		
				Self-expose	Individuality / ego				
	Affective	Emotion/ Response	n.n.	Affective		Humor	n.n.	Opinion	
						Poetic			
						Rating			
						Emotion			
	Tag avoidance	n.n.	n.n.	Generic	n.n.	unknown	n.n.	Junk	
				Filter					

2.7.3. Folksonomy and Emerging Vocabulary

Social tagging systems give users the possibility to have their own view on the resources and to express their opinion or present themselves without any restriction (cf. Marlow et al., 2006). The users do not have to learn complex and predefined schemata or syntax, and problems of controlled vocabularies can be avoided (cf. Macgregor and McCulloch, 2006a). At first glance, this seems to result in a chaotic collection of bookmarks. However, several studies have shown that among the users a common understanding and vocabulary emerges from the tags and the process of tagging.

Indeed, tagging can be considered as an act of sensemaking (Golder and Huberman, 2006; Trant, 2009; Wu and Pinsonneault, 2008)). *"Sensemaking is about labeling and categorizing; [.. i. e.] imposing labels on interdependent events in ways that suggest plausible acts of managing, coordinating, and distributing"* (Weick et al. 2005). And in the act of tagging, the users impose labels (tags) on information resources. In this way, the users establish a relation between the information resource and a concept in their mind, i. e. linking it to mental schema of interpretation. By sharing tags, the users also share their interpretation of the information resources, which again enables the community through communication (where adopting tags is a form of communication) to converge on an approximate interpretation. In this way, tags become a form of collective meaning.

As mentioned in the Section 2.7.1, tagging distributions tend to stabilize into power law or inverse-logistic distributions where the first, very frequently used tags reflect the implicit consensus of the community[12] whilst long tail tags show the opinion of individual users illustrating the variety in semantics of the terms (Halpin and Shepard, 2006; Golder and Huberman, 2006; Peters, 2009). The study by Maier and Thalmann (2008) indicates that there's convergence to a stable core set of assigned tags and a decrease of changes in the set of dominant tags over time. Similar conclusions draw Cattuto (2006) and Tonkin et al. (2008): *"This suggests that consensus does exist in the aggregate. The terms that are most common tended to provide a reasonable description of the content of the site and remained constant over time"*.

Marlow et al. (2006) present additional interesting results that support the convergence on a shared vocabulary. They compared the tag vocabulary of 2500 Flickr users pairwise both with one randomly chosen unacquainted user and one of their contacts. Their results show that the overlap in common tags is much bigger between users and their contacts than with randomly chosen users. This might reflect that sensemaking is social; *"categories [..] are socially defined [.. and] have to be adapted to local circumstances"* (Weick et al. 2005). Thus, Golder and Huberman (2006) see reasons for vocabulary stabilization in imitation and shared knowledge. That means that users tend to confirm other users' tags instead of creating new ones. Additionally, the users share some background and knowledge leading to make same choices and *"the [shared] ideas and characteristics that are represented in tags are stable"*.

Sen et al. (2006) further investigated influencing factors – specific personal tendency and community influence – on personal tagging behavior (Sen et al., 2006; Sen, 2009). Figure 2.5 illustrates the relationship between community influence and the user's personal tendency by Sen (2009). *Personal tendency* covers factors like personal preferences, knowledge, experiences and habits, based on which the users choose the tags for their assignments. Further, investment and habits from the user's past tag assignments influence the user's future tagging behavior. Precisely, once assigned tags are an investment

[12]implicit because consensus as a group decision process requires negotiated agreement

in a personal vocabulary that is costly to change mid-course. Similarly, people tend to repeat activities previously often performed. Thus, users are likely to assign the same tags they have assigned in the past. The tag assignments of all users accumulate in the community's tagging activity to a community vocabulary. *Community influence* is based on the social proof theory: *"Social proof states that people act in ways they observe others acting because they come to believe it is the correct way for people to act"*. In their study, Sen et al. (2006) show how the community's tags (previously displayed) influence the user's tag selection towards consistency and thus the emergence of a shared vocabulary. The authors furthermore studied the impact of different tag selection algorithms on the influence of the tagging community. Precisely, they compared four algorithms: (1) *shared* displays randomly selected tags assigned by others to the target resource; (2) *shared-pop* displays the most popular tags assigned by others to the target resource; (3) *shared-rec* displays based on a recommender algorithm the most commonly assigned tags of the target and similar resources; (4) *unshared* does not show any community tags but the user's own previously assigned ones. Their results reveal that the previous display of community tags positively influences the tag reuse (with the highest reuse of 1.73 users per tag for the shared-rec algorithm in comparison to 1.10 users per tags for the unshared algorithm). Also interesting is the influence on the distribution of tag types that users choose to apply. Sen et al. (2006) distinguish subjective, factual and personal tags (see also Section 2.7.2). Within their study the shared-pop and shared-rec produced a quick and strong convergence towards factual tags. Whereas there was no such clear evidence of convergence for the other two groups, the shared algorithm tend to bias towards subjective tags as the dominant tag type and the unshared algorithm towards personal tags. The model, however, does not consider the lost or gain of user interest or advances in knowledge on the user's and community's tagging behavior. We will look at this issue in more detail in Section 4.3.

Maier and Thalmann (2008) confirm this community influence by anterior display with their study in which they show that the formation of the starting tag collection that is displayed to the users influences the direction of any following tag assignments. For instance, if the starting tag collection referred to content-related tags, then the participants predominantly added content-related tags; in case of context-related start tags, they added context-related tags as well. This shows, that system features increasing awareness for already used tags like tag clouds or tag recommendation can additionally support vocabulary convergence.

Halpin and Shepard (2006) suppose sub-class, synonymous and facet relationships in the folksonomy and propose a methodology to extract sub-class relationships based on tag co-occurrence graphs (Halpin et al., 2006; Robu et al., 2009). Kipp (2009) had a deeper look not only at the tag frequencies and distributions but also at co-occurrences as well. They show that there are indeed different spellings and notations and that contextually related tags cluster together but not, as expected, all terms of a concept (i. e. not all synonyms or spelling variations). This might be traced back to differing user groups that have their own distinct terminology. Extracting semantic structures from folksonomies is an emerging field of research and will be separately discussed in Section 3.2.4.1.

2.8. Advantages and Challenges of Folksonomies and Social Tagging Systems

As we have seen, folksonomy-based systems make collecting information resource a social experience by allowing the users to share their resources with others. Not only are

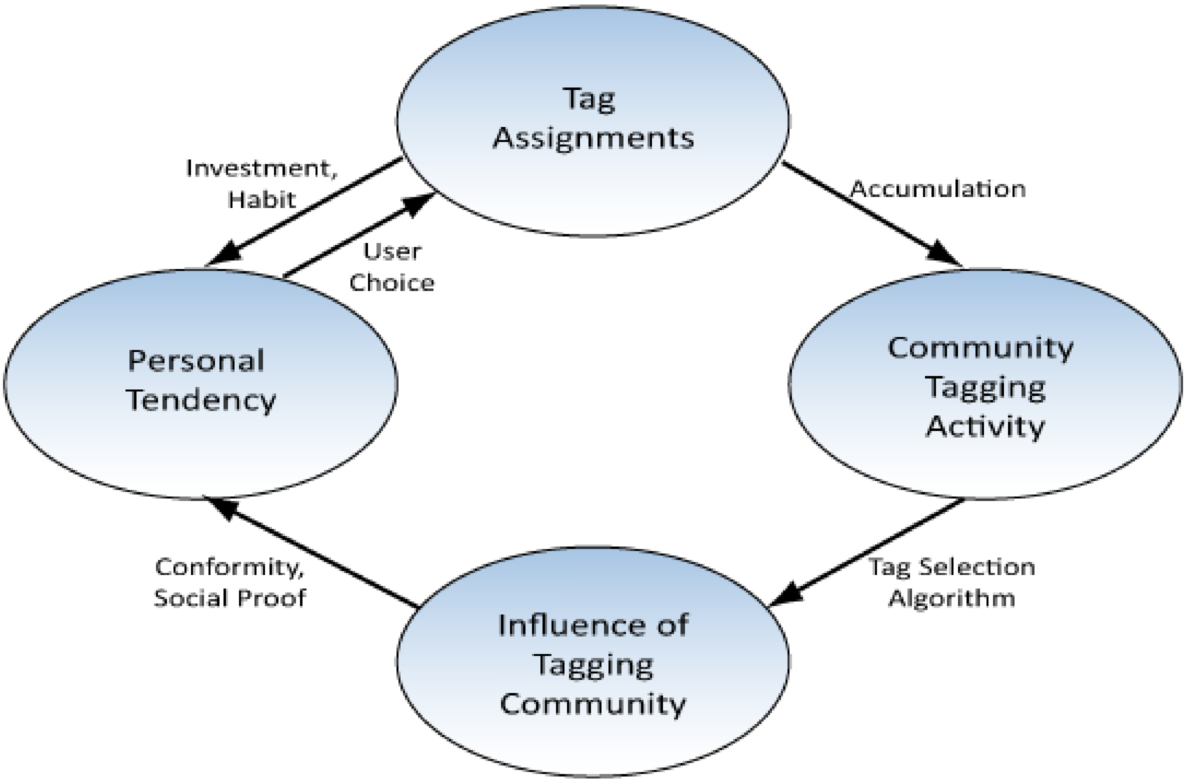


Figure 2.5.: Relationship between community influence and user's personal tendency by Sen (2009): *"tag applications* [a.n.= assignments] *accumulate to create a community's tags. Tagging systems use tag selection algorithms to select the tags shown to users. Based on these displayed tags, users form perceptions of community tagging norms. These norms influence the tags users apply. The tags users apply create an investment in a particular vocabulary"*

the resources visible to other users but also the tags used to describe them. That means, you can share your own tags and use the other users' ones. You can see which tags and annotated resources you have in common with other users or what they annotated with the same tags. In this way, you can find people with similar interests and discover new interesting resources. In the following we will summarize and discuss the advantages and challenges of folksonomies and folksonomy-based systems (cf. Mathes, 2004; Marlow et al., 2006; Shirky, 2005; Quintarelli, 2005; Kroski, 2005; Golder and Huberman, 2006; Guy and Tonkin, 2006; Hayman, 2007; Spiteri, 2007; Peters, 2009; Schmitz, 2007).

2.8.1. Advantages of Folksonomies and Social Tagging Systems

- **Meaningful words and concepts:** Folksonomies reflect the users' vocabulary and conceptual model because it's themselves who are the primary users and organizers of the resources. It matches their real needs and how they view the resources and not only the creator's ones. In contrast to controlled vocabularies, folksonomies are not excluding but incorporate everyone's words and terms regardless of background or viewpoint. Especially the long tail shows the semantic variety and alternative interpretations. In this way, folksonomies overcome the "vocabulary problem" (cf. Furnas et al., 1987) and semantic gap between different user groups (Smith, 2006).
- **Simplicity:** Entry barriers and cognitive costs are low for folksonomies. Folksonomies are easy and quickly to use and to understand. Users can create and share tags and resources in seconds without bigger investment in time or effort. Users can participate and assign tags without any restrictions on the vocabulary or any previous formal training on how to classify or index resources (Mathes, 2004). In this way, the cognitive load is lowered and "post activation analysis paralysis" avoided (Sinha, 2005). The users don't have to fear that they have chosen the wrong category for the resource and will never retrieve it again since they are free to choose as many categories as they like (Heckner et al., 2009).
- **Low Costs & Up-to-dateness:** In contrast to controlled vocabularies, folksonomies don't need to be developed in an upfront investment to their usage. They grow in an organic way and develop with their usage. The burden and costs of development and maintenance is distributed amongst many voluntaries. Additionally, folksonomies are flexible to respond immediately to changes in knowledge and terminology. They accommodate easily and quickly new concepts and neologisms.
- **Desire Lines & Emergent Semantics:** Folksonomies provide a source of information about users as they follows the users' information and discussion needs and desire lines (Merholz, 2004). Through tag reuse within the community, useful connections are formed. Tag frequencies, distributions and relations within the tripartite hypergraph can be used as a starting point and source for creating and maintaining controlled vocabularies and ontologies.
- **Serendipitous Discovery** Folksonomies broaden the access to information sources. Through browsing and exploring related users, tags and topics, users can discover and learn about unknown and unexpected resources, they would never have found through searching.
- **Community Aspects & Immediate Feedback:** Through the tags and resources, users are linked to other users with similar interests and viewpoint. In this way,

communities can emerge. When tagging a resource, users can immediately see what else is connected via the same tags. If this is not the expected result, they can modify the tags according to the community norm or retain them and try to influence the community. Thus, through the choice of tags, the users negotiate the meaning of terms and show their group cohesion. Quality control for resources takes place implicitly through the number of tag assignments and explicitly through evaluative tags.

2.8.2. Challenges of Folksonomies and Social Tagging Systems

Unsurprisingly, it's the linguistic and semantic diversity of folksonomies that is at the same time their biggest weakness because it hampers efficient information management and retrieval:

- **(Mis-)Spelling:** The most obvious problem is that tags are simply misspelled or written in different ways. There are variations in (1) plural and singular, (2) acronyms and abbreviations, compound words or word forms, e. g., 'spagetti' vs. 'spaghetti', 'noodle' vs. 'noodles', or 'spaghettiCarbonara' vs. 'spaghetti_carbonara'.
- **Multilingualism:** Tags only relate to one language. That means, especially in Europe with many different languages, users have to annotate a resource with many tags in different languages, e. g., with 'pasta', 'noodles', and 'Nudeln', in order to ensure that other users will find it later on (e. g., to promote their own great spaghetti recipe).
- **Polysemy:** Tags can have several similar meanings. This leads to search results with low precision because of irrelevant resources; e. g., with the tag 'pasta' the users can think of a dish that contains pasta as its main ingredient or of the aliment itself as shaped and dried dough made from flour and water and sometimes egg.
- **Homonymy:** The problem of homonymy is comparable to the problem of polysemy. However, in this case, one tag can have several totally different meanings. This also leads to irrelevant results as all resources that relate to these different meanings are annotated with the same tag. For instance the word 'noodle' can have the meaning of an aliment but also of a swearword for a human head.
- **Synonymy:** Resources are not found because they are annotated with another tag with the same meaning, e. g., with the tag 'vermicellini' instead of 'spaghettoni'. Similar to mulitlingualism, the users have to annotate the resources with many synonymous tags in order to ensure the retrieval by other users and likewise users searching for resources have to use different terms to find all tagged resources.
- **Mismatch of abstraction level:** Also a typical search problem emerges because tags are specified on different abstraction levels or levels of specificity, i. e. either too broad or too narrow. This problem, also known as the "basic level phenomenon" (Tanaka and Taylor, 1991), can be traced back to different intentions and expertise levels of the users. For instance, one user tags a resource on the basic level with 'spaghetti', another with 'noodles' and a third differentiates 'spaghetti' from 'bigoli' (thicker spaghetti) and 'vermicelli' (thinner spaghetti). A resource annotated with 'spaghetti', however, cannot be found with the search term 'pasta'.
- **Spam & Personal Tags:** Folksonomies are vulnerable to spam and malicious practice. There are users who use inappropriate or irrelevant tags or very popular tags

to promote their resources and in this way, produce noise and useless content. Similarly, tags with a personal meaning are of virtually no use for any other user.

- **Lack of Accuracy & Objectivity:** Due to simplicity or subjectivity, users might poorly choose and assign their tags or tag other users' resources differently than their own. Additionally, popular tags may dominate over time as users may not get aware of more precise terms available and tend to confirm existing tags.
- **Lack of Structure & Semantics:** Folksonomies are a flat set of tags without any hierarchy or structure. Similarly, the semantics when applying a tag remain unclear, e. g., if 'video' means that the resource is a or about a video. This hampers the maintenance of bigger tag vocabularies and retrieval precision and recall in general.

2.9. Conclusions

As we have seen, folksonomy-based systems make collecting information resource a social experience by allowing the users to share their resources with others. Not only are the resources visible to other users but also the tags used to describe them. That means, you can share your own tags and use the other users' ones. You can see which tags and annotated resources you have in common with other users or what they annotated with the same tags. In this way, you can find people with similar interests and discover new interesting resources. Additionally, these systems make it to minimize technical and organizational barriers: they are informal, lightweight, easy-to-use and easy to understand. This motivates users to participate in community activities and to express their opinions. On the other hand, their missing semantic precision and control hampers efficient information management and retrieval support, in particular in complex domains.

3. Ontologies and Ontology-based Knowledge Organization

In contrast to very lightweight knowledge organization systems, we have seen in the previous section, there are heavyweight systems based on ontologies that have established over the last 20 years.

3.1. Notion of Ontology

3.1.1. Ontology – Origin

The notion of *Ontology* has its origin in philosophy and denotes (Merriam-Webster, 2011):

1. *a branch of metaphysics concerned with the nature and relations of being*
2. *a particular theory about the nature of being or the kinds of things that have existence*

Dealing with the investigation of existence and the fundamental question of "what kinds of things are there?", it is concerned with general grouping and classifying of all existing things rooted in Aristoteles' *categories*[1]. For a detailed discussion of the philosophical perspective and its transformation to computer science we'd like to refer to, e. g., Smith and Burkhardt (1991); Heil (2003); Zuniga (2001) or `http://www.ontology-2.com/`. We now turn our head to our domain, which is computer science.

3.1.2. Ontology in Computer Science

Early in 1990s, ontology came up as a technical notion independently in the fields of data base systems, software engineering and artificial intelligence driven by the need for knowledge representation (Sánchez et al., 2007). Transferring the original philosophical meaning of *Ontology* to knowledge representation and computer science, information systems can benefit from the idea of ontological categorization (Grimm et al., 2011). In artificial intelligence, the term 'ontology' emerged to mean one of two related things (Chandrasekaran et al., 1999): (1) a representation vocabulary, specialized to some domain or subject matter; (2) a content theory describing some domain using a representation vocabulary. In the first case, ontology is intended in the philosophical sense as a conceptual framework at the semantic level, i. e. a synonym of conceptualization Guarino and Giaretta (1995). The second one refers to a concrete engineering artifact that encodes knowledge about a specific domain in a machine-interpretable form to make it available to information systems Grimm et al. (2011).

[1] Even though the term "Ontology" was coined later

There are several different definitions of what an ontology is since they are widely used for different purposes and by different communities (e. g., Swartout and Tate, 1999; Hendler, 2001; Jasper and Uschold, 1999; Kalfoglou, 2000). One of the first is the definition from an engineering-oriented perspective of Neches et al. (1991): *"An ontology defines the basic terms and relations comprising the vocabulary of a topic as well as the rules for combining terms and relations to define extensions to the vocabulary."*

Already Guarino and Giaretta (1995) found in 1995 at least seven different notions. At the end, they refer an ontology to an engineering artifact as *"a set of logical axioms designed to account for the intended meaning of a vocabulary"* (Guarino 1998) where the account is explicit and partial, i. e. an ontology implies some view of the world regarding the given domain. This logic-centered view excludes conceptual models without logical theory commitment.

The most cited definition is the one by Gruber (1993) that is *"an explicit specification of a conceptualization"*. Gruber (2009b) elaborates this as *"a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents"* in a domain of knowledge or discourse and defines a conceptualization as *"an abstract, simplified view of the world that we wish to represent for some purpose"* (Gruber 1993).

Whereas this definition as well as the previous one focus on explicitness and formality, Uschold and Grüninger (1996), on the other hand, emphasize the aspect of an shared understanding:

> *"Ontology" is the term used to refer to the shared understanding of some domain of interest [...] The [entailed or embodied] world view is often conceived as a set of concepts (e. g., entities, attributes, processes), their definitions and their inter-relations; this is referred to as a conceptualisation*

Studer et al. (1998) brought both definitions together[2]:

> *An ontology is a formal explicit specification of a shared conceptualization of a domain of interest. A conceptualization refers to an abstract model of some phenomenon in the world by having identified the relevant concepts of that phenomenon. Explicit means that the type of concepts used, and the constraints on their use are explicitly defined. [..] Formal refers to the fact that the ontology should be machine readable, which excludes natural language. Shared reflects the notion that an ontology captures consensual knowledge, that is, it is not private to some individual, but accepted by a group.*

In this way, the group clarifies and agrees on the concepts. These are captured and made explicit, e. g., written down in natural language. By coding in some representation language, the conceptualization is formalized in a way that it is machine interpretable and operationalizable. Complete formalization not possible. Degree of formality may vary.

3.1.2.1. Ontology Community and Commitment

Ontologies represent the manifestation of a shared understanding of a domain, i. e. how a specific community (of human and non-human agents) understands part of the world.

[2]Strictly speaking, Studer et al. (1998) refer to Borst (1997)'s notion of an *"formal specification of a shared conceptualization"* published one year after Uschold and Grüninger (1996)

This has to be agreed between the parties (Uschold, 1996). This agreement about the vocabulary, i. e. concepts and relationships used and talked about in the community, is also called *ontological commitment* (Gruber, 1993)[3]. That means, when committing to an ontology, there's an agreement on the meaning of the represented vocabulary elements and on using them in a coherent and consistent way. This includes axioms, rules, constraints etc. of an ontology. Thus, it's important that the community reaches an agreement between all parties on what their understanding and definition of the knowledge of their domain is as well as on how to represent this knowledge through an ontology.

3.1.3. Dimensions of Ontologies

Likewise there are different definitions and interpretations for the term ontology, there are various types and classifications of ontologies. Thus Uschold (1996) identified *"three key dimensions along which ontologies vary"*:

- *Subject Matter:* subject or scope of the ontology
- *Level of Formality:* degree of formality of a vocabulary and its meaning
- *Purpose:* intended application / use of the ontology

In the following, we will detail the varying scopes, formality levels, and use of ontologies.

3.1.3.1. Varying Scopes of Ontologies

An ontology can cover any topic and kind of knowledge. Nevertheless, ontologies can be classified according to the subject of the conceptualization they capture. The most established distinction is between *upper-level* (or *top-level*) ontologies, *domain ontologies* and *task ontologies*

- *Upper-level ontologies* describe very abstract and general concepts which are domain independent (e. g., space, time or event), thus usable across domains and applications.
- *Domain ontologies* attempt to capture the knowledge of a specific domain (e. g., medicine or mechanical engineering) by refining terms in upper-level ontologies.
- *Task ontologies* describe, in comparison with domain ontologies, a generic task or a sequence of problem solving steps (e. g., selling).

Mizoguchi et al. (1995) provides a very detailed typology subdividing these three main categories; e. g., domain ontologies into *object*, *activity* and *field* ontologies. Guarino (1998) adds to the three main categories *application ontologies* which bring together domain and task ontologies for knowledge-based application. Borst (1997) proposes a typology similar to Guarino (1998). van Heijst et al. (1997) provide a more detailed definition of the categories and suggest *representation ontologies* as an additional one, besides generic, domain and application ontologies, which describes representational entities without making a claim about what is represented. Uschold (1996) also uses the term meta-ontologies instead of representation ontologies, which replace the category

[3] Guarino (1998) formalized ontological commitment based on the connection of the ontology vocabulary and its terms with their conceptualization and meaning. In this way he defines it as a function that relates ontology vocabulary terms with a conceptualization

of upper-level ontologies. Fensel (2004) considers *metadata ontologies* that describe the content of online information sources as an additional category.

3.1.3.2. Varying Levels of Formality

Besides the scope, ontologies can be distinguished by their level of formality, or *"to which extent it is axiomatized by means of logical statements about the domain"* (Grimm et al. 2011). The Semantic Web community mostly speaks of *lightweight* and *heavyweight ontologies*. Lightweight ontologies include no or only few axioms, i. e. only concepts and taxonomies with relationships between concepts and properties[4]. Whereas heavyweight ontologies are lightweight ontologies to which axioms and constraints are added and extensively used; e. g., for range restrictions, disjointness or cardinalities to ensure logical consistency (Corcho et al., 2003).

Similar differentiations are proposed by Guarino (1998), Sowa (1996), and McGuinness (2003). Thus, Guarino (1998) discusses *coarse* and *fine-grained* ontologies where the former ones *"consist of a minimal set of axioms written in a language of minimal expressivity"*. Sowa (1996) speaks of *terminological* and *axiomatized* ontologies. Where an terminological ontology is *"an ontology whose concepts and relations are not fully specified by axioms and definitions"*. And an axiomatized ontology is *"a terminological ontology whose concepts and relations have associated axioms and definitions that are stated in logic or in some computer-oriented language that can be automatically translated to logic"*.

McGuinness (2003) differentiates *simple* and structured ontologies where simple ontologies need to have a *"finite controlled vocabulary, unambiguous interpretation of classes and term relationships and strict hierarchical subclass relationships between classes"*. Structured ontologies have more structure and contain for instance, apart from machine-readable encoded hierarchical relationships, information about properties and value restrictions on the properties.

Indeed, we can talk about a whole spectrum of ontologies with highly informal and low structured approaches like catalogs at one end and rigorously formalized logical theories at the other end. As we move along the spectrum, the amount of meaning and structure specified and the degree of formality increases; i. e. ambiguity decreases and support for automated reasoning increases (cf. Uschold and Gruninger, 2004).

According to Uschold and Grüninger (1996) we can identify four degrees of abstraction along the spectrum: ontologies which are

- *highly informal* if they are expressed in natural language[5],
- *structured informal* if they are expressed in a restricted and structured form of natural language,
- *semi formal* if expressed in an artificial formally defined language and
- *rigorously formal* if they provide meticulously defined terms with formal semantics, theorems and proofs of properties such as soundness and completeness

[4] Mika (2005) calls ontologies which consist in an ensemble of terms connected with a limited set of semantic relationships (broader, narrower, related) lightweight ontologies

[5] Note that a highly informal ontology would not be considered as an ontology in the sense of Studer et al. (1998) definition because it's not machine-readable

(cf. Gómez-Pérez et al., 2004).

One of the first perspectives on an ontology spectrum was the outcome of a discussion about what is referred to as an ontology at the AAAI Ontologies Panel in 1999 by M. Gruninger, F. Lehmann, D. McGuinness, M. Uschold and C. Welty (Welty et al., 1999). McGuinness (2003), Smith and Welty (2001) and Uschold and Gruninger (2004) expanded upon this ontology spectrum view; whereas McGuinness (2003)'s perspective may be the most cited one. Interestingly, even though all three of them provide more or less the same classification, they differ in the spectrum's focus. Thus, it is only Uschold and Gruninger (2004) who speak of formality, whereas Smith and Welty (2001) use a vague notion of complexity and McGuinness (2003) the notion of expressiveness.

Thus, Smith and Welty (2001) present the spectrum with seven ontology classes shown in Figure 3.3, comprising simple catalogs with pairs of unique identifier and product type (term), set of natural language text files, glossaries, thesauri, taxonomies, frame-based ontologies and logical theories (higher order, full first order or modal logics). The perspectives of McGuinness (2003) and Uschold and Gruninger (2004) are quite similar. McGuinness (2003) refined the original picture and proposed a simple spectrum of ten ontology classes. The spectrum is along the range of expressiveness shown in Figure 3.2 running from simple term lists or catalogs to complex reasoning models (e. g., with disjointness). With expressiveness, McGuinness (2003) links the classification to ontology language constructs (apart from the spectrum's left end which are indeed knowledge organization systems). Indeed, the ontology is inherently connected to its language and its expressiveness, however one can also use a highly expressive language such as OWL to represent only classes and equivalence relations between them.

In comparison to the previous ones, Uschold and Gruninger (2004) explicitly speak of a continuum of formality. Figure 3.1 shows the continuum introduced by them and slightly adapted by Guarino et al. (2009). Whereas McGuinness (2003) draw a strict line between informal and formal ontologies based on the explicit formalization of strict super-/subclass relations, it is indeed difficult to claim where the criterion of formal starts on this continuum (Guarino et al., 2009).

Obrst (2003) describe a complementary classification from the semantic interoperability point of view and with a focus on expressing hierarchical relationships. The spectrum in fig. 3.4 ranges from taxonomies with subclassifications but without strict super-/subclass hierarchies and only providing syntactic interoperability, passing thesauri (providing structural interoperability) and conceptual models to logical theory models that have the strongest semantics to provide the highest form of semantic interoperability.

Bullinger (2008) provides a classification framework, called 'OntoCube', along the dimensions of 'subject matter', 'formality' and 'expressivity'. She locates taxonomy, thesaurus and topic map as lightweight ontologies and as separate group heavyweight ontology on the dimension of expressivity. Whilst she relates the dimension of formality, on the basis of the four degrees of abstraction by Uschold and Grüninger (1996), to the knowledge representation language and divides it into *informal / natural language*, *semi-informal notation*, *semi-formal notation* and *formal language / markup*.

Indeed, we refer expressivity (or expressiveness), i. e. the required expressive power or range of constructs, to the representation language that is used to specify an ontology. For instance, we can use a list of terms and definitions in a natural language, e. g., English, to specify an informal ontology. In this way, we can compare ontologies based on the languages with the minimal expressivity that is required to define the ontologies' vocabularies. For instance, a (formal) taxonomy only requires a restricted language to specify super-/subclass relationships, even though it might be specified in a highly expressive language like OWL (Gruninger et al., 2008).

Spectra Comparison Being the same at a first glance, we can observe the approaches differ not only in their level of detail or their focus labeling (formality, complexity, expressiveness) when comparing the different spectra.

For instance, they differ in what is considered the lowest level of formality on the spectrum. Whilst Obrst (2003) and Bullinger (2008) view (informal) taxonomies with hierarchical relationships as the lowest level of formality, preceding work, e. g., by McGuinness (2003); Uschold and Gruninger (2004); Smith and Welty (2001), include lower structured vocabularies like catalogs and term lists[6].

Additionally, we can observe different perspectives on where to locate taxonomies and thesauri on the continuum. Whereas Smith and Welty (2001) consider taxonomies as heavyweight ontologies and locate them right to thesauri, Obrst (2003) locates them to the left. Indeed, we can differentiate between simple/hierarchical taxonomies and formal taxonomies[7]. Simple taxonomies structure terms thus they have a parent/ broader relation or child/narrower relation with other terms. As such taxonomies sometimes informally make use of only is-a (sub-/superclass) relations, some authors speak of informal (is-a) hierarchies. Formal taxonomies provide explicit is-a relations.

Similarly, there are two perspectives on thesauri. Thus, McGuinness (2003) and Uschold and Gruninger (2004) bear on an unspecific definition of thesauri that provide a list of terms with meanings and merely synonymous relations between the terms, even though there are additional but, in their opinion, rarely used relations. Indeed, from an information science perspective, thesauri provide, besides synonymous relations (SYN), broader (BT) / narrower (NT) relations and related (RT) relations to describe association or relatedness between terms. Thesaurus standards further specify the broader-narrower relation by specific generalization (BTG/NTG), instance (BTI/NTI) and whole-part (BTP/NTP) relations. Whilst thesauri traditionally have been represented in a term-based model, they are nowadays concept-based. With SKOS (Simple Knowledge Organisation System) (see Section 3.1.3.3) the W3C has came up with a hitherto missing formal language to model concept-based thesauri. Weller (2007) provides a modified version of the spectrum by McGuinness (2003) from the information science's perspective (see fig. 3.5) where she differentiates thesaurus with unspecific definition and thesaurus with information science definition.

We can make another interesting observation regarding formal taxonomies and frame-based ontologies. Thus, Smith and Welty (2001) attribute sub-/superclass relations, explicit instance and class distinction as well as properties and their inheritance to formal taxonomies and restrictions as a higher formality level to frame-based systems. In contrast, McGuinness (2003) makes a finer distinction. She treats explicit sub-/superclass is-a, instance relations, properties and their inheritance and restrictions as separate levels with increasing expressiveness. However she determines the explicit super-/subclass relation with the inheritance predicate based on instances even though instances are the next level. She attributes properties and their inheritance to frame-based systems but not restrictions.

[6] Bullinger (2008) state that she excluded vocabularies with lower formality level because they don't fulfill the properties of McGuinness (2003) notion of simple ontologies. Indeed, her definition of taxonomy does not follow the need for subclass relationship, i. e. strict is-a, as well.

[7] This differentiation is implicitly done by Uschold and Gruninger (2004)

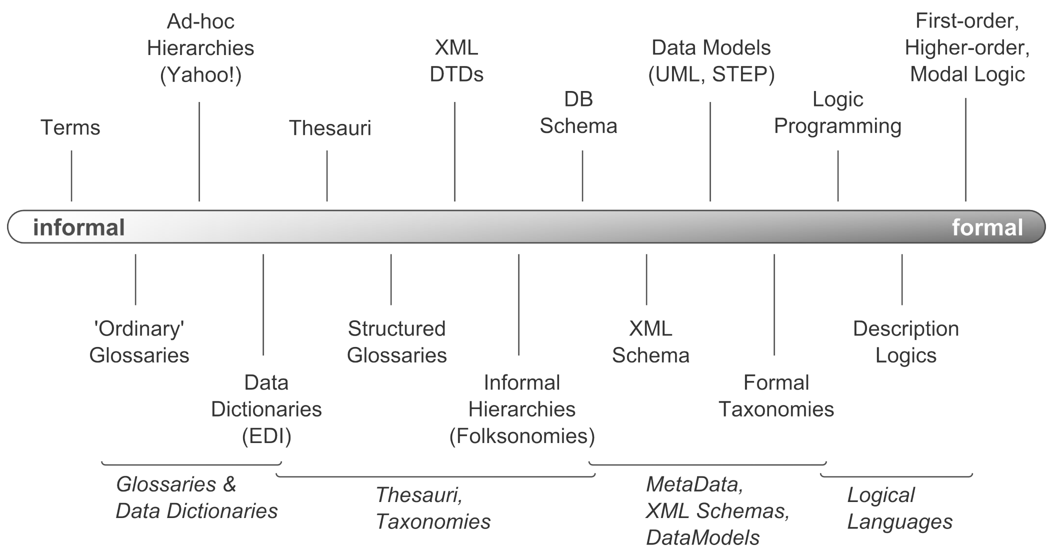


Figure 3.1.: Ontology Spectrum by Guarino et al. (2009), adapted from Uschold and Gruninger (2004)

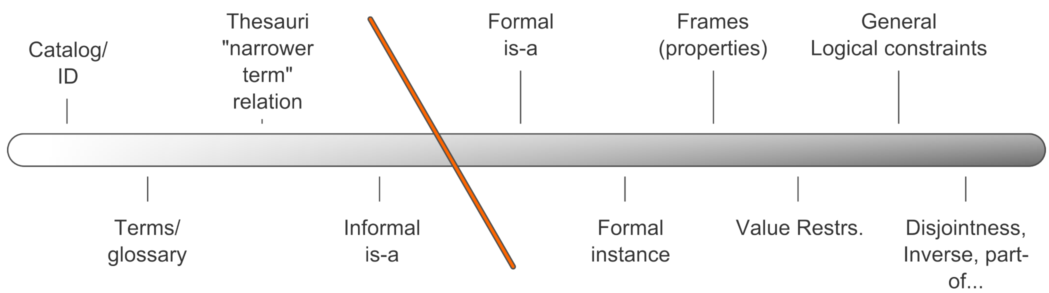


Figure 3.2.: Ontology Spectrum by McGuinness (2003)

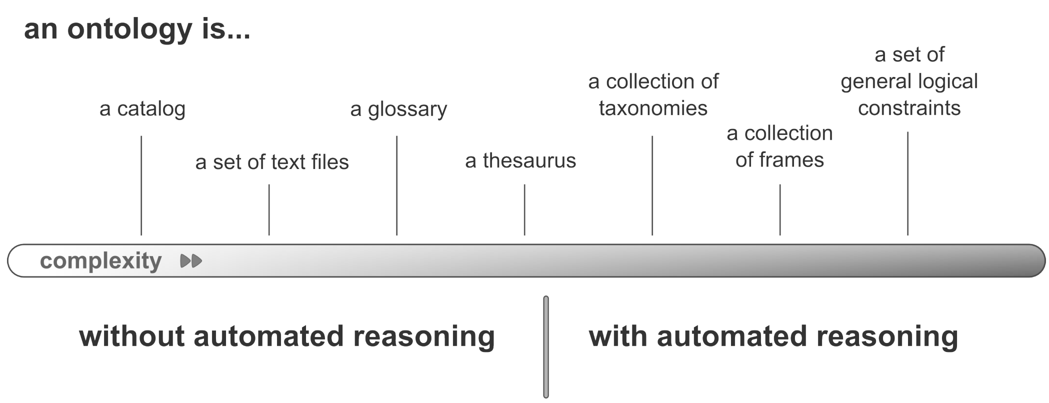


Figure 3.3.: Ontology Spectrum by Smith and Welty (2001)

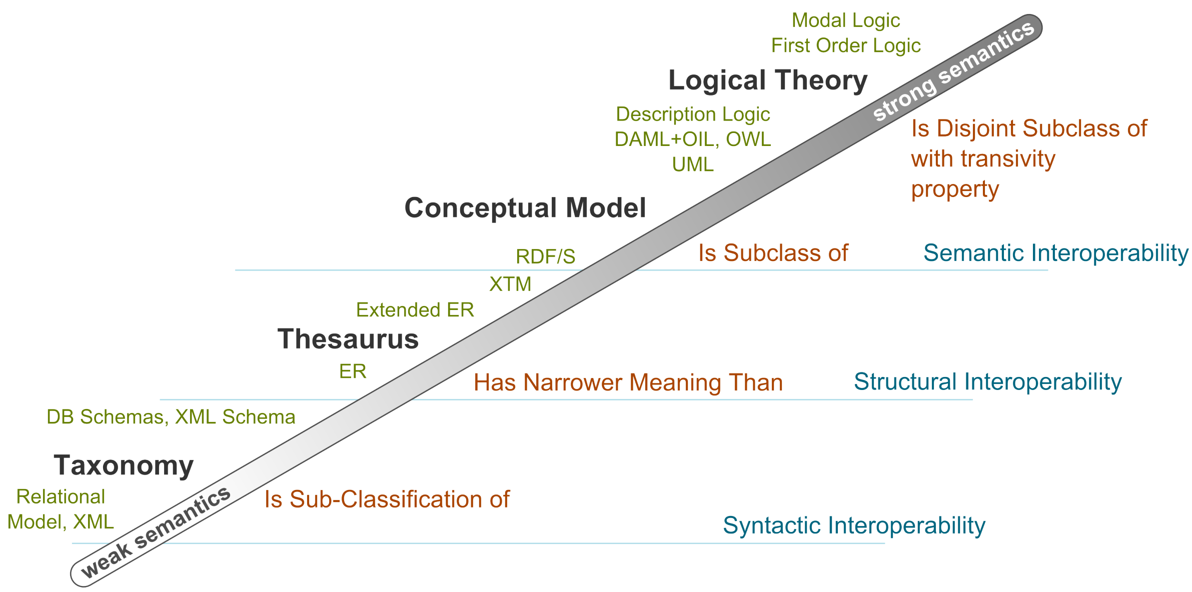


Figure 3.4.: Ontology Spectrum by Obrst (2003)

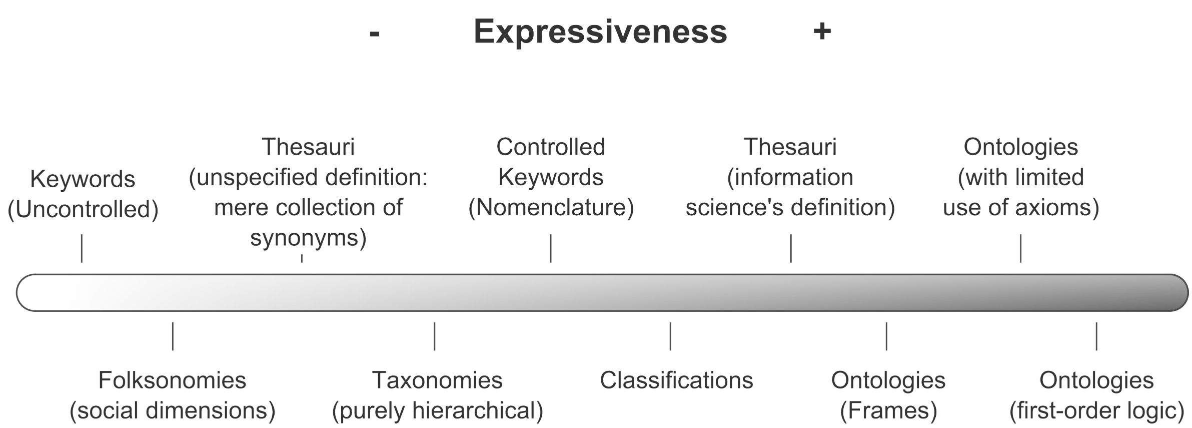


Figure 3.5.: Ontology Spectrum by Weller (2007)

Ontology Spectrum In the following we provide an ontology spectrum that consolidates the existing categorizations on the basis of Lacasta et al. (2010).

- **Controlled Vocabularies & Catalogs:** A controlled vocabulary may be seen as the simplest notion of a possible ontology. It's a finite but extensible list of terms about a certain subject that have an unambiguous definition in a form of term-identifier pairs. Catalogs are an example of controlled vocabulary in which each object has a unique identifier, e. g., a library has a list of themes and each of the them is associated with a unique code that is used to classify books.
- **Glossaries & Dictionaries:** Glossaries are an alphabetically ordered list of terms and their meanings and usually cross-references in a specific language. They are slightly more complex than catalogs as they provide human directed natural language descriptions of terms and in this way imposing some structure on text, i. e. the text is annotated with terms. However, the terms are not unambiguous, thus not machine-interpretable. Dictionaries are an enriched form of glossaries. Similarly, they consist of a list of terms and definitions but may also include synonyms, spelling variants,variant senses and other information. Dictionaries with translations in another language turn them into lexicons.
- **Subject Headings & Simple Taxonomies**: A subject heading is a set of controlled terms to describe the subject of collection items, e. g., in a library. Subject headings have a limited hierarchy resulting through the combination of terms according to strict rules into more specific concepts. Simple taxonomies structure terms (poly-)hierarchically thus they have a parent / broader relation or child / narrower relation with other terms.
- **Thesauri:** A thesaurus, with the main purpose to provide a standard vocabulary for indexing, includes besides the list of terms and their description hierarchical, associative and equivalence-based relations among the terms. These relations are a priori defined, i. e. broader / narrower, related and synonymous, and may be more granular in some cases; e. g., specifying generalization, instance and whole-part relations. Nowadays, thesauri are based on concepts with a unique identifier to distinguish them. Thus, broader / narrower and related relations aren't anymore relations among terms but concepts. Terms are associated to a concept as preferred or alternative; i. e. synonymous, labels. The labels may contain a language code which allows to manage multilingualism. The same is true for descriptions.
- **Semantic Networks:** A semantic network is a directed or undirected graph with concepts as nodes referred to by a set of terms and variable semantic relationships as edges between them. They go beyond the standard thesaurus broader/narrower and related relations and may include specific hierarchical, antonymy or causality relations. WordNet, a lexical English database, is an example of a semantic network.
- **Formal Taxonomies & Formal Instances:** Formal taxonomies include strict is-a relations, based on the notion of subsumption. Being transitive, every more specific node is also more specific to the node to which the more general one is more specific, too and an instance of a more specific node is also an instance of the more general one. Indeed, the is-a relation is not as precise as we may have thought and we can further differentiate two subtypes: an is-a relation that connects two generic nodes and an is-a relation that connects a generic with an individual node[8]. The intent of the first subtype usually means that one generic is less general than the

[8]Regarding the network, individual nodes are the leaves and generic nodes are the internal ones

other (sub-/superclass), whereas the second subtype usually is that the individual node can be described by the generic's node description (instantiation)[9] (Brachman, 1983). For example, we can associate an individual as a member or *instance* of the generic one. Therefore, the notation *instance-of* is commonly used for the individual/generic relation and leaving *is-a* for the generic/generic relation.

- **Frame-based Ontologies:** Frame-based ontologies include classes and properties (relations and attributes[10]) that are inherited by subclasses and instances. Classes are represented by *class frames*, individuals by *individual* or *instance frames* and properties by *slots*. The is-a relation between classes (*subclass-of*) has the meaning of specialization and instance-of means that the individual is an element of the class. Multiple inheritance is possible as well as a hierarchy of properties. *Facets* are used to specify cardinality and restrict values on properties.
- **Description Logic Ontologies:** Description Logic ontologies model concepts, roles and individuals. *Axioms* – logical statements that relate concepts and/or roles – play a fundamental role and is also the key difference to frame-based ontologies where these are interpreted as frame specifications that declare and completely define a class. Description Logic ontologies include much more complex constraints than frame-based ontologies such as disjointness, union of or negation.
- **General Logical Theories:** These ontologies are highly axiomatized using full first order, higher order, modal logic, or temporal logic. *" A general logical theory is not restricted to the derivation of facts at the instance level, but also captures a rich axiomatisation about classes and properties at the schema level, allowing for drawing conclusions about general situations in the domain in form of complex axioms."* (Grimm et al. 2011) Many Description Logic ontologies can be seen as a decidable fragment of first order predicate logic.

This spectrum also shows the usability or functionality of ontologies. Whereas semantically representations with formal axiomatizations provide powerful reasoning capabilities, less formal ones are much more easier to develop and maintain and provide far smaller computational costs. Indeed there are different uses of ontologies that require ontologies of different formality.

3.1.3.3. Ontology Languages

Especially in the context of Semantic Web research, there has recently evolved a broad range of specialized ontology representation languages. We will sketch the most prevalent languages only roughly but, for the purpose of this work, have a more detailed look at *SKOS* for modeling lightweight ontologies. For a detailed discussion, we would like to refer the reader to, e. g., Staab and Studer (2009); Gutierrez-Pulido et al. (2006); Gómez-Pérez et al. (2004).

The most prominent ontology languages may be *RDF(S)* and *OWL* – all developed by the W3C. RDF (Resource Description Framework, Klyne and Carroll 2004) started early in 1997 as an effort for the standardization of meta data to describe Web resources with resources – which are referred to as URIs–, properties – which define attributes or relations –, and statements – which assign values to a property for a specific resource. RDF Schema (Brickley and Guha, 2004) was built as an extension to RDF providing additional

[9]We speak of usually because there are indeed additional interpretations of each subtype. See Brachman (1983) for an intensive discussion.

[10]relations connect classes and/or instances whereas attributes involve data values

modeling primitives. Together they allow instantiation, interrelation and subsumption. Axiomatization is restricted to subclassing, typing, and domain and range restrictions. *OWL* (Web Ontology Language, Patel-Schneider et al. (2004)) is a whole family of languages with increasing expressiveness (driven by description logics) on top of RDF(S). They include includes features like disjunction and conjunction or existential and universal. *OWL 2* is the newest version.

A different group of ontology languages from logic programming form the *rules languages*. These are based on rules with an if-then-reading. Examples are *Datalog* (Ullman, 1990) or *F-Logic* (Frame Logic, Kifer et al. 1995) that integrates frame-based languages and first-order predicate calculus including objects and their identities, inheritance, polymorphic types, query methods, and encapsulation. Whilst the previous ones are not tailored for Web standards, the Semantic Web Rule Language (*SWRL*, Horrocks et al. 2004) builds rules into OWL ontologies.

There are also languages that aim to represent lightweight ontologies. In contrast to the previous expressive knowledge representation formalism these languages provide rather few but easy to use semantic features. For instance, the standard *Topic Maps ISO/IEC 13250*[11] can be used to model semantic networks. A topic map consists of topics that represent abstract subjects and refer to objects or nodes in the map (network), scopes, associations between topics, and occurrences that assign topics to external information resources. Topic types form classes to group topics together. Associations can be freely defined. Another approach is the Simple Knowledge Organisation System (SKOS) to represent Knowledge Organisation Systems. Because of its relevance for our solution, we will discuss this language in more detail.

Simple Knowledge Organisation System (SKOS) for Lightweight Ontology Representation SKOS (Brickley and Miles, 2005; Miles and Bechhofer, 2009) is a language to represent Knowledge Organisation Systems (KOS) like thesauri, simple taxonomies, subject headings etc. It is based on RDF and RDFS to publish and share the latter concept schemes in a machine-readable form for the Semantic Web. It has its origin in the thesaurus activity of the Semantic Web Advanced Development for Europe (SWAD-E) project *"as draft of an RDF Schema for thesauri compatible with relevant ISO standards"* (Miles et al. 2005)[12]. SKOS Core was presented the first time in 2003 and from then on further developed, e. g., to also represent other KOS. Since 2009, SKOS is a W3C recommendation that formally defines the SKOS data model as an OWL Full ontology and structures SKOS data in RDF triples.

The model's center is the concept, the *skos:concept*, that represents an idea or notion. Such a concept has some lexical properties and may be connected via semantic relationships to other concepts. A concept scheme (*skos:conceptScheme*) aggregates one or more concepts to represent a knowledge organization system and provides some additional metadata like the author. Both SKOS concepts and SKOS concept schemes are identified and linked to by URIs. The relation *skos:inScheme* indicates that a concept belongs to a concept scheme[13]. Vice versa the relation *skos:hasTopConcept* links the concept scheme to the topmost concepts heading the hierarchical structure within the concept scheme itself. A concept scheme may have one or more top concepts; with e. g., in a flat ontology skos:hasTopConcept relations for each concept.

[11] http://www1.y12.doe.gov/capabilities/sgml/sc34/document/0322.htm

[12] The basis provided a generic RDF schema for thesauri from the DESIRE (http://www.desire.org) and LIMBER project (Miller and Matthews, 2001)

[13] A concept may be part of several concept schemes

A concept can have labels, that are lexical strings, of any number and in any language. Only one label for each language can be related to concept as preferred label using *skos:prefLabel*. The other labels (e. g., synonyms, spelling variants, acronyms etc.) are referred to as alternative labels with the *skos:altLabel* relation. There are hidden labels (connected with *skos:hiddenLabel*) as a third label type. Hidden labels are similar to alternative labels and used e. g., to cover spelling errors of other labels. There is an optional extension, the SKOS eXtension for Labels (SKOS-XL) for more support on identifying, describing and connecting labels; e. g., to make explicit links between the labels of a concept. Notations (*skos:notation*), lexical codes, allow to uniquely identify a concept within a specific concept scheme and can be used as bridge to existing classification or identification codes, e.g the Universal Decimal Classification.

Various types of notes allow to document informal information regarding the meaning, evolution over time, examples or usage scope notes. There are seven specific properties with *skos:note* used to for general documentation purposes and the others defined as specializations of that.

Concepts may be linked with semantic relationships. SKOS provides a general relationship, the *skos:semanticRelation*, and two basic types for hierarchical and associative relationships that are specializations of the general one. *skos:broader*[14] and *skos:narrower*[15] are used to model hierarchical structures between concepts. They indicate that one concept is more specific/general than the other. *skos:broader* and *skos:narrower* are inverse but not transitive. Transitive hierarchical relationships are covered by *skos:broaderTransitive* and *skos:narrowerTransitive*. *skos:related* is used for associative relationships between concepts. It's a symmetric (but also not transitive) relationship and indicates that two concepts are connected in some way.

SKOS intentionally does not provide finer semantic granularity for the basic hierarchical relationships, e. g., instance-class or part-whole, but they can be extended by declaring sub-properties, e. g., for skos:broader:

```
skosExt:broaderGeneric rdfs:subPropertyOf skos:broader
skosExt:broaderPartitive rdfs:subPropertyOf skos:broader
skosExt:broaderInstantive rdfs:subPropertyOf skos:broader
```

These were part of the 2004 SKOS extensions (Miles and Brickley, 2004) and it is currently under discussion to provide them as official extension for the next version of SKOS.

Besides the aggregation in a concept scheme, concepts can be grouped together into labeled and/or ordered collection, e. g., to indicate that these concepts share something in common. *skos:collection* is used for general collections and *skos:orderedCollection* for collections where the order of the elements is relevant. The relationships *skos:member* for *skos:collection* and *skos:memberList* for *skos:orderedCollection* connect the contained concepts with the collection. Collections may also contain other collections.

SKOS additionally provides functionality for mapping between concepts in different concept schemes. There is a general property, the *skos:mappingRelation*, and four basic types of mapping properties: *skos:closeMatch*, *exactMatch* *skos:broadMatch*, *skos:narrowerMatch*, *relatedMatch*. These are specialization of the general one and can be used, e. g., in case the correspondence between two concepts is not exact.

[14] read as "has broader concept"
[15] read as "has narrower concept"

3.1.3.4. Uses of Ontologies

In this section we will have a look at possible uses and intended applications of ontologies. A general characteristic of ontologies, grounded in their use, is that they are descriptive. They capture the observed knowledge about the underlying structure of a domain. Unlike prescriptive models, they are used, i. e. queried, and modified at information system's runtime (Grimm et al., 2011).

Uschold and Grüninger (1996) sub-divide the space of ontology uses into three main categories; i. e. communication, interoperability, and systems engineering. Gašević et al. (2009) add education as an additional category to Uschold and Grüninger (1996), which might also be seen as a specific form of communication:

- **Communication & Collaboration**: When people work together in organizations, especially in interdisciplinary teams, they have different views on the work, tasks and problem areas because of varying backgrounds, expertise and interests. Using ontologies can help to integrate the different user perspectives to a unified knowledge reference skeleton and with this to a shared understanding, easing communication and collaboration. To come to a shared understanding requires assistance for negotiation and agreement finding.

  From an education perspective, by using ontologies experts can share their understanding of the domain. Similarly, ontologies can help to learn about a domain as they provide consensus and reference of the domain they represent.

  Ontologies may not only be used for communication between humans but also between intelligent agents or computational systems (Gruninger and Lee, 2002). In this case, using the same ontologies as world models help agents to interpret messages received in the right manner (Hadzic et al., 2009).

  The third communication form is this between humans and agents. An ontology-based environment that can bring together, i. e. represent, the different meanings in the vocabulary of different people can help in problem situations. Situations that arise when the world view of the user (and with this the vocabulary) differs from the one of the system.

- **Interoperability**: Ontologies enable interoperability between systems. For instance, they can act as auxiliary language between software systems or components, thus only one translator per system (from the system's language to the ontology) is necessary instead of translators for every system pair. Similarly, ontologies can help in integrating information in different formats or level of detail from different sources.

- **Systems engineering**: Ontologies may also support the design and development of software systems; i. e. specification, reliability and reusability. For instance, formal ontologies can provide a declarative specification of software systems that can be used for reasoning or automated consistency check to improve reliability. Informal ontologies help in identifying requirements, understanding component interplay or as a basis in manual consistency checking. Making assumptions on domain and tasks characteristics on which system components rely explicit, enables reusability.

Gruninger and Lee (2002) see the uses of ontologies, besides communication, on the one hand for computational inference, i. e. *"for internally representing and manipulating plans and planning information"* and *"for analyzing the internal structures, algorithms,* *inputs and outputs of implemented systems in theoretical and conceptual terms"* and on the other hand for reuse (and organization) of knowledge, i. e. *"for structuring and organizing libraries or repositories of plans and planning and domain information"*.

We may also take a closer look at the latter point in the light of information retrieval for what ontologies are applicable for (Nagypál, 2007b; Lacasta et al., 2010):

- *Resource categorization*: That is to simplify the creation of semantic annotations for resources for their organization and better retrieval as they can be matched with queries semantically.

- *Information search*: That is to support the user a) in query formulation, thus the user gets to know which information is available and in which connection, and b) with query expansion in order to improve the query specification and also to find additional interesting ontology entities.

- *Information visualization and navigation*: That is the ontology can act as core structure for visualizing and browsing information. Navigational elements can be provided based on the structure. Together with semantic annotations, additional or related information can be displayed making it easier to explore and understand resource contents and connections.

Grimm et al. (2011) summarize the uses of ontologies in the context of the semantic web in four generic main categories. These are knowledge & information organization, search, integration (of information from different sources) and formal processing (cf. computational inference), whereas the latter one may also take effect for the other uses. The authors map these to the four definitional main aspects of ontologies; i. e. formality, explicitness, consensus (ref. shared understanding) and conceptuality (ref. conceptualization). That's done through five previously identified basic benefits which are "*standardization of representation languages*", "*standardization of metadata schemata*", "*provision and/or standardization of domain knowledge*", "*rich declarative modeling language*" and "*machine-processable semantics*" and to which the definitional main aspects variably contribute.

3.2. Ontology Engineering

Ontology Engineering can be defined as *"the set of activities that concern the ontology development process, the ontology life cycle, and the methodologies, tools and languages for building ontologies"* (Gómez-Pérez et al. 2004). Methods and methodologies for ontology engineering provide a framework of guidelines and rules that help to break down the complexity of building and maintaining ontologies, to speed up the process and to avoid modeling errors.

There are numerous methods and methodologies for ontology learning, ontology evaluation or ontology merging (Staab and Studer, 2009). Automatic and semi-automatic methods for ontological knowledge acquisition, e. g., from texts, multimedia resources or folksonomies, support manual development approaches. In this section we will focus on manual approaches. We will present traditional methods and methodologies that are followed by collaborative approaches. Because of the thesis focus, we will give an overview of methods that extract semantics of folksonomies at the end of this section.

3.2.1. Ontology Design Criteria

Gruber (1995) developed a set of guiding criteria for the design of ontologies with an emphasis on knowledge sharing and reuse. Uschold and Grüninger (1996) further elaborated these criteria set in terms of enterprise modeling. The set consists of five criteria:

- **Clarity**: The ontology terms and meanings should be clearly communicable; i. e. possible interpretations should be restricted where needed. Definitions should be objective and documented in natural language. Where possible, they should be complete and specified in formal axioms.

- **Coherence**: Inferences need to be consistent with ontology definitions and axioms. An ontology should also be coherent with the informally defined parts such as the documentation and examples in natural language.

- **Extendibility**: An ontology should anticipate the uses and a range of tasks. When new knowledge emerges, the ontology should be easy to extend or specialize based on the existing vocabulary.

- **Minimal Encoding Bias**: The specification of the conceptualization should be at a knowledge not at symbol-level. Latter might result because of convenience of notation or implementation.

- **Minimal Ontological Commitment**: The required ontological commitment should be minimal but sufficient to support the intended knowledge-sharing activities. Thus making as few claims as possible to provide effective communication that is consistent with the conceptualization and easily extensible by third parties.

Missikoff et al. (2002) highlight *coverage*, *consensus* and *accessibility* as main principles to build usable domain ontologies:

- **Coverage:** The ontology must include domain concepts sufficient with respect to application purposes

- **Consensus:** The community must reach a consensus and common view of the domain that the ontology must reflect

- **Accessibility:** The ontology must be easily accessible, e. g., easy to integrate in an application

For further reading, we like to refer the reader to Vrandecic (2010) who analyzed ontology design criteria in detail and provides a summarized set of eight criteria subsuming all in the literature.

For our purposes, we will analyze the ontology developments methods and methodologies presented in the following by means of the following criteria derived from the five main principles to answer our research question (see 1.2). Figure 3.6 illustrates the mapping between our principles and the criteria.

- **Shared Understanding:** We will have look at if and how the methods and methodologies support the development of a shared understanding among all stakeholders

- **Role & Involvement of Users:** Regarding the aspect of active participation, we will examine the role and involvement of the different parties; knowledge engineers, domain experts and end users.

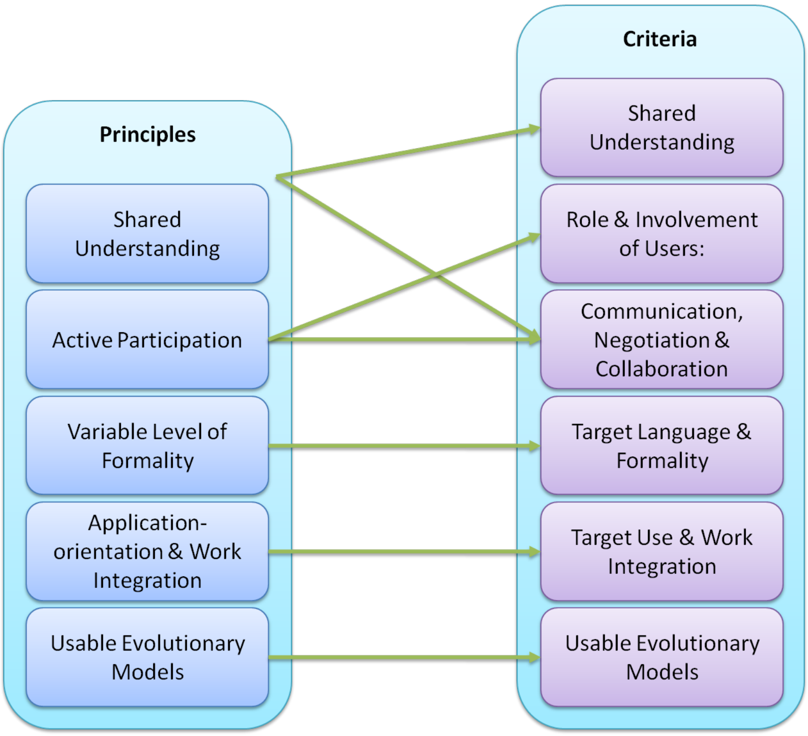


Figure 3.6.: Mapping of criteria to the five main principles.

- **Communication, Negotiation & Collaboration:** In order to achieve active participation and the development of a shared understanding, communication, negotiation and collaboration support is needed.
- **Target Language & Formality:** More participations requires lower formality. Additionally, high formality often is not needed. So we examine the language and formality the methods and methodologies target.
- **Target Use & Work Integration:** Fostering motivation and the continuous evolutionary development of the shared understanding requires an applications-oriented and work-integrated ontology development.
- **Usable Evolutionary Intermediate Models:** We will look at if and how the methods and methodologies support evolutionary ontology models that are already usable in their intermediate state.

3.2.2. Knowledge Engineer-driven Ontology Development

3.2.2.1. Generic Methodology

In the core, most methodologies comprise three core activities to engineer ontologies (Uschold and Grüninger, 1996; Gómez-Pérez et al., 2004; Pinto and Martins, 2004; Simperl and Tempich, 2006; Nagypál, 2007a; Grimm et al., 2011):

- **Domain analysis**: At first, the domain is analyzed and requirements are specified by ontology engineers and/or domain experts. The requirements include the ontology scope and purpose together with the intended end-users and level of formality.

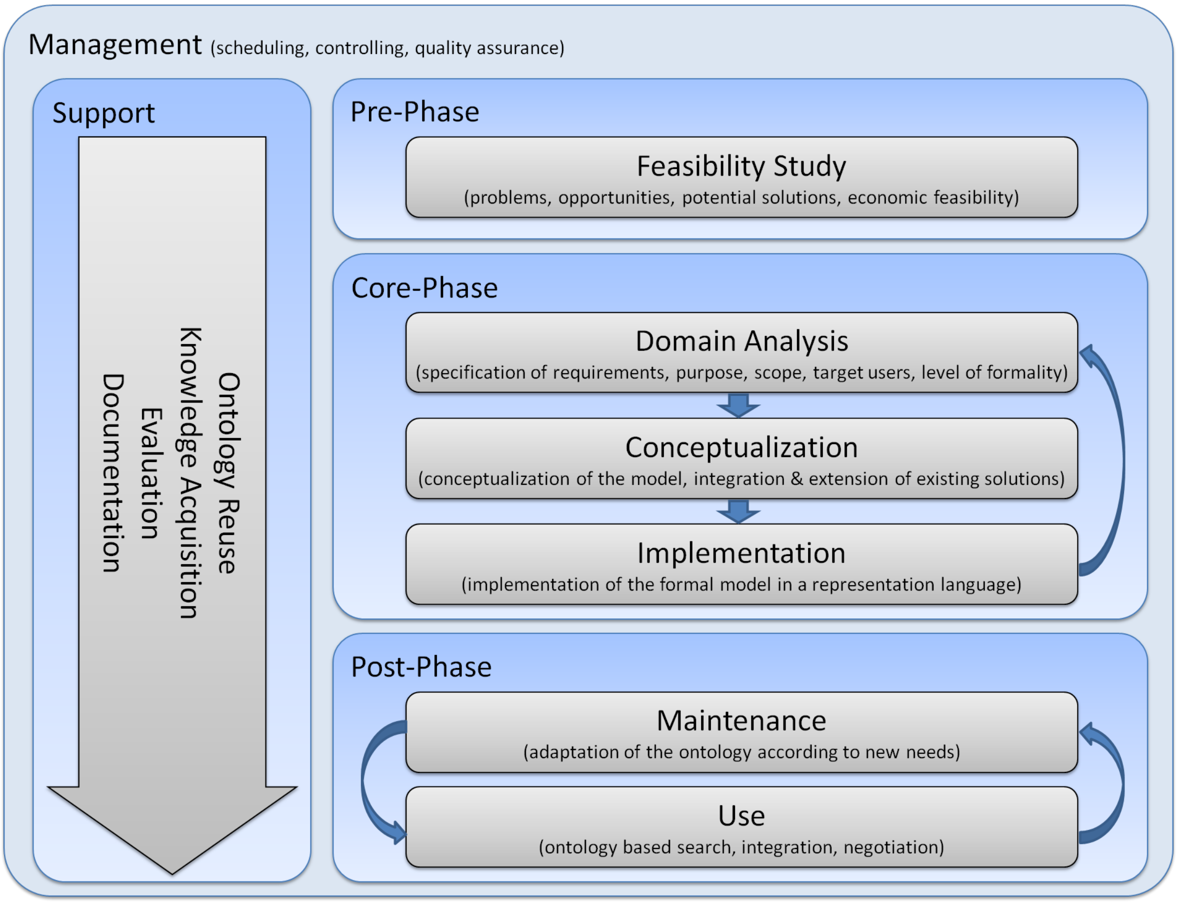


Figure 3.7.: Ontology Engineering Activities adapted from Simperl and Tempich (2006)

- **Conceptualization**: During conceptualization activities, ontology engineers and/or domain experts capture and agree on the key ontology concepts and relationships. The result of organizing and structuring the captured knowledge is an informal or semi-formal model.

- **Implementation**: At the end, the previously created conceptual model is explicitly formalized in some ontology representation language[16].

The three core development activities are framed by various other activities and tasks within the ontology engineering process. Simperl and Tempich (2006) categorize ontology engineering activities, based on Gómez-Pérez et al. (2004), into management, development-oriented and support activities (see Figure 3.7). The *management* activities scheduling, control and quality assurance form the organizational setting for the development process. *Development-oriented* activities comprise feasibility study in the *pre-phase*; the classical previously described *core development* activities of domain analysis, conceptualization and implementation; and maintenance and use in the *post-phase*. These are accompanied by *support* activities like ontology reuse, knowledge acquisition, evaluation, and documentation. As the Figure 3.7 illustrates, the ontology use is totally separated from the core development activities because traditional approaches assume ontology building to be upfront activities.

[16] Instead of codifying the conceptual model directly with an ontology language, some methodologies propose a separate formalization step in which the conceptual model is first transformed into some formalism like first order logic or description logic (see Nagypál 2007a for a detailed discussion)

3.2.2.2. Ontology Engineering Methodologies

In the following, we will briefly summarize approaches from which the basic aspects of the generic methodology originated and highlight their most important contributions. Traditional methodologies focus on centralized development mainly by knowledge engineers together with domain experts of static ontologies.

Cyc: The Cyc method originated from the experiences in developing the Cyc Knowledge Base of common sense knowledge about the world in the 1980s (Lenat and Guha, 1990). The Cyc method comprises three phases (cf. Gómez-Pérez et al., 2004): (1) manual extraction and coding of common sense knowledge; (2) knowledge codification with the aid of natural language or machine learning tools; (3) automatic knowledge acquisition and codification. In each of these phases the ontology is specified within two activities: (a) developing knowledge representation and top-level ontology of most abstract concepts; (b) representing domain specific knowledge.

Enterprise: The first ontology building method was proposed by Uschold and King (1995) and later extended by Uschold and Grüninger (1996). With the development of the Enterprise Ontology, they propose a process of four main steps: (1) identifying the ontology's purpose and scope; (2) building the ontology through ontology capturing, coding and integrating existing ontologies; (3) evaluating the ontology through making a technical judgment; (4) documenting the ontology, especially important assumptions.

For capturing the knowledge and identifying the main concepts and relationships, the authors propose three different strategies: (a) *top-down* – starting with the most general concepts and refining these in more specific ones; (b) *bottom-up* – starting with the most specific concepts, e. g., based on information resources that should be described, and generalizing these into more general ones; (c) *middle-out* – identifying the most important entities, e. g., collected in a brainstorming session, which are then specified or generalized where needed.

The authors also emphasize to take into account the required level of formality to meet the intended ontology purpose and to use informal techniques thus resulting in an semi-formal ontology or intermediate model at the end of this activity. They don't propose a specific knowledge representation language for implementation. In order to reach agreement, some guidelines are presented, e. g., term disambiguation or wording accessible to non-technical readers. The authors don't make a point on how to develop a shared understanding, nor the role and involvement of knowledge engineers, domain experts or end users, nor the target use.

TOVE: The TOVE methodology has its origins in the TOVE (TOronto Virtual Enterprise) Enterprise Modelling project with the aim to develop a knowledge-based system for query answering using first order logic (Gruninger and Fox, 1995; Gruninger, 1996). The authors propose a methodologies of activities in six steps: (1) *Capture of motivating scenarios* – describe main problem stories or examples in terms of how the ontology is intended to be applied. (2) *Formulation of informal competency questions* – informal competency questions are a set of natural language questions that are used together with the answers to specify the scope of the ontology and requirements based on the motivating scenarios. The questions should not only be simple queries but stratified; i. e. one answer can be used to as answer for a more general question as well. In this way, main concepts, their properties, assumptions and constraints can be obtained. (3) *Specification of the terminology in a formal language* – the informal competency questions are used to specify the terminology that (concepts, attributes and relations) is represented in a formal language such as first-order logic. (4) *Formulation of formal competency questions* – based on the previously specified formal terminology the informal competency questions are defined formally. (5) *Specification of axioms and definitions for the terms in the formal language* – axioms, defined as first-order sentences, limit terms and constraints in interpretations for the formal terminology. (6) *Specification of completeness theorems* – at the end, conditions have to be defined under which solutions to the competency questions are complete.

With the definition of motivating scenarios and competency questions, the TOVE methodology highlights the target use and application of the ontology. The target representation is in first order logic. The notion of informal competency questions can be seen as an intermediate model. The authors don't make a point on how to develop a shared understanding nor which role communication, collaboration or negotiations play in the development. The role and involvement of knowledge engineers, domain experts or end users is neither explicitly specified.

KACTUS: The KACTUS methodology (Bernaras et al., 1996) is, in contrast to the previous ones, an application dependent methodology and considers the ontology development synchronous with the knowledge-based system development. The methodology proposes three activity steps: (1) *Specification of the application* – specific context of an application and entities to be modeled are identified resulting in a list of terms and tasks. (2)*Preliminary design* – based on top-level ontological categories, a global model is derived using the previous list of terms and tasks as well as existing ontologies as input. (3) *Ontology refinement and structuring* – by specializing the terms and refining the structuring, the final ontology is obtained.

This approach proposes an bottom-up abstraction strategy. Starting with an ontology for a specific application, this is generalized and adapted as more applications in a similar domain are built. Besides the explicit application dependency, there's no statement regarding our criteria.

SENSUS: In contrast to the previous methodologies, the SENSUS methodology proposes a top-down strategy for building a specific domain ontology based on an existing broad coverage general ontology, the huge and highly complex SENSUS ontology, which is pruned and extended (Swartout et al., 1997). The assumption is that a common base ontology can act as a "hinge" for terminology and structure between domain specific ontologies and thus increase knowledge shareability. We can obtain the domain specific ontology within five steps: (1) *Identify seed terms* – these are the key domain terms for a particular domain. (2) *Manually link the seed terms to SENSUS* – the identified seed terms are manually linked, e. g., added as subclasses, to the existing SENSUS ontology. (3) *Add paths to the root* – all concepts from the seed terms to the root node of the SENSUS ontology are included. Irrelevant SENSUS concepts are pruned. (4) *Add new domain terms* – additional relevant terms not contained in the SENSUS ontology are manually added following the previous two steps. (5) *Add complete subtree* – for nodes having a large number of paths through them, it's to be decided to also include the subtree under each of them, because other terms in the subtree are likely to be relevant as well if many have been already found to be so.

The resulting ontology is a semantic network that is application semi-dependent because of the seed term list obtained for a specific application. The methodology itself doesn't provide any further insights regarding our criteria. However, in applying the methodology, the authors emphasize the importance of an integrated system and ontology development that includes group activities and enables system builders to adapt the ontology in a continuous way. Thus "*the ontology becomes a living document*". To provide initial support to this vision, the authors developed Ontosaurus[17], a web-based distributed browser and editor for ontologies using the LOOM representation language.

IDEF5: IDEF5 – "Integrated Definition for Ontology Description Capture Method" – was developed in 1994 by KBSI (Knowledge Based Systems, Inc.) as a software engineering method for the development and maintenance of domain ontologies. The method proposes five very detailed activities for the ontology development process: (1) *Organizing and definition* – first, the development project is organized including team formation and role assignments (project leader, knowledge engineer, domain experts, team members, reviewers). This is followed by the project definition in which the development team determines the purpose, viewpoint, scope and level of detail for the ontology. For the purpose definition, a statement of need and objectives of acquiring and maintaining the ontology should be defined. The scope and level of detail are documented by context statements and a set of examples. The IDEF5 description summary form records this information. (2) *Collect data* – the knowledge engineer gathers "raw" data in an iterative and interactive process with domain experts. This might be done by extracting data from source documents, observation of organizational activities or interviews and protocol analysis with domain experts, whilst the latter is the most commonly used method. (3) *Analyze Data* – knowledge engineers and domain experts analyze the raw data. In order to facilitate the ontology extraction, they generate a first ontology characterization by listing the relevant ontology elements, looking for boundary elements and for collections of ontology elements. (4) *Develop Initial Ontology* – based on the gathered data, the knowledge engineer develops an initial ontology. Therefore 'proto-kinds' (classes), 'proto-characteristics' (attributes), and 'proto-relations' are developed. (5) *Refine and Validate Ontology* – to finalize the development process, the proto-kinds, characteristics and relations are refined by converting them to kinds, attributes and relations. For validation the final ontology is instantiated and the resulting instantiation is compared with the ontology structure.

The methodology provides two representation language to support the development process. The is a graphical component that facilitates communication and specially supports activity four. The IDEF5 Elaboration Language is a structured textual first-order logic language and used in activity five.

This methodology clearly defines the role and involvement of knowledge engineers, domain experts etc. However, domain experts are only involved in the knowledge gathering and analyzing phase. The schematic language provides graphical assistance for communication and an intermediate ontology model. The Elaboration Language provides a first order logic target language. The target use and application is highlighted. As it is assumed that established knowledge from a specific viewpoint is to be captured, there is no support for developing a shared understanding.

METHONTOLOGY: The METHONTOLOGY methodology (Fernández et al., 1997; López et al., 1999), which was inspired by software engineering methodologies, proposes an on-

[17] http://www.isi.edu/isd/ontosaurus.html

tology development process that is embedded in an evolutionary type of life cycle and techniques to support management, development-oriented, and support activities similar to the generic methodology in Section 3.2.2.1. It's intended to build ontologies from scratch, by reusing existing ontologies or as a re-engineering process. The methodology itself underwent several adapting iterations. We refer to the version presented in Gómez-Pérez et al. (2004).

In comparison to the generic methodology, the development process misses the pre-phase activities and Use in the post-phase, whilst the core development activities additionally distinct *Formalization* that is to transform the conceptual model resulting from the conceptualization into a formal semi-computable model, e. g., first-order logic, before it is codified in some representation language in the implementation activity. For the support activities, additional configuration management activities record documentation and ontology versions for change control.

The ontology life cycle identifies phases and an ordering when the activities should be performed based on evolving prototypes of an ontology. Thus, from the management activities scheduling is done at the beginning and control and quality assurance are continuous. Support activities are also done in parallel with the development activities, but vary in their amount. Thus, knowledge acquisition, integration and evaluation are mostly done during conceptualization and then decrease.

The METHONTOLOGY methodology focuses especially on the conceptualization activity and the construction of an intermediate conceptual model. The conceptual model should be based on tabular and graph notations, i. e. not suitable for reasoning but understandable by domain experts and machine processable thus the formal model can be automatically generated out of the conceptual model. For building the conceptual model, the methodology proposes to specify various artefacts in eleven tasks by the ontology engineer. These are in the following prescribed order to build a glossary of terms, concept taxonomies, ad hoc binary relation diagrams and a concept dictionary; to describe tables of ad hoc binary relations, instance and class attributes and constants; to describe formal axioms and rules and to make a table of instances. The authors suggest to use ontology editing tools such as WebODE or ODE to support the conceptualization phase.

This methodology aims for developing formal ontologies but emphasizes the importance of intermediate models and evolving prototypes. The methodology involves knowledge engineers and domain experts where it's focused on the former ones for doing the whole modeling. The methodology is seen as application independent. Specific support for developing a shared understanding or for communication, negotiation or collaboration is not proposed.

On-To-Knowledge: The On-to-Knowledge methodology (OTKM) focuses on ontology-based knowledge management systems (Staab et al., 2001; Sure et al., 2004). It defines two orthogonal processes: the *Knowledge Process* that focus on ontology usage and the *Knowledge Meta Process* that deals with ontology building. The Knowledge Meta Process comprises five phases: (1) *Feasibility study* – OTKM adopts the CommonKADS methodology (Schreiber et al., 1999) in order to decide whether the ontology should be built or not. Therefore problems and opportunities, the focus of the knowledge management application, tools and people are identified. This serves as basis for the following kickoff phase. (2) *Kickoff* – the ontology requirements are finalized. A specification document describes the exact scope, goal and supported application, design guidelines such as naming conventions, relevant experts and knowledge sources and potential users and usage scenarios. Competency questions are collected for later validation. Besides the specification, the knowledge engineers come up with a first draft of important concepts and their hierarchical structure, the semi-formal ontology description, following a middle-out approach and with potential reusable ontologies. (3) *Refinement* – the semi-formal description is refined together with domain experts. Then the ontology engineers formalize the ontology into the target ontology using a formal language such as frame logic or description logic. OTKM suggest to use OntoEdit ontology editor (Sure et al., 2002). (4) *Evaluation* – in order to proof the ontology usefulness, three types of evaluation are defined (technology-focused, user-focused and ontology focused evaluation). Therefore the requirements and competency questions are checked, the ontology is tested in the target application environment and feedback from end users is collected, and the ontology is formally evaluated, e. g., with the OntoClean approach (Guarino and Welty, 2002). The evaluation phase is closely linked to the refinement phase and several cycles might be necessary before the target ontology can be rolled out. (5) *Application & Evolution* – this describes the ontology usage in the Knowledge Process and maintenance, i. e. responsibilities and rules for update-delete-insert processes. These should be done as part of the application system and include the feedback of the users. At the end it's necessary to decide when to initiate another ontology cycle.

The OTKM is quite similar to the METHONTOLOGY methodology but focuses on a high application dependency and aims to involve end users in the application and evolution phase.

UPON: The UPON methodology (Nicola et al., 2005, 2009) is derived from the Unified Software Development Process (UP) and differentiates itself to be use-case driven, iterative and incremental. Following UP, the process consists of cycles, phases, iterations and workflows that focus on the core ontology development activities. A cycle has four phases and produces a new ontology version. The phases are inception, elaboration, construction and transition. Each phase can have several iterations during which five workflows (requirements, analysis, design, implementation and test) are proceeded with varying relevance; e. g., the requirements workflow (capturing requirements) dominates the inception phase.

(1) *Requirements workflow* – knowledge engineer (KE) and domain experts (DE) meet to determine domain, scope and purpose of the ontology by writing a storyboard (mostly DE), creating an application lexicon (mostly DE with help automatic tools), identifying competency questions (DE & KE in interviews), and related use cases (mostly KE). The outcome of this workflow are competency questions, use-case models, and the application lexicon. (2) *Analysis workflow* – that's the conceptual analysis. DEs built a domain lexicon by acquiring and analyzing existing domain resources (with help of automatic tools). The domain lexicon and application lexicon are merged to build the reference lexicon (mostly DE). Next, DEs with the help of KEs model the application scenario using UML activity and class diagrams. At the end, DEs and KEs build a reference glossary by adding informal definitions to the reference lexicon. (3) *Design workflow* – previously gathered entities are refined and their relationships identified. Mostly KEs model concepts, concept hierarchies and domain-specific relationships resulting in a semantic network represented in a set of UML class diagrams. (4) *Implementation workflow* – the KEs select a formal language (preferably OWL) and formalizes the ontology. (5) *Test workflow* – KEs verify the semantic quality by checking the consistency with some reasoner. DEs and KEs verify the pragmatic quality by verifying the coverage and answering competency questions.

The involvement of domain experts (including end users) and knowledge engineers is clearly defined. The domain experts mainly contribute to the requirements and analysis workflows and partially to the test workflow whilst knowledge engineers are responsible for design and implementation. The workflows are very detailed, each producing an intermediate model intended only as input for the following one: ranging from lexicon, glossary, semantic network represented in a set of UML class diagrams to OWL ontology. UML diagrams are also used to model related use cases and application scenario and as a tool to achieve agreement.

3.2.2.3. Ontology Engineering Tools

There is a wide range of tools that support the different phases of ontology engineering. They are called ontology engineering, development, editor and management tools with sometimes varying feature focus. Some of these have been developed in the context of the previously presented methodologies or cover most of the phases and activities; others are methodology-independent and thus do not refer to any specific phase or course of activity.

Seremeti and Kameas (2010) differentiate specialized ontology engineering tools supporting some activities of the ontology life cycle from integrated ontology engineering environments and provide an entire classification. In their core, the tools provide at least some support for constructing, editing and saving ontologies. Weller (2010) identified as typical components: (a) basic editing functionalities like creating and deleting ontology elements; (b) import and export functionalities to import and store ontologies of different formats; (c) inference and reasoning functionalities for consistency checks; (d) visualization for graphically representing the ontology structure and (e) task management to be aware of planned and upcoming tasks. Additional functionality might include change tracking and versioning, graphical editing or to some extend already multi-user support. However, this support is often limited just to access management.

There are several surveys, e. g., by Corcho et al. (2003); Gómez-Pérez et al. (2004); Denny (2004); Fensel (2004); Waterfeld et al. (2008); Mizoguchi and Kozaki (2009); Seremeti and Kameas (2010); Weller (2010), which extensively discuss different ontology development tools and their functionalities.

Ontolingua: One of the first tools in the mid-1990s for editing and managing ontologies is the Ontolingua Server (Farquhar et al., 1997). Based on top of the Ontolingua KR system, it provides a simple Web interface to edit and browse ontologies in the Ontolingua language. The Ontolingua language is built on top of the knowledge interchange format and a combination of frames and first order logic for representing classes in taxonomic structures, relations, functions, formal axioms and instances. HTML forms support the user in creating new ontology elements. However knowledge of Ontolingua language and KIF are crucial to fill the forms; e. g., relation constraints must be directly typed in. Additional features comprise graphically browsing the class taxonomy and relations but also first approaches for collaboratively and modularly developing ontologies, e. g., ontology sharing in user groups; however without any notification or provenance information. It does not target any methodology.

Ontosaurus: Developed in parallel to Ontolingua and quite similar, OntoSaurus supports the development of LOOM ontologies (Swartout et al., 1997). It additionally includes some inferencing support (concept classifier). OntoSaurus does not support user groups but locking mechanisms to prevent updates when someone else is editing the ontology. Similarly to Ontolingua, it does not follow any methodology.

WebOnto: WebOnto is a Web-based ontology editor for OCML ontologies (Domingue, 1998). Using Java applets WebOnto provides a fully graphical interface; e. g., drag ontology elements from the icon bar and drop and connect them in the graphical edition area. Complex operations need to be directly written in OCML. Additionally, it provides capabilities for consistency checking. The broadcast and receive mode allows one user to lock the ontology and broadcast his/her changes thus other users being in the receiving mode can view the changes. However, there is no export functionality. The additional tool Tadzebao supports collaboration via synchronous and asynchronous discussions.

ODE & WebODE: ODE (Ontology Design Environment) is a standalone ontology editor and WebODE its web-based descendant (López et al., 1999; Vega et al., 2001). They are language independent; i. e. support several different ontology languages for import and export, and explicitly support the METHONTOLOGY methodology, e. g., with administration, planning and documentation modules. Based on the intermediate representations of METHONTOLOGY, WebODE includes concepts and attributes, disjoint concept sets, concept taxonomies, class partitions, ad-hoc binary relations with properties, constants, formal axioms and instances of concepts and relations. The ontology editor provides an HTML form-based interface for ontology term editing, a graphical user interface for taxonomic and ad-hoc relations and an axiom builder for first order logic axioms and rules. There are additional tree-based browsing, clipboard and details features as well as services for merging, inferencing, evaluation and documentation. The latter one is used to represent intermediate models e. g., as HTML tables. There is also collaborative editing support through defining user groups for an ontology and synchronization mechanisms.

OilEd: OilEd (Bechhofer et al., 2001) is a standalone editor for description logic based languages (first OIL, later DAML + OIL and finally adapted to OWL) and can be connected to reasoners like FaCT, e. g., for consistency checks. OilEd organizes the editing of the individual ontology components (classes, properties, individuals, axioms and some general information and namespaces) in tabs. Classes can be described with a name, description, super-classes and property restrictions including cardinality and type restrictions. Classes are listed in alphabetical order. There is no class hierarchy because this is computed by the connected reasoners. Property editing comprises documentation, domain and range, super-properties and inverses. OilEd is aimed to be a simple tool for description logic ontologies. It does not provide any methodological support or additional collaboration or graphical functionality.

KAON: KAON is part of the KAON tool suite, an ontology management infrastructure, and provides different modules for ontology creation and management (Bozsak et al., 2002). These are the OI-Modeler and KAON Portal as frontends and the KAON API and its implementations for managing ontology repositories as the core for programmatic access. The OI-Modeler is the ontology editor for KAON ontologies, a proprietary extension of RDF(S). It is possible to model concepts, properties with relations and attributes, and instances of concepts and properties. Concepts and properties are organized in hierarchies. We can define property cardinalities as well as symmetric, transitive and inverse relations. The editor provides am interface for navigating the ontology graph using the TouchGraph graphical library[18]. With the search and query function, keyword-based search for ontology elements and KAON queries can be executed. There is an additional form-based user interface to manage concepts, properties and their instances. For concepts, we can edit and browse the hierarchy, its properties, labels, documentation, synonyms, word stems and instances. Similarly for properties, we can additionally edit domains and ranges. The ontology evolution service provides support on implications of changes, e. g., handling orphaned concepts. KAON also provides change reversibility options and a locking and transaction protocol for collaborative editing.

Swoop: Swoop is a standalone browser and editor for OWL ontologies (Kalyanpur et al., 2006). Originally developed by the MIND lab at the University of Maryland, it has been continued as an open-source project that is however, similar to the tools above, no more under active development. Swoop was intended to provide an interface that resembles a frame-based website in a Web browser. So, it provides a navigation sidebar showing a list of available ontologies and the class and property hierarchies for each ontology whilst the center pane displays the ontology's or its entities' details and links to edit. The address bar allows to load not only local but also ontologies from the Web that can be modified and then locally maintained separate from the origin. History buttons allow to navigate to the previous or next perspective. Nevertheless, it resembles classical editors. Via its plug-in architecture, renderers (e. g., for graph visualization) and reasoners (e. g., for ontology debugging) can be used. Swoop also provides change tracking with a simple undo functionality and change annotation based on the Annotea framework to exchange and discuss ideas (but not collaboratively work on ontologies). It does not follow any methodology.

Protégé: The most well-known ontology development tool is probably Protégé[19] (Noy et al., 2001; Gennari et al., 2003). Originally developed by Stanford Center for Biomedical Informatics Research at the Stanford University School of Medicine for creating knowledge bases and ontologies in biomedical informatic research, it is now an open-source desktop-based framework available in two version – one to build frame-based ontologies in a format specific to Protégé and the other to build OWL and RDF(S) ontologies (Knublauch et al., 2004). The quite complex tool provides, besides basic functionalities to create, visualize and modify ontologies, a plug-in infrastructure for additional components, e. g., for version control, inferencing, search or different visualization tools[20]. Protégé does not support any specific methodology barring the Ontology Development 101 Guide by Noy and McGuinness (2001) that was built upon Protégé. A plug-in to support multiple users has recently been developed and will be discussed in Section 3.2.3.2.

TopBraid Composer: Since 2006, TopQuadrant provides the commercial ontology editor TopBraid Composer[21] for developing ontologies in OWL and RDF(S). TopBraid Composer is based on the Eclipse platform and resembles the Protégé user interface (as it has its roots in Protégé-OWL) with own features for e. g., ontology mapping or geodata

[18] http://www.touchgraph.com
[19] http://protege.stanford.edu
[20] http://protege.stanford.edu/download/plugins.html
[21] http://www.topbraidcomposer.com

mapping to Google Maps. Recently, it also provides multi-user support based on shared Sesame server to which the users commit their changes on their local view.

OntoStudio: OntoStudio is one of the most popular commercial editor and the descendent of OntoEdit provided by ontoprise GmbH[22](Weiten, 2009). It supports building ontologies especially in F-Logic as well as OWL and RDF(S) by expert users. It is the front-end counterpart to the inferencing machine OntoBroker. Therefore OntoStudio provides specific support to rule-based tasks; e. g., graphical creation and debugging of rules. It is based on the Eclipse Framework following the common interface design with an ontology navigation sidebar showing a list of ontology projects with subfolders for concepts with its hierarchy, attributes, relations, rules, queries, meta models and mappings, a sidebar for the instances and the center pane for displaying and editing the entity properties. Additional features allow to import UML 2. or Microsoft Excel and Outlook. Theres some collaboration support through the backend OntoBroker collaboration server.

3.2.2.4. Discussion

Table 3.1 shows the summary of comparing the knowledge engineer driven ontology development approaches based on our criteria. None of the approaches supports the development of a shared understanding. Indeed, the ontology is considered as an engineering artefact based on the common assumption that shared knowledge and understanding is given and only needs to be externalized.

All methodologies focus on a centralized development by knowledge engineers and mainly involving domain experts only to gather requirements or knowledge. Only IDEF5, On-To-Knowledge and UPON actually include end users in the process.

Besides TOVE, all approaches propose some tool support or at least some guidelines and techniques. However, there is practically no support of communication, negotiation and collaboration processes. Only IDEF5 propose to use the IDEF5 Schematic Language as a graphical component for communication or UPON to use UML activity and class diagrams for achieving agreement.

Two third of the approaches propose to build intermediate models. However, these are only intended as input for the next development step but not to make any other use of. At the end, all besides SENSUS and Enterprise aim for developing highly formal ontologies. The Enterprise methodology emphasizes to consider the formality level that is really needed. The resulting ontology in SENSUS is a semantic network.

In one third of the methodologies, the ontologies are developed independent of any target use or application. Indeed all methodologies assume an upfront development of the ontology that is established, no more changing shared understanding. Thus, none of the approaches considers to integrate the development into work and application.

In conclusion, only Swartout et al. (1997) mention a vision similar to us. They emphasize the importance of an integrated system and ontology development that includes group activities and enables end users to adapt the ontology in a continuous way. However, their methodology does not reflect this vision.

[22] http://www.ontoprise.de/en/home/products/ontostudio

Table 3.1.: Comparing knowledge engineer driven ontology development approaches

	Shared Understanding	Role & Involvement of Users	Communication, Negotiation & Collaboration	Usable Living Intermediate Models	Target Language & Formality	Target Use & Work Integration
CYC	Non-supported	Knowledge engineer centered; no further statements	Proposes Cyc tools; no explicit communication, negotitation or collaboration support	No intermediate models; use non-supported	CycL representation language based on first order logic	No target use; no work integration
Enterprise	Non-supported	Knowledge engineer centered; no further statements	Very vague guidelines, e.g. techniques for brainstorming or term disambiguation	semi-formal ontology as result welcome; use non-supported	Considering needed formality emphasized; no specific language	No target use; no work integration
TOVE	Non-supported	Knowledge engineer centered; no further statements	Non-proposed	Competency questions as intermediate model; use non-supported	First order logic	Target use reflected in motivating scenarios & competency questions; no work integration
KACTUS	Non-supported	Knowledge engineer centered; no further statements	Proposes KACTUS toolkit; no explicit communication, negotitation or collaboration support	No intermediate models; implicitly assumes life cycle synchrounous with target application development	Frames with restrictions	Application dependent; development synchonous with target application development; no work integration
SENSUS	Non-supported	Knowledge engineer centered; no further statements	Proposes Ontosaurus browser & editor; no explicit communication, negotiation or collaboration support	No intermediate models; use non-supported	Semantic networks; LOOM language used in Ontosaurus	Target use reflected in seed term list; no work integration
IDEF5	Non-supported; explicitly assumes established knowledge	Explicit role allocation (project leader, knowledge engineer, domain experts, team members, reviewers) , domain experts only for knowledge gathering & analysis phase	Proposes IDEF5 Schematic Language for communication to develop initial prototype version	Explicit initial prototype version; use non-supported	IDEF5 Schematic Language as graphical component to hide formality; IDEF5 Elaboration Language as first order logic language	Target use and application highlighted; no work integration
METHONTOLOGY	Non-supported	Involves knowledge engineers & domain experts; modeling done by knowledge engineers	Proposes ontology editors (WebODE or ODE) and tools to automatically transfer intermediate model into formal ontology; no explicit communication, negotiation or collaboration support	Evolving prototypes; use non-supported	Focuses on conceptualizatio n actitivites but targets formal ontology, e.g. first order logic; no specific language	No target use; no work integration
On-To-Knowledge	Non-supported	Involves knowledge engineer for modelling, domain experts for gathering refinements and end users in evaluation & application	Proposes OntoEdit editor and to collect feedback from domain experts and end users	Explicit semi-formal description of concepts and hierarchical structure; use non-supported	Formal ontology, e.g. first order or description logic	Application dependent; no work integration
UPON	Non-supported	Explicit role allocation; knowledge engineer for design and implementation, domain experts & end users for requirements capturing, conceptual analysis and partially testing	Proposes UML activity and class diagrams to achieve agreement	Explicit iterative and incremantal methodology life cycle from lexicon, glossary, semantic network as UML class diagrams to OWL ontology; use non-supported	Formal ontology; suggests OWL	Use-case driven; no work integration

3.2.3. Collaborative Ontology Engineering

Ontologies provide consensual knowledge for a specific domain of interest. Intensive research in the Semantic Web community along with Web 2.0 developments have directed the attention to the hitherto mostly neglected consensus building process – i. e. how a group of stakeholders or a community of practice agrees upon a common view of a domain of interest – and brought up collaborative engineering methods (Simperl and Luczak-Rösch, 2011).

3.2.3.1. Collaborative Ontology Engineering Methodologies

Methodology of Holsapple & Joshi: Holsapple and Joshi (2002) presented the first collaborative approach to ontology design. It uses a Delphi-like method (Lindstone and Turoff, 1975) with expert panels and questionnaires to structure the consensus building process. An initial ontology, developed by synthesizing existing ontologies, is the starting point for the collaborative design and revision process. The process consists of four phases: (1) *Preparation* – design criteria are defined and boundary conditions and standards against which to evaluate the ontology are determined. (2) *Anchoring* – knowledge engineers develop a first initial ontology for seeding and orientation purposes of the participants. (3) *Iterative improvement* – to adapt the seed ontology, it is provided to an expert panel together with questionnaires. The knowledge engineers collect and consolidate the feedback and hand it back together with an revised ontology addressing the critique for another feedback loop. This is iterated until the experts consensually agree on all design issues. (4) *Application* – the actual uses of the ontology is demonstrated in a specific context.

This methodology clearly defines the involvement of domain experts and knowledge engineers; knowledge engineers are responsible for the modeling and application whilst domain engineer provide feedback on the presented ontology. The methodology supports the development of a shared understanding and negotiation processes among the involved domain experts through feedback loops with questionnaires. The authors don't make a point either at the target representation language and formality or target use. Even though iterative, the development process is not work-integrated. There are intermediate models within the development process which are used to collect feedback but not for the target application. Indeed the methodology assumes the upfront development of ontologies.

Consensus-based Ontology Engineering Approach: The methodology by Karapiperis and Apostolou (2006) is based on the collaborative ontology design approach by Holsapple and Joshi (2002) with a focus on iterative cycles of consensus building. It additionally follows Noy and McGuinness (2001) ontology building steps to develop the initial ontology and the Nominal Group Technique (Delbecq and VandeVen, 1971) to reach consensus. The methodology comprises four phases: (1) *Definition of the design criteria* – following the design criteria by Uschold and Grüninger (1996) (Section 3.2.1) except for the criterion minimal encoding bias that is dismissed because no formal representation language is directly used but Protégé with its automatic encoding into OWL. (2) *Designing the initial ontology* – the knowledge engineer builds an initial ontology version compliant with the design criteria following the steps by Noy and McGuinness (2001) and based on gathered information from a group of selected domain experts. (3) *Initial structure evolution according to Nominal Group Technique* – the domain expert team evaluates in a face-to-face meeting the initial ontology by the dint of evaluation sheets and through voting following the Nominal Group Technique. The evaluation sheets to be completed in every evaluation cycle comprise rating the usefulness and ambiguity of ontology elements as well as feedback for additions of new ontology elements, deletions of elements and substitutions of properties and relations. Based on the overall rating, ontology elements are either accepted, detailed by the knowledge engineer and presented for revoting, or presented for substitution. New elements are presented and voted for. Then a new cycle starts with the brainstorming, presentation and voting for the mutual agreed ontology version. This is repeated until the participants do not propose new additions, deletions or substitutions anymore which means consensus is reached and the final ontology is accepted and reflects the participants shared view. (4) *Ontology application* – the ontology is applied by answering indicative simple competency questions using the Queries Protégé plug-in tool.

The evaluation regarding our design criteria is similar to the methodology of Holsapple & Joshi except for the target representation language and formality. At this point the authors state that they do not use any specific formal representation language but rely the development on using Protégé and its OWL encoding functionality.

DILIGENT: DILIGENT (Pinto et al., 2004; Tempich, 2006) is a methodology for Distributed, Loosely controlled and evolvInG Engineering of oNTologies that grounds in the OTK methodology (see Section 3.2.2.2). The authors identified four main requirements to be satisfied in contrast to classical knowledge engineer centered methodologies: decentralization – i. e. geographically but also expertise and number of involved parties related distribution, non-expert builders – participation of domain experts and end users in the modeling process, (partial) autonomy – adapt a shared ontology to personal needs, and iteration – interleave ontology development and use towards an iterative evolution process. Consequently, the methodology specifically aims in supporting non-modeling experts in developing and evolving ontologies in a distributed manner by using an argumentation model-based approach.

In the DILIGENT methodology, it's assumed to involve several participants geographically distributed, with different background and expertise but common interest in a domain and building one ontology collaboratively. There are four different types of participants: (1) domain experts knowing about the domain, (2) ontology engineers knowing about ontology modeling, (3) knowledge engineers knowing about knowledge modeling and ontology-based information systems and (4) end users using the resulting ontology. The DILIGENT process comprises five main stages looped in several iterations: (1) *Build* – a small group of all kinds of participants build an initial ontology version following established methodologies like OTK complemented with argument provision. (2) *Local adaptation* – the initial ontology is distributed to the end users who start using and adapting it to their own needs in their local environments, e. g., for local knowledge organizations. The local changes are logged for future analysis together with arguments users provide to explain their decisions and requirements, but the original ontology shared by all remains unchanged. (3) *Analysis* – a control board that is a small subgroup of the participants with special role and responsibility analyzes the change requests and decides which changes to be introduced in the shared ontology. (4) *Revision* – based on the board's decisions, ontology engineers update the shared ontology. The board makes regular revisions according to the users' evolving requirements and in order to avoid bigger divergence between the local and shared ontologies. (5) *Local update* – once a new version of the shared ontology is available, the users can update the entire or parts of their own local ontology.

Each stage is additionally detailed in major roles, input, decisions, actions, available tools, and output information. After the local update, a new iteration might start with the local adaptation stage and the (not-)accepted changes are indicators for the board's review in the next analysis stage. To support collaboration and argumentation provision especially in the analysis and revision stage, the DILIGENT argumentation framework is provided that consists of an argumentation process and a formal model (Tempich et al., 2007). The argumentation process consists of five activities: (1) The participants *choose a moderator* who organizes the decision process but does not contribute. This can be any participant and may change. (2) The participants *choose a decision procedure*, i. e. a voting mechanism and triggers for a new round, to reach agreement during discussions. (3) The participants initiate new discussions by *specifying issues* that correspond to domain or application requirements. During the discussions, the issues are elaborated and new ones may be added by any participant. Later, the participants prioritize the issues and treat them accordingly. (4) To structure the discussions around issues, the participants *provide arguments and ideas*. Once an issue is agreed to be relevant, the participants suggest natural language ideas to formalize it. Other participants provide arguments of agreement or disagreement and alternative supporting or weakening ideas specified in the argumentation ontology. (5) The participants *decide on issues and ideas*. Agreed ideas are integrated into the shared ontology, non-agreed are postponed for further discussion.

The argumentation ontology specifies the kinds of arguments and main concepts in discussions. Therefore, it relies on the Issue-Based Information System (IBIS) argumentation model (Kunz and Rittel, 1970) to represent argumentation processes, integrating the concepts Issue, Idea, Argument, Challenge and Justification, and the Rhetorical Structure Theory (Mann and Thompson, 1987) to define and restrict the types of arguments that can be used during the discussions to Elaboration, Evaluation and Justification, Example, Counter Example and Alternative.

For tool support, the authors provide a plug-in for the OntoEdit ontology engineering environment (Sure et al., 2002) and suggest to use in addition wiki and online chat tools for the argumentation process. In the continuation of their work, a wiki-based system was developed that can be used to collaboratively develop an ontology and that supports the argumentation through extended discussion page functionalities.

Similarly to the previous collaborative methodologies, DILIGENT clearly defines the role and involvement of different participant groups. Reaching a shared understanding is supported by the argumentation framework. The users can start using and adapting intermediate ontology versions within their local environments. In this way, target use is reflected. However, it is not obvious if these local adaptations can be done in a work-integrated way. Grounded in the OTK methodology, it also targets formal ontologies.

NeOn Methodology: The NeOn methodology aims to support the development of networked ontologies and focuses specifically on reuse (del Carmen Suárez-Figueroa et al., 2011; del Carmen Suárez-Figueroa et al., 2008, 2009; del Carmen Suárez-Figueroa, 2010). The methodology is based on nine scenarios for building ontologies and ontology networks. Each scenario encompasses different processes and activities to be carried out. These are defined by the NeOn Glossary of Processes and Activities. The methodology provides guidelines for the different processes and activities, e. g., ontology requirements specification or reuse and re-engineering of (non-) ontological resources, that are described with filling cards, workflow and examples. The scenarios consider collaborative situations which are reuse-oriented or in which participants actively work on ontologies. For the former case, the methodology suggests methods inspired by extreme programming methods to design ontologies with design patterns. For the latter case, the methodology emphasize supporting argumentation and documentation processes. The scenarios are further related to ontology (network) life cycle models. The methodologies proposes two main models: the waterfall ontology network life cycle model with different versions and the iterative-incremental model.

In general the process and activities guidelines are tailored towards ontology engineers; involving end users mostly in ontology specification and evaluation activities. For the collaborative ontology development the methodology proposes an editorial workflow approach (Palma et al., 2008). Therefore, ontology engineers are assigned either the role of subject experts who can add and modify ontology elements or the role of validators who can approve, revise or reject ontology changes by the subject experts and publish the ontology. Furthermore, the whole ontology and each ontology element have a status. When a subject expert wants to insert or update an ontology element, the status "Draft" is assigned to the element until the subject expert is confident with the change. Then the element receives the status "To Be Approved". If the subject expert wants to delete an element, the status "To Be Deleted" is assigned. The requested changes are then sent to the validators who reject or approve and thus actually apply the changes. The whole ontology can have four states: "Draft" – changes are performed, "To Be Approved" – all changes are in the To Be Approved state, "Approved" – all changes are approved, and "Published" – validators published a version; therefore the ontology had to be in Approved state.

Tool support is focused on the NeOn toolkit[23] that provides an Eclipse based ontology engineering environment for F-logic and OWL ontologies with several plug-ins to support e. g., ontology management, modularization, reuse etc. The Workflow Support and Oyster plug-ins support the editorial workflow approach. The ontology engineers can either work on local copies of the ontology or on a centrally stored version from a distributed version repository. Conflict resolution mechanisms are provided for concurrent change commitments.

The Cicero plug-in (Dellschaft et al., 2008) is a further development of the DILIGENT argumentation framework (Tempich, 2006) to specifically support argumentation processes and track discussion. Discussions take place within a semantic wiki environment and are then connected to ontology elements in the toolkit environment. Supporting the development of a shared understanding is not a primary issue. And the editorial workflow approach does not consider argumentation processes. Besides the ontology state assignments, the methodology does not make any point on using intermediate models. The methodology is seen as application independent.

HCOME: The Human-Centered Ontology Engineering Methodology (HCOME) provides a methodology for engineering ontologies in an organizational context in a decentralized way (Kotis et al., 2004; Kotis and Vouros, 2006; Vouros et al., 2007). Ontology development is seen as a dynamic process that is integrated in the daily activities of knowledge workers. The approach focuses particularly on argumentation-based ontology evolution. The methodology assumes a decentralized engineering model with three different knowledge spaces: personal spaces in which the individual parties formalize their own ontologies, a shared space in which the individual ontologies are published, merged and collaboratively further developed, and an agreed space in which agreed ontology versions are stored for browsing, import into personal space and use in applications.

[23] http://neon-toolkit.org

The methodology consists of three main phases, which are closely linked to these separated spaces: (1) *Specification* – teams of knowledge worker collaborators establish and collaboratively specify scope and aim of the ontology and agree on requirements recorded in a requirements specification document. (2) *Conceptualization* – the knowledge workers start to develop the ontology in their personal space. Therefore they may perform tasks like importing existing ontologies; consulting generic top ontologies for better understanding; improvising ontologies, i. e. building from scratch; managing, mapping and merging of ontology versions; comparing different ontology version for ontology evolution and identifying merge candidates; and adding documental information. (3) *Exploitation* – the collaborators push their developed ontologies into the shared space where they inspect and compare the different versions and publish their feedback. Structured discussions are used to achieve a common understanding and produce an agreed ontology that is published in the agreed space. The users may also inspect and use shared and agreed ontologies in their local space, thus new ontology versions might be generated.

The methodology is supported by the HCOME-3O framework (Vouros et al., 2007) and HCONE tool suite (Kotis and Vouros, 2006). The framework provides three meta-ontologies for capturing administrative information, change operations and the rationale for these in order to facilitate the management of the collaboration processes. HCONE provides tools and features to develop and manage shared ontologies such as WordNet consultation, argumentation dialogs based on the IBIS argumentation model or email notification about status of discussions or new ontologies in the shared space. An editor that is similar to Swoop (Section 3.2.2.3) additionally provides a graphical tree representation and predefined natural language dialogs to support the users in specifying their conceptualization. The formal specification of the ontologies in NeoClassic Description Logic is automatically done in the background. In continuation of their work, the authors recently extended their approach by integrating techniques of ontology learning from query logs and semantic wiki technologies for argumentation tasks (Kotis and Papasalouros, 2010; Kotis, 2008).

DOGMA-MESS: DOGMA MESS (de Moor et al., 2006; Spyns et al., 2008; Leenheer, 2009) is an extension of the DOGMA methodology, which is a database-inspired approach (Jarrar and Meersman, 2002; Spyns et al., 2002), to support inter-organizational and collaborative ontology engineering. In the DOGMA approach, ontologies follow the so-called "double articulation", i. e. they consist of an ontology base of language lexons – context-specific binary fact types that specify the conceptualization of the domain – and of ontological commitments – application-specific semantic constraints. Spyns et al. (2008) divide their methodology into two main parts: preparatory phases and actual ontology engineering phases. The preparatory phases comprise: (1) *Formulation of vision statement* – the stakeholders develop a shared vision of the purpose and scope of the ontology. (2) *Feasibility study* – the vision statement is refined and checked against costs, benefits and technological feasibility. (3) *Project management* – project management activities are initiated; i. e. time, resource and team planning. (4) *Preparation and scoping* – scaling down the problem domain this includes the definition of user requirements and purpose, the identification of domain experts and the compilation and scoping of knowledge resources. The ontology engineerings phases comprise (5) *Domain conceptualization* and (6) *Application specification*. The domain conceptualization is the core phase in which the domain of interest is analyzed resulting in a DOGMA-style ontology (domain fact types). This might involve five activities which are knowledge discovery, knowledge elicitation, knowledge negotiation and knowledge breakdown.

- *Knowledge discovery* is done with the help of techniques for ontology learning from text. Knowledge engineers validate discovered terms, concepts, relations etc. together with domain experts and finally create a formal representation.
- By *knowledge elicitation* activities, which are brainstorming, abstraction exercises and the compilation of baseline taxonomy, knowledge engineers gather knowledge from domain experts to produce a conceptualization based on their expertise.
- The *knowledge breakdown* activity produces a hierarchical structure using linguistic segmentation and highlighting techniques. Therefore knowledge engineer together with domain experts first verbalize elementary sentences and extract elementary facts and then engineer lexons.
- The *knowledge negotiation* embodies the actual DOGMA-MESS approach to collect feedback from domain experts regarding concept meanings and context dependencies in a conversational manner. Therefore it differentiates between several roles and meaning layers. Core domain experts, who are authorities in their domain, create initial domain templates that describe a common knowledge definition. Therefore concepts from an upper ontology and from a general upper concept type hierarchy are used. Then domain experts, who represent a certain organization or community, specialize the templates with respect to their organization into divergent organizational ontologies and organizational concept type hierarchies. By meaning negotiation between core domain experts and relevant domain experts on relevant differences, the organizational ontologies are aligned, retaining the most relevant and agreed conceptualizations in a lower common ontology. Knowledge engineers are mainly responsible for assisting the (core) domain experts in defining and analyzing the ontologies.

The *application specification* is the final phase. It includes structuring the application domain, adding application-specific constraints to the domain conceptualization and the validation of the ontology.

Tool support is focused on DOGMA Studio consisting of a Workbench based on the Eclipse Rich Client Platform and a JBoss Server that is extended by version, community and perspective managing modules. For instance, the perspective manager provides support to view and browse conflicts. There is also a web interface. However, how to reach agreement on the shared common ontology is not not detailed.

GM methodology: The GM methodology aims to support the development of bio-medical ontologies (Castro et al., 2006). The methodology focuses especially on the knowledge acquisition phases in which it makes use of Concept Maps that were considered to be useful for capturing and sharing knowledge and formalizing use cases. A fourth issue that is emphasized is collaboration support in distributed environments. The development process comprises six steps and so called milestones that represent the output: (1) identification of purpose, scope, competency questions and scenarios with questions and scenarios as milestones; (2) identification of reusable ontologies with reusable ontologies as outcome; (3) domain analysis and knowledge acquisition that is the linguistic phase which terminates with the baseline ontology; (4) iterative building of informal ontology models (glossaries) with domain experts resulting in a refined ontology; (5) formalization by constraining classes and adding instances; and (6) evaluation ending in a formalized ontology.

When applying the methodology, the process is guided by knowledge engineers who sit together with domain experts in several either face-to-face or virtual meetings. CMAP tools are used to capture knowledge and test the representation provided by domain experts, Protégé OWL is used to aim for description logic as final knowledge formalism. Even though iterative, the development process is not work-integrated.

RapidOWL: The RapidOWL methodology (Auer, 2006; Auer and Herre, 2006) aims to support the iterative refinement, annotation and structuring of a knowledge base for information integration tasks or establishing shared classification systems, vocabularies and conceptualizations. It therefore adopts methods from agile software engineering. It focuses on involving a larger number of domain experts into the ontology development process and therefore provides a number of guidelines and identifies simple strategies and techniques. The methodology does not propose a specific process model or ontology life cycle. According to agile methodologies, the methodology's main components are values, principles and practices: (1) *Values* – the long-term goals following the philosophy of eXtreme Programming being Communication and Feedback merged to Community as crucial enabler for collaboration and stakeholder-driven evolution; Simplicity as facilitation for maintainability of the ontology and knoopewledge base; and Courage as driver to overcome potential modeling dead-ends. (2) *Principles* – guide the ontology development over time. Principles are based on Ward Cunningham's[24] Wiki systems design goals including among others incremental and organic changes, uniform authoring methods for modeling as well as schema and instance representation, observable development and rapid feedback and an open-world assumption. (3) *Practices* – ten best practices to organize the engineering process in daily life include among others Joint Ontology Design between knowledge engineer, domain experts and users; View Generation with domain specific views for individual stakeholders; Information Integration for capturing needed conceptualizations; Modeling Standards for reusability and interoperability; Ontology Evolution for smooth modeling and instance data migration; and Short Releases with quick and frequent ontology publishing.

The methodology differentiates between domain experts, experienced domain experts and knowledge engineers. It further aims for encouraging domain experts to instantly participate, e. g., to express any worthwhile facts by simple RDF statements. Working on the ontology should be observable by other domain experts giving the opportunity to comment or add/delete knowledge fragments. Experienced domain experts are supposed to consolidate and restructure collected data; e. g., detect and merge duplications to which (semi-) automatic approaches might assist. Knowledge engineers are supposed to assist the domain experts community with modeling and evolution methods and strategies, to extend the ontology with logical axioms or to extract OWL DL or lite conform parts.

The methodology is tailored to establish shared vocabularies and conceptualizations. Following the Wiki principles, the methodology can support the development of a shared understanding with communication and feedback as main values. The methodology does not provide any concrete methods, techniques or tools. It proposes an open, incremental and organic development resulting living intermediate models. However it is not immediately obvious if these models are already usable. The methodology relies on simple knowledge models on the basis of RDF statements. It does not prescribe a specific degree of formality but should be appropriate to the actual querying and reasoning needs. Similarly, the development does not focus on any target use or application or work integration.

[24] http://c2.com/cgi/wiki?WikiDesignPrinciples

Ontoverse Ontology Life Cycle: The Ontoverse approach (Paulsen et al., 2007; Mainz, 2009) aims to support scientific communities in collaborative ontology engineering and knowledge management with a wiki-based approach. It differentiates three kind of participants; domain experts providing their knowledge and expertise, knowledge engineers called ontology designers with modeling skills and project administrators being responsible for discussion coordination and final decisions. For each ontology being built, a specific project is set up. Any user, who needs a domain ontology e. g., for knowledge management purposes, can join an ontology project either as domain expert or ontology designer. The user is also free to either actively contribute or passively view and use collected data. If a user starts a new ontology project, the user automatically becomes the project administrator.

In the first phase of the ontology engineering process the ontology's scope, intended goal and target user group is determined. Therefore an ontology requirement specification document modified from the OTK methodology (see Section 3.2.2.2) is set up in a wiki. It comprises sections of motivation and domain, goal, design guidelines, technical requirements, available knowledge sources and competency questionnaire (Weller, 2010). When the document reaches a sufficient mature state, the project administrator locks it for further editing. So that it can be used as a fixed guideline for the following engineering process. This is followed by the conceptualization phase in which a recruited team of project administrator, ontology designers and domain experts collect so called proto-ontological data and hold thematic discussions. Proto-ontological data is non-formalized knowledge and may be simple glossaries, concepts or concept collections, notes or external resource references etc. that domain experts enter in a wiki. Afterwards, ontology designers start with the cooperative editing or formalization phase using a separate ontology editor for formal ontology engineering in OWL. This editing may take place synchronously or asynchronously. The Ontoverse approach proposes a CVS[25]-like system that is tightly coupled with the wiki. That means, domain designers can check out the public shared ontology version to their private workspace (a user-specific branch), make amendments and when finished commit their versions to the shared one. For the synchronous collaboration, users can share their private workspace with other users. Locking of ontology elements and immediate visibility of changes prevent conflicts. At the end, the private workspace can be again committed to the public ontology workspace. For further extending the ontology, the ontology designers might use information extraction techniques or ontology mapping and merging techniques to reuse and integrate existing ontologies. In the last phase, the ontology is evaluated by the domain experts. Therefore it is presented in the wiki in a semi-structured way where discussions and suggestions for modification are directly recorded.

3.2.3.2. Collaborative Ontology Engineering Tools

Already tools of the first generation of ontology engineering, as we have seen in the previous sections, provide some features like locking/unlocking mechanisms to support multiple users, e. g., Ontolingua (Farquhar et al., 1997), Ontosaurus (Swartout et al., 1997), WebOnto and Tadzebao (Domingue, 1998), OntoEdit (Sure et al., 2002) or additional ones like APECKS (Tennison and Shadbolt, 1998), Co4 system (Euzenat, 1996) or Hozo (Kozaki et al., 2002, 2007). As the list of Norta et al. (2010) shows, many tools (16 of 39) have not been under active support or never exceeded their sketchy state. And what's more, most tools' support is limited and targeted to knowledge engineer-driven development without considering the management of negotiation (Díaz et al., 2006). In

[25]Concurrent Versioning System

the following, we will have a closer look at tools that aim to support communities in developing ontologies and consensus building.

Collaborative Protégé: Collaborative Protégé is an extension of the Protégé ontology editor (see also Section 3.2.2.3) whose development has started in 2007 with the Collaborative Knowledge Construction (CKC) Challenge (Tudorache et al., 2008b,a, see also Section 7.1). The user interface of Collaborative Protégé graphically extends Protégé with collaboration tabs for (1) annotating ontology elements and changes – with typed comments like example or question; (2) discussion threads – with similar features like the annotations, however without referring to a certain ontology element but the whole ontology; (3) change tracking – displaying the change log history for a certain ontology element; (4) filtering and search – in annotations by different criteria, e. g., date or author; (5) chats – exchanging messages and links to ontology elements with other online users. Annotations and changes are represented as instances of the CHAO ontology (Changes and Annotations Ontology, in Noy et al. 2006).

There are two different modes to use Collaborative Protégé. In multi-user mode, multiple users can work on one ontology, which is hosted on a Protégé server, at once. All edits are immediately visibly to other online users. In standalone mode, multiple users work consecutively on the ontology, i. e. one user modifies, stores and shares the ontology and then the next user can access and work on the ontology and works on it. Concurrent work is not possible. Similarly, the collaborative features are limited in use, e. g., the chat is not usable.

To leverage the collaboration functionalities to the Web, Web-Protégé[26] has recently been developed as a web front-end for the Collaborative Protégé server (Tudorache et al., 2008c). It allows to browse and edit ontologies to a limited extent (e. g., moving and renaming classes but no creation of restriction). The user interface design follows a portal style, i. e. it is composed of portlets, e. g., class and property tree, property details or notes, whose layout is customizable by the user.

COE: CmapTools Ontology Editor (COE) is an extension of IHMC's CmapTools for ontologies focusing on the provision of visual representations and graphical means for ontology development by "subject-matter experts and small teams" (Hayes et al., 2005a,b). CmapTools is a software environment that enables users, individually or collaboratively, to model, browse and share their knowledge using concept maps (Cañas et al., 2004). In this way, COE can be used to import, view and edit OWL and RDF(S) ontologies as concept maps. In order to represent OWL syntax as concept maps, COE introduces specific notations and convention; e. g., subclass-relations are blue edges with the label "are", domain and range are dotted edges, and textual labels like "at least One" indicate property restrictions. There are templates for often used OWL structures like restriction or intersection that the user can drag from a side panel and drop to the editing area. Additional features are an ontology viewer to navigate through large ontologies and a text-based concept search to find concepts with specific names in previously imported ontologies whose result also displays clusters of related concepts using an external service. Generated concept maps can be exported to OWL/XML files; not translatable parts are ignored. COE's collaboration support is based on CmapTools' functionality; i. e. using the client-server-installation with the Cmap Server to remotely store the concept maps. CmapTools provides support for synchronous and asynchronous collaboration. In synchronous mode,

[26] http://webprotege.stanford.edu

users can simultaneously create and modify ontology elements and discuss changes in chats. In asynchronous mode, users can add annotations and threaded discussions to the ontology.

3.2.3.3. Discussion

Table 3.2 shows the summary of comparing the collaborative ontology development approaches based on our criteria. All methodologies support to some extent the development of a shared understanding. However, except for HCOME, RapidOWL and Ontoverse Ontology Life Cycle, this support is limited to a specific group of participants and does not involve all stakeholders.

Similarly all methodologies provide a clear role and involvement of different participating user groups with the methodology of Holsapple & Joshi, the consensus-based ontology engineering approach and NeOn still being knowledge engineer centered. So that domain experts and end users do not have the possibility to apply their needs for changes in a timely manner.

All of the methodologies provide some tool support or guidelines for communication, negotiation and collaboration processes. Three are based on the IBIS argumentation model, Holsapple and Joshi (2002) and Karapiperis and Apostolou (2006) are based on consensus building techniques with expert panels. These negotiation processes result in intermediate models (except for NeOn methodologies). The intermediate ontology versions are only supposed to be used by DILIGENT and HCOME. Both propose a decentralized process in which users can adopt and further develop agreed intermediate models into their private local space. However this decentralized manner bears the challenge of additional efforts, e. g., when updating the own local version with the shared one with potentially loosing the individual local organization. Additionally, HCOME does not indicate when and which intermediate models are applied and integrated into daily activities. This makes the benefit of collaborating unclear when first a private ontology version should be developed and applied.

More than half of the methodologies aim for developing highly formal ontologies. DOGMA-MESS results proprietary lexons ontologies that are controlled vocabularies for inter-organizational communication to which application-specific semantic constraints are separately added. It is only the RapidOWL methodology that emphasizes simple lightweight knowledge models appropriate to target querying and reasoning needs.

Similarly, in more than half of the methodologies, the ontologies are developed independent of any target use or application and only HCOME emphasizes an integration into daily activities. However the detailed integration remains undefined. Thus, none of the approaches provides obvious work integration of the development process.

3.2.4. (Semi-) Automatic Ontology Learning Techniques

(Semi-)Automatic techniques for ontology learning support the development of ontologies. The general techniques usually support knowledge engineer driven approaches for ontology development. Therefore we won't detail them and refer the reader to Zhou (2007); Buitelaar and Cimiano (2008); Cimiano et al. (2009) who give an overview of this area.

Table 3.2.: Comparing collaborative ontology development approaches

	Shared Understanding	Role & Involvement of Users	Communication, Negotiation & Collaboration	Usable Living Intermediate Models	Target Language & Formality	Target Use & Work Integration
Methodology of Holsapple & Joshi	Supports shared understanding development among domain experts in the expert panel	Knowledge engineer for modeling & application; domain expert panel provides feedback	Delphi-like method with feedback loops with questionnaires	Iterative development with intermediate models used for feedback collection; target use non-supported	No target language or formality	No target use; no work integration; assumes upfront development
Consensus-based Ontology Engineering Approach	Supports shared understanding development among domain experts in the expert panel	Knowledge engineer for modeling & application; domain expert panel provides feedback	Uses Protégé; Nominal Group techniques with evaluation sheets and voting for negotations among domain experts	Intermediate models used for feedback collection; target use non-supported	Make use of automatic encoding into OWL	No target use; no work integration; assumes upfront development
DILIGENT	Supports shared understanding development among control board members	Explicit role allocation (knowledge engineer, ontology engineers domain experts, end users)	Proposes OntoEdit plug-in or wiki-based system; IBIS-based argumentation framework	Intermediate shared revisions are locally used and adopted	Formal ontologies	Target use reflected in local ontology versions; no obvious work integration
NeOn Methodology	Not primarily supported	Ontology engineer centered; end users involved in specification and evaluation activities	Proposes NeOn toolkit and Cicero plug-in as further development of DILIGENT argumentation framework	No intermediate models; use non-supported	Formal ontologies in OWL or F-logic	No target use; no work integration
HCOME	Supported	Knowledge worker centered	Proposes HCOME-3O framework & HCONE tool suite with IBIS-based dialog support, personal/shared /agreed spaces, ontology and external resource alignment	Intermediate ontology versions are locally used and adopted	NeoClassic as description logic language	Target use reflected in personal ontology versions; emphasizes daily activity integration
DOGMA-MESS	Supports shared understanding development among core and relevant domain experts	Core domain experts define domain templates; domain experts specialize templates; knowledge engineers do assist	Proposes extended DOGMA Studio system; meaning negotations in a conversational manner	Evolving lower common ontologies; use non-supported	Controlled vocabularies for inter-organizational communication using lexons	Application dependent, no work integration
RapidOWL	Supports shared understanding development with communication and feedback as values	Domain experts provide facts by RDF statements; experienced domain experts do consolidate and restructure; knowledge engineer assist and add logical axioms	General guidelines with long-term values, mid-term principles and concrete practices	Open, incremental, organic evolving prototypes; use non-obvious	Targets simple lightweight knowledge models appropriate to target querying and reasoning needs; based on RDF(S) or OWL	No target use; no work integration
Ontoverse Ontologly Life Cycle	Supported	Particpants are free to choose their role: domain experts provide knowledge; ontology designers do modeling; project administrator coordinate discussions & decisions;	Provides wiki to collect & discuss „proto-ontological data" and formal ontology editor with collaboration support; both closely coupled	Intermediate ontology versions through CVS-like system; use non-supported	Formal ontologies in OWL	Target use reflected in OSDR; no work integration

Recent work aim for deriving ontologies from folksonomies. These approaches usually analyze the structure of folksonomies with statistical means. Therefore, with cleansing and preparation mechanisms (e. g., stemming algorithms) they first attempt to get rid of inappropriate content within the tag space (e. g., dissolving plurals or compound words) and afterwards to discover semantic relations such as synonymous, homonymous or hierarchical relations between tags. Some approaches additionally use external knowledge bases such as online dictionnaires, Wordnet[27] or DBpedia[28] to map meaning and structure information to tags and in this way semantically enrich folksonomies. In the following, we will give a short overview on these approaches. We would also like to refer the interested reader to the related discussion by Ramezani (2011).

3.2.4.1. From Folksonomies to Ontologies – Extracting Semantics of Folksonomies

Mika (2005) was one of the first who proposed to exploit social tagging systems and their underlaying social semantic network to extract lightweight ontologies following the idea and vision of emergent semantics (Aberer et al., 2004b,a). The author suggests a tripartite model of Actors (users), Concepts (tags) and Instances (annotated resources) and in this way to extend the traditional ontology model of concepts and instances by the social dimension of actors. Based on reducing the tripartite folksonomy hypergraph, the resulting graphs model associations (relations) between users and tags, tags and resources, and users and resources as explained in Section 2.5.1.

In general, most of the approaches that aim for deriving ontologies from folksonomies rely on co-occurrence models. For instance tag co-occurrence frequency helps to detect similarity between tags (Begelman et al., 2006; Specia and Motta, 2007) but also the co-occurrence of tags and users or users and resources to explore the interests of users or the co-occurrence of tags and resources to identify a description of a certain resource (Ley et al., 2009). In order to cluster possibly related tags, additional tag similarity and distance measures are quite popular as well such as the probability-based Matching, Overlap, Jaccard and Cosine similarity or distance-based the Euclidean and Hamming distance. Markines et al. (2009) evaluated different similarity measures for tag-tag and resource-resource similarity by comparing the results with WordNet for tag similarity (called semantic grounding) and Open Directory Project for resource similarity. Similarly, Cattuto et al. (2008) analyzed five tag similarity measures – co-occurrence, a graph-based measure called FolkRank that is based on PageRank (Hotho et al., 2006a) and three distributional measures (tag context, resource context and user context similarity). With semantic grounding in WordNet, they conclude that tag and resource similarity are suitable to discover synonyms and FolkRank and co-occurence for hierarchies.

Most approaches also cleanse and prepare the tag space with filtering mechanisms. This may involve to get rid of unusual – e. g., tags with numbers for a more general applicability – or infrequent and isolated tags – e. g., tags occurring less then a certain number of times – (Specia and Motta, 2007) or meaningless tags – e. g., tags like 'a' or misspellings by using algorithms like the string edit function (Xu et al., 2008). Grouping morphologically very similar tags helps to tackle minor morphological variations or misspellings using stemming algorithms or the Levenshtein distance with a high threshold (Specia and Motta, 2007).

[27] http://wordnetweb.princeton.edu/perl/webwn
[28] http://dbpedia.org

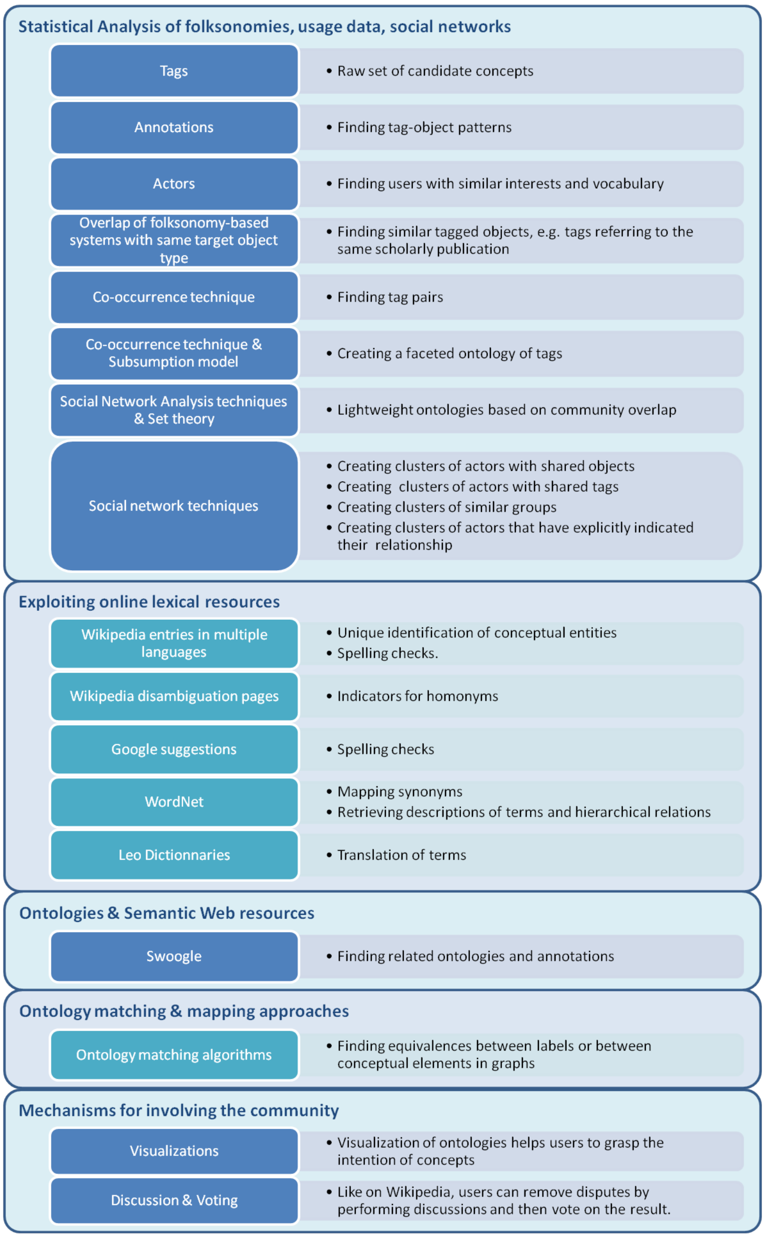


Figure 3.8.: Techniques and resources and their type of contributions to derive ontologies from folksonomies; adopted from Van Damme et al. (2007)

Van Damme et al. (2007) analyzed and summarized related work on different methods and techniques and external knowledge sources that can contribute to deriving ontologies from folksonomies. These comprise (1) statistical analysis of folksonomies, usage data, social networks, (2) exploiting online lexical resources, (3) ontologies and Semantic Web resources, (4) ontology mapping and matching approaches, and (5) mechanisms for involving the community (see also Figure 3.8). Regarding ontology matching algorithms, the authors refer to the formal classification theory presented by Giunchiglia et al. (2005). Schmitz (2006) discusses the combination of co-occurence between tags and a subsumption-based model. Similarly, the authors adopt the approach of combining social network analysis with set models presented by Mika (2005).

Based on this analysis, Damme et al. (2008) propose a 6-step methodology for deriving ontologies from folksonomies by integrating multiple techniques and resources. These techniques comprise language specific stop words filters, Levenstein metric to identify similar tags, co-occurence and conditional probability to find broader-narrower relations and transitive reduction and visualization to involve the community of users. In future work, they also want to include other existing resources like Google, WordNet, Wikipedia, ontologies and other Semantic Web resources for mapping.

Likewise, Angeletou et al. (2009b,a) try to automatically enrich folksonomies using existing resources with their FLOR folksonomy enrichment algorithm. They propose two strategies, one based on WordNet, the other using online ontologies, in order to map meaning and structure information to tags. The FLOR algorithm comprises four phases: (1) *Lexical Processing* – involving lexical isolation and normalization to filter non-processable tags, e. g., non-English tags or with numbers/special characters, and to map between tags used in folksonomies and external knowledge sources; (2) *Sense Definition and Semantic Expansion* – involving sense disambiguation and expanding tags with synonyms and lexical variations using WordNet; (3) *Semantic Enrichment* – discovering and selecting existing entities from online ontologies that correspond to tags; (4) *Semantic Aggregation* – the enriched tag sets are used as instantiations of a final schema and relations are discovered using the SCARLET relation discovery algorithm (Sabou et al., 2008).

Heymann and Garcia-Molina (2006) aim for creating tag hierarchies. Therefore they use the cosine similarity between tag vectors. Every new tag in the system is categorized as the narrower of the most similar tag except the similarity value is less than a predefined threshold. Then the new tag forms a new category as a new child of the root. However because of lacking broader-narrower heuristic, any new tag is considered as the narrower of the most similar tag in the system even though it might be more general than the other tag.

Marinho et al. (2008) enrich a folksonomy with a domain expert ontology; i. e. they make a mapping between an ontology and a folksonomy and include additional triples as tag assignments of the expert. Based on the enriched folksonomy, they build a tag taxonomy (without multiple heritance) using frequent itemset mining. A frequent itemset is a set of frequently co-occurring items; applied to folksonomy this means tags that are often assigned together to a resource. They take up the idea of Schmitz et al. (2006); Jäschke et al. (2008) who mine association rules of the form *Users assigning the tags from A to some resources often also assign the tags from B to them*. Based on this, Schmitz et al. (2006) suggest if resources are often tagged with tag t_x and tag t_y and there are many resources tagged with t_x but not with t_y then t_x is a super-concept of t_y. In this way Marinho et al. (2008) build the taxonmy tree by assuming that (1) more popular tags are more general and thus at a higher level within the taxonomy. Thus, a tag t_x is a super-concept of a tag t_y if there are frequent itemsets containing both tags such that $support(t_x) > support(t_y)$; (2) in case that itemsets have approximately the same support, t_x and t_y in a large itemset are considered to be closer related than in a smaller itemset; (3) itemsets with higher support are more priviledged than the itemsets with lower support.

Monachesi and Markus (2010) also rely on an existing ontology. However, vice versa the previous approaches, they aim to enrich the existing ontology with extracted tags from social tagging systems by integrating the existing techniques such as similarity measures, disambiguation algorithms or DBpedia as external knowledge resource.

Barla and Bieliková (2009) also generate hierarchies with parent-child relations if the usage of t_x is significantly higher than t_y and with sibling relations if both tags are more or less equally used. They additionally extend this approach by applying spreading activation (Crestani, 1997) to find new relations between tags. Similarly, Tang et al. (2009) aim to learn concepts; i. e. synonymous tags that can be merged, and hierarchical relations between tags. Therefore, they first generate a tag-topic model on the basis of tagged documents using generative probabilistic models. The underlying assumption is that a tag having a high distribution on only one specific topic probably has one specific meaning whilst a tag with similar high distributions on multiple topics is likely to have several submeanings and thus a general tag. Possible relations between tags are determined based on four divergence measures – of the dissimilarity between two topic distributions of t_x and t_y, of the likelihood of t_x being a super-concept of t_y, of the likelihood of two tags describing the same concept or different concepts.

3.2.4.2. Discussion

Despite the various approaches that have been developed in recent years, the results from automatic ontology creation are still not meant to be usable for ontology end users and need additional manual validation. Additionally, the use of external knowledge resources is limited for very specific communities and organizational contexts. Because they have their own very specialized terminology for which public resources are often too general. Similarly, all approaches, except for Damme et al. (2008), have worked on huge public web datasets and their applicability to smaller contexts like in enterprises still has to be proven. Another issue is that most approaches aim to construct the ontologies from scratch based on the whole time period of their available dataset. They do not consider either dynamics within the analyzed dataset or subsequent evolution once the ontology is created. Nevertheless, we can use techniques and methods of presented approaches to support gardening activities (see also Section 6.4.2).

3.3. Semantic Annotation

Semantic annotation adds semantic extra information or metadata to resource contents that reflect their meaning in order to facilitate their organization and retrieval. It *"can denote both the process of annotating and the result of the process"* (Oren et al. 2006). Uren et al. (2006) describe it by *"semantic annotation formally identifies concepts and relations between concepts in documents, and is intended primarily for use by machines"* and Siorpaes and Simperl (2010) refer to "ontology population". This is what Oren et al. (2006) define as *ontological annotation*. They additionally differentiate *formal annotation* that is machine-understandable, i. e. using URIs, but not using ontological terms.

Semantic annotation approaches can be roughly split into two categories: on the one hand approaches that mostly rely on automatic or better to say semi-automatic annotation and on the other hand those that rely on manual annotation (Reeve and Han, 2005). (Semi-)Automatic approaches use natural language processing and information extraction techniques to automatically create annotations for documents based on a training set. Manual approaches support a user in creating annotations with respect to an ontology. Mostly being textual documents, there are also approaches that aim for annotating multimedia contents. Siorpaes and Simperl (2010) additionally summarize approaches for the semantic annotation of web services; all of them based on manual annotation. We may also differentiate web-based annotation tools and desktop annotation tools (Hunter, 2009).

Uren et al. (2006) provide a requirements framework to assess semantic annotation approaches in document centric knowledge management. The seven requirements include: (1) *standard formats* – using standard formats for describing ontologies such as OWL and for annotations such as RDF; (2) *user-centered / collaborative design* – providing usable, collaborative and work-integrated annotation interfaces for knowledge workers; (3) *ontology support* – being able to support multiple ontologies and to cope with ontology changes; (4) *support of heterogeneous document formats* – dealing with multiple document formats besides word processor files like spreadsheets or graphics files; (5) *document evolution* – being able to support document and annotation consistency as documents changes; (6) *annotation storage* – storing annotations separately from or as integral part of the original document; and (7) *automation* – integrating techniques such as natural language processing to facilitate automatic document (collections) annotation.

Based on their requirements framework, Uren et al. (2006) provide an extended review on semantic annotation systems. Additional surveys can be found in Kashyap et al. (2008); Siorpaes and Simperl (2010). In the following, we will give a short overview on some of the approaches with a particular interest for our purposes on the requirements of user-centered/collaborative design as well as ontology support regarding coping with ontology changes. We will focus on document-based approaches. For the specificity of multimedia approaches we would like to refer to detailed discussions by Walter (2010); Hunter (2009); Simperl et al. (2011) Especially multimedia annotation is a recent and growing field on its own; e. g., with the solely dedicated W3C Multimedia Semantic Incubator Group[29].

3.3.1. Manual Semantic Annotation Approaches

Tools for semantically annotating web resources or textual documents and sharing the annotations have existed for over a decade. Besides individual tools, there are two bigger frameworks for annotation in the Semantic Web field. These are Annotea (Kahan and Koivunen, 2001; Koivunen, 2006) and CREAM (Handschuh et al., 2001).

Annotea & Extensions: Annotea is a metadata standard for annotating web documents (HTML or XML based) with a free text description together with information like creator or creation time (Kahan and Koivunen, 2001; Koivunen, 2006). Annotations might be typed e. g., as comment, rating etc. and either locally or publicly stored; with the latter aiming for the collaborative use of annoations. It is implemented in a number of

[29] http://www.w3.org/2005/Incubator/mmsem

tagging tools and server applications, e. g., Amaya[30] and Annozilla[31]. Annotea and its implementations have been developed by the W3C. However, the Annotea project itself has not been under active development for some years.

Koivunen (2006); Hunter et al. (2008); Schroeter et al. (2007) extended the Annotea standard towards ontology-based annotations. HarVANA (Hunter et al., 2008) and Vannotea (Schroeter et al., 2007) aim for supporting the annotation of images respectively videos with pre-defined ontologies. Koivunen (2006) introduced bookmark and topic objects. Users can create and share bookmarks for web documents and attach topics from a topic hierarchy. Users may also create their own topics. However it remains unclear if and how users can share and maintain the topic hierarchy collaboratively.

CREAM: In CREAM (CREAtion of Metadata) annotations are generated either manually or semi-automatically in RDF or OWL (Handschuh et al., 2001). The framework includes besides the annotation interface components for document management, annotation inferencing, information extraction and deep annotation of database generated web pages. Annotation might be stored separately on a server or integrated in a web page. S-CREAM, OntoAnnotate, OntoMat Annotizer are different implementations of the CREAM framework. OntoMat Annotizer (Handschuh and Staab, 2002) supports authors of web pages in creating and maintaining OWL metadata for their pages. It provides a browser and editor to display web pages and the users can highlight parts of text and add OWL mark-ups via drag and drop from a pre-defined ontology, i. e. the users create instances of concepts, of their attribute values and of relations. The users can explore the ontology and its annotation instances via an ontology browser. Bloehdorn et al. (2005) have developed M-OntoMat-Annotizer to support image and video annotation. A more recent re-design of M-OntoMat-Annotizer is the K-Space Annotation Tool (KAT) (Saathoff et al., 2008) with plugins for content analysis or ontology browsing. S-CREAM (Handschuh et al., 2002) is an extension to support the semi-automatic annotation creation based on the Amilcare information extraction system (Ciravegna and Wilks, 2003) and a training set of manually annotated web pages. There is no support for collaboration nor ontology modification.

SHOE & SMORE: Another early system for embedding annotations in HTML web pages is the SHOE (Simple HTML Ontology Extensions) Knowledge Annotator (Heflin and Hendler, 2001). Similar to OntoMat Annotizer but without browser functionality, annotations refer to ontology concepts and relations in the SHOE language. SHOE was the basis for the more complex RDF annotator SMORE (Semantic Markup, Ontology, and RDF Editor) by Kalyanpur et al. (2003) that allows to annotate images and emails besides text and HTML. SMORE supports OWL ontologies and stores the generated RDF annotations separately from the web pages. In addition, users can add new classes and properties to the underlying ontology. However, there is no collaboration support.

COHSE: The COHSE (Conceptual Open Hypermedia Services Environment) Annotator (Bechhofer and Goble, 2001) aims for annotating text passages in web pages and creating navigation links between annotations. It consists of three services: the Ontology Service supports the interaction with the ontologies, the Resource Manager manages resources and retrieves links to other annotations based on concepts and the Distributed Link

[30] http://www.w3.org/Amaya
[31] http://annozilla.mozdev.org

Service adapts the web pages depending on the ontology service's and resource manager's output. Available as browser plug-ins in earlier versions, it is now provided as portlet to avoid installation (Bechhofer et al., 2006). With the portlet, user groups have been introduced. As far as we can see, the users or user groups automatically share their annotations. However, adapting the underlying ontology is not considered. An additional interesting point of COHSE is the gained insight that looser knowledge models are more appropriate and the decision to shift from OWL ontologies to SKOS (Bechhofer et al., 2008).

Open Ontology Forge: Open Ontology Forge (OOF) (Collier et al., 2003) is an ontology editor that integrates annotating web documents. It differentiates between domain managers, domain experts and domain users. The domain manager is a representative who sets up an ontology project and community of experts and users. Domain experts privately develop a RDFS ontology with the OOF client. The domain manager decides when the ontology is ready to be released and publishes it on the Ontology Forge Server public area. Domain users then take the public ontology and start annotating web documents by loading web documents into the OOF client and dragging and dropping selected items to ontology classes in the ontology browser. Annotations can be uploaded and the domain manager copies the annotations from the private to the public server area. Using published annotations as input, an information extraction system is trained to support subsequent annotation processes. It is also possible to annotate images and to extend the ontology by adding new classes to the root. However, the changes seem to be valid only locally. It is not obvious if and how these changes are distributed within the community.

Desktop Integrations: The importance of integrating the semantic annotation process into the users' desktop environment and the document authoring process has also been recognized by annotation environments like AktiveDoc (Lanfranchi et al., 2005) to write and annotate text and WiCKOffice (Carr et al., 2004) or the commercial application OntoOffice (Moritz Weiten, 2006) with extensions to the Microsoft Office Suite. The SATIN (Support for Annotation Integration using Transparent Interfaces in P2P Networks) tool aims for handwritten or keyboard-based in-place annotation and sharing not only of Microsoft Office documents but also PDF documents (Braun et al., 2007; Braun and Hefke, 2009) and follows the (Social) Semantic Desktop vision (Decker and Frank, 2004; Sauermann et al., 2005) that is to link, share and access a user's desktop data across applications and different desktops.

Semantic Desktop Systems: Several Semantic Desktop systems have recently been developed, being the most well-known Nepomuk[32], Haystack[33], Chandler[34], OpenIris[35], DeepaMehta[36] and DBin[37]. For more details we would like to refer to Schandl (2009) who presents an extensive discussion of the different approaches. The importance of collaboration is acknowledged and most systems incorporate collaborative mechanisms, e. g., to publish and share information. Nevertheless, the systems' main focus is to improve the personal information and knowledge management. We will discuss additional

[32] http://nepomuk.semanticdesktop.org
[33] http://freshmeat.net/projects/haystack
[34] http://chandlerproject.org
[35] http://www.openiris.org
[36] http://www.deepamehta.de
[37] http://dbin.org

approaches in the field of social semantic applications, e. g., for annotating wiki content or blog posts, separately in Section 9.1.

3.3.2. Automatic Semantic Annotation Approaches

The best known examples for the mostly automatic approach are the KIM platform (Popov et al., 2003), MnM (Vargas-Vera et al., 2002) and SemTag (Dill et al., 2003). ConAnnotator aims to support cooperative working (Hu and Du, 2006; Wu et al., 2006).

KIM: The KIM platform provides an infrastructure for automatic semantic annotation, indexing and retrieval of textual documents (Popov et al., 2003). It aims for the extraction of and annotation with named entities like persons, locations and organizations from text. It is based on the GATE framework for information extraction and relies on the pre-defined KIMO or PROTON ontology holding the knowledge. This pre-defined ontology is also KIM's limitation.

MnM: MnM supports the automatic and semi-automatic annotation of web pages (Vargas-Vera et al., 2002). Using the Amilcare information extraction engine, it requires training period and based on the training corpus it takes over more and more the annotation process. It provides an HTML browser to display documents and an ontology browser to view the ontology.

SemTag: SemTag is based on IBM's text analysis platform Seeker and follows an annotation process of three steps (Dill et al., 2003): Spotting, Learning and Tagging. The Spotting step retrieves and tokenizes source documents from Seeker and finds labels matching the public TAP taxonomy. For each label match, the surrounding text of ten words to both sides is kept. The Learning step examines a representative corpus sample to determine the distribution of the taxonomy terms. The Tagging step scans the kept text passages from Spotting and disambiguates the matches using a Taxonomy-Based Disambiguation algorithm based on similarity functions. The tool is not targeted to knowledge workers (Uren et al., 2006).

ConAnnotator: Supporting cooperative working and linking ontology development with annotation processes is the aim of ConAnnotator (Hu and Du, 2006; Wu et al., 2006). It uses the Google Web API to crawl web resources that undergo a process of tokenizing and part of speech tagging. Using support vector machines ConAnnotator automatically annotates the whole resource with domain ontology concepts. In a second step it performs information extraction at a lexical level; i. e. extracting basic information about title or author and keywords candidates. The results are then presented to the user to validate the annotations. Additionally, users can relate extracted keywords to concepts or add them to the ontology. ConAnnotator is a component of cooperative ontology development environment CODE (Hu et al., 2005) that incorporates the role-based collaborative development method (RCDM) (Li et al., 2005). There are five kinds of roles with increasing number of members and decreasing privileges, which are Knowledge Manager – main authority of the whole ontology, Knowledge Expert – adjusts the ontology structure when confirmed by knowledge manager, Knowledge Engineer and Knowledge Proposer – enrich the ontology; the former with immediate effect, and Knowledge User – only browses and queries the ontology. Based on these roles, developers have different ontology views and privileges to work concurrently on the ontology. A semi-automatic weighted statistical algorithm supports the resolution of conflicts. Similarly, ConAnnotator based on a user role model. It differentiates annotation manager, annotation expert, annotation engineer, annotation proposer and annotation user by privileges in the annotation creation process. Despite the role system, it remains unclear how the collaboration takes place in detail.

3.3.3. Discussion

Even though semantic annotation approaches do exist for over a decade, they have not found widespread adoption yet. To a large extent, this might be rooted in their perspective on the annotation process (i. e. , the use of the ontology) and the creation of the ontology as two separate processes (Hepp, 2007). As we have seen maintaining and adapting the underlying ontology to the current users' current annotation needs is poorly or not at all supported; based on the assumption that end users won't be willing to create and edit their ontology and rather use existing ontologies (Uren et al., 2006; Kashyap et al., 2008). Thus, different sets of people perform the annotation process and the creation of the underlying ontology; the latter being done by dedicated knowledge engineering specialists. Similarly, the importance of collaborative working is more and more acknowledged but limited to sharing annotations and changes to the ontology remain only personally valid.

More recent approaches focus on investigating on multimedia annotation and how to incorporate information extraction, machine learning and natural language processing techniques to support the annotation process. Usually best results seem to be achieved with training data based approached and where the user has to supervise and validate the machines results (Siorpaes and Simperl, 2010). Still the requirements of user-centered/collaborative design as well as ontology support regarding coping with ontology changes are not sufficiently dealt with. The recently upcoming field of social semantic applications, in which this thesis can be located, provides a more user-centered and integrated view. We will discuss related work separately.

3.4. Advantages and Challenges of Ontologies & Ontology-based Knowledge Organization

Ontology-based systems help to better organize and retrieve relevant resources through machine processable background knowledge. This background knowledge that represents the users shared understanding and in this way enables applications to better fulfill the users' needs. However in practice, ontology-based applications still struggle with its breakthrough whilst folksonomy-based systems made their widespread adoption. In the following we will summarize and discuss the advantages and challenges of ontologies and ontology-based systems.

3.4.1. Advantages of Ontologies and Ontology-based Knowledge Organization

Already lightweight forms like controlled vocabularies have a number of advantages, e. g., in its application of information resource annotation. These approaches restrict the index terms that can be assigned to information resources to a controlled set of terms. With these approaches users cannot use just any term to annotate a resource, but are restricted to the controlled vocabulary. We can summarize the advantages of a controlled vocabulary (cf. Macgregor and McCulloch, 2006b):

- It controls the use of (near-) synonyms by establishing the term that is to be used to represent a word.
- It discriminates between homonyms, i.e. it enforces that every term used has only one well defined meaning.
- It controls lexical anomalies such as grammatical variations or the use of terms without relevance for the information retrieval task (such as leading articles or prepositions)
- A structured vocabulary also facilitates the use of codes and notations that are mnemonic, predictable and language independent.
- In physical environments a controlled vocabulary facilitates the filing, storage and organization of resources.
- It may point the user to closely related, more suitable terms by indicating the presence of broader, narrower or related terms.

Additional benefits provide semantic annotation approaches that rely on a semantically described controlled vocabulary. They make subject, creation, usage, relational or other context of resources explicit and machine understandable through the use of some standardized formal language for representing the ontology, such as RDF, SKOS or OWL.

- **Better Retrieval:** The formally represented relations between the concepts in the ontology can be used to offer superior browse or query facilities. In the case where a powerful language like OWL is used, queries may even be answered using reasoning algorithms.
- **Better Use of Annotation:** The availability of machine understandable context for the used annotation terms can be utilized to make better use of the annotation; e. g., information that some annotations represent geographic locations for which a latitude and longitude is known can be used to show the annotated document in a map or to make them available based on the users current location.
- **Better Quality Assurance:** The information contained in the ontology about concepts used for annotation can enable checks on whether an annotation is likely to make sense; this can help to catch errors early. Also changes in the ontology can be checked whether they violate its integrity.
- **Better (Semantic Web) Integration:** The ontology that is used in the annotation is usually assumed to be also used in other systems and the common usage of the ontology can enable the integration of data created and managed in these diverse systems. Another related aspect is, that semantically annotated data can become part of the Semantic Web and then Semantic Web aware agents and applications can make use of it.
- **Better Support of Vocabulary Management:** Through the use of standardized languages to represent the ontologies, these approaches can rely on a landscape of tools that is available to create, manage and evolve these ontologies.

3.4.2. Challenges of Ontologies and Ontology-based Knowledge Organization

As we have seen, the ontology development approaches, either collaborative or knowledge engineer driven, as well as the semantic annotation approaches view the use of the ontology, for the latter the annotation process, and the development of the ontology as two separate processes, performed by different set of people. Especially the semantic annotation approaches assume the ontology to be predefined and stable, mostly created by dedicated knowledge engineering specialists. However, as stated by Gruber (2009b): *"one can say that ontology is a tool and product of engineering and thereby defined by its use. From this perspective, what matters is the use of ontologies to provide the representational machinery"* (Gruber 2009b). Thus separating the use and the creation of the ontology and involving ontology engineering specialists is causing a number of problems:

- **High Cost:** Knowledge engineers are highly paid specialists (cf. Barker et al., 2004), and their effort comprises not only the actual implementation of the domain ontology, but also learning about and understanding the domain of interest. While in many Web 2.0 scenarios a large amount of work is done for free by users interested in the result, this is unlikely to work when knowledge engineers with little innate interest in the domain in question are involved.

- **Domain Errors:** Knowledge engineers are specialists for the domain of knowledge formalization - not for the domain that is being formalized. For this reason they will not have an understanding of the domain comparable to that of domain experts, this limited understanding may cause errors in the resulting ontology (cf. Barker et al., 2004).

- **Heavyweight Process and Upfront Investment:** Because annotation cannot start without an available ontology, there needs to be an upfront investment to finance the development of this ontology, which includes a systematic requirements elicitation phase. During the usage phase of the ontology, there also needs to be an accompanying process to collect newly emerging requirements, bugs and other change requests and to implement them into a newer version of the ontology.

- **High Time Lag:** There will always be some time lag between the emergence of a new concept and the time when it is included in the ontology and can eventually be used. This time lag is relatively large, when the users of the ontology cannot make the change themselves but must rely on knowledge engineers understanding the requirement, implementing it and finally rolling out the new version of the ontology. In fast moving domains this time lag can quickly get so big that the ontology as a whole becomes unusable (cf. Hepp, 2007).

- **Low Appropriateness and Understandability:** An ontology is appropriate for a task if it enables the users to reach their goals more quickly. However, having different people using and developing the ontology makes reaching appropriateness of the ontology much harder. A particular challenge is to ensure that the ontology is at the right level of abstraction to be understood by the domain experts. Otherwise, if the ontology is used against the intended shared commitment, it may turn useless (cf. Guarino et al., 2009).

3.5. Conclusions

Following the vision of emergent semantics (Aberer et al., 2004b,a), we envision to have *"a community of self-organizing, autonomous [..] agents co-operating in dynamic, open environments, each organizing knowledge (e. g., document instances) according to a self-established ontology"* (Mika 2005). However, communities are rarely homogeneous. There are always varying backgrounds and expertises leading to a different labeling and problems like different abstraction levels. So it's necessary to understand each others' labels and notions and how everything is connected and in this way to develop a shared understanding. At this point, folksonomies are insufficient. According to Decker et al. (2000) ontologies *"provide a shared and common understanding of a domain that can be communicated across people and application systems"*. However, as we have seen, the widely distributed perspective on ontologies, ontology development methodologies and systems making use of ontologies is focused on highly formal ontologies often neglecting the interrelation of the ontology and its modeling purpose and especially the aspect of the shared understanding. And indeed, *"The problem is not in what ontologies are, but how they become shared formal specifications of a domain, and be made operationally relevant and sustainable over longer periods of time"* (de Moor et al. 2006) .

In fact, the shared understanding of a domain is not given but has to be developed by the community. This is a learning process which needs active participation. Active participation fosters the acceptance of the ontology. Additionally, active participation supports the up-to-dateness. And we often don't need highly formalized ontologies (Uschold and Gruninger, 2004). As stated by Fluit et al. (2004): *"Our experiences to date in a variety of Semantic Web applications (knowledge management, document retrieval, communities of practice, data integration) all point to light-weight ontologies as the most commonly occurring type"*. Especially, as soon as we involve people, high formality can be counterproductive in developing a shared understanding.

On the other hand, structures are needed and with this a certain level of formality as they support later use, e. g., for retrieval, and especially the accessibility of the vocabulary, e. g., to explore how existing tags / concepts are connected or how own notions relate to the others' ones. However, we need structures and with this a level of formality, users can understand. It has been shown that normal users already have difficulties with generalization and instantiation. Therefore, we need to lower the level of formality and start with informal "is-a" that is incrementally developed and higher formalized for the relevant (from the community's perspective) parts.

Part II.

Solution

Overview

The goal of this part is the introduction of our methodological and technical framework. The methodological framework consists of the Ontology Maturing model and conceptual design. Figure 3.9 presents an overview of our solution.

The Ontology Maturing model is developed to address the requirements and shortcomings of collaborative ontology development and to support the development of socio-technical systems that enable knowledge workers to construct and maintain ontologies in a collaborative and continuous way, integrated into the usage of these ontologies within their daily work. The model provides a cross-domain description with the metaphor of "maturing". It structures this maturing into four characteristic phases, ranging from emergence of ideas, consolidation in communities via formalization up to axiomatization. We instantiate this model for the maturing of competence ontologies for exemplary purposes (Chapter 4).

The conceptual design framework complements the Ontology Maturing model. Its goal is to support developers in deriving and realizing such socio-technical systems for the collaborative ontology development. It considers technical and non-technical aspects and provides methods and functions that scaffold and guide ontology maturing. Due to the criteria of application-orientation & work integration, the conceptual design framework is oriented towards the application of social semantic bookmarking and semantic people tagging; i.e. the application and use of the ontology for organizing and sharing web resources respectively the knowing on who knows what (Chapter 5).

Finally the technical framework SOBOLEO is an implementation of the conceptual model and the design framework. It is a system that enables knowledge workers to construct and maintain ontologies in a collaborative and continuous way, integrated into the usage of these ontologies within their daily work. SOBOLEO and two instantiations – one for supporting Social Semantic Bookmarking and the other for Semantic People Tagging – are presented in Chapter 6.

We have developed the conceptual model and design framework based on literature and implemented the SOBOLEO framework as supporting system. Each of the three have been refined during their design and development process based on the evaluation by case studies and expert groups (see Part III).

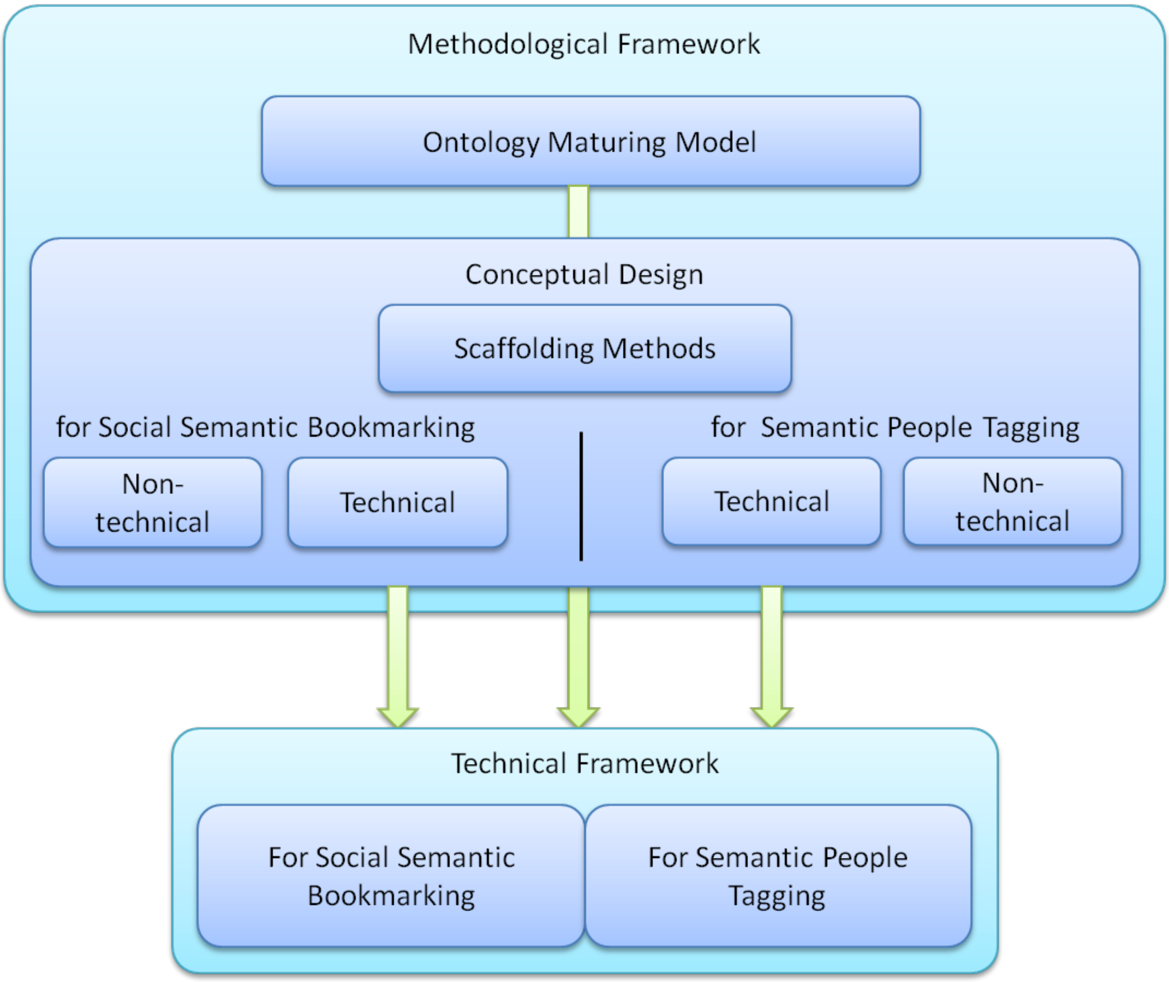


Figure 3.9.: Solution overview

4. Ontology Maturing

Starting point of our approach were the shortcomings of the usual separation of creation and usage processes of ontologies (see 3.2.2.4 and 3.2.3.3 for a detailed discussion). While this might be possible in rather static domains, it is not acceptable for dynamic domains where contents change fast and the ontology requires a permanent update to cover the available contents.

In real world setup, this leads to frustrating situations when users cannot extend the used ontologies by themselves in a **work-integrated** way, e. g., when they require them for the semantic annotation of resources. Instead, they are forced to ask ontology experts for the extension and wait for the update of the underlying ontologies, which – in very dynamic domains – can even last until the ontology element has become obsolete again Hepp (2007).

This led us to rethink ontology engineering as a collaborative and work-integrated activity. In this view, users within their communities themselves modify the underlying ontology of a semantic application, e. g., add new ontology elements or modify existing ones, and make immediately use of these modification. Such an **active participation** fosters the acceptance of the ontology and supports the up-to-dateness.

However, communities are rarely homogeneous. There are always varying backgrounds and expertises. Therefore, the **shared understanding** of a domain is not given but has to be developed by the community. As we have seen, the widely distributed perspective on ontologies, ontology development methodologies and systems making use of ontologies assume an established shared understanding and focus on highly formal ontologies often neglecting the interrelation of the ontology and its modeling purpose. In fact, the development of a shared understanding and with this the development of an ontology is rather a social and collaborative learning process (see also Allert et al., 2006).

However, as soon as we involve people, high formality can be counterproductive in developing a shared understanding. And we often don't need highly formalized ontologies (Uschold and Gruninger, 2004). Indeed, the formalization of ontologies is not possible completely from scratch. In particular for emerging ideas and concepts, it is not possible to directly integrate them into an ontology as they are not clearly defined, yet. Therefore, we need **variable levels of formality** where we start with a lower level of formality that is incrementally developed and higher formalized for the relevant (from the community's perspective) parts.

That means, ontologies need to become **continuously evolving but usable models** where different levels of formality might co-exist. The outcome is an adequate level of formality in the ontology, avoiding both overformalization and the inability to apply reasoning.

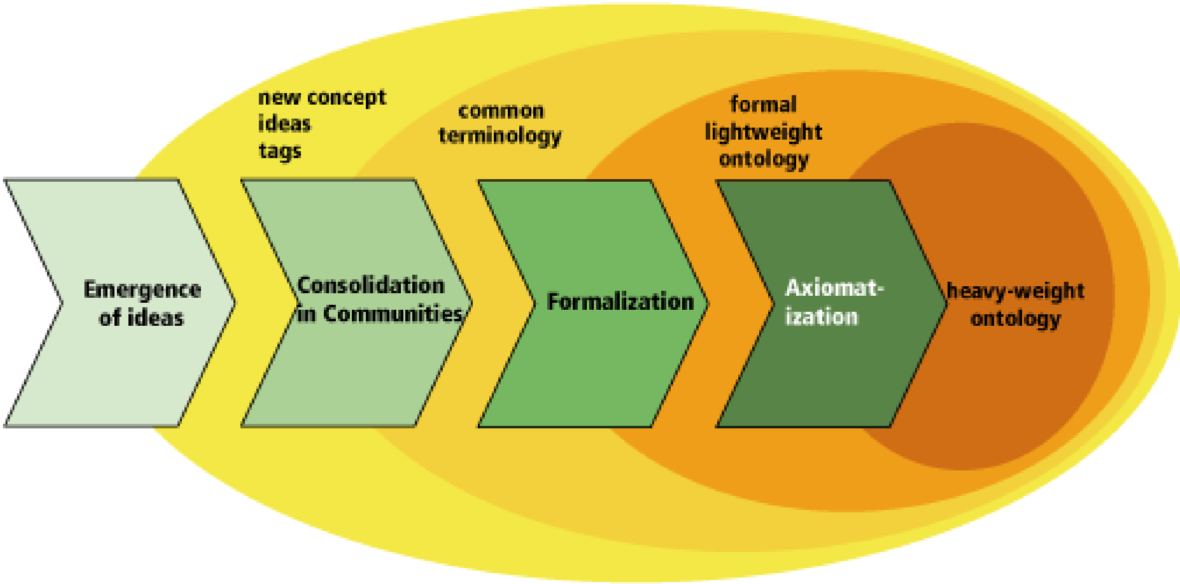


Figure 4.1.: The four phases of the ontology maturing model

4.1. The Model of Ontology Maturing

To address these requirements and shortcomings we have developed the Ontology Maturing model. This phase model aims to structure truly collaborative ontology development processes. Together with the supporting conceptual and technical design frameworks it aims to guide the development of systems enabling knowledge workers to construct and maintain ontologies in a collaborative and continuous way, integrated into the usage of these ontologies within their daily work a reality.

Therefore we combine the benefits of social tagging with those of semantic annotation in order to address their respective weaknesses. Starting with simple tags, each user should be enabled to contribute to the collaborative development of ontologies. For this purpose, we integrate the creation process of ontologies into their usage process, e. g., search and annotation processes. Each community member can contribute new ideas (tags) emerging from the usage to the development of ontologies. The community picks them up, consolidates them, refines them, and formalizes them with semantic relations towards lightweight ontologies.

These observations are similar to those made in Schmidt (2005) about how new ideas develop in the context of knowledge management and e-learning to become reusable training material. This development process was described with the metaphor of maturing and structured into five phases as the so-called knowledge maturing process. This process is viewed as a macro model for interconnected individual learning processes.

The Ontology Maturing model identifies characteristic maturing transitions and structures the ontology engineering process into four phases (see Fig. 4.1):

1. **Phase Emergence of ideas.** New ideas emerge and are introduced by individuals as new concept ideas or informal tags. These are ad-hoc and not well-defined, rather descriptive, e. g., with a text label. They are individually used and informally communicated. For instance, while annotating or seeking for resources, we recognize that the tag we want to use does not exist. Accordingly, we introduce a new tag like "spaghetti" and "tomato sauce".

2. **Phase Consolidation in Communities.** Through the collaborative (re-)usage of the concept symbols (tags) within the community, a common vocabulary (or folksonomy) develops. The concept ideas are refined, useless or incorrect ones are rejected. For instance, when comparing currently envisioned tags with previously used ones or with tags from other people, we discover similarities and differences that allow for creating concepts from tags; e. g., we realize that "spaghettoni" noodles are the same as "vermicellini" noodles. We establish a link between these terms in our understanding and thus can merge synonyms into concepts. The emerging vocabulary, which is shared among the community members, is still without formal semantics.

3. **Phase Formalization.** Within the third phase, the community begins to organize the concepts into relations. These can be taxonomical (hierarchical) ones as well as arbitrary ad-hoc relations, e. g., in the course of becoming aware of different abstraction levels such as "pasta" being broader than "spaghetti". This results in lightweight ontologies that rely primarily on inferencing based on subconcept relations.

4. **Phase Axiomatization.** In the last phase the adding of axioms allows for inferencing processes, e. g., in query answering systems. For instance, spaghetti are only made of seminola and water and salt and therefore fine for egg white allergy sufferer. This step requires a high level of competence in logical formalism so that this phase is usually done with the aid of knowledge engineers.

It is important to note that ontology maturing does not assume that ontologies are built from scratch, but can be equally applied to already existent core ontologies used for community seeding. Likewise, this model must not be misunderstood as a strictly linear process; rather real ontology development processes will consist of various iterations between the four different phases.

4.2. Different Contributors in Ontology Maturing

If we translate this to corporate practice, we can distinguish between different types of contributors with different levels of involvement and expected skills (cf. Figure 4.2):

- In the first phase of **Emergence of Ideas**, we mainly rely on a large number of individual contributors with little or no knowledge about modeling. They are mainly concerned with their task at hand and use tagging practices to find resources or people later on more easily. They can align themselves with other peers through observing their tagging behavior and potentially through tag suggestions (which essentially is a system-mediated observation of the behavior of others).
- In the second and third phase of **Consolidation in Communities** and **Formalization**, we rely on gardeners – people who devote themselves especially to cleaning and structuring the knowledge base; often because of their personality. They partially and incrementally consolidate the tags, usually focused on areas with a high volume and heterogeneity of tags. These gardeners are usually not in a special centralized function unit, but rather emerge from their community (but can be equipped with organizational legitimation from their superiors) because of their interest and mission. They play the important role of facilitators of the consolidation process although they do not accomplish the task alone. These gardeners do

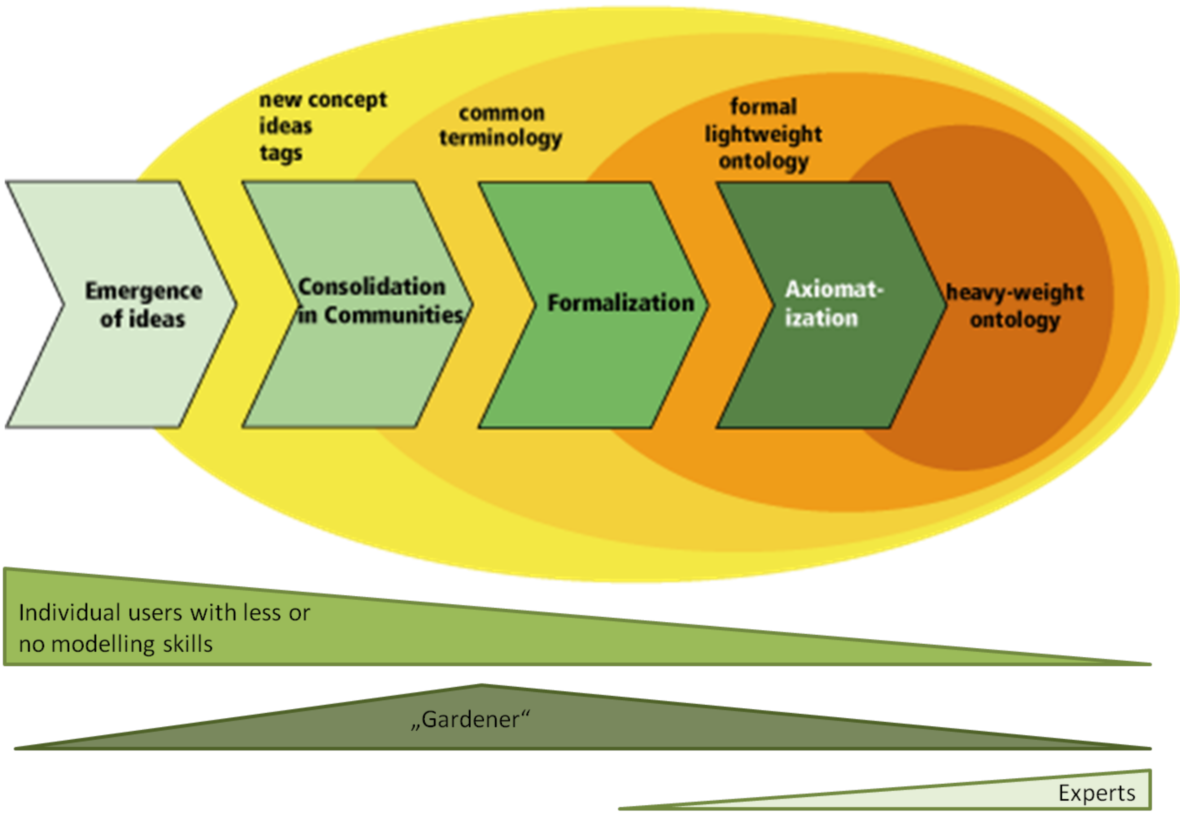


Figure 4.2.: Relating the involvement of different types of contributors to the four phases of the ontology maturing model

not necessarily need to be experts in modeling, but need basic semantic modeling know-how to discover problems and suggest solutions.

- In the last phase of **Axiomatization**, mostly experts or expert groups will be responsible for core thematic areas, which are important to formalize. These are similar to the knowledge engineers in the traditional approaches, but as opposed to those, they are now informed about what is considered important by employees as part of their daily activities.

4.3. The Artifact, Knowledge and Social Dimensions of Ontology Maturing

Our evaluations (which are described in III) have shown, that concentrating on the development of the ontology as a mediating artifact is not sufficient to prepare for sustainable community-driven use and evolution of ontological knowledge structures. Beyond the mere construction of an artifact, we have to consider that users have different levels of understanding of parts of the domain (e. g., identified by interest in background knowledge to improve their own understanding, asking for help, or taking the lead within a group) and that this understanding also evolves within usage processes. Furthermore, the social dimension of community-driven sites has to be addressed, e. g., which instruments are needed to support a growing community. As a consequence, we need to describe ontology maturing in three different dimensions: the artifacts, knowledge and social dimensions.

The artifact dimension is concerned with the created ontology elements, the knowledge dimension with the maturing and alignment of knowledge, and the social dimension with the development of competencies and social structures.

- **Artifact Dimension:** Artifacts are "something viewed as a product of human conception" that are touchable or visible items. In folksonomies, tags are the product of human conception. In semantic applications, ontologies are considered as a product of formalized human conception. Using our ontology maturing model, artifacts mature from simple tags over common terminologies to formalized or even axiomatized ontology elements as described in the previous section. Thus, the artifact dimension identifies the available ontology elements and their relations. This dimension has (naturally) been the focus of semantic technology research so far.
- **Knowledge Dimension:** Users can only model appropriately what they have sufficiently understood, and the process of modeling usually involves a deepening of the understanding of the real-world topic. That is knowledge in a narrow sense, i.e., domain knowledge in a non-tangible form, including "know-what" and "know-how". Within the knowledge dimension, we need to distinguish between individual knowledge and the abstraction of collective knowledge. On the level of the individual, augmenting and changing that knowledge is what is usually called (individual) learning. Here, we need to consider alignment processes that bring forth a sufficient level of shared understanding of the domain. On the collective level, it is an aggregation of individual pieces of knowledge. Here, in contrast to the individual level where an individual might just learn what others have learnt before, this is about the active construction and development of an understanding as such that advances knowledge on the collective level.
- **Social Dimension:** Viewing ontology development as collaborative learning processes, e. g., interaction, communication and coordination among the individuals, we have to consider the social structures and processes in the social dimension. Users can only build a shared understanding, shared artifacts and methods to create these if they learn to act in a social context and perform collaborative actions on the individual as well as on the collective level. This includes shared values, 'unwritten' normative orientations as well as regulating norms for actions. Learning on the individual level comprises a general willingness and capacities to interact with others, to communicate, negotiate, compromise and accept rules – a process of socialization.

When we want to design systems that enable knowledge workers to construct and maintain ontologies in a collaborative and continuous way, integrated into the usage of these ontologies within their daily work, we need to be aware of these three dimensions and their complex interactions. Thus, they will turn up again implicitly and explicitly in our conceptual design framework: for instance, the artifact dimension is represented by the ontology itself; the knowledge dimension is inherently bound to the provided collaborative nature by which the community develops its knowledge and shared understanding (in contrast to knowledge engineer-driven development approaches) and the social dimension e. g., manifests in rules and norms.

4.4. Ontology Maturing for Competence Ontologies

In this section, we want to exemplify our hitherto abstract ontology maturing model by the maturing of competencies in companies that is a key domain of application. Knowledge about competencies and capabilities of its employees is an essential need for an organization and its development. This encompasses activities like finding the right person to contact, team staffing or identifying training needs. This has been addressed by approaches to expert finding and competence management like Schmidt and Kunzmann (2007); Biesalski and Abecker (2005); Mcdonald and Ackerman (2000); Becerra-Fernandez (2006). However, it has turned out that these approaches (which are more or less based on a top-down philosophy) fail to live up to their promises because data on competencies and expertise are not up-to-date or competence catalogs don't contain the information relevant to the users.

We aim at overcoming these difficulties by a higher degree of participation, i.e., by broadening the scope of people maintaining competency data and the competence catalog. Therefore, we combine Web 2.0-style bottom-up processes with organizational top-down processes: through Web 2.0 oriented bottom-up processes we make use of the "wisdom of the crowd" effect and collect the collective view of the community of employees on the competencies of the individual. Employees assign tags to each other referring to expertise or interests. The organizational processes take up and guide these bottom-up developments towards organizational goals.

We expect to bring benefit in: (1) a higher up-to-dateness and completeness of the employee profiles, (2) a more realistic assessment of competencies and expertise than with self-assessment, and (3) an additional awareness for the tagged person who can see his/her colleagues' perspective.

However, we need to adapt Ontology Maturing in certain aspects to consider competence management specific aspects. Also we give more concrete instantiations of the different formality levels and phases:

- **Shared vocabulary for comparability.** Competencies usually have an integrating function in the enterprise, bringing together strategic and operational levels, and human resources, and performance management aspects. This means that competencies are not limited to an individual or to a group, but these notions have to be shared by the whole organization (in the ideal case): in consequence we need a shared vocabulary.
- **Legitimation and commitment by the organization.** If competencies are to play an important role in diverse organizational processes, ranging from team staffing, via human resource development process, up to organizational competence portfolio management, it is important that the resulting competency profiles and competency catalogs are not only derived from the "wisdom of crowd", but have also the commitment of the organization. This is a main difference to the open world of the Web of individuals. Major decisions depend on the appropriate identification of competencies and competency profiles so that the organization must decide at some point to which extent it relies on the result of collective bottom-up processes and to which extents it defines certain binding aspects.

4.4.1. Phases of Competence Ontology Maturing

When applying Ontology Maturing to competence management, we realize that there are indeed two different strands of knowledge that mature but are inherently interwoven: the knowledge about people's individual expertise and the knowledge how to describe people's expertise by means of semantics, i.e. the ontology. Both strands are closely

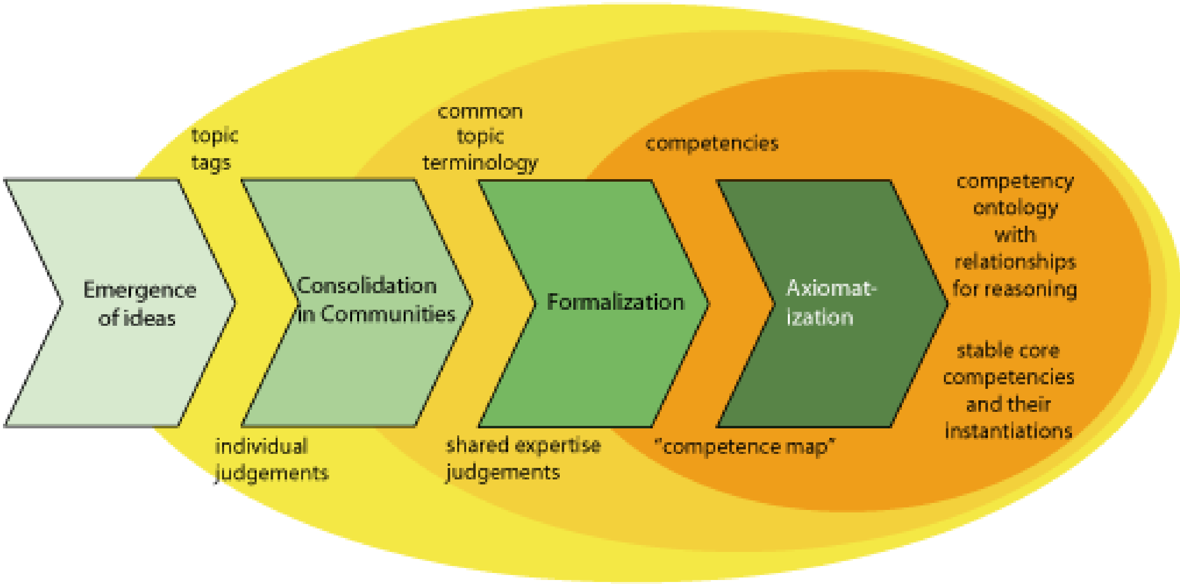


Figure 4.3.: Ontology Maturing for Competence Ontologies

interrelated as higher maturity of (collective) knowledge about others' expertise requires a matured vocabulary to describe it (see Figure 4.3).

1. **Phase Emergence of ideas.** At this phase new topics are taken up, e. g., from news, by surfing the net, daily client requests, or team meetings. Or an employee might get into contact with another person s/he has not known yet, e. g., based on another colleague's recommendation or an event such as a conference. By annotating the other person with any topic tag, the user starts to document his/her newly gained knowledge, e. g., about new topics or contact data that are deemed useful for later retrieval. In this way new topic ideas emerge; especially very recent or very specialized topics. These topic tags as well as judgement on the person are rather personal and reuse is restricted to the "inventor" because other users are not yet aware of the new topic tag or person.

2. **Phase Consolidation in Communities.** A common topic terminology evolves through the collaborative (re-)usage of the topic tags within the community of employees. Topic tags are distributed and negotiated through the reuse and structuring of tags. They are defined and refined, useless or incorrect ones are rejected. Similarly, expertise judgements are shared through the reuse of other people's knowledge about individual experiences.

3. **Phase Formalization.** Within the third phase, usually special members of the community (who might be additionally legitimated by the organization by assigning "gardening" tasks) begin to organize the topic terminology into competencies by introducing relations between the topic tags. These relations can be taxonomical (hierarchical) ones as well as arbitrary ad-hoc relations, expressing similarity (e. g., Java Programming and C# Programming). That results in new or updated competency notion, i. e. lightweight ontologies, which allow primarily for inferencing based on subconcept relations. Similarly, the agreement about people profiles results in a "competence map".

4. **Phase Axiomatization.** In the last phase, we formalize measurable definitions of competencies on a department or unit level. This usually goes along with the introduction of an explicit and defined HR development process. Modelling experts

add axioms for exploiting relationships for reasoning. This includes in particular precise composition relationships. That allows for complex inferencing processes, e. g., subsumption of competencies for the purpose of competency gap analysis, or competency-based selection of learning opportunities (cf. Schmidt and Kunzmann, 2007). Going on, a pure HR development topic may be extended to a strategic company topic with company-wide definitions, including core competency definitions and prioritization. Similarly regarding expertise knowledge, we may have competence development from an HR development perspective on a department or unit level with explicit and defined processes that may be extended towards explicitly defined processes for strategic competence management to reach stable core competencies and their instantiations.

This can serve several purposes and use cases (see also Braun et al., 2010b):

- Colleagues can find each other more easily, e. g., for asking each other for help.
- Employees become aware of other colleagues with similar interests or experience to stimulate the formation of communities.
- For team stuffing purposes, it provides information about the employees' actual profiles for matching them with the requirements.
- It supports human resource development by providing information about the aggregated needs (e. g., by analyzing search requests) and current capabilities of the employees (aggregated tagging data) to make the right decisions about training required.

4.4.2. Variable Level of Formality in Competence Ontology Maturing

In the following we show how the different phases of ontology maturing result in different levels of formality. These different levels of formality co-exist within a single model. For representing these different levels of formality so that we can also exploit the information, we build upon the conceptualization of competencies as part of the Professional Learning Ontology[1] Schmidt and Kunzmann (2006). This conceptualization has three basic levels: topics (as weak notions), competence types (without differentiation) and competencies (with levels). These relate to each other as shown in Figure 4.4 (see Appendix A.1 on the notation).

In this way, we can (1) represent all four phases of the ontology maturing process and (2) deal with less formal statements if needed. Especially, the latter is important for the different use cases of competence models (Braun et al., 2010b; Schmidt and Kunzmann, 2007):

- **Topic tags.** As many Web 2.0 sites or Farrell et al. (2007b) show, tags are sufficient to provide a basic level of useful search and retrieval functionality and similarity between the tagged resources. Precise tag definition would help, but are not needed.
- **Competence types.** For basic profile matching, we need well-defined competency notions and taxonomic relationships to allow for different levels of abstraction by using broader-narrower relationships. For instance, <OWL Modeling> has broader competence <Ontology Modeling>. We can also perform basic competency gap analysis (by exact matching).

[1] http://professional-learning.eu/competence_ontology

- **Competencies (with levels).** This allows for a more extended version of profile matching as you can have different degrees of fulfillment for individual competencies. For instance, <OWL Modeling Beginner>, <OWL Modeling Intermediate>, and <OWL Modeling Expert>. This can also form the basis for describing the objectives of learning opportunities (trainings, learning objects).
- **Competency relationships.** If we have precise is-a semantics, or composition of competencies in the competence model, we can introduce the notion of competence subsumption. For instance, if <Protégé-OWL Editor Expert> and <Ontology Engineering Methodology Beginner> are part of <OWL Modeling Intermediate> then <OWL Modeling Intermediate> subsumes <Protégé-OWL Editor Expert> as well as <Ontology Engineering Methodology Beginner>. Similar is true for the competency levels: for instance, <OWL Modeling Expert> is higher than – thus subsumes – <OWL Modeling Intermediate> and is higher than <OWL Modeling Beginner>. This allows for more sophisticated competency gap analysis (as in Schmidt, 2008), and competency-based selection of learning opportunities.

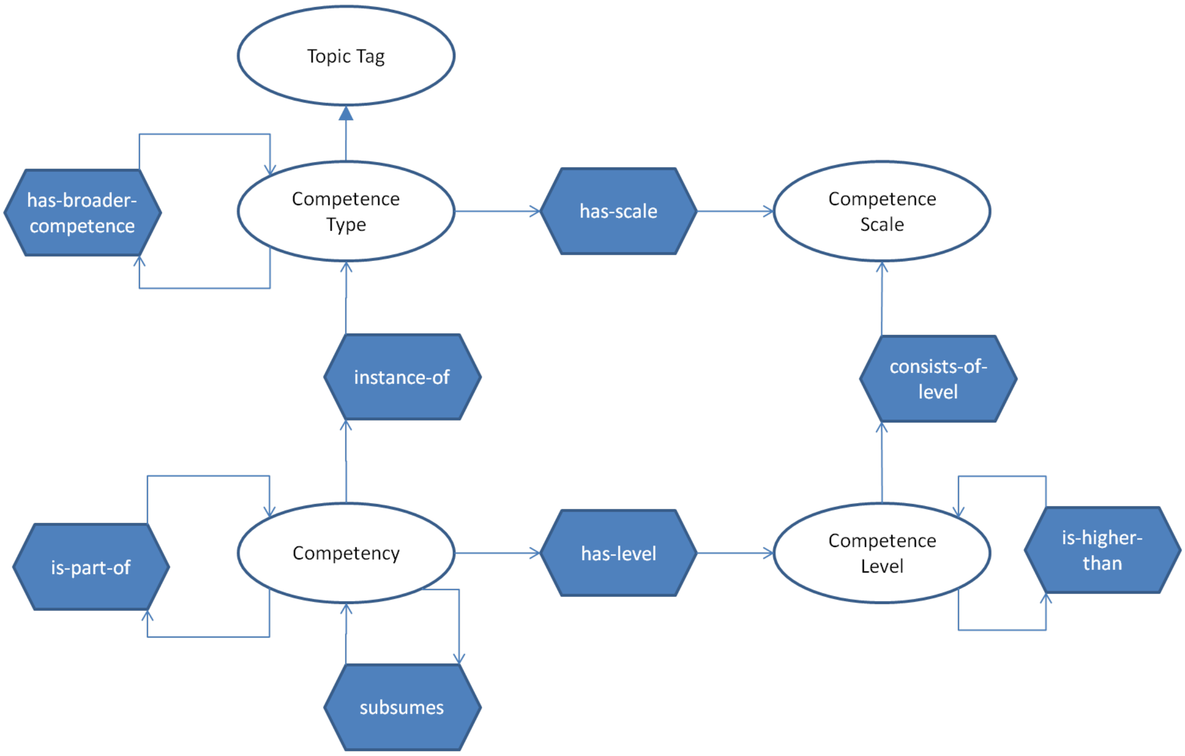


Figure 4.4.: Core competence model adapted from Schmidt and Kunzmann (2006)

5. Conceptual Design Framework

The previous chapter introduced ontology maturing as a conceptual phase model to structure the collaborative evolving and use of ontologies. In itself, however, the Ontology Maturing model is not sufficient as a guideline for creating applications that support this process. Therefore, we need any notion of how systems can support the different phases. We also need to consider technical as well as non-technical aspects. Thus, we need to provide guidelines to adapt the process and systems to a given organizational setting in order to design a working socio-technical system.

This chapter introduces the conceptual design framework that complements the Ontology Maturing model to derive and realize such socio-technical systems for the collaborative ontology development.

We start with the Seeding – Evolutionary Growth – Reseeding model (SER model) that shows and emphasizes in particular the necessity for reseeding activities to achieve "cleaned" and higher formalized knowledge units and structures and the necessity for scaffolding and guidance. Based on this model we collect, describe and categorize methods and functions.

We then proceed to the design framework for social semantic tagging that enables and fosters ontology maturing. That is further concretized for the application and use of the ontology for organizing and sharing web resources – Social Semantic Bookmarking – and the knowing on who knows what – Semantic People Tagging.

5.1. Scaffolding Ontology Maturing

5.1.1. Seeding, Evolutionary Growth and Reseeding

Having a closer look at the individual maturing phases, we notice that the theory of *Seeding, Evolutionary Growth and Reseeding* (SER) by Fischer et al. (1994, 2001) is applicable (cf. Schmidt et al., 2009, p.83). The SER model was developed to describe the evolution of complex systems. An initial seed of knowledge units, structures and capabilities form the starting point for a complex, socially driven and evolutionary development process. The community of users develop the original units, structures and capabilities over time through the interaction and use of tools for combination, analysis and change. This activity leads to evolutionary, undirected (and often confusing) growth of the seed.

At some point in time, the growth may stagnate. The evolved system may loose order and usefulness, e. g., because inconsistencies, dislocation and disorganization etc. crept in. Reseeding is necessary in order to organize, generalize and formalize the knowledge base and in this way to keep the system manageable. This reseeding can happen in a form of consolidation and negotiation processes in which the variety of units, structures, and capabilities are assessed, pruned and restructured .

New seed may be necessary to stimulate the next cycle of evolutionary growth and reseeding. Therefore, Fischer et al. (2001) argues not only for providing the community supporting tools for evolutionary growth activities but also for participating in the reseeding phase; e. g., through mechanisms making the users aware of activities within the system or tool suggestions how to formalize the knowledge.

5.1.2. Ontology Maturing and SER

The SER model is important for Ontology Maturing because it emphasizes the necessity for reseeding activities to achieve "cleaned" and higher formalized knowledge units and structures that are the more and more stable basis for further developments and the necessity for scaffolding and guidance support; i.e. the need for embedding communication with and history of an artifact during the evolutionary growth and the need for tool support for knowledge formalization during reseeding.

For example, if we consider the maturing phase of *Consolidation in Communities*, we may seed the collaboration space of the community with an initial knowledge structure of concept ideas (tags) and relations together with their knowledge units, e. g., annotated documents. As McAfee (2006a) stated: *"one of the most surprising aspects of Enterprise 2.0 technologies is that even though they're almost completely amorphous and egalitarian, they appear to spread most quickly when there's some initial structure and hierarchy"*. This also goes along with our experiences from evaluations. Indeed, as we do not assume absolute correctness or completeness, we may intentionally provoke users to contribute through the shortcomings of the seed.

For the evolutionary growth through use, we have to provide the community not only tools to use the instantiated knowledge structures, units and capabilities, but also to extend, combine, and change these by themselves according to their needs and tools to comment and give feedback. *"Talking about an artifact requires talking with the artifact. Therefore later interpretation of the discussion requires that the discussion be embedded in the context in which it was originally elicited."* (Fischer et al. 2001) Analysis tools enable the community to monitor and guide its activities.

At some point, when a certain level of maturity is reached, we need to decide whether to take the concept ideas to the level of *Formalization*. If the development of the concept ideas stagnates, we initiate reseeding activities such as pruning the current knowledge base, tool suggestions to use existing concepts or add a relation between these, introducing new ideas or people into the community or changing the topic.

However, this does not mean that we may equate one SER cycle with an ontology maturing phase where maturing takes place in well defined steps and strict order of growth and reseeding. Indeed, we aim for the domain experts and users themselves organizing, generalizing, and formalizing the knowledge and not only with the assistance of knowledge engineers. Therefore, growth and reseeding activities are rather interwoven and may take place in parallel. For instance, integrating a new piece of knowledge may cause the users to add this into the existing knowledge structures or to restructure their mental models in order to accommodate it.

Due to this interweaving and emphasis on domain experts and users, it is important to provide continuous triggers to engage in reseeding activities and recommendations what and where it is worthwhile to garden the knowledge base.

Additionally, Fischer et al. (2001) do not further consider to support the development of a shared understanding that is essential for ontology maturing; especially because different understandings cause inconsistencies and disorganization that require reseeding. Therefore, it is also important to provide guidance during the growth phase that helps to prevent inconsistencies and disorganization creeping in and in this way prolongs the evolutionary growth and need of reseeding.

5.1.3. Scaffolding Methods and Functions

5.1.3.1. Nudge, Agree, and Aggregate

In the following we will present and categorize methods and functions that scaffold and guide ontology maturing in its growth and reseeding phases. Based on existing systems for collaborative knowledge creation and formalization (cf. Tekin, 2009), we can additionally identify three main classes of methods and functions – provided in a technical or non-technical form: (a) *Nudge* and (b) *Agree* support evolutionary growth and (c) *Aggregate* specifically supports reseeding.

- **Nudge:** nudge methods guide the users during the creation and input of new knowledge and information resources; e. g., during the tagging process naming conventions within the community as a non-technical form or auto-completion and disambiguation as a technical form support shared (re-)use of (existing) vocabulary elements.
- **Agree:** agree methods scaffold the unification and consensus reaching process; e. g., through structured discussions or voting and rating.
- **Aggregate:** aggregate methods provide suggestions based on data analysis and aggregation; e. g., concept candidates to merge.

5.1.3.2. Categorizing Methods and Functions

Table 5.1 provides an overview on methods and functions that scaffold and guide ontology maturing in its seeding, growth and reseeding activities. This overview brings together examples from existing systems and what has been perceived to be useful, e. g., from our evaluations (see Part III) but might be extended for additional application and organizational contexts.

Seeding

- **General:** To avoid in particular the cold start effect of an empty collaboration space, *import functionalities* help to provide a seed with existing vocabulary, resource and user information. Especially in organizational contexts, there are usually preexisting data from other sources to build upon or integrate with. Therefore, we may apply ontology learning techniques that *extract candidate concepts* from community produced documents[1] (see e. g., Maedche, 2002; Missikoff et al., 2002).

[1] What Missikoff et al. (2002) refer to as *implicit consensus*.

Scaffolding Ontology Maturing		Methods & Functions
Seeding	General	- Import functionality for ontology, resource and user information, and annotations - Candidate concepts extraction from texts

Evolutionary Growth	General	- Easy adding of new knowledge
		    -Automatic concept creation
		    -Annotation bookmarklet
		    -Everywhere add/edit options
		- Tag Recommendations based on resource content
		-Make aware of new knowledge
		    -Automatic publication
		    -Search, Browse, Feed notification etc.
		-System feedback on own contributions
		    -Make social use visible

	Nudge	-Auto-completion -Disambiguation -Typing error correction -Blacklist -Tag recommendations from vocabulary; .e.g. own popular tags, others' tags -Social norms, e.g. naming convention -Training, Tutorials, Examples
	Agree	-Discussions & Dialogues -Find the right discussion partners -Feedback mechanisms, e.g. rating, comments -Voting -Mutual editing/amending/complementing -Revision management
Reseeding	General	-Easy vocabulary manipulation -History -Subscriptions to changes -Blacklist
	Aggregate	-Make aware of (un-)used elements, discussion status -Identify inconsistencies, redundancies, or gaps -Recommendations for ontology enrichment - Knowledge completion based on usage data analysis

Figure 5.1.: Categorizing scaffolding methods and functions
Categorizing scaffolding methods and functions that support ontology maturing w.r.t. seeding, evolutionary growth and reseeding

Evolutionary Growth

- **General:** Knowledge structures and formalizations are used to make data better accessible to the computer and to enable more powerful processing; however, their value and with that their potential to grow only becomes visible when they are used. Therefore, we need methods and functions that make even small steps in formalization visible. For instance, in systems like the Semantic Media Wiki: as soon as just a few attribute values have been specified, these can be used to create tables and overview pages that before had to be maintained manually. Hierarchical organization of tags allows for more effective maintenance of the tag repository as well as for more effective navigation and retrieval. This works after having just one such relation. With an *advanced search* function, adding just one synonym for a tag/concept will already improve the search experience: searching for this synonym will then also consider the documents annotated with the concept.

  Therefore, we also need methods and functions to easily *add new knowledge* to the system during its utilization. For instance, *annotation bookmarklets* allow to quickly tag, store and share a bookmark for the currently opened web resource from within the browser. Another function is the *automatic take up of tags / concept ideas* hitherto not existing/used into the vocabulary during the tagging. As well during the tagging, *tag recommendations for new tags / concept ideas* e. g., gained by content mining techniques for text based resources or the analysis of user behavior may help to extend the vocabulary. *Everywhere add or edit options*, similar to systems like Wikipedia for each page, provide the opportunity to contribute new contents from various places within the system.

  For the growth and especially the consolidation in the community, the newly added knowledge should be available for the other users. Here, making new knowledge elements *automatically public and visible* to the community can help. However, this publication does not imply that the ideas and knowledge are distributed. It's necessary to *make others aware* of the existence of this new knowledge so that others can reuse it and it is accepted in the community. For example, we can say that users get aware of a person and associated tags, e. g., via search, browsing or feed notification, when they access the person's profile, when they add additional tags or approve the already assigned tags. Therefore, we provide simple features to access a person or any information resource in order to see which topics are how often assigned to and to approve tags.

  When the effort of ontology engineering is distributed over a large user base of a system then a large part of any contribution is not realized by the contributor itself but by other people who benefit from the contribution; e. g., more information on a concept improving the search. Hence the effects and value of an individual user's changes (e. g., the increased usefulness of a concept, more visitors and possibly a better rating) is – as a start – invisible to the contributor. However, making this *social use visible* can motivate further contributions. Therefore, methods and functions are needed to also make this social use visible; e. g., the number of people that used a tag/concept or information resource a particular user has contributed to.

  Additional explicit *system feedback*[2] can foster user contribution and motivation

[2] We understand "system feedback" broadly as an output of a software system in response to user actions with the goal of helping the user to understand the effects of his or her actions (Niegemann et al., 2008). Explicit system feedback means outputs designed specifically as personal feedback to support a user and verify or summarize his/her actions. For instance, a message like "Didn't you actually meant to

(Cheshire and Antin, 2008; Ling et al., 2005; Mazarakis et al., 2011). So we may use feedback mechanisms like (a) gratitude – i. e. displaying simple "thank you" messages, (b) historical reminder – i. e. showing the number of contributions of an individual user so far, (c) relative or (d) social ranking – i. e. illustrating the percentage of the contributor's edits in relation to the contributions of the other users or respectively the number of contributions in comparison with other users with numbers in proximity – to stimulate user contributions.

- **Nudge:** In order to nudge the users to reuse and consolidate ontology elements during the tagging activity, we may provide *auto-completion* support or *tag recommendations of existing tags* from the vocabulary (cf. Sen et al., 2006). Functions like *typing error correction*, *disambiguation* or even checks against a *blacklist* help to prevent inconsistencies and disorganization.

  Besides technical functions, we may also develop and establish non-technical guidelines within the user community; e. g., in the form of *social norms* how to work together and use the system and thus prevent disorganization – similar to naming conventions such as CamelCase practice in social tagging systems like Delicious or the "Wikiquette"[3] of Wikipedia. *Tutorials, trainings and examples* may additionally help to establish common understanding and practices and to manage the collaboration space in this way.

- **Agree:** To give space to complex negotiation processes among potentially conflicting viewpoints and to reach consensus, *discussion and dialog* support forms the foremost function; like for instance provided by the DILIGENT argumentation framework (see Section 3.2.3.1). As mentioned by Fischer et al. (2001) it is helpful to embed discussions in the context of the artefacts. On the one hand, this makes decisions and with this their rationale accessible for later use. On the other hand, it provides the opportunity to easily and directly engage in discussions. For instance, Wikipedia provides a discussion page for each article. In this way, it is also important to find the right people for discussing (see Section 7.2).

  To express agreement or disagreement, we may apply *feedback mechanisms* like comments and ratings. Additional *voting* mechanisms can be used for decision taking like it is for instance implemented by the Cicero plug-in or consensus-based ontology engineering approach(see Section 3.2.3.1 and Section 3.2.3.1).

  Besides verbose negotiations through discussions and dialogs, users may also reach consensus and agreement through *collaborative and mutual editing and complementing* of knowledge structures and units as particularly wiki systems have shown. *Revision management* helps to roll back unwanted decisions.

Reseeding

- **General:** Reseeding the vocabulary and especially performing gardening activities, e. g., through directly amending the vocabulary, implies that the supporting function and methods should facilitate *vocabulary manipulation*, e. g., through an ontology editor to manage semantics.

search for this keyword instead" that is specifically created to support the user in the interaction with the system, is explicit feedback, whilst a search result page gives implicit feedback on the formulation of a user's query, but it is not designed specifically for this goal.

[3] http://en.wikipedia.org/wiki/Wikipedia:Wikiquette

History functions for knowledge elements, information resources or users help to monitor community activities. Making user activities transparent additionally facilitates identifying knowledgeable people; e. g., as discussion partners. As an additional push method, users may *subscribe* to elements and their changes; e. g., via RSS or ATOM feeds. Checks against a *blacklist* may additionally support freeing the knowledge base from potential inappropriate or unwanted contents.

- **Aggregate:** Aggregating and *making aware of the use and discussion status* of knowledge elements and information resources, help to set focus for reseeding activities. For instance, which knowledge elements or resources might be removed because they are insufficiently used which might be further improved because they are intensively used.

  We may support the reseeding and gardening by *identifying inconsistencies, redundancies or gaps* and providing suggestions; e. g., concept candidates to merge. Methods to extract semantics from folksonomies (see 3.2.4.1) can be used to provide *recommendations for the enrichment of the ontology*. In order to come from lightweight to heavyweight ontologies, we may use algorithms as presented by (Lacasta et al., 2010, pp. 99) that help for instance to identify a broader/narrower relationship between two concepts as an is-a relationship.

  By the analysis of usage data, e. g., tags used, related resources, etc., we may gain additional information that helps us to complete the knowledge base. For instance, we may deduce a person's expertise based on his/her activities and therefore augmenting the person's profile by the tags used for annotations.

  From a guidance perspective, it is not only about agreement e. g., on people profiles but also to monitor what knowledge is requested and thus needs to be developed. For example, showing the topics searched for in comparison with the topics used for annotation within the last month. Based on this, decisions can then be drawn accordingly.

Even though some of the presented functions and methods seems to fit naturally to a specific ontology maturing phase at a first glance, our evaluations have shown that a mapping of these methods and functions explicitly and exclusively supporting a specific phase is not feasible.

5.2. Social Semantic Tagging

In the following, we present a framework for the socio-technical design of a *Social Semantic Tagging System* that enables and fosters ontology maturing. Before, we will explain the concept of Social Semantic Tagging and provide a formal model.

We have seen that (linguistic) tagging approaches, while popular, struggle with problems such as polysemy, multilingualism or abstraction level mismatches. At the other end many current semantic annotation approaches struggle (like most approaches building on controlled vocabularies of some kind) with the problem of timely updates and appropriateness of the controlled vocabulary as well as affordable creation. Social Semantic Tagging combines the positive aspects of semantic annotation with those of collaborative tagging in order to address their respective weaknesses. Social semantic tagging allows for the annotation of resources with tags extended by semantic definitions and descriptions that also evolve (collaboratively) within the same system. Similar to tagging approaches, new tags can be created whenever a need arises. Unlike these approaches, tags can have powerful descriptions and can be interlinked; for example allowing the system to understand that 'swimming bath' and 'swimming pool' are synonyms for the same concept. These powerful tag descriptions are similar to those used in traditional semantic annotation, but social semantic tagging allows for adding and changing these concepts permanently and easily within the same system and at the same time as these concepts are used.

5.2.1. Social Semantic Tagging Model

We extend the formal model of folksonomy $F := (U, T, R, A)$ presented in Section 2.5.1 to $SST := (U, ST, R, STA)$ with semantic tags $ST = \{st_1, ..., st_l\}$ that are concepts from a shared ontology O, i.e. $ST \subset O$, and semantic tag assignments STA as a set of ternary relations between users U, semantic tags ST and resources R; i.e. $STA \subseteq U \times ST \times R$. For Social Semantic Bookmarking $SoSeBo := (U, ST, B, STA)$, resources R are restricted to bookmarks $B = \{b \,|b \in R\}$ that are references to online web documents identified by the URL. For Semantic People Tagging $SePT := (U, ST, P, STA)$, resources R are restricted to persons $P = \{p \,|p \in R\}$ representing real people.

5.3. Design Framework

Our social semantic tagging model provides the basis; being the users U, the resources R, the semantic tags ST respectively the ontology O, and the semantic tag assignments STA the main objectives of design. We will discuss different design options and their implications regarding these main elements[4]. The design options and their implications may vary depending on the application and use context. Therefore, we will first discuss the general options for each element in Section 5.3.1 before looking closely at the application and use for semantic people tagging (brief SePT) in Section 5.3.2 and social semantic bookmarking (brief SoSeBo) in Section 5.3.3. Table 5.2, 5.3, 5.4, and 5.5 provide a condensed overview of the design framework[5].

We have developed this design framework based on the design of folksonomy-based systems in the literature (see Section 2.6.1), the analysis of social semantic bookmarking systems (cf. Braun et al., 2009b) and the evaluation of semantic people tagging with Human Resource experts (see Section 8.2) and field experiments (see Section 8.1). Our experience (see also the evaluation part of this thesis) has shown that there is no universal design for social semantic tagging, especially for the sensitive subject of people tagging, but any implementation has to be flexible enough to adapt to a particular organization. In this way this framework brings together of what has been perceived to be useful with focus on tagging links to web resources and people but might be extended for further application and organizational contexts.

5.3.1. General Design Options

We identified three main aspects to design the elements R, O and STA of a social semantic tagging system for ontology maturing. These aspects are: a) Who – regarding rights & control to edit and view the elements; b) What – regarding type & properties of

[4] In the following we will look at semantic tags within the scope of the ontology.

[5] empty cells mean that there are not particularities for semantic people tagging or social semantic bookmarking and the general options do apply

Users				
Design Aspect		**General Options**	**SePT**	**SoSeBo**
Types & Properties	Role	Tagger, creator, provider, owner, gardener, administrator	Additional: taggee	
	Social connectivity	(typed/directed) links, groups, none	Refinement: organizational units	

Figure 5.2.: Design Aspects for Users

the objectives; and c) How – regarding support & tools to implement the objectives. The users U are distinct because the first and last aspects are not essential or reasonably applicable for design, e. g., rights & control may be doubled because the users edit and view the other objectives. Therefore, we focus on types & properties, for which we consider:

- **Role[6]:** The most obvious and important role in social semantic tagging systems is the *tagger*, who creates the associations between tags and resources. Regarding the specificity of rights related to other objectives we may distinguish additional roles. For instance, by assigning tags to resources the tagger may also become the *creator* of new tags or *provider* and *owner* of new resources. A special role may be taken by *gardeners*; people who devote themselves especially to cleaning and structuring the knowledge base. They can establish themselves out of the community or explicitly be nominated by the organization. *Administrators* may be responsible for the configuration of the different design options.

- **Social Connectivity:** Systems may allow their users to connect each other. We can differentiate social connectivity in forming *links*, *groups* or *none*; whilst links may additionally be *typed* (e. g., friends and family) and/or *directed* (e. g., follower and following). Such social structures may additionally support the adoption and convergence on the used tags (ontology) (cf. Marlow et al., 2006) and influence the users' motivation (cf. Section 2.7.1.2).

5.3.1.1. Resources

Regarding the types & properties of resources, we consider the type of object, the source of material and the resource connectivity in detail (cf. Marlow et al., 2006):

[6] There might be additionally applicable roles e. g., from the organizational structure or surrounding system landscape but we do not dwell on at this point.

- **Type of Object:** Systems may support different kind of contents as resources being tagged, e. g., *bookmarks*, *people*, *multimedia objects*, *events etc*. This might influence the nature and type of assigned tags.

- **Source of Material:** The resource content being tagged may be *user-contributed*, *system-supplied* or *global*, e. g., bookmarks link to web pages taken from another web source. This might have an impact on the users motivation to tag as well as on the nature and type of assigned tags (see Section 2.6.1).

- **Resource Connectivity:** Similar to the social connectivity of users, systems may allow to connect resources via direct *links* or *groups* beyond the users' *tags*. This might lead to connected resources being similarly tagged.

Regarding the aspect of rights & control, we can further subdivide editing in the creation and the deletion & modification of resources resulting in:

- **Creation:** By the creation of resources, we refer to the creation of the representation of the object within the system; i.e. the users are not necessarily the creator of the resource content but rather its provider. The creation options are closely connected to the previous aspects type of object and source of material. If the source of material is system-supplied, it is only the *system* that can create the resource representation within the system. If users can provide resources, this may be done by *any user* or limited to a *specific group* with extra privileges. The restriction to the system and/or a responsible group might avoid scattering and provide better strategic guidance. On the other hand, this might negatively influence the users' motivation to participate and contribute in cases they cannot annotate the resources they want to.

- **Deletion & Modification:** Besides the same options as for the creation, we additionally consider the *resource creator/provider* (viewing him/her as owner) for the deletion and modification of a resource or *none*; i.e. a resource once added cannot be modified or deleted anymore. However, restricted or no modification and deletion options can hamper the maintenance of the knowledge base.

- **Visibility:** With the creation of a resource, it has to be decided to whom it shall be visible. We may have a public visibility to *any user*, a private one limited to the *resource creator*, or a restricted one limited to *specific group* of people, e. g., friends or team colleagues of the resource creator. The latter option can help to disseminate information to a specific target group and in this way to avoid "spamming" the whole organization or to avoid disgrace and exposing oneself, a barrier particular to early maturing phases (Barnes et al., 2010). On the other hand, public visibility fosters serendipitous discovery of interesting and useful information.

Content control mechanisms or social norms may support to implement the previous aspects:

- **Content Control:** *File type filter* can regulate the type of objects. Thus only certain types may be used as taggable resources. Similarly, methods of *web filtering* such as URL and domain filters may support the control of the source of material from any other web source (e. g., regarding bookmarks). Systems may control the creation and deletion of resources based on *word filters* or other more advanced techniques.

- **Social Norms:** The user community may implicitly or explicitly develop and establish rules to control resources, e. g., to use only resources of *specific type* or *from* *specific sources* or that are *only work-related* topics. For instance on CiteULike[7] the community only organizes literature references even though it is possible to add any URL.

The community may also develop guidelines how to publish and distribute resources or how to modify and delete them; particularly when any user can view and delete them. This may be done for instance by special tags for the former case or by notifying every use who used the resource for the latter case.

Resources				
Design Aspect		**General Options**	**SePT**	**SoSeBo**
Types & Properties	Type of object	Bookmarks, people, multimedia objects, events etc.	Refinement: internal/ external contacts Additional: opt in vs. opt out	Bookmarks
	Source of material	User-contributed, system-supplied, global	Refinement: self-contributed, external profiles	Open for any web resource
	Resource connectivity	Links, groups, none	Refinement: organizational units	
Rights & Control	Creation	Any user, restricted group, system		
	Deletion & Modification	Any user, resource creator/provider, restricted group, system, none		
	Visibility	Any user, resource creator/provider, restricted group		
Support & Tools	Content Control	File type filter, word filter, web filtering etc.		
	Social Norms	From specific sources, specific types, only work-related	Refinement: Internal/external contacts	

Figure 5.3.: Design Aspects for Resources

[7] http://www.citeulike.org

5.3.1.2. Ontology

The types & properties of the ontology crucially depend on the chosen ontology language. We take SKOS (see Section 3.1.3.3) as a starting point and look at semantic tags (i.e. concepts) and relations.

- **Semantic Tags:** One or more *labels* and *descriptions* may denote semantic tags. They may be multilingual, e. g., to support users with different linguistic background. Each semantic tag may be connected through relations with one or *multiple* other tags.

- **Semantic Relations:** Semantic relations may interlink semantic tags and/or tags with the resources. There might be unspecified associative relations or different types of relations[8]. On the one hand, the latter provides higher expressivity and utilization for e. g., reasoning. On the other hand, the user community may have difficulties to understand and use the finer differences and therefore unspecified associations may be preferred.

  Attributes may further specify semantic relations. These are mainly *reflexivity*, *symmetry* or *transitivity* with their inverses. These may be utilized for retrieval purposes.

The semantic tags and relations may be edited and controlled in the following way:

- **Creation:** Besides the same options as for the creation of resources, we additionally list the *gardener* to emphasize his/her position. We already discussed the difficulties with a completely controlled vocabulary; i. e. if the users cannot use or modify their own tags. On the other hand, as our expert interviews have shown (see Section 8.2), it might be necessary from an organizational and strategic perspective to fix a specific part of the vocabulary as a core, for instance the top level categories, around / beneath which the users continue developing the vocabulary.

  The right of creation may also be applied to different level of details. Depending on the options of semantic tags and relations, any user may create new tags and define their meaning with description and different labels. However, they may not create relations, because these are only automatically extracted by the system; e. g., relatedness based on co-occurrence analysis like it is done by BibSonomy. Or the users may connect the tags only with relations from a set of relation types predefined by the system; e. g., SKOS basic types for hierarchical and associative relationships. This can help to reduce the complexity for the user community.

- **Deletion & Modification:** Similarly to the creation options, we consider, besides *any user*, *gardener*, *restricted group*, and *system*, the *tag creator* or *none* for the deletion & modification at the different level of details of the ontology. Where the latter option may mean that possible mistakes cannot be corrected anymore.

- **Visibility:** It has to be decided to whom the creation, deletion & modifications are visible and effective. We may have a public visibility to *any user*, a private one limited to the *tag creator*, or visibility to a *restricted group* of people. For instance, on BibSonomy you share your tags but not the relations you have defined between the

[8] There are various studies on classifying types of relations but no standard ontology of relations. Markowitz et al. (1992) propose a taxonomic hierarchy of relations that distinguishes at the top level *taxonomic classification*, *semantic markers*, *parts, wholes, and aggregates*, *generic typical case frame filters*, and *ordering and measuring*. (Weller, 2010, pp. 170)

tags; whereas on Fuzzzy semantic relations are immediately visible and effective system wide for every user.

On the one hand, the public visibility makes similarities and differences in the used vocabulary and understanding among the users transparent. However, it may also show mistakes of the individual. Therefore, some users may prefer leaving semantic tags and relations private until they reached a certain confidence to make them public.

Tools like word filtering with blacklists, tagging support, ontology enrichment and social norms may help to implement the previous aspects:

- **Blacklist:** Additional control might be achieved based on a blacklist; e. g., in order to avoid inappropriate tags. Such a blacklist of tags can be implemented by automatic system checks that block or delete the tags. This can be done either a priori before its creation or a posterior in the course of gardening processes.
- **Social Norms:** What are *appropriate tags and relations* to use or not to use or what are appropriate ways to modify and delete may also be established by community rules.
- **Tagging Support:** Regarding the control of the ontology, it might be influenced during the tagging process. We can differentiate three forms of tagging support: (a) *blind tagging* where users cannot see tags assigned by others to the resource; (b) *viewable tagging* shows already assigned tags and thus can foster the reuse of tags, and (c) *suggestive tagging* where the system recommends possible tags for annotation. Especially auto-completion or suggestions of existing *tags from the vocabulary* can foster reuse and consolidation (cf. Sen et al., 2006), whereas recommendations of tags e. g., extracted from document contents might encourage including new tag ideas. New tag ideas are additionally supported if it is possible to simply use new tags during the tagging thus they are *automatically created* and integrated into the ontology.
- **Ontology Enrichment:** In order to automatically create and enrich the ontology, systems may use *lexical resources*, e. g., dictionary for adding translations, or *statistical analysis*, e. g., co-occurrence techniques for relatedness. However, as we have seen in Section 3.2.4.1, these require a manual review, e. g., by the gardener.

Ontology					
Design Aspect			**General Options**	**SePT**	**SoSeBo**
Types & Properties	Semantic Tags		Labels, description, multilingual, (multiple) connected		
	Relations		Types, reflexive, symmetric, transitive		
Rights & Control	Creation	Who	Any user, gardener, restricted group, system		
		Level of detail	Tags, tag meaning, relations, relation type		
	Deletion & Modification	Who	Any user, tag creator, gardener, restricted group, system, none		
		Level of detail	Tags, tag meaning, relations, relation type		
	Visibility	To whom	Any user, tag creator, restricted group		
		Level of detail	Tags, tag meaning, relations, relation type		
Support & Tools	Blacklist		Blocking , deleting; a priori , a posteriori checking		
	Tagging Support		Blind tagging, viewable tagging, suggestive tagging; tags from vocabulary, automatic creation		
	Social Norms		‘appropriate’ tags and relations, ‘appropriate’ deletion and modification	Refinement: professional tags, non-professional tags, negative tags	
	Ontology Enrichment		Exploiting lexical resources, statistical analysis		

Figure 5.4.: Design Aspects for the Ontology

5.3.1.3. Tag Assignments

Regarding types & properties of tag assignments, we look at the aspects aggregation and semantics of the tagging:

- **Aggregation:** Systems may support two different kinds of aggregation: (a) one tag may be *multiply assigned* to the same resource by different users, or (b) the community collectively annotates an individual resource; i. e. one tag can only be *once assigned* to the same resource. Where the latter fosters a higher community interaction and communication, the former allows to exploit aggregated statistics; e. g., analyzing tag frequencies or distributions on the level of resources.

- **Semantics of Tagging:** The semantics of assigning a tag to a resource again technically depends on the chosen ontology language. On the one hand, dealing with unspecified semantics (like, e. g., the resource is associated with a topic) might facilitate the assignment of tags because of its lower cognitive effort (Sinha, 2005). The users can simply choose the tags they associate with the resource without being required to think about the differences in semantic relations. On the other hand, differentiated semantics, e. g., 'is about', 'has location', 'has author' etc., allow for more differentiated reasoning or fine granular retrieval support.

For the rights & control of tag assignments we consider again the creation, deletion & modification and visibility at different levels of detail; i. e. the semantic specificity of the tagging:

- **Creation:** Besides *any user* (that is also called "free-for-all tagging"), a *restricted group* and the *system* based on mining techniques, it may also only be the resource owner/provider who is allowed to tag a resource, like e. g., it is done on Flickr[9]. However, this may hamper the sharing of information within the community.

- **Deletion & Modification:** This aspect is similar to the ones of the ontology whereas extended by the resource owner/provider and the *tagger* as the creator of a tag assignment.

- **Visibility:** The visibility of tag assignments can reach from an unrestricted visibility for any user over a limited group of people towards only the tagger or resource owner/provider being able to see the assignments. We may also think about time aspects; i. e. if tag assignments are immediately visible or need to be approved before being published, e. g., by the resource owner/provider because of certain owner or quality claims. However, the upfront approval requires additional effort for the resource owner/provider and might result in a time lag till tags get visible, which again is frustrating for the taggers. Tag assignments may also be valid and visible only for a specific time period. This can support up-to-dateness and regular maintenance.

  There is also a close coupling with the question of the visible level of detail and particularly the anonymization. Having an anonymized view, i.e. information about the person who made a tag assignment is not visible, can make the taggers feel more confident and lead to more tag assignments on the one hand. On the other hand, aggregated and anonymous tag assignments might hamper evaluating their expressiveness. For instance, people rather might trust more the assessment of experts and thus their tag assignments than the ones of beginners.

[9] The resource owner can additionally authorize other users for a resource.

Tools like blocking based on blacklists, tagging support, and social norms may help to implement the previous aspects:

- **Blacklist:** Blacklists can help controlling tag assignment similarly to the ontology and additionally be used for an automatic approval mechanism in order to block potentially inappropriate or unwanted tag assignments.
- **Tagging Support:** Regarding the control of tag assignments, we can again differentiate the three forms of tagging support *blind tagging*, *viewable tagging*, and *suggestive tagging*. This may be supplemented by a *seeding from external sources* that can help especially at the beginning to overcome cold start difficulties. However, our field experiments have shown (see Section 8.1) that people might feel biased by the displayed tags towards confirming these instead of adding new ones.
- **Social Norms:** What are *appropriate tag assignments* or what are appropriate ways to modify and delete assignments may also be established by community rules. If the semantics of tag assignments are not further specified, the user community may also develop a common understanding on how to interpret them.
- **Mining Techniques:** Mining techniques, like *content mining* for text based resources or the *analysis of user behavior* for people, may help to automatically create tag assignments.

Tag Assignments					
Design Aspect			**General Options**	**SePT**	**SoSeBo**
Types & Properties	Aggregation		Multiple assignments, one time assignment	Multiple assignments	
	Semantics of tagging		Without any further semantics, with specific system-/user-defined semantic relationship	'interested in', 'has competence', 'occupied with', weighting (on level XY)	
Rights & Control	Creation	Who	Any user, resource owner/provider, restricted group, system based on mining techniques	Refinement: self-tagging	
		Level of detail	Semantics of tagging		
	Deletion & Modification	who	Any user, tagger, resource owner/provider, gardener, restricted group, system based on blacklist, none	Refinement: taggee	
		Level of detail	Semantics of tagging		
	Visibility	Time	Need for approval, blacklist blocking, immediate visibility, validity date		
		To whom	Any user, tagger, restricted group	Refinement: taggee	
		Level of detail	Aggregated and anonymous, tagger is visible, semantics are visible, details visible for restricted group	Refinement: highlight self-tagging, semantics with weighting	
Support & Tools	Blacklist		Blocking , deleting; a priori , a posteriori checking		
	Tagging support		Blind tagging, viewable tagging, suggestive tagging; seeding from external sources		
	Social Norms		'appropriate' tag assignment, Interpretation of semantics	'interested in', 'has competence', 'occupied with' etc.	
	Mining techniques		User behavior analysis, content mining	User behavior analysis, content mining	Content mining

Figure 5.5.: Design Aspects for Tag Assignments

5.3.2. Design Options for Semantic People Tagging Systems

Going into details for Semantic People Tagging Systems, we can look at the following particularities for the four main objectives.

- **Users:** For semantic people tagging, we additionally have to consider the role of the *taggee* that is the person being tagged and who may have e. g., particular rights regarding tags assigned to the taggee's profile.

  Regarding the aspect specificity social connectivity, we have to take into account the *organizational structures* and *units* by which the users are additionally connected. These may influence the users' motivation and tagging and general behavior. For instance, the transparency of discussions and changes can interfere with organizational hierarchy; as our expert interviews have shown (see Section 8.2). Apart from this, we may use this information e. g., for visibility settings to define the restricted group or tag suggestion purposes.

- **Resources:** The type of object in focus of semantic people tagging are apparently *people*. We may further refine this and decide – either technically or as social norm – on which people can be tagged; only *internal*, i. e. colleagues within the organization or a specific department, or also *external* contacts. Including external contacts might enrich the search for potential contact persons on the one hand; on the other hand people might not want to share their personal social network as it is seen as an important capital of their own. Another issue to be clarified is if any colleague can be tagged from the very beginning, but who might later *opt out*, or only system users who first *opted in*. The latter one might limit the use when semantic people tagging is introduced in a smaller group with the aim for incremental spreading. The former one provokes that everyone is forced to participate at the beginning or when you start on a small scale that people are tagged without having the opportunity to object.

  Depending on the decision who can be tagged, the *users* may provide the persons to tag – particularly if external contacts should be recorded as well, or they may *contribute "themselves"* if they have first to opt in. The *system* may further supply a predefined set of taggable people – particularly if the focus is on a restricted group of internal contacts – or it may be open for *external profiles* as source; i. e. if the principle of social bookmarking is applied, people are tagged based on their personal websites or profiles in the Web like on Facebook or LinkedIn.

  Similarly to the users' social connectivity, the taggees may be grouped through the organizational structures in *organizational units*.

- **Ontology:** From our field experiments (see Section 8.1) the regulation and control of what are "appropriate" tags and relations and what are not turned out to be crucial. This particularly pertains to *non-professional* or *negatively assessing* tags. Therefore it is necessary to specify if such tags are acceptable or not, i. e. if only *professional tags* are allowed, either by social norms or organizational guidelines. Additional control might be achieved based on a blacklist in order to avoid inappropriate and especially malicious tags.

- **Tag Assignments:** Concerning aggregation, it is necessary to go for *multiple assignments* because the frequency of a tag assigned to a person is an important indicator for the person's knowledge/expertise.

When answering the question how semantically differentiated tagging is, we can on the one hand deal with *unspecified semantics* like, e. g., “a person is associated with a topic”. That might facilitate the assignment of tags because the users can simply choose the tags they associate with the other person without being required to think about the differences between competences and skills etc. or different levels. On the other hand, differentiated semantics, e. g., *’having a competence’* vs. *’being occupied with’* a topic or ’being on the *level* of beginner’, ’. . . intermediate’, ’. . . expert’ etc., allow for more differentiated reasoning.

Regarding the question who is allowed to assign tags, the options can reach from any user being allowed to tag, over a limited group of persons like friends, team colleagues or the one who contributed the contact, to only *self-tagging*. That means that only the users themselves can assign tags to their individual profiles, like it is very common on social networking sites like LinkedIn or Xing. Besides, it may also be decided if self-tagging is appreciated or not in general; i. e. when others can tag. On the one hand, self-tagging may be used for seeding purposes in order to avoid e. g., cold start difficulties. The study by Raban et al. (2011) also indicated that self-tagged users are the more productive contributors (see also Section 9.2.3). On the other hand, it may be seen as distasteful self-advertising. Limiting the group of authorized people might avoid that people can tag who only know little about the other person or about the topic they assign (thus limiting the impact of their assessment).

An additional issue to particularly consider is that some people fear the transparency – especially of being associated with topics they don’t feel confident enough – as our field experiments have shown (Section 8.1). So the visibility of tag assignments may reach from an unrestricted visibility for everyone over a limited group of people towards *only the taggee* being able to see and eventually approve the tag assignments. We already detailed the pros and cons of an upfront approval mechanism.

Having a closer look at the visible level of detail and particularly the anonymization issue for people tagging, an anonymized view can lead to more ’honest’ and in general more tag assignments on the one hand. On the other hand, anonymity might also foster mobbing. Sometimes, we may also go for a mixed visibility of the details; i. e. showing an aggregated and anonymized view to any user and a more detailed view *with weighting* (e. g., beginner, intermediate, expert) only to the taggee or a specific group of people. If self-tagging is possible, we may also highlight self-assigned tags in contrast to tags by others.

To automatically create tag assignments or to provide useful tag suggestions, we may use *user behavior analysis*. For instance, if the user annotates often with certain tags and makes very differentiated assignments, i. e. not only using very general topics, then the user may be knowledgeable about these tags/topics as well. If the principle of social bookmarking is applied, we may additionally use *content mining* techniques to extract keywords from the website.

5.3.3. Design Options for Social Semantic Bookmarking

Whereas we can identify several particularities for semantic people tagging because of its natural sensitivity, the general options mostly apply to the design of social semantic bookmarking.

With respect to resources, the type of object in focus are of course *bookmarks* that link to web pages taken from the *global* source of *any other web resource*. To automatically create tag assignments or to provide useful tag suggestions, we may use *content mining* techniques to extract keywords from the website.

5.3.4. Search Heuristics Using Social Semantic Tags

Finding the right resource or person based on the query of a user is a non-trivial task. The ranking algorithm incorporates heuristics that are based on certain assumptions and depend on the previous design decisions, which may apply in one organizational context, but not in the other. In the following we present a list of indicators and the underlying assumptions and reasons.

- **Tags of the Resource:** This is the most obvious assumption: if a resource gets tagged, we assume that the resource is somehow associated with the tag and thus relevant if someone searches for the tag. For persons, we can additionally differentiate between self-assigned tags and tags by others.
- **Frequency of tags:** The more often a tag is assigned, the more relevant the resource is for a specific tag. This leverages collective review.
- **Time stamp of the tag assignment:** The more recent a tag assignment is, the more relevant it is. This is especially true for people since a person could have thematically reoriented.
- **Background knowledge on the structure of tags:** If a resource is assigned with broader or narrower or related tags, it gets less weight in comparison with exact matching tags.

Besides the previous indicators that tackle characteristics of our main objectives Resource, Semantic Tag Assignment and Ontology, we may further consider less ostensible indicators that relate to the social dimension.

- **Tagging activities of the tagger:** If the tagger is a highly active user and makes differentiated assignments, then the tagger's assignments and resources get more weight.
- **Additional activities of the tagger:** If the tagger created, edited or interacted with documents or contributed to discussions etc. for a specific topic, then the tagger's assignments gets more weight.
- **Social connectivity between searcher and tagger:** Tagger who are closer connected with the searcher; i. e. have a shorter or strategically preferred path in the searcher's social/organizational network, get higher ranking because they are likely to have more commonalities like similar interests and viewpoints.

5.3.4.1. Semantic People Tagging Specific Aspects

For the application of semantic people tagging, we may consider the following additional aspects relevant for our search heuristics:

- **Tags of the tagger:** If the tagger is tagged with the same tag as he/she assigns to another person, then it gets more weight. Example: If a Google Web Toolkit (GWT) expert assigns the tag GWT to someone else, this is a more meaningful judgment than that of a person, who hardly knows what GWT means, assigning this tag.
- **Tagging activities of the taggee:** We may conclude on the taggee's expertise from his/her tagging activities. Therefore, if the taggee is a highly active user for a specific topic and makes differentiated assignments, then it gets more weight.
- **Additional activities of the taggee:** Similarly to the activities of the tagger, if the taggee created, edited or interacted with documents or contributed to discussions etc. for a specific topic, then it gets more weight.
- **Current availability of the taggee:** If the request is urgent, taggees who are available in time and/or location get a higher ranking.
- **Social connectivity between searcher and taggee:** Taggees who are closer connected with the searcher; i. e. have a shorter or strategically preferred path in the searcher's social/organizational network, get higher ranking because it is easier to establish the contact.
- **Social connectivity between tagger and taggee:** The social relationship between tagger and taggee might influence the tag assignments. Example: If close colleagues assign a tag, it is more meaningful, because they may know the taggee better than other colleagues with loose contact.

5.4. Concluding Remarks

For the organizational implementation it is necessary to analyze and specify the different design aspects according to the respective organizational context in the sense of a system-culture-fit. This should include an organization analysis. Hereby we have to consider that the different design possibilities are partly dependent on each other and might limit the utilization of the information afterwards, e. g., not storing the information of the tagging person excludes the tags of the tagger as an indicator for the search.

6. Technical Framework

6.1. Overview of the Technical Framework

In this chapter we present our technical solution. The SOBOLEO framework is an implementation of the conceptual model and the design framework. It is a system that enables knowledge workers to construct and maintain ontologies in a collaborative and continuous way, integrated into the usage of these ontologies within their daily work.

Specific example settings from our evaluation case studies for SOBOLEO are:

- An HR department organizing data about available experts. Here SOBOLEO supports an HR department in maintaining a competency ontology and using this ontology to keep track of – and search within – the database of experts. Note that the competency ontology is never really finished – as new possible skills are emerging all the time (q. v. Section 8.3).

- A learning course that is jointly developing an understanding of a domain, creating an ontology and discussing and interlinking it with relevant documents. SOBOLEO supports this use case through the management of all three – the ontology, the discussions and the relevant documents. Note that here the ontology is also evolving throughout its use - always reflecting the current knowledge of the course participants (q. v. Section 7.6).

- A group of experts jointly collecting the state of the art in a scientific domain (with links to both documents and experts). Here, too, the ontology will constantly evolve to reflect both the knowledge of the experts as well as the current scientific consensus (q. v. Section 7.2).

We describe the SOBOLEO framework and its configuration for the two instantiations – one for supporting Social Semantic Bookmarking and the other for Semantic People Tagging. Figure 6.1 shows an overview of this chapter's structure. We will start with the architecture and the main data models. We will show afterwards how it is made use of the models by zooming in the different components of the architecture. We will explain the ontology editor & management as well as the dialog & discussion parts before going into the details of the two instantiations of social semantic bookmarking and semantic people tagging, because these are relevant and used in both instantiations. Afterwards, we want to recap how the presented system functions and methods enable and support the particular phases of ontology maturing in both the application of social semantic bookmarking and semantic people tagging, before we conclude this chapter with a summary on how the SOBOLEO framework technically realizes the design options and scaffolding methods and functions of the conceptual design framework (cf. Section 5.3 and Section 5.1.3.2).

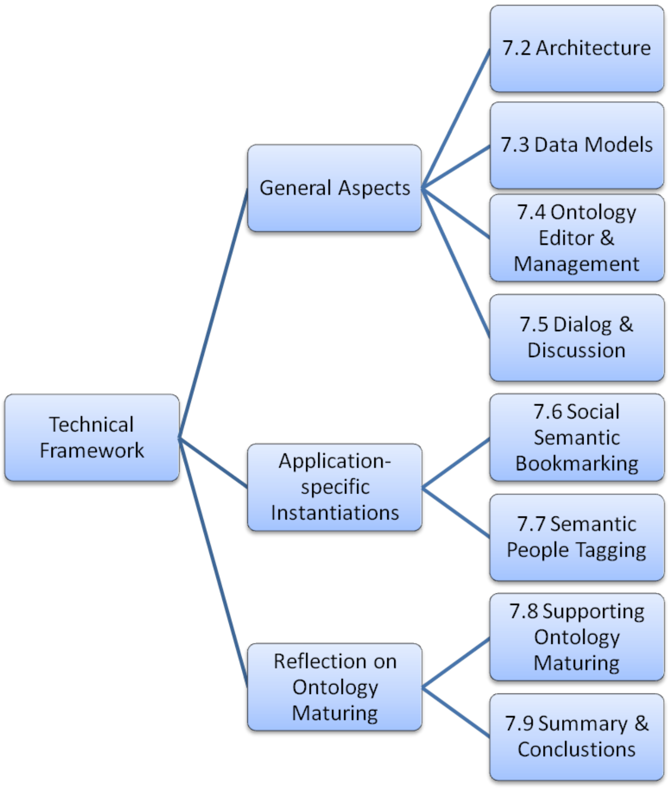


Figure 6.1.: Overview of the technical framework

6.2. Architecture

SOBOLEO is fundamentally organized around the concept of *Collaboration Spaces*. A Collaboration Space is the virtual space in which collaboration between people with a shared goal takes place. In this way, different user groups with different application cases and organizational requirements can be supported.

Each Collaboration Space has one SKOS ontology (see also Section 3.1.3.3) and may have information about documents and/or people that are annotated and organized with this ontology. All information in a space is jointly edited by the members of this space (spaces can also be configured to allow anonymous users to read and write).

Each SOBOLEO installation supports an arbitrary number of collaboration spaces and these are (except for user data) completely independent of each other. All collaboration spaces within the SOBOLEO system are isolated and share nothing but user information.

SOBOLEO is built in a four-tier architecture (five, when counting the AJAX application running in the browser). An overview of the architecture is shown in Figure 6.2. From bottom to top the layers are:

- **Data and Storage.** User data and most collaboration space data are stored as RDF data in a triple store. There is one RDF triple store for the user data and one for each collaboration space – each storing its data into files.

  The text content of the space is stored and accessed with a text search and index – again with one index per space. The structure of the data tier means that we can represent each collaboration space as one folder (which can be moved between installations).

- **Application.** The service components in the application level receive domain specific commands from the event bus, change the data as needed and return a result via the event bus. The most important service components in this layer manage (people and document) annotations, the ontology itself and the text index.

  This means, for instance, that annotations to a person are stored or deleted together with the timestamp and information of the tagger (see Section 6.3 for detailed model information). In case the person is not yet known to the system, a new representation of the person has to be generated or in case of a new key word, a new concept may be generated and added to the ontology – depending on the system configuration (see Section 6.2.3 for details). If a webpage is annotated or associated to a person, the text indexing component fetches its content and stores it together with additional information like URL or title in a text index.

  Further service components manage the history and logging, gardening recommendations, and dialogues. There is one instance of each service component for each space, although some service components may be present only for certain space configurations. It is important to note, that the different service components only communicate through the event bus (see also Section 6.2.2 for the detailed event bus communication).

- **Communication.** The event bus in the communication layer manages communication between the different parts of the application logic as well as between the application logic and the presentation layer. In addition to the pure communication aspect it also enforces a total ordering of events (to ensure a consistent state across all components) and it enforces the rules on which parties are allowed to read and write. There is one event bus for each collaboration space.

- **Presentation.** On the presentation layer we provide a web interface and web services that are called by SOBOLEO's AJAX parts and by other applications. The XML and SOAP web services expose the entire non-administrative functionality offered by SOBOLEO - through these interfaces a remote application can function almost like the application components integrated in SOBOLEO. Additional interfaces provided are for the export of the RDF/XML data and Atom feeds.

- **Client.** Finally, e. g., for the editor and annotate components, parts of the SOBOLEO web application is realized as an AJAX application that runs in the client's browser.

6.2.1. Implementation Frameworks

The SOBOLEO framework is implemented by means of the Apache Maven build manager[1] in Java 6 on top of the Apache Tomcat 6.0 application server[2]. The web interface is realized through Java servlets and Java Server Pages (JSPs) together with Google's Web Toolkit framework[3] for the AJAX parts. For the storage of RDF data we use Sesame 2.3.0[4] and the text index is built on Apache Lucene[5].

[1] http://maven.apache.org
[2] http://tomcat.apache.org
[3] http://code.google.com/webtoolkit
[4] http://www.openrdf.org
[5] http://lucene.apache.org

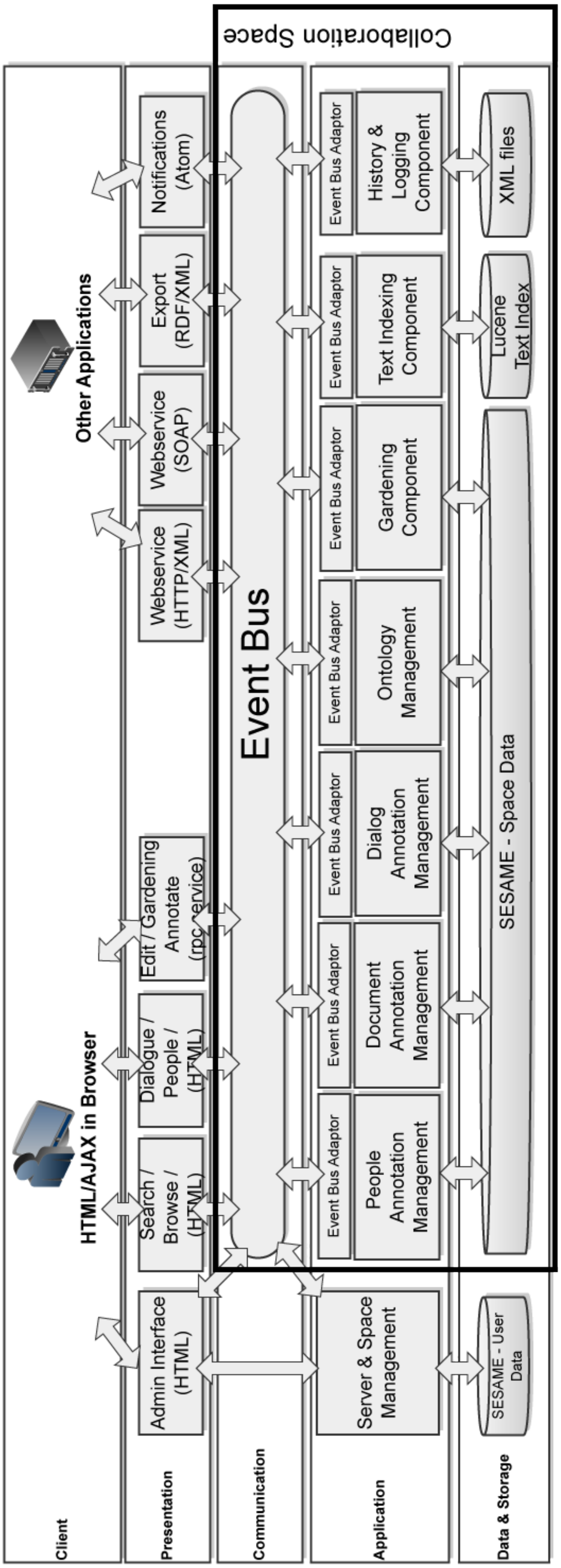


Figure 6.2.: SOBOLEO architecture

6.2.2. Event Bus Communication

Technically each collaboration space is represented by an **Event Bus** component that manages the communication within the space. Different parts of the functionality are realized as services (both local within the server and remote) that are registered to the event bus. These services have access to a shared SESAME triple store and to disk space to store files (e. g., for the text index and logging information). On the server each collaboration space is represented as one folder that can even be moved between SOBOLEO installations.

Communication within the space is organized around the concept of **Events**. There are three different types of events:

- **Command Event:** any change is represented as a command event object; for instance the request to create a concept sends the command event *CreateConceptCmd* containing an initial name to create a new concept for.
- **Query Event:** a query is represented as a query event; for instance a query to search for persons sends the query event *SearchPersons* containing the query string.
- **Notification Event:** any notification is represented as notification event; for instance to notify about an user opening a tagged person's profile sends the notification event *BrowseProfile* containing the URI of the tagged person whose profile is opened.

Each event contains additional standard information that are creation time, an id unique within a collaboration space for increasing and establishing an order within the events, and sender information. The latter three are only set by the event bus and the first automatically when creating an event object.

The event bus routes these events (and results) between the requester (mostly in the presentation layer) and the services that can process it. For example, a delete-concept command event is processed in the following way:

1. The event bus asks all registered command processing services (part of the individual components of the application layers) to extend this command event with implied commands. In this example implied commands include the removal of relations that start or end in the deleted concept.
2. Next a different class of services is asked whether this event is permitted – this tests the user credentials sent with the event as well as the adherence to SKOS integrity constraints.
3. Command processors actually execute the change.
4. All registered event listeners, e. g., history logger, are notified of the changes that have been done.

All events exist as Java objects, as JavaScript objects (allowing to create and receive them within AJAX applications), as XML serializations (allowing to create and receive them by applications written in any language) and as SOAP methods (allowing to create and receive events from any SOAP client). Thanks to these interfaces, almost all conceivable remote clients can do arbitrary changes to the collaboration space. The interfaces also allow to poll for recent events – enabling remote clients to stay up to date with the changes to the information space. We also supply a (Java) client library that takes care of the communication details and manages a local copy of the ontology (kept up-to date by polling the server in customizable intervals). In addition to these custom interfaces, SOBOLEO supports the export of space data as RDF/XML and notification about changes to the information space in the ATOM format.

6.2.3. Configuration

The SOBOLEO framework is configurable in order to realize the different aspects and options of the conceptual design framework previously presented. Via the space configuration service that is part of the server & space management component, each collaboration space can also be configured separately. Configuration options include read and write rights (open to the world or given on a per-user basis), options to enable only parts of the functionality (e. g., disabling people or (web-) document tagging, dialog support, document rating or logging) and configuration options that determine e. g., the aggregation (multiple or one time) or control and visibility of the tag assignments (q. v. Section 5.3). Some configuration options are available for the whole SOBOLEO instantiation on source code basis, e. g., ranking, kind of ontology gardening computations or the provision of SOAP web service interfaces. Table 6.1 gives an overview on the configurable parts, the options together with remarks; i. e. UI or code based.

6.3. Data Models

This section shows the main data models with their attributes and relationships and how they relate with other formats that are SKOS and its extension, FOAF and CommonTag. In accordance to our social social semantic tagging model, the main data models are the users, the ontology with semantic tags (hereafter Concept), the resources being documents and people (hereafter TaggedDocument and TaggedPerson), and the tag assignments (hereafter DocumentTag respectively PersonTag).

URIs are cited in an abbreviated form and should be expanded as follows:

- **soboleo:** or no namespace prefix http://soboleo.com/ns/1.0#
- **skos:** http://www.w3.org/2004/02/skos/core#
- **skosExt:** http://www.w3.org/2004/02/skos/extensions#
- **foaf:** http://xmlns.com/foaf/0.1/#
- **ctag:** http://commontag.org/ns#

Figure 6.3 shows the overall model for tagging a document which additionally comprises dialogs as special document type and ratings and Figure 6.4 the overall model for tagging a person that may be complemented by the overall competence ontology model in Figure 6.5 (see also Section 4.4.2 and Section 6.4.1; see Appendix A.1 on how to read the notation). The details of the individual main concepts are illustrated in Figure 6.2, 6.3, 6.4, 6.6, 6.7, 6.8, 6.9 and 6.10. The subsequent sections will show how it is made use of these models.

Table 6.1.: Overview on the configurable parts, options and type of the SOBOLEO framework

Configurable Part	Options	Remark
People tagging	enable/disable	code based
Web document tagging	enable/disable	code based
Office document tagging	enable/disable	code based
Dialog support	enable/disable	UI based
Document rating	enable/disable	UI based
System out Logging	-enable/disable - by exceptions/ commands/queries/events	UI based
Event History Logging	-enable/disable - by event type	-UI based - Code based
User settings	-set email -set name -set password -set default space -set default language	UI based
Space members	add/remove users	UI based
Space access	Only space users can read & write / only space users can write / anyone can read and write	UI based
Space language	Set default language	UI based
People tag assignments	enable/disable deletion of others' tags to own person	UI based
Auto-completion	Enable/disable	Code based
Tag recommendations	-enable/disable -existing concepts and/or new keywords -external service	Code based
Tag assignments	-multiple/one time assignments to same resource -enable/disable display of existing assignments during annotation -enable/disable implicit annotations	Code based
Taggeable people	-any person/persons with email in a pre-defined domain/only space members	Code based
Ontology gardening	-enable/disable -SKOS Analysis and/or concept relationship analysis and/or usage analysis -external service	Code based
Competence ontology editor extension	-enable/disable	Code based
SOAP web services	enable/disable	Code based
Languages in the ontology	Any language	Code based
Explicit system feedback mechanisms	-enable/disable	Code based
Expertise Analytics	-considered time frame -max. of displayed competencies -threshold for less used competencies	Code based
Document Search	-web documents and/or office documents and/or dialogs -weight of tags of the resource, frequency of tags, time stamp of tag assignments and background knowledge	Code based
Refinements & Relaxations	- Number of documents to base on	Code based
People Search & Profiles	-weight of Tags of the taggee, Frequency of tags, Time stamp of the annotation, Tags of the tagger, Tagging activities of the tagger, Tagging activities of the taggee, Additional activities of the tagger, Additional activities of the taggee	Code based

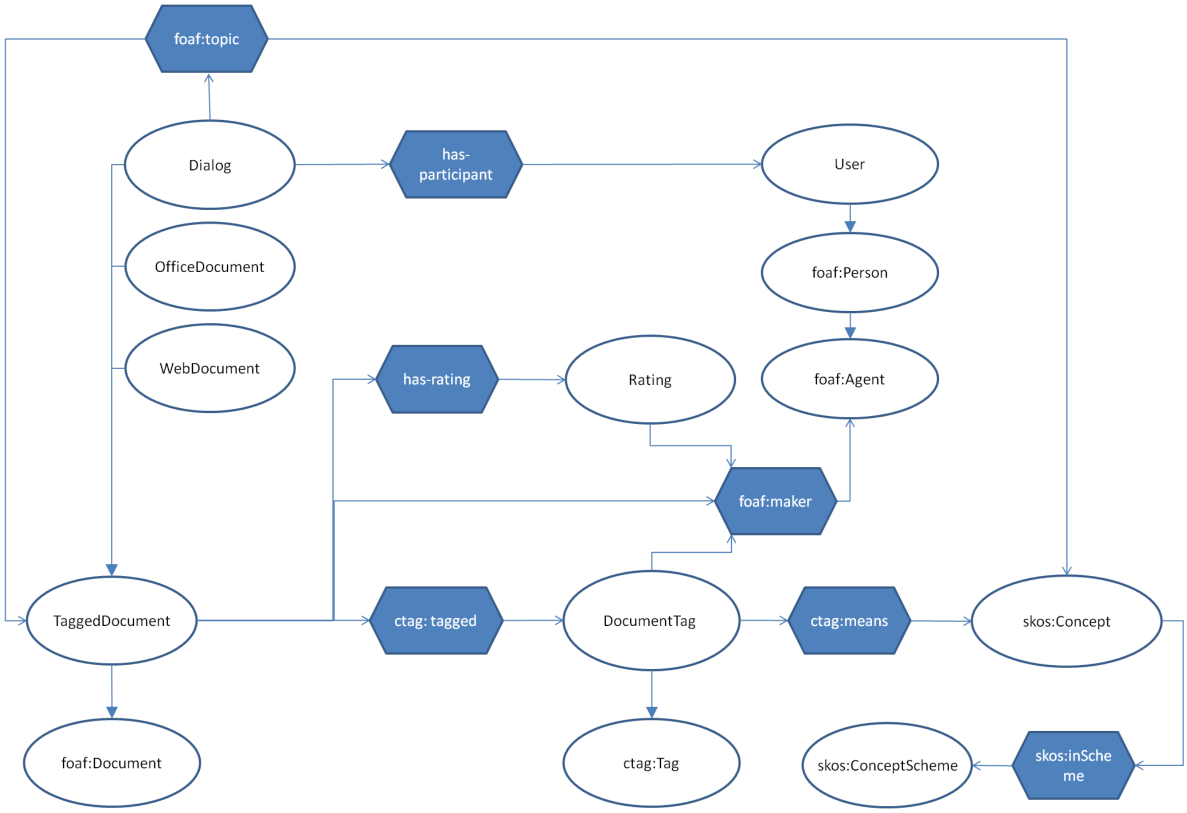


Figure 6.3.: Document tagging model

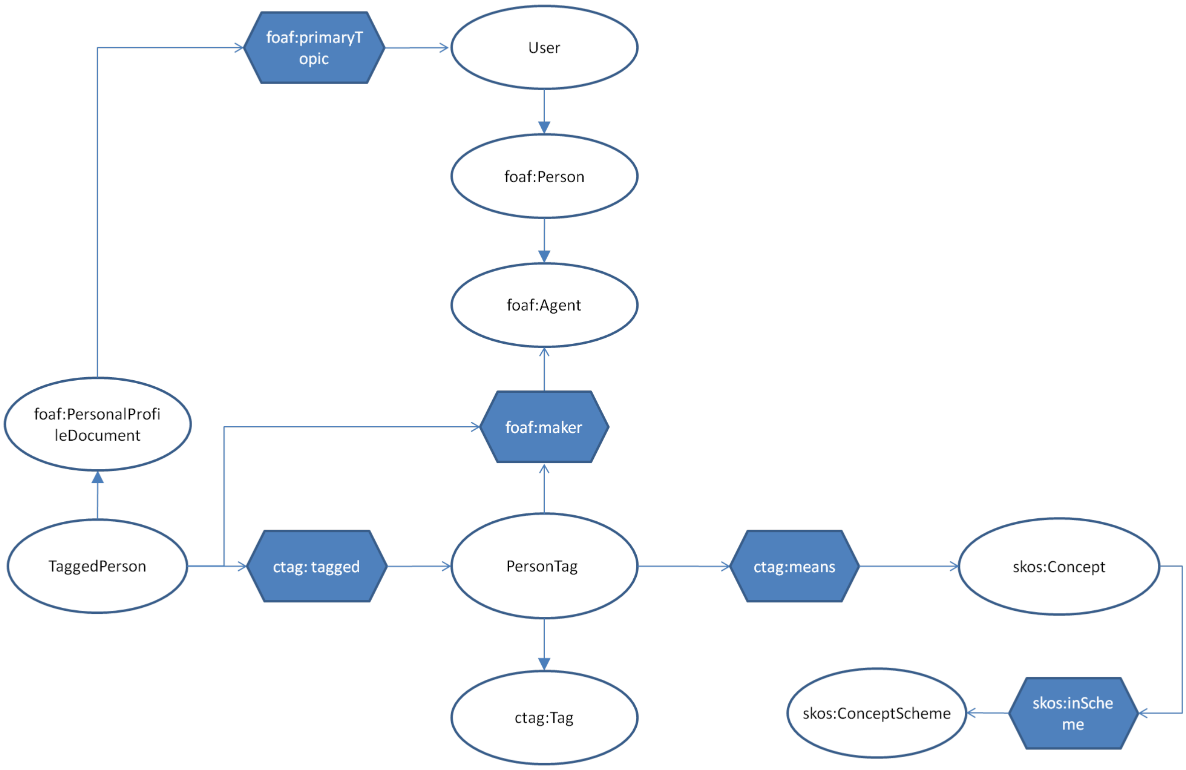


Figure 6.4.: Person tagging model

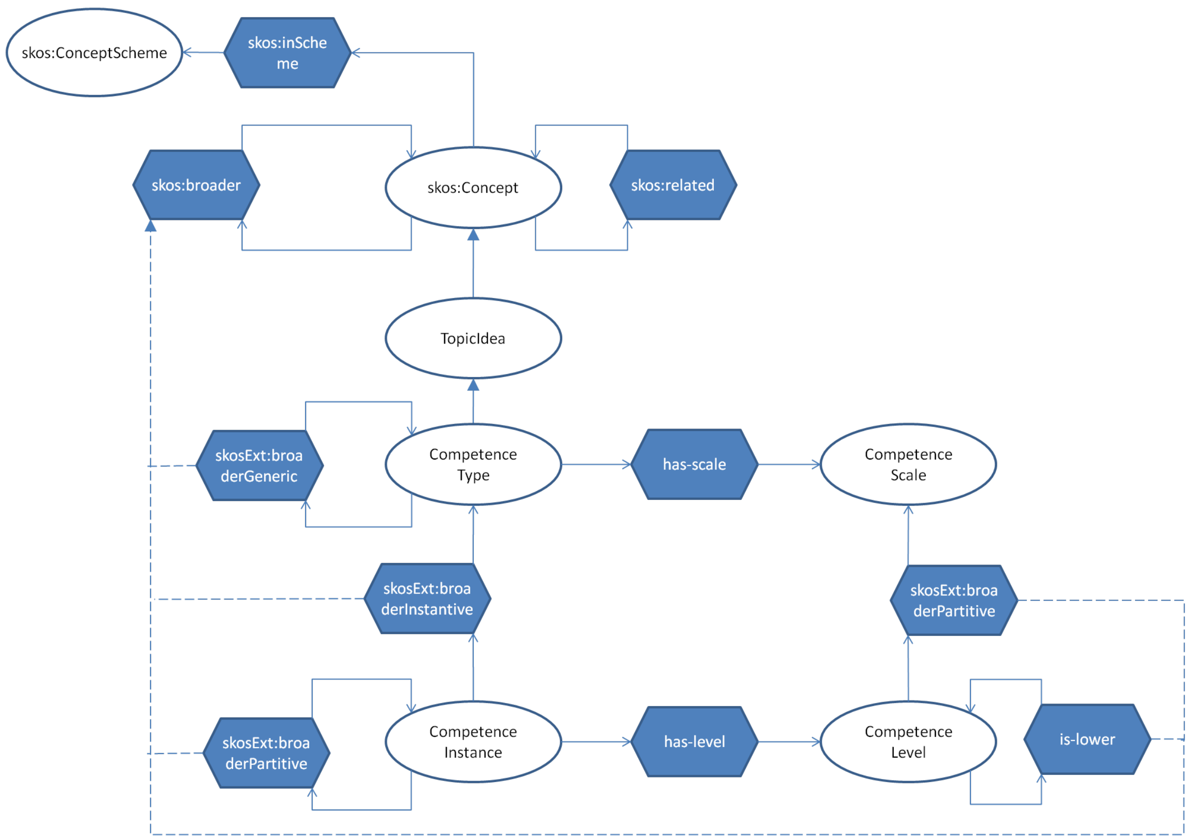


Figure 6.5.: Competence ontology model. (Narrower relations exist vice versa to broader relations but were skipped for readability.)

Table 6.2.: SKOS Concept Model Details

Property	Example	Cardinality
skos:Concept		
skos:prefLabel	„preferredLabel@en“	1..*
skos:altLabel	„alternativeLabel@de“	0..*
skos:hiddenLabel	„hiddenLabel@de“	0..*
skos:broader	<conceptURI>	0..*
skos:narrower	<conceptURI>	0..*
skos:related	<conceptURI>	0..*
skos:note	„Description@en“	0..*
skos:inScheme	<ontologyURI>	0..*

Table 6.3.: Competences Model Details

Property	Example	Cardinality
TopicIdea		
CompetenceType		
skosExt:broaderGeneric	<competenceTypeURI>	0..*
skosExt:narrowerGeneric	<competenceTypeURI>	0..*
skosExt:narrowerInstantive	<competenceInstanceURI>	0..*
soboleo:has-scale	<competenceScaleURI>	1
CompetenceInstance		
skosExt:broaderPartitive	<competenceInstanceURI>	0..*
skosExt:narrowerPartitive	<competenceInstanceURI>	0..*
skosExt:broaderInstantive	<competenceTypeURI>	1
soboleo:has-level	<competenceLevelURI>	1
CompetenceScale		
skosExt:narrowerPartitive	<competenceLevelURI>	1..*
CompetenceLevel		
soboleo:is-lower	<competenceLevelURI>	0..*
soboleo:is-higher	<competenceLevelURI>	0..*
skosExt:broaderPartitive	<competenceScaleURI>	1

Table 6.4.: User Model Details

Property	Example	Cardinality
User		
foaf:name	„firstname lastname“	1
foaf:mbox	<mailto:me@soboleo.com>	1
soboleo:has-password	„pwdhash“	1
soboleo:has-key	„randomString“	1
soboleo:has-space	<spaceURI>	0..*
soboleo:default-space	<spaceURI>	1
soboleo:default-language	„en“	1

Property	Example	Cardinality
TaggedPerson		
foaf:name	„firstname lastname“	1
soboleo:has-email	<mailto:me@soboleo.com>	1
foaf:primaryTopic	<userURI>	0..1
foaf:maker	<agentUri>	1
soboleo:has-created-date	„2009-07-29T20:54:19“	1
soboleo:dateLastModified	„2009-07-29T20:54:19“	1
ctag:tagged	<personTagURI>	0..*

Figure 6.6.: Tagged Person Model Details

Property	Example	Cardinality
PersonTag		
foaf:homepage	<http://mypage.de>	0..1
ctag:taggingDate	„2009-07-29T20:54:19“	1
foaf:maker	<agentURI>	1
ctag:means	<conceptURI>	1

Figure 6.7.: Person Annotation Model Details

Property	Example	Cardinality
TaggedDocument		
foaf:has-title	“document title“	1
foaf:maker	<agentUri>	1
soboleo:has-created-date	“2009-07-29T20:54:19“	1
soboleo:dateLastModified	“2009-07-29T20:54:19“	1
ctag:tagged	<documentTagURI>	0..*
soboleo:has-content	“text content”	0..1
soboleo:has-rating	<ratingURI>	0..*
WebDocument		
foaf:homepage	<http://anypage.de>	1
OfficeDocument		
soboleo:has-type	[“pdf”, “doc(x)”, “xls(x)”, “ppt(x)”]	1
Dialog		
soboleo:has-type	[“CDR-DG”, “OM-DG”]	1
soboleo:has-participant	<userURI>	1..*
foaf:topic	[<conceptURI>,<documentURI>]	1..*

Figure 6.8.: Tagged Document Model Details

Property	Example	Cardinality
DocumentTag		
ctag:taggingDate	„2009-07-29T20:54:19“	1
foaf:maker	<agentURI>	1
ctag:means	<conceptURI>	1

Figure 6.9.: Document Annotation Model Details

Property	Example	Cardinality
Rating		
soboleo:has-created-date	“2009-07-29T20:54:19“	1
foaf:maker	<agentURI>	1
soboleo:has-score	[0-5]	1

Figure 6.10.: Rating Model Details

Figure 6.11.: SOBOLEO Ontology Editor

6.4. Ontology Editor and Management

The central building block of the SOBOLEO framework for both the Social Semantic Bookmarking and Semantic People Tagging instantiations is the ontology editor and management part. The ontology visualization and editing UI service together with the ontology management service component allows users to structure the concepts (semantic tags) with hierarchical relations (broader and narrower) and to indicate that concepts are "related" (see Figure 6.11). Concepts can have a (multi-word) preferred label and a description in multiple languages; they can have any number of alternative and hidden labels (according to the SKOS format which allows us to seamlessly work with half-formalized domains and it is relatively easy to understand for non-modeling experts).

The ontology management service component controls the ontology elements. If a command event requests a change for an ontology element, it checks the adherence to SKOS integrity constraints; e. g., for potential cycles produced by adding new relations. Eventually, it executes the request including the responsibility for the execution on data level.

The collaborative editor on the UI side can be used by several users at the same time. Changes by any user are immediately visible and effective to all users[6] and the ontology's usage (e. g., for search). There are no privileges for specific users (groups) on the creation, modification and deletion or visibility or level of details.

[6] where any user resp. all users either refers to all members within the space or any user of the web depending on the space configuration

The editor visualizes the ontology in a tree display. It shows the concepts' preferred labels in the user's default language and their narrower and broader relations (forming the tree structure). Users can drag'n'drop concepts on other concepts to create new broader relations or while pressing shift to remove them. The operations cut, copy and paste (and their shortcuts) are also supported.

Creating or removing a broader relation between concept c_i and concept c_j automatically creates/removes the inverse narrower relation between c_j and c_i. Similarly, the symmetry of c_i skos:related c_j entails c_j skos:related c_i.

Buttons above the ontology tree allow for the creation and deletion of concepts. Via the search component above, the users can search for a specific concept in the tree (supported by auto-completion), which facilitates the navigation. Underneath "prototypical concept", any new concept, added during the annotation process, is automatically collected. This supports the uptake of new topic ideas (see also Section 5.1.3.2).

In two tabs in the center, the editor displays details of the currently selected concept. On the "Labels and Descriptions" tab, users can add, change and delete the concept's labels and description in multiple languages. On the "Relations" tab, users can see and change the current concept's relations. Any changes made are automatically visible in the tree. Clicking on one of the broader, narrower or related concepts, the user can jump to this concept in the tree.

The "log" message pane to the right displays automatically generated messages about recent changes that is currently a short history of the last 30 events and any upcoming change made when the user is using the editor. The "chat" tab opens a simple chat pane where users can send chat messages to other users currently using the editor.

If dialog support is enabled (see Section 6.5), the editor also provides the possibility to see performed dialogs related to a selected concept or start a new dialog, where the user can define a title to give the discussion an initial direction.

6.4.1. Variable Level of Formality in Competence Ontology Maturing

With SKOS we can implement first variable levels of formality; i. e. starting with informal tag ideas that are developed towards concepts with different labels and descriptions and further formalized with hierarchical and ad-hoc relations. However, for some competence management purposes (see Section 4.4.1 and Section 4.4.2) we need to support additional levels. Therefore, we provide an optional extended ontology editor & management version using the SKOS extension (Miles and Brickley, 2004).

As presented in Section 4.4.2, we have topic tags and competence types, both sub-classes of skos:Concept. Topic tags are concepts hitherto not existing and automatically created during the annotation process (see also Section 6.7.1) and/or positioned under the container concept "prototypical concepts" within the tree. When topic tags are removed from the "prototypical concepts" container and placed within the ontology, they become competence types. In order to represent hierarchical relations between competence types, we use *skos:broaderGeneric* and *skos:narrowerGeneric* respectively. These relations are automatically created when the users models the tree structure in the editor (see Figure 6.12).

Competencies (with levels) are instances of competency types. They are needed to have different degrees of fulfillment for individual competencies. A competence instance relation is represented by *skos:broaderInstantive* and the inverse *skos:narrowerInstantive* between a competency (with level) and a competence type:

<*competenceInstanceURI, skos:broaderInstantive, competenceTypeURI*> and

<*competenceTypeURI, skos:narrowerInstantive, competenceInstanceURI*>

where *competenceInstanceURI* is the URI to identify a specific competence instance and similarly competenceTypeURI the URI to identify a specific competence type.

Additionally, a competence instance has a certain level: <*competenceInstanceURI, hasLevel, CompetenceLevel*>. We provide by default the generic competence level 1 to 5; being 5 the highest one. The levels are connected with is-lower/is-higher relations that are subproperties of skos:narrower/skos:broader. We use generic levels to allow the extension with different competence scales. By clicking on the button "Activate Competence Levels" in the competence details view in the editor's center part, the instances are automatically created.

So, users can annotate either with a competence instance and a specific level for differentiated statements or with competence types; e. g., if the specific degrees of a type are not yet known or if the user cannot make a more precise statement about the person to be tagged.

Between different competence instances, i. e. competencies with level, compositions relations can be defined. These are modeled using *skos:broaderPartitive* and *skos:narrowerPartitive*. In the user interface, the creation of competence compositions is supported by the additional composition box in the upper right corner of the editor. There, the users can choose a competence instance of the competence type selected in the tree display and add or remove any other competence instance for composition.

6.4.1.1. Implicit Annotations

Through compositions, we may introduce implicit or system-generated annotations. For instance, let us assume there might be defined the competence instance "GWT Programming, Level 2" that is composed of "Java Programming, Level 3" and "AJAX Programming, Level 1" and "Web Development, Level 2". And "AJAX Programming, Level 1" again is composed of "XHTML Programming, Level 2", "XML Programming, Level 2" and "JavaScript Programming, Level2". When now a person is tagged with "AJAX Programming, Level 1", we may automatically generate annotations for "'XHTML Programming, Level 2", "XML Programming, Level 2" and "JavaScript Programming, Level2". In detail we have to consider the following cases for a competence instance (c, l) where c is the competence type and l the level, that is composed of $\{(c_1, l_1), (c_2, l_2), \ldots, (c_m, l_m)\}$:

- **Case 1:** If a person is tagged with (c, l), then the ontology management component adds the implicit annotations $\{(c_1, l_1), (c_2, l_2), \ldots, (c_m, l_m)\}$ and recursively additional implicit annotations if (c_i, l_i) itself is a composition.

  Example 1: Going back to our example from the beginning of this section that means if we tag a person with "GWT Programming, Level 2" then its parts "Java Programming, Level 3" and "AJAX Programming, Level 1" and "Web Development, Level 2" are added in a first step. Afterwards, iterating through the newly added parts, the parts of "AJAX Programming, Level 1" are added as well; i. e. "XHTML

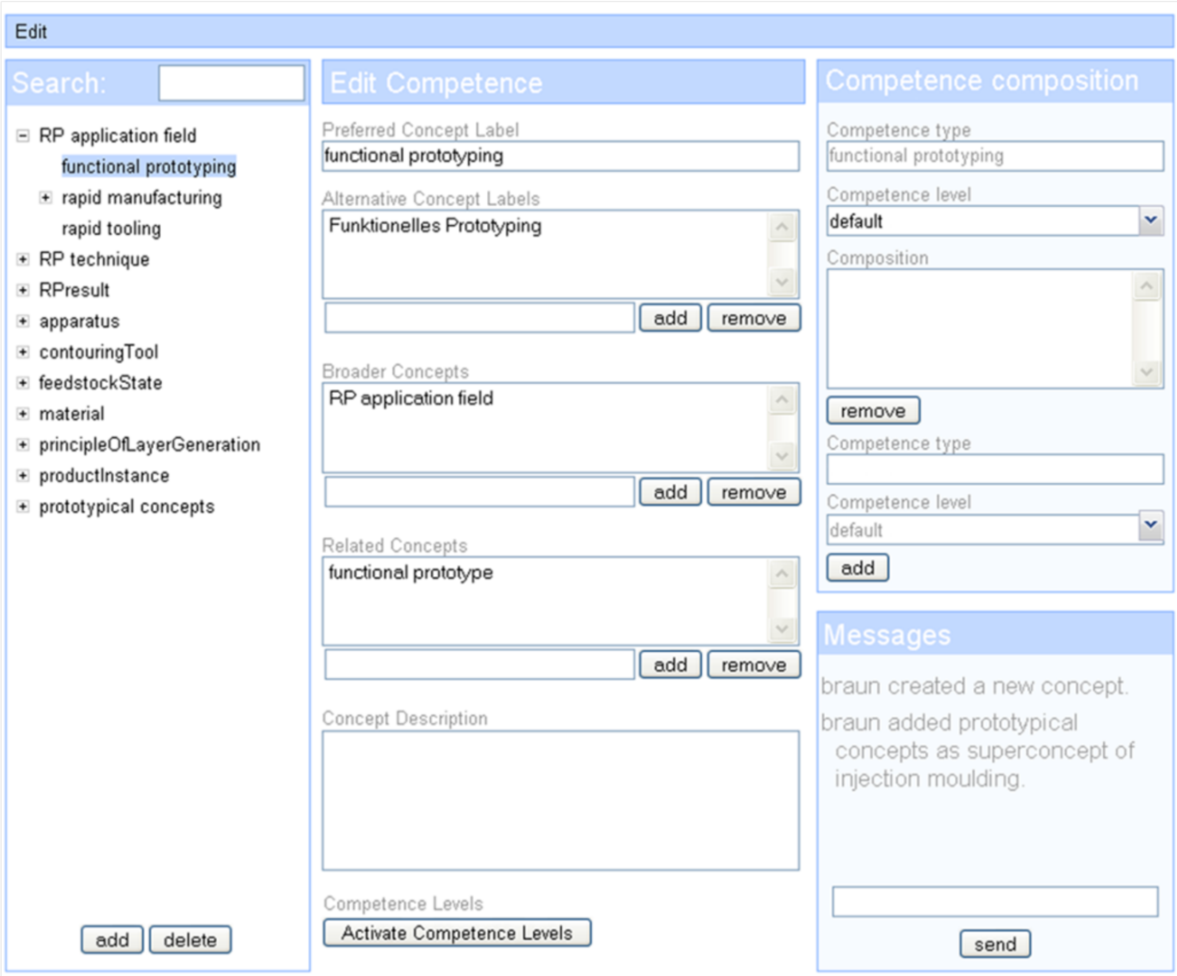


Figure 6.12.: SOBOLEO competence ontology editor

Programming, Level 2", "XML Programming, Level 2" and "JavaScript Programming, Level2".

- **Case 2:** If a person is tagged with any $\{(c_1, l_1), (c_2, l_2), \ldots, (c_m, l_m)\}$, then the ontology management component adds the implicit annotation (c, l) and recursively additional implicit annotations if (c_i, l_i) itself is a composition[7].

  Example 2: Let us assume, a person is tagged with "Java Programming, Level 3" and "Web Development, Level 2". The person is also tagged with "XHTML Programming, Level 2", "XML Programming, Level 2". If the person now is newly associated with "JavaScript Programming, Level2" then the implicit annotation of "AJAX Programming, Level 1" is added automatically in a first step and followed by the additional annotation of "GWT Programming, Level 2".

- **Case 3:** If the annotation with (c, l) is removed from a person, then the ontology management component also has to remove the implicit annotations $\{(c_1, l_1), (c_2, l_2), \ldots, (c_m, l_m)\}$ and any additionally dependent implicit annotations.

  Example 3: Let us assume, a person is tagged with "GWT Programming, Level 2" and thus has the implicit annotations "Java Programming, Level 3" and "AJAX Programming, Level 1" and "Web Development, Level 2" and "XHTML Programming, Level 2", "XML Programming, Level 2" and "JavaScript Programming, Level2" (cf. Example 1). If the "GWT Programming, Level 2" is removed then the implicit annotations "Java Programming, Level 3" and "AJAX Programming, Level 1" and "Web Development, Level 2" are removed as well in a first step, followed by the removal of the (implicit) composition parts of "AJAX Programming, Level 1".

- **Case 4:** If the composition (c, l) is extended by (c_n, l_n) to $\{(c_1, l_1), (c_2, l_2), \ldots, (c_m, l_m), (c_n, l_n)\}$, then the ontology management component adds for case 1 the implicit annotation (c_n, l_n) and dependent implicit annotations and respectively removes for case 2 the implicit annotation (c, l) and dependent implicit annotations if the person is not explicitly tagged with (c_n, l_n).

  Example 4: For instance, programming with new GWT version now requires to know about the development for the Google AppEngine; i. e. the competence "GWT Programming, Level 2" is extended by "Development for AppEngine, Level 1". Then, for any case 1, i. e. when people are tagged with "GWT Programming, Level 2", the people's implicit annotations are updated and "Development for AppEngine, Level 1" is added as well. For any case 2, where people have the implicit annotation "GWT Programming, Level 2" based on the composition parts "Java Programming, Level 3" and "AJAX Programming, Level 1" and "Web Development, Level 2" and do not additionally have the competency "Development for AppEngine, Level 1", the implicit annotation "GWT Programming, Level 2" is removed.

- **Case 5:** If the composition (c, l) is reduced by (c_m, l_m) to $\{(c_1, l_1), (c_2, l_2), \ldots, (c_l, l_l)\}$, then the ontology management component removes for case 1 the implicit annotation (c_m, l_m) and dependent implicit annotations or respectively adds the implicit annotation (c, l) and recursively additional implicit annotations if a person is tagged with any $\{(c_1, l_1), (c_2, l_2), \ldots, (c_l, l_l)\}$.

  Example 5: For instance, "GWT Programming, Level 2" is defined as the composition of "Development for AppEngine, Level 1" and "Java Programming, Level 3" and "AJAX Programming, Level 1" and "Web Development, Level 2" and the

[7]Based on the assumption that the competencies are fully decomposed. For a detailed discussion of this issue, we would like to refer the reader to Schmidt (2009)

competency “Development for AppEngine, Level 1” is removed from the composition. Then, for any person, who has the implicit annotations “Development for AppEngine, Level 1” and “Java Programming, Level 3” and “AJAX Programming, Level 1” and “Web Development, Level 2” based on “GWT Programming, Level 2”, the implicit annotation “Development for AppEngine, Level 1” is removed. Vice versa, for any person tagged with “Java Programming, Level 3” and “AJAX Programming, Level 1” and “Web Development, Level 2” the implicit annotation “GWT Programming, Level 2” is added.

If compositions are modified, the ontology management component has to take into account the principles of subsumption when creating or removing implicit annotations. Similarly, the search engine can make use of the subsumptions. For instance, competence instance (c, l_1) subsumes another competence instance (c, l_2) if $l_1 > l_2$. Thus, searching for people tagged with at least (c, l_2), we also retrieve people tagged with the competence instance (c, l_1) $(l_1 > l_2)$; such as searching for people who are advanced in Java programming will include experts in Java programming in the result set.

6.4.2. Ontology Gardening

The ontology gardening service component supports the user in cleansing and improving the ontology. It analyses the ontology and provides a list of suggestions where the ontology might need improvement. For instance, it refers the user to concepts with missing description or with identical labels which might indicate to merge these or finding the right place for a new concept.

The recommendations for gardening the ontology are available in the ontology editor via the menu entry *Improve It!*. The users can ask for recommendations related to the whole ontology or to a specific concept; i. e. the one currently selected in the tree. The recommendations are grouped by type of heuristics. By clicking on the concept’s name, the user can jump to the concept in the ontology tree and thus immediately realize the recommended modifications (see Figure 6.13).

For computing the recommendations the service uses several different heuristics and works with the help of diverse other service components; e. g., ontology management, history& logging, annotation management or may make use of external services – depending on the configuration. With its generic interface for recommendations the service is meant for extensions; e. g., extensions based on the ontology evolution heuristics by Stojanovic (2004).

Currently the three different types of heuristics, presented in the following, are implemented: the *SKOS Analysis* provides suggestions for changes induced from the ontology (structure) itself, the *Concept Relationship Analysis* from the ontology structure and the application of concepts, and the *Activity Analysis* from patterns of ontology usage. Each of them being a representative of the three different variants from which to induce ontology change suggestions – *structure-driven*, *data-driven*, and *usage-driven change discovery* – that were identified by Stojanovic (2004).

The need for ontology gardening and maintenance support is shown in Part III. The heuristics presented here have been proven outside the case studies. Due to the case studies differing focus, we could only collect selected statements on the quality of the recommendations in using the framework. The Ontology Relationship Analysis and Activity Analysis have been evaluated separately from the case studies on large real life

Figure 6.13.: SOBOLEO ontology editor showing recommendations to garden the ontology.

datasets; i. e. Wikipedia data (Ramezani et al., 2010) and Delicious data (Hagelauer, 2011).

6.4.2.1. SKOS Analysis

The service analyzes a SKOS ontology with concepts $C = \{c_1, \ldots, c_m\}$ for potential redundant or missing information. It checks three anomalies that indicate possible ways to mature the ontology. These are (1) multiple concepts that have the same label (indicating concepts that could be merged), (2) concepts with a missing description and (3) concepts that lack a preferred label in the users' language. It can be configured based on the anomalies it detects, per default it detects the three mentioned before. In detail:

1. **Identical Labels:** to find concepts with identical labels, the service analyzes the ontology for string identicalness of the labels in languages $L = \{l_1, \ldots, l_n\}$ of two concepts; i. e. if for any two concepts $c_i, c_j \in C$ and $c_i \neq c_j$, and the language $l_k \in L$[8]:
   - $preferredLabel(c_i, l_k) = alternativeLabels(c_j, l_k)$
   - $preferredLabel(c_i, l_k) = hiddenLabels(c_j, l_k)$
   - $alternativeLabels(c_i, l_k) = alternativeLabels(c_j, l_k)$
   - $alternativeLabels(c_i, l_k) = hiddenLabels(c_j, l_k)$

[8] Two concepts with the same preferred label for the same language are not permitted according to the SKOS reference (see also Section 3.1.3.3) and therefore not checked.

   - $hiddenLabels(c_i, l_k) = hiddenLabels(c_j, l_k)$

2. **Missing Description:** to find concepts with missing description, we analyze the ontology for any $c_i \in C$ for which $description(c_i, l) = \emptyset$ for all languages $l \in L$.

3. **Lacking Preferred Label:** to find concepts that lack a preferred label in the user's language $l_k \in L$, we analyze the ontology for any $c_i \in C$ for which $preferredLabel(c_i, l_k) = \emptyset$.

6.4.2.2. Concept Relationship Analysis

The service analyzes an existing concept hierarchy and the application of concepts; i. e. used for tagging, to compute a list of potential broader or related concepts for certain given concepts. Such gardening recommendations assist the users in improving the ontology structure, e. g., finding the right place for a new concept. Therefore, it computes for given concepts a list of potential broader or related concepts (depending on the configuration) that is ranked by a recommendation confidence value. This can be done either based on string similarity, k-nearest neighbor or a hybrid combination of both. Therefore, a simplified and for related relations extended version of the recommendation algorithms by Ramezani (2011, pp.72) was integrated.

The result is a list of top n broader or related concepts with the highest confidence value or higher than a configurable threshold value. No broader concepts or confidence values lower than the threshold might be interpreted by the recommendation service to add the given concept as independent root concept at the top of the hierarchy.

6.4.2.3. Activity Analysis

The service provides recommendations to improve the ontology based on the analysis of the activity of ontology elements. Therefore, we may differentiate three different types of semantic tags, that are "inactive", "dead" or "living" tags (see also Hagelauer, 2011). Inactive semantic tags are those that are used, e. g., for annotation, very rarely, e. g., only once, or not at all because they are so specific or inadequate for later reuse or retrieval that even the creator has not reapplied them. Such semantic tags might be candidates to be deleted. Dead semantic tags are those that were, after a phase of increased activity, only rarely or no more used; e. g., because the concept/topic became unpopular[9] or were replaced by other synonymous ones. Particularly the latter ones would be candidates for merging. Living semantic tags are those which are neither inactive nor dead. While living tags may show significant levels of activity through their life, the way they are used may change. So, it might be useful to focus gardening and higher formalization on those elements that are important for the community; e. g., those often used for annotations or searched for. Similarly, Stojanovic (2004) use the users' querying and browsing activities in an information portal scenario for ontology change suggestions.

Therefore, the service identifies ontology elements that have not been used for annotation for a specific time period of the last N days or not at all. If enabled, it also takes into account the discussion status; i. e. those ontology elements that have not been discussed at all. Such elements are recommended as candidates to be deleted.

[9] Often to be observed for event related tags.

Similarly, the service identifies ontology elements that are often applied and searched for within a specific time frame of N days backwards. Such ontology elements are recommended to focus on; especially if they have not been edited over a longer time and missing details.

6.5. Dialog and Discussion

The system supports discussions for ontology development on the one hand with a simple chat functionality integrated in the ontology editor. On the other hand structured dialogs may be enabled for each collaboration space for extended discussion support. Therefore, we integrated an adapted version of InterLoc5[10] (cf. Ravenscroft et al., 2010b).

InterLoc5 is a web-based tool that aims to support purposeful Digital Dialog Games for learning and critical and creative thinking in groups. It is used to address relatively generic learning problems and opportunities related to the need for collaborative reasoning and discourse such as among knowledge workers. Essentially, these dialog games realize engaging and structured rule-based interactions that are performed using pre-defined dialog features (such as dialog Moves, Locution Openers and a model of turn-taking). These dialog features are specifically designed to foster thinking and learning (Ravenscroft, 2007; Ravenscroft and McAllister, 2008).

All contributions or replies are made using these Move categories (Inform, Question, Challenge, etc.) and scaffolded through using specific Locution Openers ("I think. . . ", "I disagree because. . . ", "Let me elaborate. . . " etc.) that have to be used to perform the dialog. Similarly, rules about the legitimate and logical responding openers, based on the specific openers that are replied to, are offered selectively, but these can be overridden where necessary. The model of turn-taking is incorporated to ensure that the dialogs support: 'listening' to others contributions; fairly balanced patterns of contribution; and, generally, the sort of coherent sequencing that results in reasoned discourses.

By integrating our framework with InterLoc, we aim to supplement the chat component (see Section 6.4) with a specially designed dialog for ontology maturing, where we stimulate users to have a dialog with and about the developing ontologies to specify, clarify and refine the semantic features or degrees of certainty about their classification. This allows both individual users and the community to have dialogs with and about the ontology, to construct more understandable and meaningful representations. Allowing the community to engage in collaborative dialogs about the ontologies in this sort of way will catalyze ontology maturing and social learning in relation to the domain and the users who are continuously developing their understanding of it. In other words, having a structured dialog about the development and use of the ontology will actually help to 'bring it to life' and make it more useful.

Based on our evaluations (see Section 7.6), there is additional need to discuss not only the ontology but also resources, how to rate and annotate these or topics in general. And above all, that records of dialogs are linked to subjects of discussion and later accessible (in keeping with other collaborative ontology engineering approaches; cf. Section 3.2.3). Therefore, structured dialogs can additionally be used to discuss, besides (a) the ontology and its development process, (b) the resources (e. g., reflecting and debating the correctness and quality), (c) the resource classification (according to the ontology), and (d) the topic in general (e. g., initiated by a member's question). At the end, performed

[10] http://www.interloc.org.uk

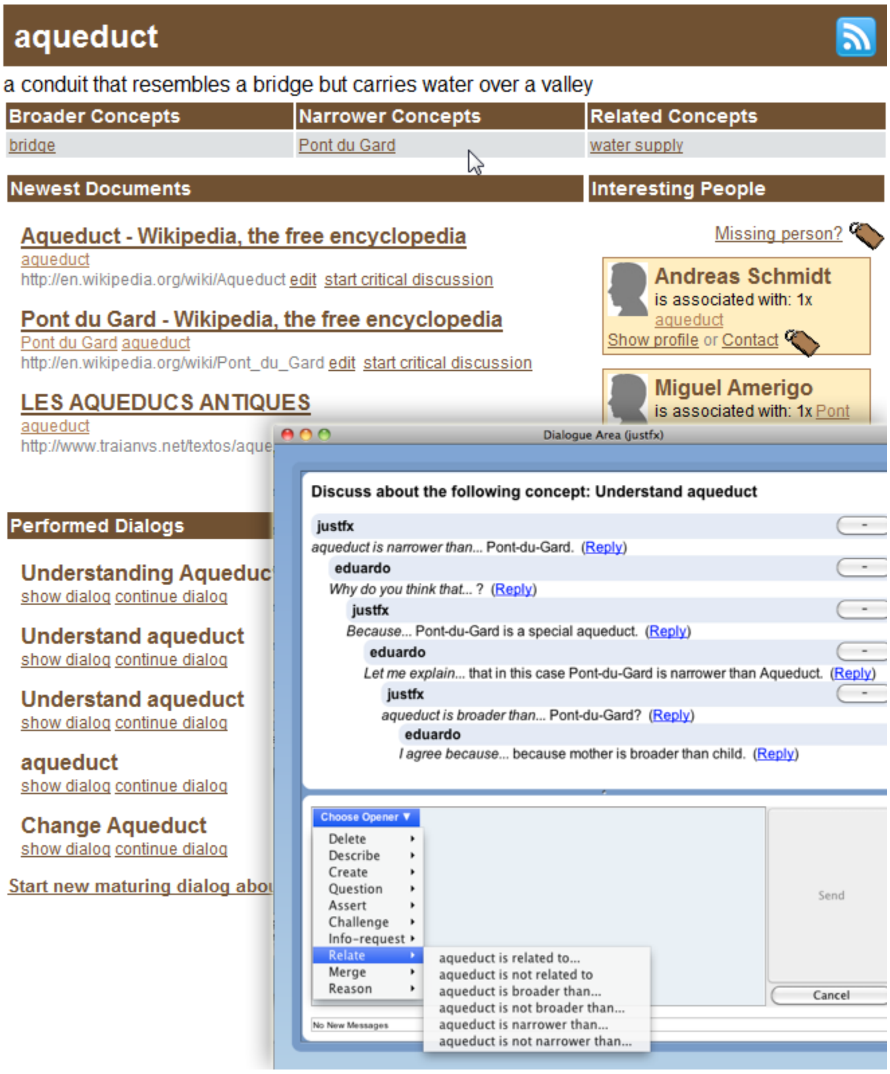


Figure 6.14.: An ontology maturing dialog about the concept 'aqueduct'.

dialogs become a resource themselves, which are linked to the related resources and are accessible and searchable.

So, in order to link dialog and semantics to promote ontology maturing we have developed an original ontology maturing dialog (OM-D) to stimulate and support discussions about ontology development (e. g., changes), which is shown in Figure 6.14 (showing resources and a maturing dialog about 'Aqueducts'). This has been achieved through specifying the pre-defined Moves and Openers of the dialog (see Table 6.5) in terms of the key semantic interaction actions within SOBOLEO (e. g., "Add narrower than", "Add broader than", "Add related to") and supplementing these with attested and more argumentative, or critical, ones from existing dialog games (e. g., "I think. . . ", "Is it the case that. . . ", "What if. . . "). This is complemented by an adapted critical discussion and reasoning dialog (CDR-D) that is aimed at reasoned discussion about a knowledge domain or particular perspectives and resources related to it.

We provide the opportunity to start a maturing (OM-D) or critical discussion (CDR-G) dialog from various points within SOBOLEO, e. g., during the bookmarking and annotating process, when editing the ontology, or browsing through the ontology and bookmarks. So, when a user decides to initiate a dialog, a command event either to start an OM-D or CDR-D is sent with information about subject of discussion, title and initiator that

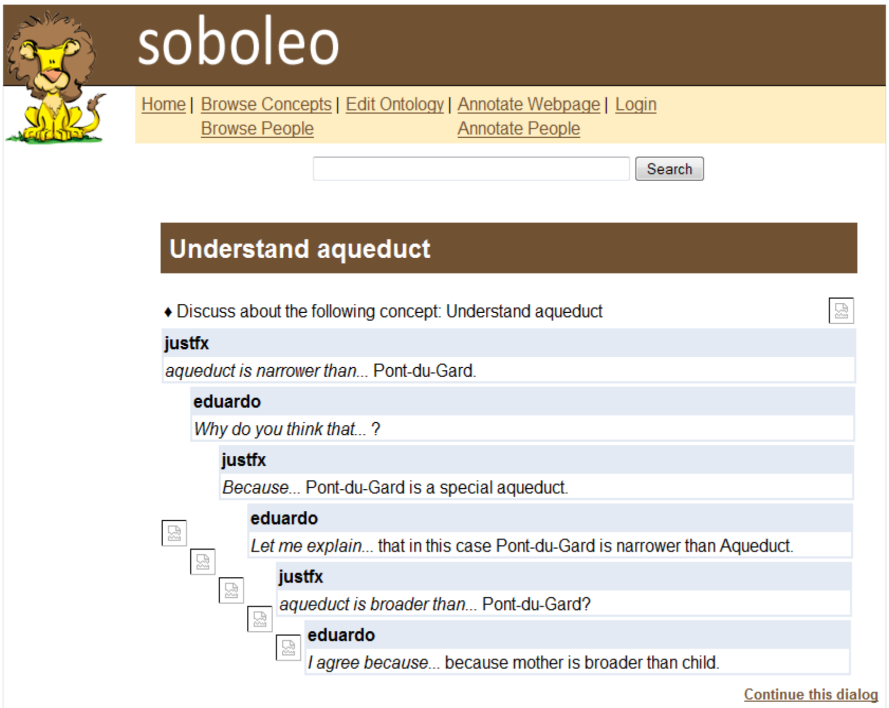


Figure 6.15.: Detailed transcript of a maturing dialog about the concept 'aqueduct'.

the event bus routes to a WidgetServer, a messaging environment developed for mashing up and integrating independent software clients (Nelkner, 2009) which we use for integration. This creates and opens a new dialog in InterLoc; other users are invited to join. When the users finish the discussion, a message with the content of the performed discussion and the participants is sent back and the dialog annotation management service then stores the transcript together with information about participants, creation and modification date and links it to the related resource or concept.

In this way, performed dialogs are automatically linked to their related objects and available for later retrieval. For instance, when navigating to a concept in the browse area, the users can see a list of all performed dialogs about this concept (see also Section 6.6.3).

The user can have a look at the dialog's transcript by clicking on one of the listed dialogs. Thus, users may understand the rationale of changes to those resources. It visualizes a performed dialog similarly to the InterLoc dialog component (see 6.15).

Additionally, it is also possible to continue an existing dialog; e. g., if they do not agree. When a user (re-)starts a dialog, the user who originally created the discussion and all formerly participating users are asked for participation in this dialog. Any other users get aware by the update of the dialog list in InterLoc and are free to participate. Again, when closing a session, the content is stored by the dialog annotation management service.

Table 6.5.: Dialog Moves & Openers

Moves	Openers	Moves	Openers
Create	"Create..."	Assert	"I think..."
Delete	"Delete..."		"I read that..."
Describe	"Rename..."		"Let me explain..."
	"Modify..."		"I agree because..."
	"...is an alternative to..."	Question	"How can..."
	"...is not an alternative to..."		"Is it the case that..."
	"...is a variant spelling to..."		"Why do you think that..."
Relate	"Add...related to..."	Reason	"Are you saying that..."
	"Remove...related to..."		"That is valid if..."
	"Add...broader than..."		"Because..."
	"Remove...broader than..."	Challenge	"I disagree because..."
	"Add ...narrower than..."		"What if..."
	"Remove...narrower than..."		"Please give a reason..."
Merge	"Merge...with..."		"An argument against that is..."
		Info-Request	"I'm not sure..."
			"Can you tell me..."

6.6. Social Semantic Bookmarking

6.6.1. Bookmarking and Annotating Documents

Users can add semantically annotated documents – either web or local office documents – to the shared collaboration space through a document annotation interface that is available both as a web page and a browser bookmarklet – particularly useful for annotating web resources. When using the bookmarklet for web documents, the web document annotation UI service fetches the URL and title of the page that is open in the browser – in this way we my use any web resource as source of material – and opens a pop up window with this information already filled in. Similarly, when uploading a local office document – which is possible and checked by a file type filter for PDFs, MS Word, Excel and Powerpoint both the 2004 and 2007 formats – the document title is fetched.

For annotating the resource, the user can use any concept from the collaboratively developed ontology or arbitrary (multi-word) tags. New tags are automatically added to the ontology as "prototypical concepts" (supported by the ontology management service); so users can consolidate and move them within the ontology later. In this way we allow for the seamless gathering of new concept ideas when they are occurring.

During the annotation the user is supported with auto-completion of entities in the ontology to create awareness for and encourage users to reuse these. Auto-completion not only includes the concepts' preferred labels but also alternative labels. For instance, there is the concept C_i with the preferred labels "building"@en and "Gebäude"@de and alternative label "structure"@en. Then, one user may use "building", the second user prefers "structure" and the third one, who is German, uses "Gebäude" for annotation but all of them referring to the same concept.

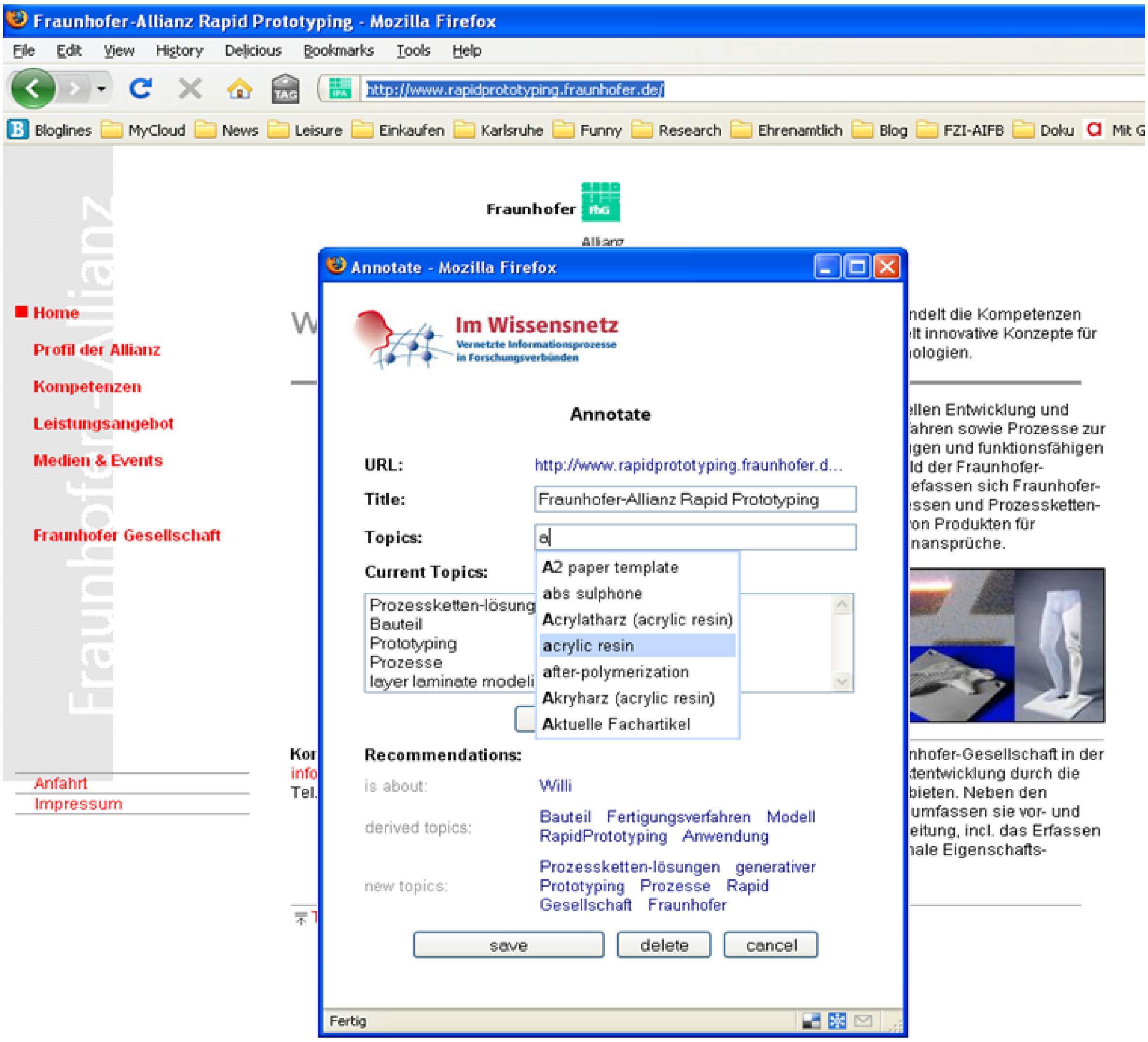


Figure 6.16.: SOBOLEO annotation tool with different types of tag recommendations.

Additionally, the system provides recommendations how to annotate the resource using external services[11]; e. g., services like TagThe.Net[12]. For ontology maturing purposes, we may differentiate two major classes of recommendation services: services that recommend (1) existing concepts from the ontology and (2) new topics/keywords not yet contained in the ontology but potential concept candidates. Whilst the first ones again support reuse and consolidation of existing ontology entities, the latter ones help to incorporate and grow the vocabulary by new topics. For instance, Figure 6.16 shows three different types of recommendations using the information extraction component of empolis:Information Access Suite by Attensity formerliy empolis[13]. Recommendations are (a) *is about* – ontology entities occurring in the document content; (b) *derived topics* – concepts not explicitly contained in the document but inferred based on rules on existing ontology entities, e. g., making use of broader/narrower relations; (c) *new topics* – that are based on statistical analysis of the document content.

Additionally, when enabled, the bookmarking and annotation tool allows the instant

[11] Tag Recommender is an active area of research that is not in this work's focus. For further reading we would like to refer to Ramezani (2011) or Wetzker (2010)

[12] http://www.tagthe.net

[13] http://www.attensity.com

initiation of discussions about the resource and/or its classification – i. e. its annotations – supported by the structured dialog service (see Section 6.5). The users may also express their opinion on the resource through the 5-star-rating functionality. This functionality can also be enabled or disabled within the individual collaboration spaces.

When saving the bookmark and annotations of a web resource, the web document annotation management service automatically sends out a crawler to fetch and index the contents of the page. The crawler is able to parse HTML, PDF and both the 2004 and 2007 MS Office formats. Similarly, we index the contents of local office documents during their upload. Whenever the annotate pop up is opened for a particular URL, the web document annotation management service checks whether there are already annotations. If so, the annotations are displayed.

Depending on the space configuration for tag assignment aggregation, annotations are "personal annotations" or "community annotations". For personal annotations each user adds and maintains his annotations separately. That is equivalent to the multiple tag assignments approach in most common social bookmarking systems (see Section 2.6.1). With community annotations all users share the same annotations; e. g., user u_i can change the annotation made by user u_j. That is, at a first glance, similar to one time tag assignment approach, but in detail we applied a mixed mode. That means, in the backend and data level, we handle and store the assignment of the same tag/concept to the same resource by several users as separate assignments; e. g., for user $u_i$and u_j $\{(u_i, st_m, r_n), (u_j, st_m, r_n)\}$. On the user interface side, we display the annotations as one time assignments (st_m, r_n); i. e. only the tag/concept with no frequency indication (or frequency 1), and that are all editable by the user. If one user, e. g., u_i, decides now to remove a tag/concept from the annotations, then any assignment of this tag/concept to that resource by any user is deleted on the backend and data level; i. e. (u_i, st_m, r_n) and (u_j, st_m, r_n). In this way, we can foster community interaction and communication whilst retaining the information what user used which concept for which resource.

When saved, the (new) web or uploaded resource together with its rating and annotations as well as possible newly created concepts are immediately publicly visible and effective within the collaboration space.

6.6.2. Searching for Documents

The semantic document search service (see Figure 6.17) allows users to search and retrieve annotated documents (and performed dialogs). The users can type their search terms into a text field – similar to common Internet search engines. A semantic auto-completion supports the users when they formulate their input.

The search engine (supported by ontology management service) combines semantic search with a keyword-based full text search[14]. Therefore, the search engine analyzes the entered search string for occurrences of concepts from the ontology. If it recognizes references to concepts, it searches for documents – web and/or local and/or dialogs – annotated with these concepts or narrower ones. At the same time it also searches the full text of all annotated documents.

[14]The proposed search provides a novel approach of an hybrid flexible search combining keyword-based and semantic annotation-based search. A similar approach also showing the effectiveness of hybrid search has only recently been published by Bikakis et al. (2010). As semantic search has not been the focus of this thesis, we would like to refer to work by, e. g., Nagypál (2007b); Walter (2010); Thanh (2011).

Figure 6.17.: SOBOLEO page displaying results of the semantic search

In detail, when the Query Event is sent, the ontology management service as a registered query processing service extends the Query Event by identified concepts and their narrower ones. To identify the concepts, the ontology management service checks if the search string or parts of match any labels of the concepts. Then the text indexing component executes the actual query.

Therefore we index each document (including dialogs) with the fields: f_{TITLE} – the title of the document, $f_{CONTENT}$ – the extracted text content, f_{URL} – URL of web documents, f_{DATE} – when the document is added or modified, f_{TYPE} – web or local document or dialog and $f_{CONCEPTS}$ – concepts the document is annotated with. Then we create a BooleanQuery with t_{ss}, t_c and t_{c++}, where t_{ss} is the search string, t_c the identified concept and t_{c++} their narrower concepts. Then we match t_{ss} with $f_{CONTENT}$ and f_{TITLE} with boost of 2.0 for the latter one and t_c and t_{c++} with $f_{CONCEPTS}$ with a boost of 10.0 for the exact concepts t_c and a boost of 4.0 for the narrower ones. Together we get the resulting score for each matching document based on which the documents are ranked.

Thus, the semantic document search makes use of the main indicators *Tags of the resource*, *Frequency of tags*, *Time stamp of the tag assignment*, and *Background knowledge on the structure of tags* (see also Section 5.3.4).

On the result page the users get feedback how many documents were found and on which concepts it understood the query to be referencing. The resulting documents are shown with their title, annotated concepts, a short excerpt of the document's content with search terms highlighted, the exact URL and the overall rating – if enabled – together with its number of votes. Clicking on the title either opens the web page, starts the download of the local document or shows the dialog transcript. Clicking on one of the concepts opens the concept details in the browse or ontology-based navigation area (see Section 6.6.3). An "edit" link for each result entry lets users change or delete the bookmark, uploaded document or dialog, change their rating, and add or remove concepts.

Additional links let the user start/continue discussions about the documents or navigate to the discussion contents.

6.6.2.1. Query Refinements & Relaxations

Depending on the search string, the results and the ontology, the system may support the user with query guidance by proposing a number of query refinements or relaxations. That take into account the narrower concepts of the concepts identified in the search string for refinements and the broader ones for relaxation. These are combined with the concepts co-occurring in the annotations of the first N documents of the search result.

For refinements we change the query to get a more useful *subset* of results; i. e. we restrict results to documents annotated with a particular concept. A good refinement is (a) relevant i. e. related to the information need of the user and (b) discriminatory i.e. significantly reduces the number of results.

Therefore we define:

- **Relevant:** A concept is relevant if it is related to the query or used to annotate some of the current results.
- **Discriminatory:** A concept is discriminatory when it reduces the number of results to half or less.

Therefore, we compute the count of the narrower concepts of the concepts identified in the search string and the concepts that are also used to annotate the first N results. Then we propose the concepts with the highest counts but occur in less than the half of the documents (otherwise they would not significantly reduce the result).

For relaxations we change the query in a way that it returns *more* relevant results; i. e. we add another concept and also return results annotated with this concept. A good relaxation is (a) relevant i. e. related to the information need of the user and (b) gainful.

Therefore we define:

- **Relevant:** A concept is relevant if it is related to the query or used to annotate some of the current results.
- **Gainful:** A concept is gainful when it increases the number of results.

Therefore, we compute the count of the broader concepts of the concepts identified in the search string and again the concepts that are also used to annotate the first N results. Then we propose the concepts with the highest counts.

6.6.3. Ontology-based Navigation through the Shared Document & Dialog Space

The semantic browsing interface enables users to browse the directory of annotated documents and performed dialogs (supported by the (web) document and dialog annotation and ontology management services). The system guides the user with an ontology-based navigation that also informs about ontology details at the same time.

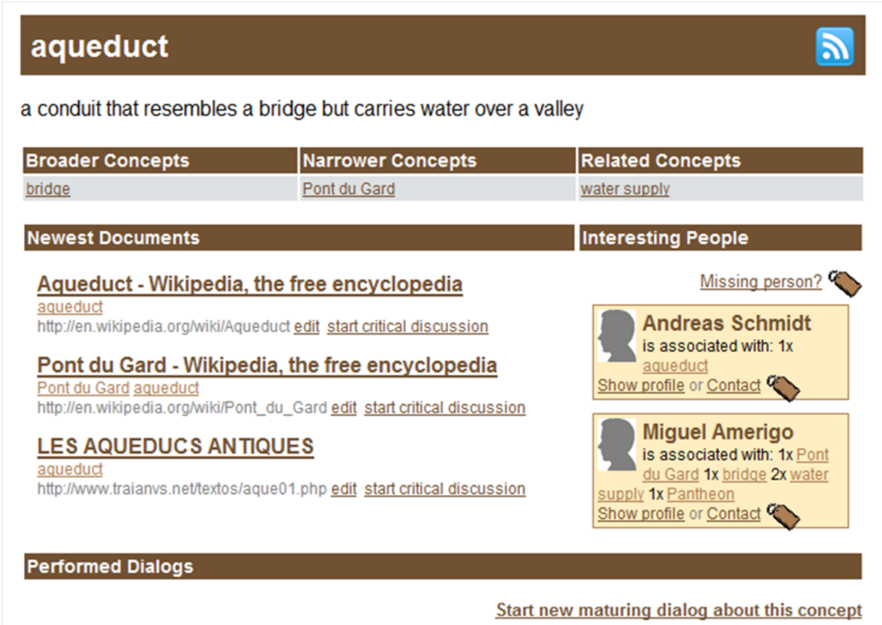


Figure 6.18.: Browsing web documents.

Starting from the root concepts, the users can click through the ontology concepts. Selecting a concept (see Figure 6.18 for the concept 'aqueduct'), the users can see the concept details; i. e. its preferred and alternative labels and its description (in the users' default language). Additionally, all its broader, narrower and related concepts are displayed as links for further navigation. If the users encounter some inconsistencies or possibilities to improve the concept/ontology, they can click on the "edit concept" link. The link opens the ontology editor for that concept and the users can immediately change the ontology.

Underneath the concept details, there is a list of annotated documents, related to the current concept; i. e. either annotated with this concept or one of its narrower concepts. The documents are ranked by the date they were collected or last time edited, with the most recent resources at the top.

Therefore, similar to Section 6.6.2, the ontology management service as a registered query processing service extends the Query Event by the narrower concepts of the concept to show. Then the text indexing component executes the actual query by matching t_c and t_{c++} with $f_{CONCEPTS}$ with a boost of 10.0 for the concept to show t_c and a boost of 1.0 for its narrower ones and ranking the matching documents by f_{DATE}.

The document entries are displayed similarly to the search results entries, only the content excerpt is left out. Via the edit link, the users can again immediately change the document details, annotations or rating. Again if enabled, the users can start/continue the discussion about each document or further navigate to the discussion's content. If a document is missing, the users can easily add a new document via the link "Add new document". The concept is already added as annotation.

At the bottom, there is a list of all performed dialogs related to the current concept; e. g., dialogs performed for ontology development purposes that involved this concept. Again the user can continue one of the dialogs, have a look at the contents or start a new maturing dialog.

In order to stay up to date, the system provides an ATOM Feed service for all new documents or just the documents annotated with one specific concept. In a nutshell, the ontology-based navigation structure helps to get an overview and better understanding how everything is connected. The document list makes aware of recent entries. The add and edit links all over the place provide the possibility to easily add and improve the knowledge base.

6.7. Semantic People Tagging

6.7.1. Bookmarking and Annotating People

We transfer the social bookmarking paradigm to people; i. e. the primary idea is to annotate a person (identified by his/her email address) via his/her personal web page, e. g., in the intranet, on the company's website or on a social network site. Similarly to bookmarking documents (see Section 6.6.1), there is a people annotation interface that is available both as a web page and a browser bookmarklet – the latter especially to work on top of existing people/employee directories.

If the system already knows the web page and the person, the people annotation management service will fill the annotation interface (see Figure 6.19) with the known information like name, email address and concepts the person is already annotated with. Otherwise it asks for the person's email address and – if not known and adding new persons is configured to be allowed – the person's name. In this way new contacts are easily taken up – if intended. Persons for which no web page is easily available can also be added and annotated directly via the annotate page in SOBOLEO by only providing their email address. When entering the person's email address or name, the system supports the users with auto-completion. Depending on the configuration, only existing persons may be tagged or person's with an email address in a specific domain.

Similarly to annotating documents, the users can use any concept from the ontology or arbitrary (multi-word) tags for annotating a person and the system provides auto-completion support for existing ontology entities. Existing annotations of other users are shown in aggregation in the box of "People's Topics"; i. e. the concept's preferred label together with its frequency is visible – the latter one is displayed as tooltip when hovering the label with the mouse – but not the taggers. The users may adopt the concept's assigned by others to their own annotations by just clicking on the concept label. If the person to be tagged is the user his-/herself, s/he can delete the concepts assigned to him/her by other users. Any other user can only remove his/her own annotations but not the ones of the others. Similarly, it is not possible to delete the whole person representation.

Furthermore, the system can provide tag/concept suggestions. If the person that is to be tagged is a user of the system, we can analyze the person's activities as indicator of potential expertise. Such activities can be grouped in: any concept editing, searching a concept, browsing a concept or using a concept for annotation that might be additionally restricted to a specific time frame. Currently, we apply as default a user's performed annotations limited to document annotations (if people and document annotation are enabled) as otherwise this might unintentionally reveal whom the user has tagged.

Each tagged person that is represented by one page and can also be tagged directly on this page. We will detail this hereafter. When saved, the (new) person together with its annotations as well as possible newly created concepts are immediately publicly visible and effective within the collaboration space.

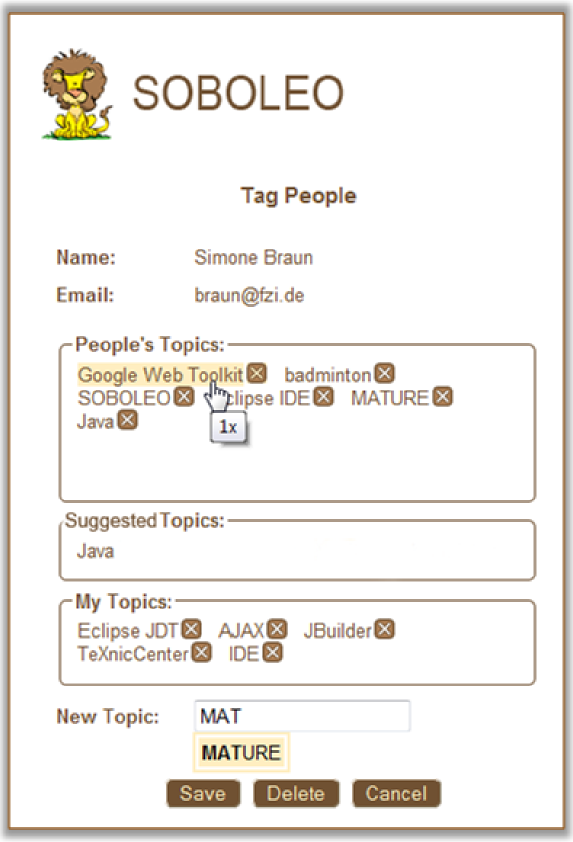


Figure 6.19.: Annotate People interface.

6.7.2. People Directory and Profiles

6.7.2.1. People Directory

The People Directory interface gives an overview of any tagged person and user within the collaboration space by dint of the people annotation management service (see Figure 6.20). The list shows the people with their name, profile picture, and annotated concepts (if some are assigned) in alphabetical order. The users have the possibility to directly annotate or contact the person or to navigate to the person's profile to get detailed information. If a person is missing, the users can easily add a new person via the link "Add new person" (if the option of adding additional persons is enabled). A search function that is similar to Section 6.7.3 can be used to search within the people directory and in this way to reduce the list.

Summarized, the people directory helps to get an overview on any person in the collaboration space. The add and edit links all over the place provide the possibility to easily add and improve the knowledge base.

6.7.2.2. People Profiles

The people profile user interface shows detailed information of a person, such as the person's name, profile picture or email address – by which to directly contact the person – or associated personal web pages; i. e. on a social network site (see Figure 6.21).

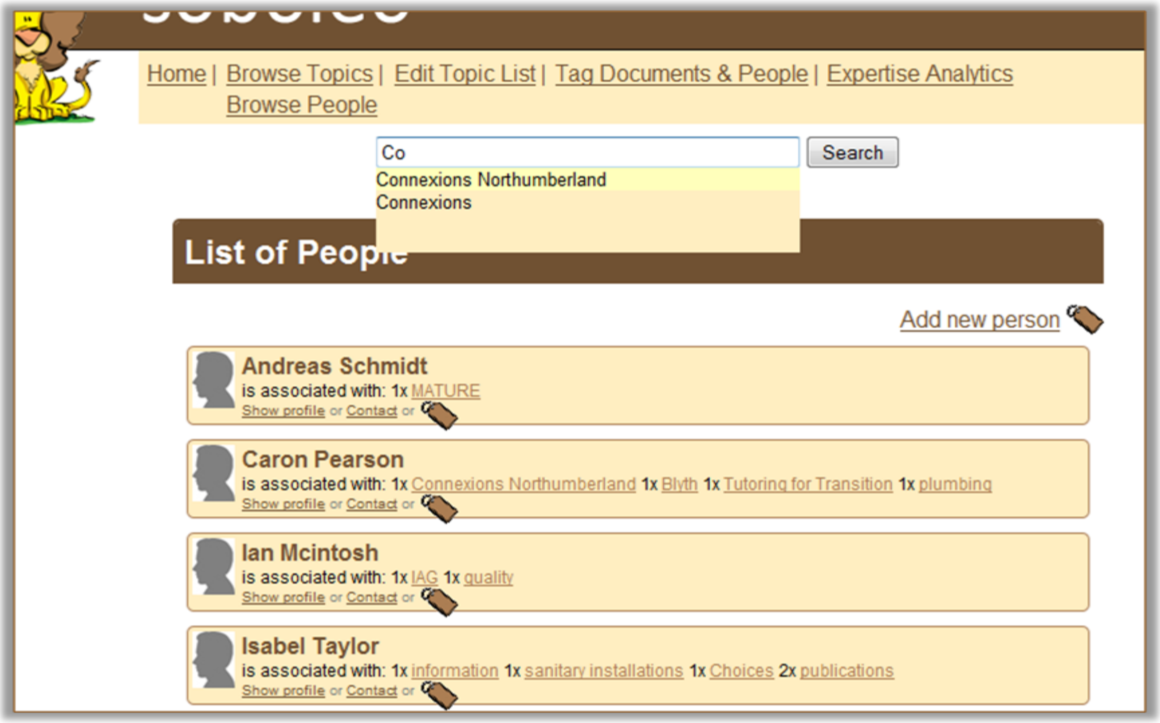


Figure 6.20.: People Directory interface.

Additionally, it visualizes the expertise and competency profile of the person. By dint of several other components (people annotation management, ontology management, history etc.) this service generates an expertise and competency profile for a person based on the information available about the person and the relations in the vocabulary. This information can include for example the concepts the person was tagged with, the concepts the person used to described resources or the tags the person searched for – if the person is a system user as well. This information is combined to give an aggregated picture of the user.

In general, the algorithm calculates a score for each assigned concept (tags) to a person as a weighted sum of the following indicators (see also Section 5.3.4). It is configurable how these additional indicators influence the overall score.

- Tags of the taggee
- Frequency of tags
- Time stamp of the annotation
- Tags of the tagger
- Tagging activities of the tagger
- Tagging activities of the taggee
- Additional activities of the tagger
- Additional activities of the taggee

For the profile visualization, the default setting skips the activities of the tagger and looks at the annotations of the taggee separate from the taggee's activities. That means, the people profile interface visualizes the concepts explicitly assigned to the person in one tag cloud and – if the person is a system user as well – topics/concepts extracted

Figure 6.21.: Profile of an individual person.

from the person's activities in a second tag cloud (shown as "Activity Tags"). The basis for the activity tags may be – besides using a concept for annotation and the time stamp of the annotation – as additional activities of the taggee any concept editing, searching a concept, browsing a concept; as well additionally restricted to a specific time frame. In this way, the system provides a possibility to compare the collective view of the person within the community with the topics the person is actually dealing with. And thus the system also helps to complete the picture of the person.

Hovering the individual labels in the tag clouds with the mouse, shows the exact number of frequency. The tag symbol lets users directly annotate the person; e. g., adopt the activity tags.

Related resources and people also gained from usage data analysis additionally completes a person's profile. Related resources are resources the person has annotated. Related people are people with a similar profile; i. e. currently by default people who are similarly annotated. This supports increasing awareness and potential community formation. For the computation of related people, the people ranking algorithm in Section 6.7.3.1 is applied with the concepts the person is annotated with as input for the search-topic-cloud. This is further extended by broader tags and the person-topic-cloud includes tags assigned to the person.

6.7.3. Searching for People

The semantic people search service allows users to search and retrieve annotated people. Similarly to the semantic document search service, the users can type their search terms into a text field supported by a semantic auto-completion and the semantic search engine (supported by the ontology management service) analyzes the entered search string for occurrences of concepts from the ontology. If it recognizes references to concepts, it searches for people annotated with these concepts or narrower ones.

On the result page the users get feedback on how many people were found and which concepts it understood the query to be referencing. Depending on the search string, the results and the ontology, the system may also propose a number of query refinements or relaxations.

The resulting persons are shown with their names, profile picture, and annotated concepts. The users can look at the profiles via the “Show profile” link. The “contact” link opens an email program for direct contact; the tag symbol lets users directly annotate the person.

Hereafter, we describe the search and ranking process in detail.

6.7.3.1. People Search and Ranking Algorithm

Based on heuristics, the people search service provides a ranking for a set of persons associated with a certain concept/tag. It analyses tagging data for that purpose, but also exploits higher levels of formality.

The algorithm works as follows:

1. Based on the query, i. e. based on the concepts extracted from a search string, the service generates a *search-topic-cloud* that includes the exact concepts from the query and, depending on the configuration, their broader concepts with a lesser weight. If the latter one is activated, the ranking also includes persons associated with a broader concept than the searched concepts.

2. In a next step, the service generates *person-topic-clouds* based on tagging data of each person. Depending on the configuration, the service extends the person-topic-cloud by broader concepts with a lesser weight. By including the broader concepts, the effect on the ranking is vice versa; i. e. this will also include persons who are associated with a NARROWER tag of the searched concepts. It is also configurable if and how to consider different types of tagging data; for instance only taking into account concepts that a person is tagged with or also including concept the person has used, e. g., with a lesser weight.

   The service retrieves a ranking score by matching search-topic-cloud and person-topic-cloud. Thus we have the search-topic-cloud $STC = \{(t_1, w_1), \ldots, (t_m, w_m)\}$ where $t_1, \ldots, t_m$ are the concepts extracted from the query and if applicable their broader concepts and $w_1, \ldots, w_m$ are their respective weights. And the person-topic-cloud $OTC = \{(t_1, w_1), \ldots, (t_n, w_n)\}$ where $t_1, \ldots, t_n$ are the tags applied to the person and if applicable their broader tags and $w_1, \ldots, w_n$ are their respective weights; i. e. the frequency of application or for broader concepts a third of the frequency. Thus we calculate the score $rs = \sum_{0 \leq i \leq m; 0 < \leq j \leq n} \frac{w_i + w_j}{2}$ if $t_i = t_j$.

3. The resulting matching score may be further adjusted by name match or taking into account current availability of the taggee and the social relationship between the searcher and the taggee. It is configurable how these additional indicators influence the overall score.

4. The result set is then sorted according to the adjusted scores. It is configurable if results with an overall score lesser than a threshold shall be dismissed.

By default the search-topic-cloud is not extended by broader tags but the person-topic-cloud is extended by broader tags. The person-topic-cloud includes tags assigned to the person. Exact tags are weighted with 1.0, extended tags with 0.3. We apply a threshold of 0.2. If there is a name match the person is pushed to the top of the list.

6.7.4. Ontology-based Navigation through the Annotated People Space

The semantic browsing interface enables users – similarly to browsing documents and dialogs – to navigate through the ontology and the directory of annotated people (supported by the people annotation management service). Starting from the root concepts, the browsing interface presents similar people for the current user using the people search & ranking algorithm similar to the related people on a person's profile. The aim is to make the user aware of people similar to him/her and thus to foster communication and community building.

When the users can click through the ontology concepts and their details, there is a list of all people underneath the concept details who are related with the currently selected concept. It depends on the configuration, which indicators are included or not to define the relatedness respectively the ranking. By default these are all people who are either annotated with this concept or with one of its narrower ones.

The people are displayed similarly to the search result. If a person is missing, the user can directly add a person to a concept, i.e. annotate the person with the selected concept, via the "Add new Person" link. If enabled, there is the list of all performed dialogs related to the current concept.

6.7.5. Expertise Analytics

From a guidance perspective it is not only important that a consolidated view on people profiles is achieved but also to monitor what knowledge is requested and thus maybe needs to be developed. Therefore, the expertise analytics service provides an aggregated overview and comparison of available and requested expertise based on annotations and search query analysis within a certain time frame. This services aggregates (1) the total set of topics possessed by people stored in the database and (2) the total set of topics queried for through the search interface. Based on this data it generates an output that summarizes (a) the expertise and competencies present and (b) the expertise and competencies queried. The topics are outputted together with a score indicating their frequency.

This service can be configured based on the time that is considered, the maximum number of competencies displayed and on the threshold for disregarding competencies that are used rarely.

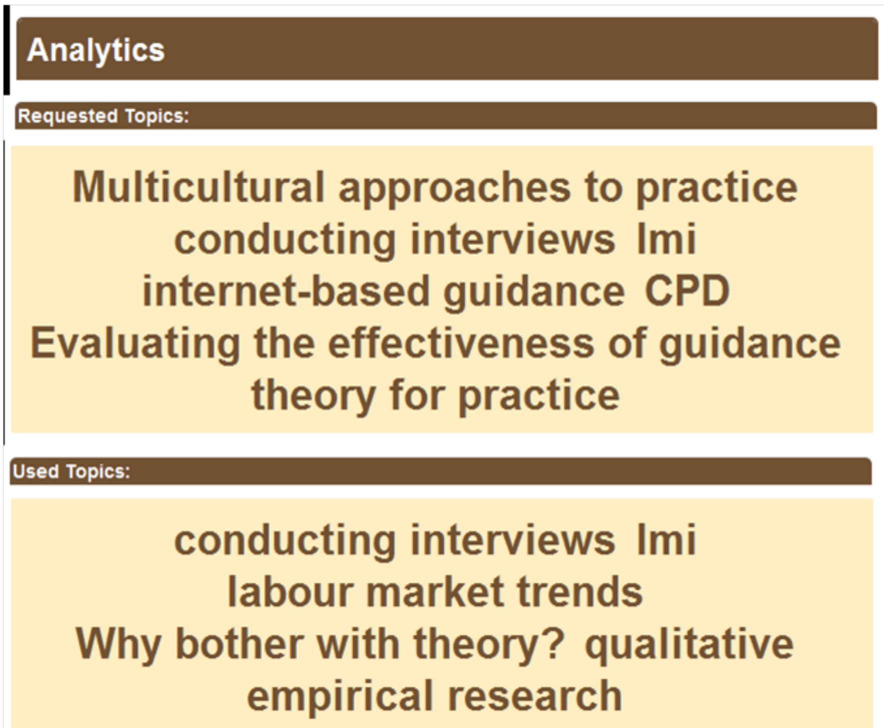


Figure 6.22.: Guidance overview showing topic tags the users requested, i. e. searched for, and actually used for annotation

- **Available Expertise:** For each concept assigned to a person within the given time frame, the algorithm calculates a score as a weighted sum over all tagged persons in the space.

- **Requested Expertise:** For each concept extracted from search queries within the given time frame, the algorithm calculates a score as a weighted sum over all tagged persons in the space.

Both result sets may be sorted according to the scores. It is configurable if results after a certain index or with an overall score lesser than a threshold shall be dismissed. Per default it is configured to consider a time frame of the last 30 days, to return a maximum of 20 competencies and to consider no threshold. The guidance overview or expertise analytics interface visualizes both the available expertise in comparison to the requested expertise in a tag cloud (see Figure 6.22). Hovering with the mouse over one label shows the score; clicking on it navigates to the details in the browsing area.

6.8. Supporting Ontology Maturing with SOBOLEO

In the following we want to illustrate how the SOBOLEO framework enables and supports the particular phases of ontology maturing as well as transitions between phases – with the emphasis being on phases I to III – on its application of social semantic bookmarking to an online community of practice in Classic Roman Civil Engineering (see also Section 7.6) and its application of semantic people tagging to a career advising organization (see also Section 8.3).

Figure 6.23.: Bookmarking and annotating a resource in SOBOLEO

6.8.1. In the Application of Social Semantic Bookmarking

- **Emergence of Ideas:** New ideas (topics) emerge from bookmarking and the collaborative development of ontologies by each community member's addition. This also leads to greater awareness of a relevant topic or asset within the community. Figure 6.23 shows a screenshot of how to bookmark and annotate resources in SOBOLEO.

  Similarly, when the functionality for structured dialogs is enabled, the proposal to initiate a dialog signals the desire to express further ideas and develop the collaborative understanding in relation to the domain. The annotation plug-in allows the instant initiation of a critical discussion. Thus, initiating a dialog also advances ontology maturing from this phase to the second, which is more focused on distributing and consolidating ideas and concepts within a community of practice.

- **Consolidation in Communities:** First the community becomes actively aware of new resources, through refined ontological classifications or collaboratively 'working with' the ontology. This can be achieved through directly using or modifying the ontology, or indirectly, through e. g., specific dialogs. Figure 6.14 presents a screenshot of SOBOLEO and InterLoc dialog game, which emphasizes this. Users are enabled to change the ontology in the editor and become aware of resources annotated with this concept. Moreover, they can initiate and participate in discussions about a concept. The current status of dialogs about the concept is also visible.

  Thus, performing critical dialogs about resources raises the level of the community's involvement and engagement with these assets, which also leads to greater formalization of these assets as legitimate materials of value to the community,

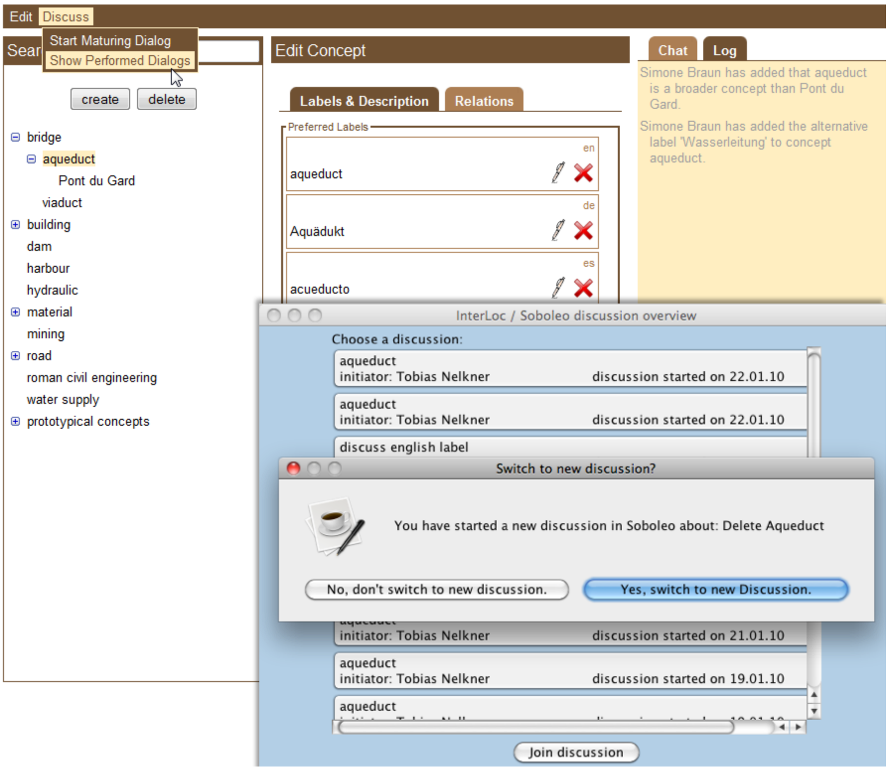


Figure 6.24.: Starting an ontology maturing dialog about a concept

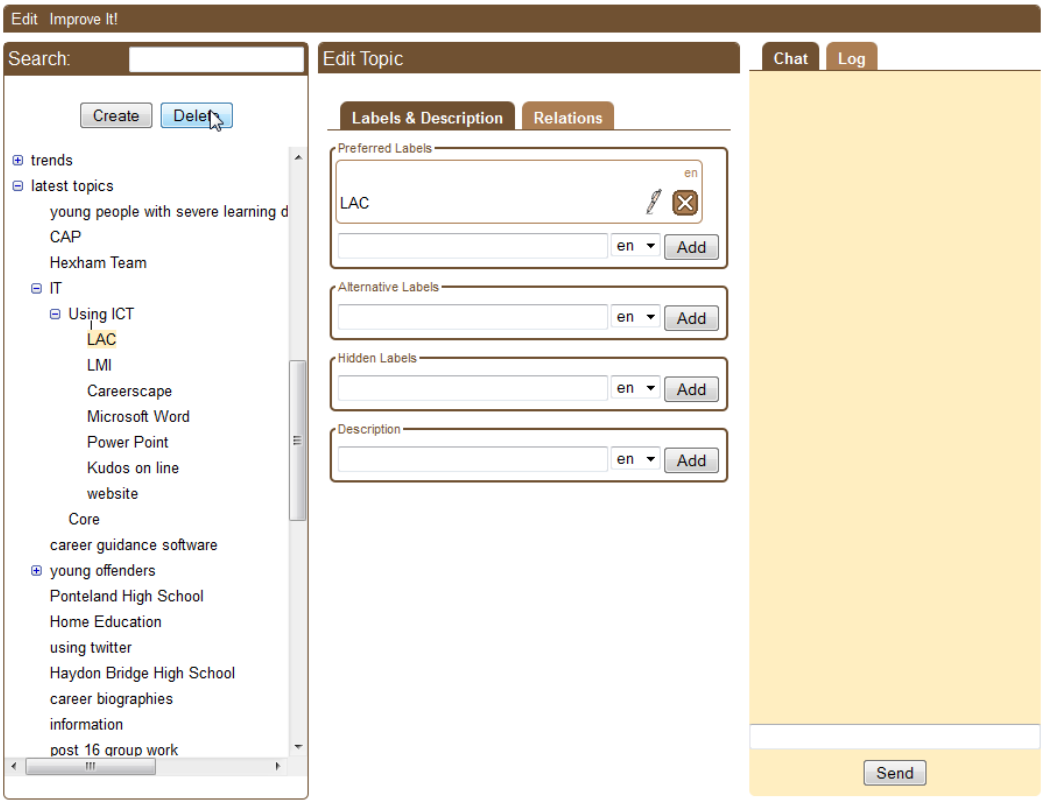
or what may be considered as a sort of 'acceptance' and 'embedding' of discussed resources within the community.

- **Formalization:** Formalization is realized once the community has performed gardening activities, through directly amending the ontology or performing dialogs that lead to the development of a mutually understood and shared ontology. Additionally, there will be a logical 'deepening' of the collective understanding about the resources within the ontology through the performance of critical discussion dialogs.

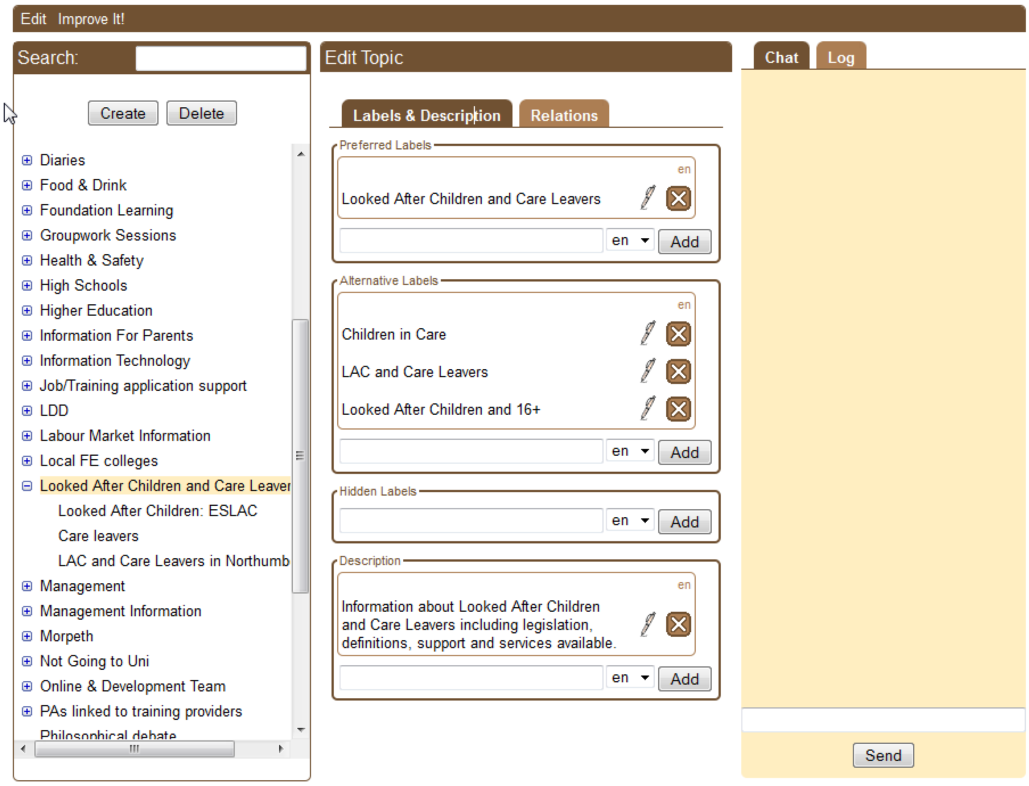
  Figure 6.24 depicts the ontology editor, which allows the collaborative changes to ontologies and provides the possibility to start a dialog related to a concept ('Aqueduct' in this case), where the user can define a title to give the discussion an initial direction. Awareness of, and reflection about, changes are promoted through amendments being logged and presented. Once maturing dialogs stop being performed, we can assume that a relatively stable and consensual ontology has been developed. Similarly, once the community has decided that a critical discussion dialog (about a particular resource) has been completed and its text is ready for publication as a resource itself, we can assume this is a consensual and formalized asset.

6.8.2. In the Application of Semantic People Tagging

- **Phase Emergence of Ideas:** At this phase new topics are taken up, e. g., from news, by surfing the net, daily client requests, or team meetings. On the other hand, a user might get into contact with another person s/he has not known yet, e. g., based on another colleagues' recommendation. The user starts to document



(a)



(b)

Figure 6.25.: Showing the maturing of the tag 'LAC' towards the finally formalized concept 'Looked After Children and Care Leavers'. In 6.25a the concept idea 'LAC' has been automatically collected during the annotation process. After a while in 6.25b, the community placed the new concept in the ontology, refined it by renaming it, adding alternative labels and a description, and finally adding further relations.

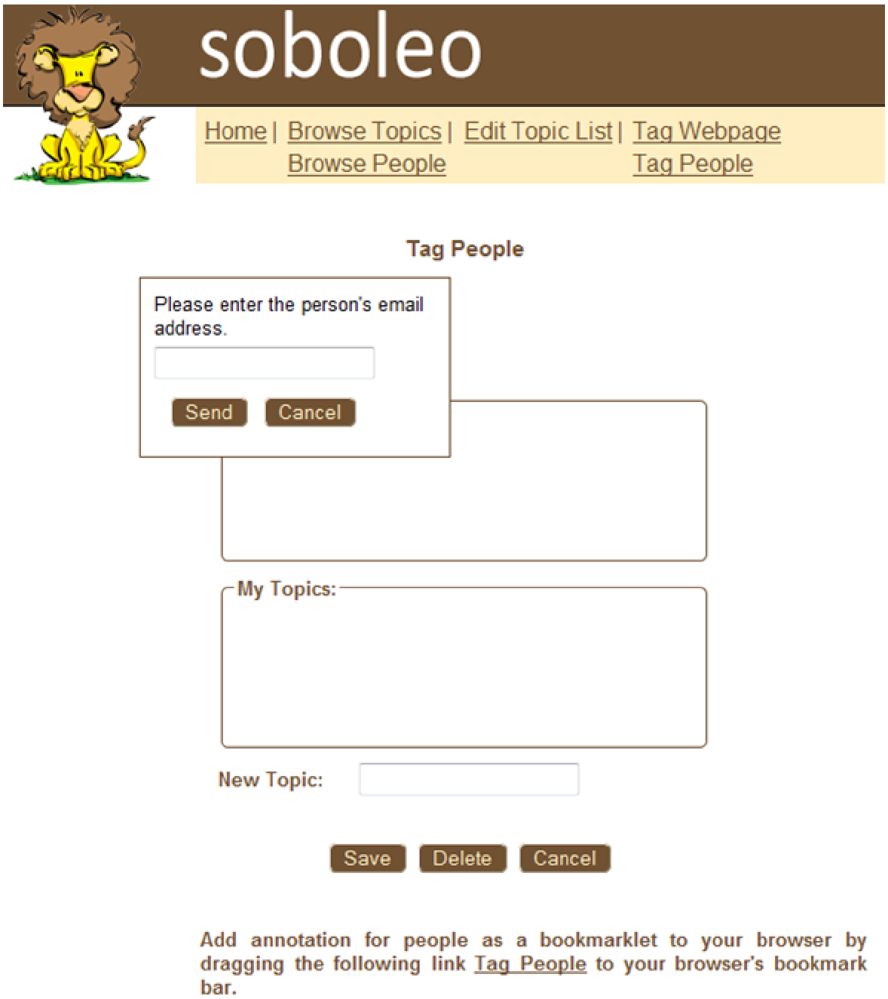


Figure 6.26.: Adding a new person to the collaboration space.

his/her newly gained knowledge, e. g., about new topics or contact data, in the system. Links to new contacts, external resources or tags for emerging topics that are deemed useful for later retrieval are added to the system (e. g., in Figure 6.25a the tag 'LAC' is automatically collected as 'latest topics'). This is a form of appropriation. The user may further appropriate the knowledge about a certain person by adding additional tags at a later stage. Similarly, s/he may reuse his/her newly added topic tag for associating other persons/resources with this new topic. At this stage the knowledge is still rather personal. Reuse is restricted to the "inventor" because other users are not yet aware of the new topic tag or person. Figure 6.26 shows how a new person can be added to the system by simply entering the person's name and email address. New topic tags can be added during the annotation process by entering the tag and saving the annotation. New ideas and topics may also be brought in through bookmarking interesting web resources or uploading local documents. This may additionally be supported with tagging suggestions based on text analysis (see Figure 6.19).

- **Phase Consolidation in Communities:** The core actions of transition to and within this phase are (1) increasing the awareness of expressed ideas and (2) reusing and adapting these ideas, i. e. the knowledge about the individuals' expertise and the knowledge about how to describe the expertise.

  New persons, (web) resources, tags and tag assignments are automatically made public and visible to the community. However, this publication does not imply that the ideas and knowledge are distributed. It's necessary to make others aware of the existence of this new knowledge so that others can reuse it – it is accepted in the community.

Again, we need indicators that signal the increase of awareness and acceptance. For example, we can say that users get aware of a person and associated tags, e. g., via search, browsing or feed notification, when they access the person's profile, when they add additional tags or approve the already assigned tags. This knowledge gains maturity when more and more people add and confirm tags or select and contact a certain person. The system also provides simple features to access a person, to see with which topics and how often s/he is tagged and to approve tags or contact the person.

Similarly, we can conclude that users get aware of new topics and tags, when they access a bookmark or newly added tag, when they add additional tags or rate the bookmark respectively, and when they add additional information to the tag, e. g., description. This knowledge about how to describe topics and expertise gains maturity when more and more people reuse tags, work collaboratively on tags (ontology) or tag a document with the same tags as other document. Again the system can act supportively, e. g., by providing auto-complete functionality or tagging suggestions for existing tags.

- **Phase Formalization:** Achieving agreement about the common topic terminology results in shared and structured competencies for expertise. Similarly, the agreement about people profiles results in a "competence map". We reach this phase, once the community has performed gardening activities, e. g., through directly amending the vocabulary, removing insufficiently used resources, etc. Therefore, the ontology editor facilitates the ontology manipulation.

  Figure 6.11 depicts the ontology editor as the core functionality to manage semantics. It allows the collaborative changes to the vocabulary and provides the recommendations to improve it. Support for the generation of ontologies is provided by the possibility to organize tags in hierarchical order and to define related tags. Awareness of, and reflection about, changes is promoted through amendments being logged and presented.

  Concerning the knowledge about available expertise, additional information based on the analysis of usage data, e. g., tags used, related resources, etc., helps us to complete a person's profile (see Figure 6.21).

  From a guidance perspective, it is not only about agreement on people profiles but also to monitor what knowledge is requested and thus needs to be developed. Figure 6.22 presents an analytical overview of trends. Decisions can then be drawn accordingly. The analytical overview can analyze historical usage data, extract useful information and display the information to the management in an integrated way. For example, showing the topics searched for in comparison with the topics used for annotation within the last month.

- **Phase Axiomatization:** Based on the meanwhile deepened understanding within the community, it is suitable to differentiate abstract competences (competence types) into competences with levels (competence instances). Besides these generalization relations, composition information made available provide, for instance, a better search support. Again the ontology editor with its extension for compositions provides the core functionality.

While the activities hitherto concentrate on rather informal aspects of competence management, it can also support the more formal process, e. g., the human resource development process. Indeed, finding people and human resource development require different levels of maturity of the used concepts. While for finding people all parts of the competence catalog can be used (even those who were just assigned by a single individual), for HR development we use only those that have a certain level of maturity that is manually assigned to the stable parts of the competence catalog.

So the competence catalog becomes a boundary object for both the operational as well as the strategic level that operate on.

6.9. Summary

As mentioned above, the SOBOLEO framework represents the technical instantiation of the conceptual design framework presented in Section 5. It does not implement the entire range of any possible design variable, which was down to the fact that (a) some options are not technically implementable, because they are concerned with the social dimension like norms and rules and (b) constraint resources meant that work had to focus on the functionality and the options required to support actual user demands. The technical framework is, however, build in a modular fashion such that all options can be added quickly to the system.

Hence, we summarize in the following which design options and scaffolding methods and functions of the conceptual design framework (cf. Section 5.3 and Section 5.1.3.2) are technically realized by the SOBOLEO framework. A comparison with related social semantic bookmarking systems together with overview tables can be found in Section 9.1.2.

- **Users:** The SOBOLEO framework explicitly differentiates between administrator and users and taggees. The other roles of tagger, creator/provider, owner and gardeners are inferred based on the models and actual usage data. For instance, the creator of a tag assignment is defined by foaf:maker. A user applying ontology gardening recommendations may be identified as a gardener.

  Users are socially connected through the concept of collaboration spaces of which users are members of one or more.

- **Resources:** The SOBOLEO framework supports either bookmarks, office documents or people or all together to be added and annotated as resources. Users can link to any web resource in the web or contribute any office document. Merely office documents are limited by file type filter to PDF or MS Office 2004 and 2007 formats. Persons to be tagged may be either provided by the users or the system. For the latter, taggable people are restricted to the space members. For the former, taggable people can be limited to people with an email address in a pre-defined domain, e. g., to only organization internal people, or for external contacts a bookmarklet may support the take up of external profiles as source. Resources are no more connected than by the users' tags.

  There are also no limitations to a specific group with extra privileges regarding the creation, deletion and modification or visibility of either type of resource. That means, any change is immediately publicly visible to any user and any document may be deleted or modified by any user (either system/web or space wide). Only people may not be deleted but by the administrator.

- **Ontology:** According to the SKOS format, users can give concepts (semantic tags) a (multi-word) preferred label and a description in multiple languages and any number of alternative and hidden labels. Currently, the ontology editor user interface supports English, French, German, Italian and Spanish as languages. Concepts may be connected with (multiple) other concepts through broader, narrower or related relations where the latter is symmetric and the former two inverse symmetric. Only for retrieval and navigational guidance purposes, the system interprets broader and narrower relations to be transitive[15]. For the maturing of competence ontologies, we additionally differentiate between competence types, instances with levels and compositions that are connected with respective subproperties skos:broaderGeneric/skos:narrowerGeneric, skos:broaderInstantive/ skos:narrowerInstantive and skos:broaderPartitive/narrowerPartitive according the SKOS extension (cf. Section 6.4.1). Any user can create, modify and delete any concept, its labels and meaning descriptions and relations to other concepts. Only the relation type is limited to broader, narrower and related and the generic, instantive and partitive specific one are automatically set by the system when the user (dis-)connects two competence types, creates/deletes competence instances or compositions. Any other potential change identified by the system is only provided as recommendation and executed on user approval. Changes to the ontology are again immediately publicly visible and effective.

- **Tag Assignments:** Regarding the aggregation, the SOBOLEO framework supports both multiple assignments and one time assignments to the same resource – where the option multiple assignments is requisite for people tagging. One time assignment is the default for bookmarking documents but indeed realized as a mixed mode; i. e. handled as multiple assignments in the backend but used as one time assignment on the user interface. In this way, we can foster community interaction and communication whilst retaining the possibility for analyses and aggregated statistics. Annotations do not have any further explicit semantics and technically are represented using commonTag:tagged as relation type.

  Any user may create annotations for any resource. For people tagging, self-tagging is supported and the system may create additional (implicit) annotations based on e. g., the analysis of usage data or competence compositions. System generated annotations are marked as such and separately handled; e. g., as "activity tags" in the people profile (see Section 6.7.2).

  The deletion and modification of annotations of documents depends on the aggregation mode. Either any user changes any annotation in the one time assignment aggregation or only the taggers can delete/modify their own annotations in the multiple assignment aggregation. The latter one is also true for people tag assignments. Additionally, it is configurable if the taggee can delete others' annotations. Any annotation or change of is immediately publicly visible without a need for approval – for people tagging in an aggregated and anonymous form; for documents the tagger is visible, e. g., on the user's profile.

- **Scaffolding Methods:** To support the seeding at the very beginning, the SOBOLEO framework supports the import of existing knowledge structure in the SKOS format.

[15] see also the note on skos:broader and Transitivity in the SKOS reference (http://www.w3.org/TR/2009/REC-skos-reference-20090818/#L2413)

For the general growth, the framework allows the easy creation and adding of new knowledge. There are various links distributed over the whole system that enable the user to add or edit either the ontology or any resource and their annotations; e. g., editing the annotations of documents or persons from the search result or similarly, to engage in a dialog. Browser bookmarklets support the take up and annotation of web documents or people with external profiles. During the tagging the users can use new keywords. When saving the annotation data, the annotation services automatically take up and store the new keywords as “prototypical concepts” within the ontology. As well during the tagging, the framework can provide tag recommendations for new tags/concept ideas based on external services that analyze the text content of the documents or user activities.

New persons, documents, tags and tag assignments are automatically made public and visible to the community as well as effective; e. g., when searching or browsing the collaboration space. Feed notification is also available to additionally make aware of new elements.

To foster user contribution and motivation, there are the explicit system feedback mechanisms gratitude, historical reminder and relative and social ranking available.

To enforce reuse of concept from the ontology, SOBOLEO provides auto-completion together with disambiguation support and tag recommendations of existing tags/-concepts from the ontology during the tagging activity. Auto-completion and disambiguation support is also provided for formulating a search query and for query relaxation and refinement or in the editor for adding concept details.

To support the agreement process, there is a chat integrated in the ontology editor to discuss ontology development related issues in particular. If enabled, users can engage in structured dialogs either about the ontology, its development and gardening process or about resources (e. g., reflecting and debating the correctness and quality) and their annotations from various points in the system. To preserve the rationale, there is an additional record of dialog performance that is linked to the related content and are accessible and searchable. In this way, discussions are also embedded in the context of the artefacts.

Rating may be enabled for documents as a feedback mechanism. Then the users can see a document’s overall ratings together with the number of given ratings.

In general, collaborative and mutual editing, amending and complementing is set by default for any resource, annotations and ontology element. For gardening activities, the ontology editor provides an easy-to-use tool. Changes are recorded in a history and for instance shown when accessing the ontology editor. The user may also subscribe to specific resources, persons or ontology elements in order to monitor community activities.

The ontology gardening recommendations provide a range of services that analyze and aggregate the collaboration space and its data to provide support for reseeding and gardening activities. The services include the analysis of a SKOS ontology, potential concept relationships and usage (cf. Section 6.4.2). Similarly, if enabled, the system analyzes usage data and user activities to infer a person’s expertise and provides the results, for instance, as tag suggestions or “activity tags” in the person’s profile, thus to complete the community’s knowledge base.

The expertise analytics service component provides for guidance purposes a visualization the topics searched for within the collaboration space in comparison with the topics used for annotation during a specific configurable time frame.

Part III.

Usage and Assessment

Overview

In this part we present the third pillar of our contributions – a framework of empirical application-oriented insights gained from the usage and assessment in the formative evaluation within nine case studies (Bortz and Döring, 2006). The nine case studies were conducted during the design and development process following the design science research principles of iterative cycles of design, implementation, evaluation and redesign. Figure 6.27 gives an overview on the chronology of the case studies between 2007 and 2010 from the bottom up.

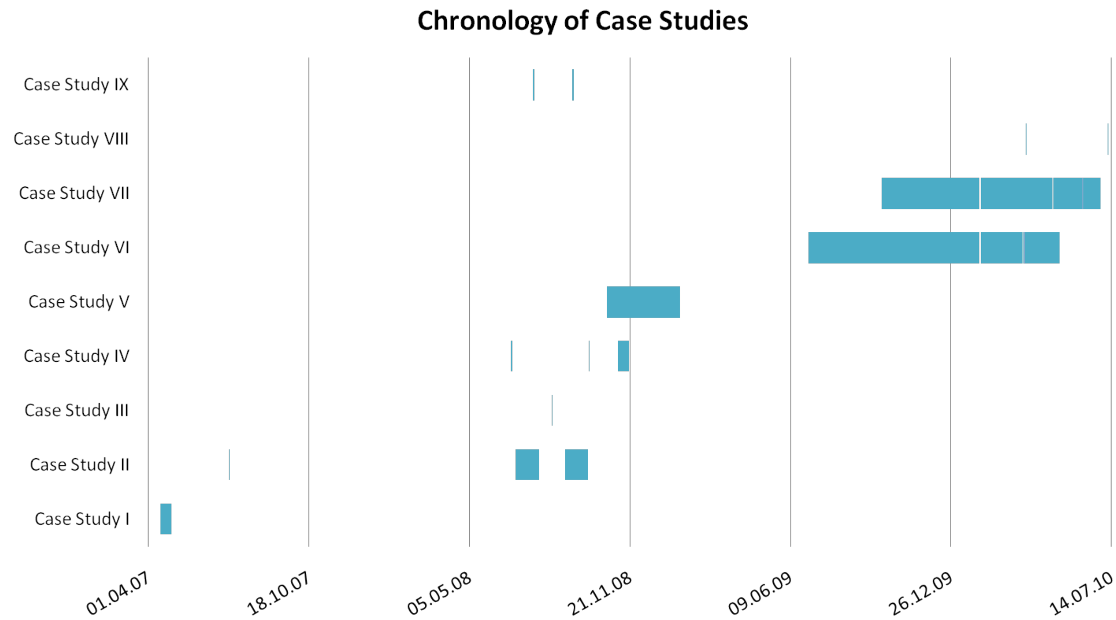


Figure 6.27.: Timeline of case studies

Six case studies focus on social semantic bookmarking and three on semantic people tagging due to the design and development process. Each of the evaluations fed back into our development of the technical framework and refinement of conceptual models and design. They further show the feasibility and usefulness of our approach and that our technical framework is actually usable.

The case studies are of variable length and intensity and originally conducted with different purposes and objectives. For instances, two case studies were carried out not by ourselves but research partners (q. v. 7.4 and 7.5).

Starting with the evaluation of our approach in the application of social semantic bookmarking the first case study, case study I in Section 7.1, took place in 2007 as part of the challenge on collaborative knowledge construction that aimed to bring together the available applications to construct ontologies in a social and collaborative manner. Case study II (Section 7.2) was the first evaluation with "real", non-technical users of a rapid prototyping community[16].

[16]Here, rapid prototyping does not relate to rapid software prototyping but the field of generative or additive manufacturing concerned with the computer-aided construction of tangible three dimensional objects; see e. g., http://www.rtejournal.de

The subsequent case studies took place more or less in parallel to the activities with the rapid prototyping community in 2008. Case study III (Section 7.3) describes a workshop part of a summer school in which we aimed to pass on understanding design processes as learning process and to experience with our approach a distributed knowledge modeling process as such a learning process. Case study IV and V (Section 7.4 and 7.5) describe the experimental studies which our application partners conducted in the context of their university courses by means of our SOBOLEO framework. The last case study more focused on social semantic bookmarking – case study VI in Section 7.6 – aimed at dialogs for the developing a shared understanding in a Spanish then newly built up community on classic roman civil engineering.

For the evaluation of semantic people tagging, we conducted two field experiments with pen-and-paper prototypes (case study VII in Section 8.1). These were targeted at the individuals' perspective and acceptance in order to explore the potentials and risks prior to the implementation and evaluation in a real enterprise setting. Additionally, we aimed with case study VIII (Section 8.2) for a conceptual validation with the method of experts focus group (Krug, 2000; Nielsen, 1994). This was targeted at the organizational perspective. The last case study IX (Section 8.3) describes the evaluation of semantic people tagging in the real enterprise setting of British career guidance.

7. Social Semantic Bookmarking

7.1. Case Study I: The Collaborative Knowledge Construction Challenge – Where Everything Began

The Collaborative Knowledge Construction (CKC) Challenge took place in 2007 as part of the Workshop on the Social and Collaborative Construction of Structured Knowledge co-located with the 16th International World Wide Web Conference in Banff, Canada[1]. The challenge's aim was to get users trying out then nascent social semantic web applications that allow to construct structured knowledge and ontologies in a social and collaborative manner and in this way to evaluate the tools and to understand what is the users' expectation of such tools. Researchers and developers were invited to contribute their tools and finally six tools in total – BibSonomy (Hotho et al., 2006b), Collaborative Protégé (Tudorache et al., 2008a), DBin (Tummarello et al., 2006), Hozo (Kozaki et al., 2002), OntoWiki (Auer and Riechert, 2007) and SOBOLEO – were selected to take part in the challenge. The workshop details and challenge results have been published by the workshop organizers in Noy et al. (2008).

7.1.1. Procedure

The challenge ran for two weeks from April 16 to 30 2007 prior to the workshop and the task was *"to construct structured knowledge for a portal that would provide information about research"* (Noy et al. 2008), e. g., research topics, groups, conferences, publications etc. We provided a basic ontology to facilitate getting started and to give thematic orientation for the participants. This ontology was tailored to the research domain as a whole with concepts like 'research topic', 'people', 'institution', 'publication', and 'event'. Everyone was free to participate and contribute information about their research domain. At the end, the participants were asked to provide feedback in a little survey. Altogether, 49 users registered and 33 contributed actively to the challenge.

7.1.2. Results

During this evaluation, the participants added a total of 202 new concepts and 393 concept relations to the ontology. Further, they collected 278 web resources, which they annotated with 3 concepts per resource on average. None of the users had the opportunity to meet other users using SOBOLEO at the same time. Thus, the chat functionality was barely used; only for testing. More detailed statistical information shows Table 7.1.

[1] http://km.aifb.uni-karlsruhe.de/ws/ckc2007

Table 7.1.: Statistical data overview of the Collaborative Knowledge Construction challenge

Target	Activity	Number
Concepts	created	202
	deleted	104
	changed label (preferred /alternative)	306
	changed description	11
Relations	added relation (broader, narrower, related)	393
	removed relation	159
Documents	added	278
	removed	n/a°
	updated	n/a°
Annotations	added	660
	removed	n/a°
Misc	browsed concepts	n/a°
	browsed key concepts	n/a°
	performed searches	n/a°
	sent chat messages	18

° data not available due to then limited logging

Summarizing the feedbacks, the participants appreciated the ease-to-use of SOBOLEO and having a shared ontology. They emphasized in particular the editor's real-time nature. The users further enjoyed the simple way for annotating web resources with concepts or tags, which are then automatically added. Thus, to have the possibility to integrate not yet well defined concepts but something like *"starter concepts"* and, in this way, to *"get the ontology building almost for free"*.

For improving SOBOLEO, the users pointed out several times that they missed a personal view on the data, i. e. on the own annotated resources but also on the ontology (especially in case of a growing and dispersing user base). Although the users appreciated the messages/chat pane informing about changes and for communication with other users, the users expressed the wish to have more possibilities to discuss and be informed about modification (on 'own' data) by other users. Thus, to gain more translucence and awareness, especially as they could not experience working together simultaneously. A further aspect was to have better support for identifying or suggesting conflicts, synonymous concepts and broader-narrower relations in order to facilitate the maintenance of the ontology.

7.1.3. Discussion

This was the first of our case study series and we could show that the users appreciate our approach and that the users indeed do create semantic tags and structures for minor investment. The suggestions for improvement gave us the first indications for the importance of appropriate discussion and ontology gardening support for ontology maturing scaffolding (q. v. 5.1.3.2) that we later tackled in the following way:

- **Personal view on data:** We provided a personal view on data through each person's profile that shows any information related to that specific person; among others the documents the person annotated and/or added to the system (q. v. 6.7.2).
- **Awareness of changes:** We integrated various different functions to provide awareness of changes. These are ATOM feeds (see 6.6.3), recent changes of the ontology when entering the ontology editor (q. v. 6.4) and the history page either for any type of events or selected ones.
- **Extended discussion:** These suggestions led us to provide – besides the simple chat functionality in the ontology editor – specifically designed structured dialogs supported by integrating the InterLoc5 dialog game tool and linking making performed dialogs available for later access (q. v. 6.5).
- **Ontology maintenance support:** We started with the concept of gardening and gardeners and have developed a modular and extensible ontology gardening recommendation service component (q. v. 6.4.2).

7.2. Case Study II: The Rapid Prototyping Research Community – Linked Information Processes in Research Networks

In this section, we present a case study that was part of the German national funded research project "Im Wissensnetz – Vernetzte Informationsprozesse in Forschungsverbünden"[2], which aimed to support efficient interdisciplinary knowledge-adding processes within e-Science.

Research is likely to be the most knowledge-intensive environment. An empirical analysis of existing (cooperation-)processes, information and knowledge exchanges, and instruments for the preservation of knowledge accomplished in the application domain "rapid prototyping" revealed that scientific work is characterized by high variability, dynamics and unpredictability as well as by high significance of social interactions and communication.

Especially in applied research that requires an increasing interdisciplinarity, the know-how of the individual researcher plays an important role. The researchers are lacking a comprehensive overview on existing findings beyond the own domain. They can only make limited use of available competences within and beyond research networks and risk delayed finding processes and inefficient project handling.

So, in the domain of rapid prototyping, linking people with individual expert knowledge and contents from various disciplines like plastics, ceramics, and mechanical engineering is one of the most important challenges. For instance, one major problem is searching and retrieving adequate contemporary resources. This process is very tedious. The users have to access many various data sources with different interfaces, but also the Internet with common search engines like Google.

In the area of plastics and their market these high dynamics are particularly obvious. New materials or new forms of existing ones frequently enter the market; brand names and manufacturers are permanently changing and hardly trackable—attributes of a chemical substance retrievable using its brand name today, are very hard to find once it's sold under a different label. There is also no general up-to-date database which lists manufacturers and brand names of currently available forms of plastics. Thus, the

[2] http://www.im-wissensnetz.de/

users rely on search engines like Google in order to find the wanted material of which only the old brand or chemical name is known. However, common search engines provide many irrelevant results because of their missing focus on the domain. For instance, when looking for "nylon" you receive lots of results for stockings.

At this point, using annotation and retrieval tools can help; e. g., when a colleague already found the new brand name of a product and tagged it with the old one you are looking for. With such tools being further semantically enriched with background knowledge and domain ontologies, it is possible to find out the search context and thus to extend or refine the search in order to reduce irrelevant results and to guide the user. However, this also requires that (1) the users collaboratively build up and maintain a shared understanding and terminology, (2) these activities are embedded into their information seeking activities and (3) this understanding is expressed formally enough to enable ontology-based query refinement.

Even though the project's scope was the area of e-science, these characteristics and difficulties, such as the prevailing lack of (adequate) information, are also true for the industrial sector.

By analyzing the rapid prototyping domain, the subsequent issues emerged to be the most challenging ones:

- The flexible design and support of collaborative processes
- The qualitative improvement of distributed search over different sources of information
- Methods and technologies for more efficient identification and support of interdisciplinary expert teams
- The collaborative development of shared vocabularies.

The implementation and evaluation with the rapid prototyping research community was divided in two main series of user tests and participatory design activities we will present subsequently.

7.2.1. Procedure User Tests 1

The first series of user tests took place on July 10 2007. The aim of that user test was to explore and assess the current status of system development by end users under controlled conditions on the basis of typical use cases. The user test focused on interface and human computer interaction design as well as on user support including user satisfaction according to DIN EN ISO 9241 "Ergonomic requirements for office work with visual display terminals" and DIN EN ISO 13407 "Human-centred design processes for interactive systems" (cf. DATech).

The user tests were conducted within two sessions of a 1-day-workshop[3]. Four pilot users tested the system based on a scenario that represented a usual situation in the end users' daily work and on guiding tasks. Half of the users were researchers of the rapid prototyping domain and half of them patent experts for German research. All of them were unexperienced in ontology development. We provided a basic ontology with 31 concepts to start with that was thematically tailored to the rapid prototyping domain.

[3] We did the preparation of the material and workshop as well as the analysis together with our evaluation partner Fraunhofer IGD Rostock. The detailed test materials and analysis are reported in Oertel and Schulz (eds.) (project internal report)

Similarly, the web resources to work with were specific for rapid prototyping. There was one test iteration; i. e. the second test session, with a similar scenario to assess learnability respectively learning supportiveness during usage.

For data collection we used five different methods:

- **Thinking Aloud:** We asked the users to comment their actions and reactions during the task performance by "thinking aloud"; e. g., by describing what's happening, what's going wrong, why they do an action. We motivated the thinking aloud by a list of questions.
- **Observation:** We systematically captured the user behavior by observation and note taking. We recorded the individual events based on a system of observation categories.
- **Questionnaire:** We used questionnaires with 7-point Likert scale to collect aspects of user expectations, learning supportiveness, task appropriateness, and user satisfaction.
- **Interview:** subsequent to the test situation we interviewed the users based on an interview guideline to dwell on arisen problems.
- **Screen Recording:** We additionally used screen recording together with sound and video recording of the test users as a special form of user behavior observation.

Before the test sessions, we gave a presentation to introduce the SOBOLEO system; i. e. its purpose and the main functionalities to be tested being (1) the collaborative development of ontologies, (2) the collaborative collection and organization of bookmarks, (3) the annotation of bookmarks with concepts of the ontology, and (4) semantic search within the space of annotated web resources.

Each test session lasted one hour. The participants worked in parallel in the same room on separate computers. Each participant had his/her personal test instructor, who guided the sessions, did the observation and conducted the questionnaires and interviews. There were instructions for the test instructors and test users to keep the sessions constant.

At the beginning of the first session, the test instructors welcomed their assigned test user, gave an overview on the test procedure and informed about recording. Then they conducted the first interview on the users' expectations and previous experiences. The interview comprised five open-ended questions about the users' current work processes, expectations on the test system, opinions on advantages and disadvantages and other known systems.

After this first interview, the test users received a short description of the test system, the scenario and the tasks to read. The tasks to do were:

- Annotating three online articles from the RTejournal website[4]
- Integrating two concepts newly added during the annotation into the ontology
- Browsing the bookmark collection and searching web resources for a specific topic.

[4] http://www.rtejournal.de

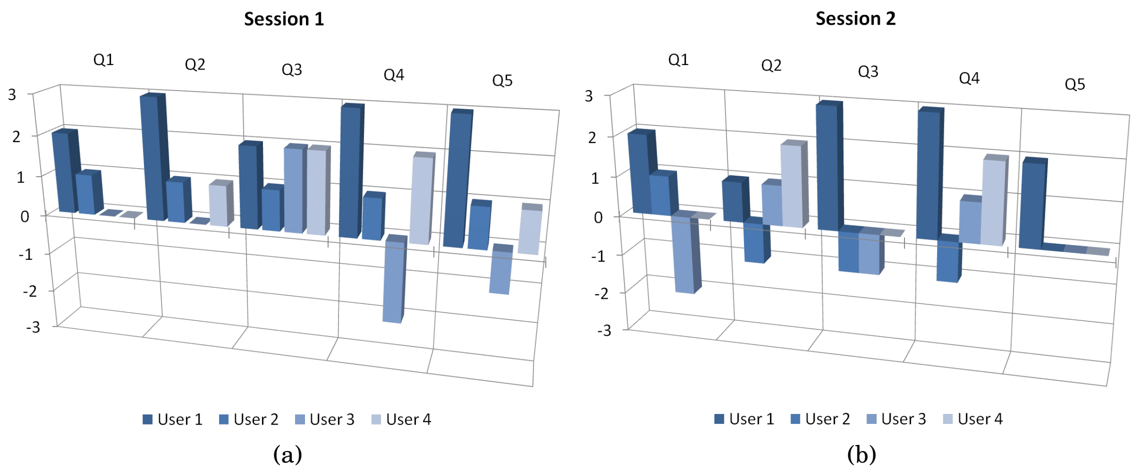


Figure 7.1.: (a) and (b) show the task appropriateness in session 1 & 2 on Q1: The system is easy to use; Q2: It provides all functionalities in order to efficiently accomplish the arising tasks; Q3: It provides good opportunities to automate repetitive process steps; Q4: It does not require redundant user input; Q5: It is tailored to the work needs.

Before the users started to work on the three given tasks, there was a short briefing of the work place and the thinking aloud method.

During the test session, the instructors took minutes including working time per task, errors, consultations etc. and stimulated the users thinking aloud with questions about the users' goal, reasons for an action, expectations, unsureness, understanding or ideas of improvement.

After finishing the tasks, the instructors asked about the test users experiences on the system with two questionnaires on their satisfaction and human computer interaction design. The former one comprised five 7-point Likert scale questions and the latter one each with five 7-point Likert scale questions about the system's task appropriateness and learnability. The users' overall impressions were collected within an interview of eight open-ended questions. We concluded the session with a questionnaire on personal data.

This procedure was repeated in the second session with a different but similar set of tasks that focused on collaborative aspects. We left out the briefing, the interview on expectations and the questionnaire on personal data.

7.2.2. Results User Tests 1

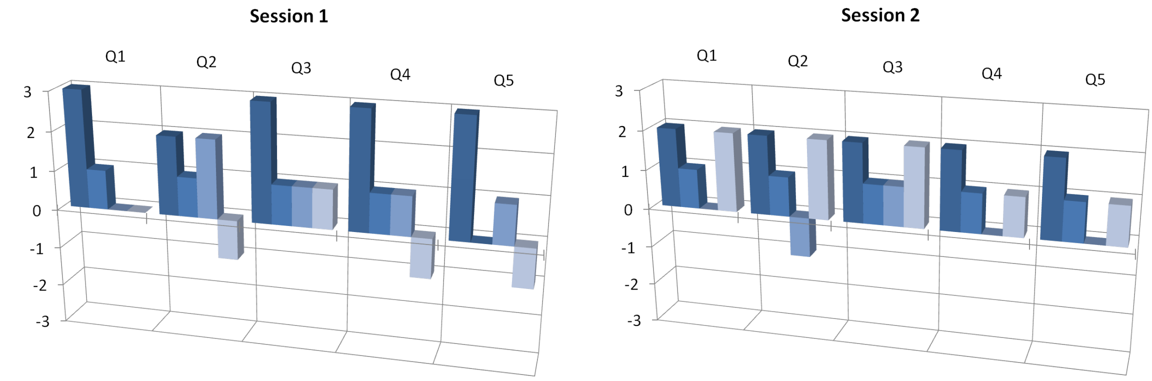
The given tasks were tailored to familiarize and gain orientation within the ontology by e. g., letting the users place or add synonyms to existing concepts. In total the users added 6 concepts to the ontology with 11 alternative labels and 21 concept relations (q. v. Table 7.2). They did not add any description. In total, the users added 42 web documents that were annotated with 104 concepts. The amount of generated data is limited due to the tight task structure.

The diagrams in Figure 7.1 show the detailed results of the user satisfaction and the human computer interaction design questionnaires. Table 7.3 presents the answers of the interview on the users' overall impressions.

Table 7.2.: Statistical data overview of the user test

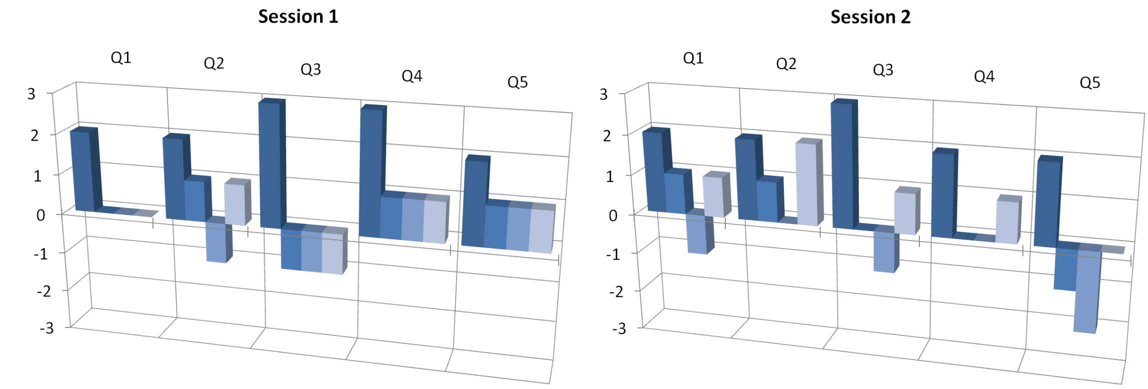
Target	Activity	Number
Concepts	created	6
	deleted	2
	renamed	5
	added alternative	11
	removed alternative	2
	changed description	0
Relations	added broader/narrower*	20
	removed broader/narrower	8
	added related	1
	removed related	0
Documents	added	42
	removed	n/a°
	updated	n/a°
Annotations	added	104
	removed	n/a°
Misc	browsed concepts	n/a°
	browsed key concepts	n/a°
	performed searchs	n/a°
	sent chat messages	18

* not including automatically created relations with the container "prototypical concepts" when a new concept is used during the tagging process
° data not available due to then limited logging



(c) (d)

Figure 7.1.: (c) and (d) show the learnability in session 1 & 2 on Q1: The system needs little time to learn; Q2: It encourages trying out new features; Q3: It does not require remembering many details; Q4: It is designed in a way that it is easy to memorize what the user once has learned; Q5: Learning the system does not require assistance.



(e) (f)

Figure 7.1.: (e) and (f) show the satisfaction in session 1 & 2 on Q1: The system is appealing; Q2: Working on the tasks was easy; Q3: Starting with the system was straightforward; Q4: The system is appropriate for the tasks at hand; Q5: It is made for daily use.

Table 7.3.: Results from the interview on the users' overall impression

Question	Session	User 1	User 2	User 3	User 4
What did you like particularly of the system?	1	-Few, logical functionalities -System improves, the more links are added	-annotating is straightforward	-Deleting annotated concepts -Typos quickly changeable	-n/a
	2	- Very easy, especially annotating	-Collaborative chat	- Quick overview of the ontology	-n/a
What did you like not at all of the system?	1	-nothing	-no modifications of entered key words after the annotation	-no feedback of success	-n/a
	2	-there is no search functionality for the ontology; only for the bookmark collection	-complexity of ontology levels -there were no alternative structures or ontologies -system crashes	-"edit" is cumbersome -inputs are difficult, not intuitive -Drag'n'Drop missing	-meaning of categories needs to be obvious at a glance -simplify or filter the taxonomy -confusing Drag'n'Drop -target not highlighted when moving
What was difficult to understand when starting with the system?	1	-big leap between explanation and real use	-open SOBOLEO and installing the Annotate-Button -Annotate pop-up hides the actual text of the website	-interpretation of the dialog	-start similar difficult like other tools; there is no intuitive software
	2	-"alternative label" difficult to enter	-orientation problems	-center part of "Edit"	-start is difficult -highlighting of different categories
For what kind of tasks do you think the system is well-suited?	1	-for collecting links – it is very easy and you can do it simultaneously with your work	-integrating literature with a knowledge base	-annotating, storing, retrieving information for a specific topic	-structure document terminilogy
	2	-Basically for everything -Sharing possibilities are good	-locating the right expert	-organizing web resources once found	-organize documents with the help of a schema
For what kind of tasks do you think the system is unsuited?	1	-for none of the given ones	-annotated concepts are not highlighted in the document	-comprehensive (literature) research	-for scrolling through big document; no support -search within documents
	2	-for none of the given ones	-for building up/finding an ontology due to complexity	-not suitable for collaboration; chat not comfortable	-navigation in big taxonomies
What caused you the most difficulties during task execution?	1	-operation control not a problem but task definition not always clear	-editing the ontology -clashes with other input structures	-English/German -locating where to enter own annotations	-starting with the system; understanding the main functionalities
	2	-finding a specific concepts in the ontology	-could not work on task no. 4 because of system crash -the tasks of the colleagues contained too little suggestions for improving the ontology	-browsing the tree -moving the browser window upwards	-orientation in the tree structure -make the display 2-dimensional as landscape; e.g. STN, Delphioni
What would you like to be changed?	1	-n/a	-annoate pop-up should not hide the text behind	-feedback	-facilitate investigation and search within documents -highlighting of search terms
	2	-search within the ontology	-the system should suggest and display different kinds of ontologies	-better structure of the message pane	-display of the tree; maybe as a landscape -Drag'n'Drop with highlighting similar to Windows Explorer
Do you see problems when introducing the system in your work context? If so, which one?	1	-n/a	-structural clashes in the concept hierarchy with other people	-data source integration	-in general: annotating usually works with "post it"
	2	-the first interaction	-the system tempts people to chat without focus	-edit area is unfamiliar	-n/a

Concluding, the results of the human computer interaction design criteria show that the users positively affirmed learnability/learning supportiveness as well as task appropriateness in the first session. In contrast, user satisfaction showed only average results in the first session; especially the "starting was straightforward" has been negatively assessed. Comparing the first and second session, we can find a slight positive trend regarding the assessment of the task appropriateness.

The results from the observation and screen recording can be found in the Appendix in Table B.1. Some of the given tasks also included to work with other tools than SOBOLEO. These results have been left out. In total, there were no further remarkable results from note taking. Similarly, the post-analysis of the screen recording only provided one additional issue that was a difficulty with login.

But one additional interesting observation made especially during the second session was that the chat turned out to be an essential utility for simultaneous working. For instance, two users had problems in placing concepts in the given ontology because they had only basic knowledge of the rapid prototyping domain. In consequence, they began to ask their colleagues for help via the integrated chat functionality. Nevertheless, the chat appeared to be too simple. For improvement, the users wished to have a better integration of what is discussed and where the changes are done.

7.2.3. Procedure Participatory Design and User Tests 2

Participatory design activities followed the user tests in which our application partners regularly tested the SOBOLEO system and gave us feedback either in phone calls or face-to-face meetings. During these activities, the application partners representatives also acted as mediator to make SOBOLEO known to broader audience.

In the course of the Im Wissensnetz project, the SOBOLEO system was additionally brought together with the other partners' applications to a combined tool suite. The other partners' application were:

- the e:Information Access Suite (e:IAS) by Attensity formerly empolis[5] for among others information extraction or classification of ontological entities
- the SATIN tool by Fraunhofer IGD providing desktop application integrated annotation and sharing of documents.

The "Im Wissensnetz" portal provided the entry point to any functionality and tool to support individual and collaborative innovation and research processes. Here the Sesame data store on the semantics layer, which implements semantic services for processing and interpreting contents, is the most important element for persistence and integration. It is used for the common data storage of the ontology and meta data by all applications. For instance, the information extraction component of the e:IAS provides tag recommendation for annotating web resources based on the shared ontology that is developed by the community with the SOBOLEO ontology editor. Similarly, the same ontology is used to make hand written annotation locally in MS Word.

In summer 2008, remote user tests took place under the lead of our evaluation partner Fraunhofer IGD Rostock. There, the four application partners representatives had the opportunity to use the "Im Wissensnetz" portal and combined tool suite for their daily incoming tasks at their own work place; i. e. the users could freely choose any task and execute it using the tool suite.

[5] http://www.attensity.com

By dint of questionnaires sent out end of July 2008, the users assessed the current status of software development. The questionnaires focused again on aspects of human computer interaction design; i. e. quality of system behavior getting obvious in the dialog with the users during task performance, and aspects on user support; i. e. the quality of task performance as well as user satisfaction. The detailed test materials and analysis are reported in Schulz (ed.) (project internal report).

Additionally, during July and September 2008[6], we could observe the use of the "Im Wissensnetz" portal and combined tool suite by the rapid prototyping community.

7.2.4. Results Participatory Design and User Tests 2

One activity resulting from the first user tests and the following user feedback meeting was the need to provide a seed ontology to start to work with. So the community's collaboration space was seeded with a basic rapid prototyping ontology that provided the basis for the users' activities. For instance, one application partner modified and complemented the ontology with concepts and structures related to systems engineering and ceramic materials. He collected and organized with the ontology especially relevant Web resources related to machine manufacturers (e. g., of laser sintering machines or 3D-printers), research funding, and research findings of ceramic rapid prototyping.

In our collected feedback and observations three main issues became apparent besides the seeding issue:

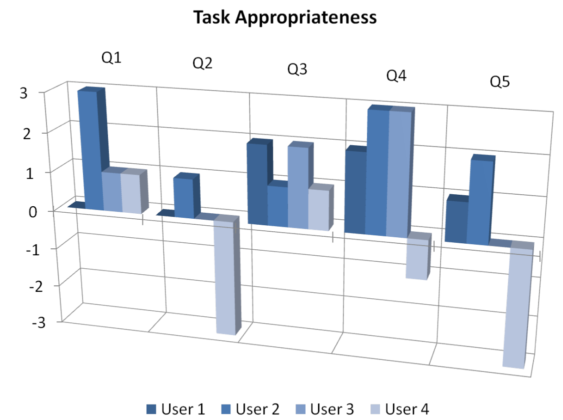
- **Information & support material:** Support material was requested that facilitates starting to work with the system.
- **Gardening support:** After a certain time, the ontology began to 'frazzle'; for instance concepts gradually collected under "prototypical concepts" which the users structured to some extent but did not sort into the overall ontology.
- **Search for people:** The users not only liked to collaboratively organize, share and retrieve documents but also people with their expertise to build up communities and to stay up-to-date on their activities.

The diagrams in Figure 7.2 show the detailed results on user satisfaction and human computer interaction design of the remote user tests. These results are not specific to SOBOLEO as the whole combined tool suite has been evaluated.

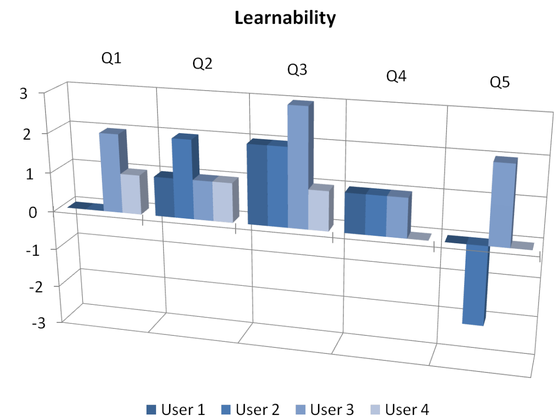
Regarding aspects of human computer interaction design, the results show that both learnability/learning supportiveness and task appropriateness were positively assessed except for one of each. Similarly, inquiry related to the general satisfaction resulted in positive statements. According to the comments the negative assessments were due to a too technically written manual on the one hand and too little adaptation to patent experts needs on the other hand. The users particularly liked the integration and combination of the different tools and systems, however criticized its late provision.

In comparison with the previous user tests, there is slight tendency to a better evaluation of learnability, task appropriateness and satisfaction. However, a comprehensive direct comparison is not feasible because of the tool integration. Nevertheless, the evaluation concluded an advancement in development because there are only very few serious problems observed.

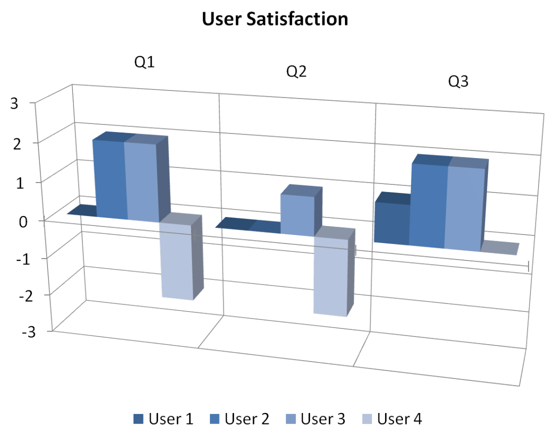
[6] August was omitted because of vacation season



(a) Q1: The system is easy to use; Q2: It provides all functionalities in order to efficiently accomplish the arising tasks; Q3: It provides good opportunities to automate repetitive process steps; Q4: It does not require redundant user input; Q5: It is tailored to the work needs.



(b) Q1: The system needs little time to learn; Q2: It encourages trying out new features; Q3: It does not require remembering many details; Q4: It is designed in a way that it is easy to memorize what the user once has learned; Q5: Learning the system does not require assistance.



(c) Q1: The tasks could be well worked on; Q2: There were no serious problems; Q3: The system was appealing.

Figure 7.2.: Questionnaire results: (a) shows the task appropriateness, (b) shows the learnability, and (c) shows the satisfaction of the remote user tests.

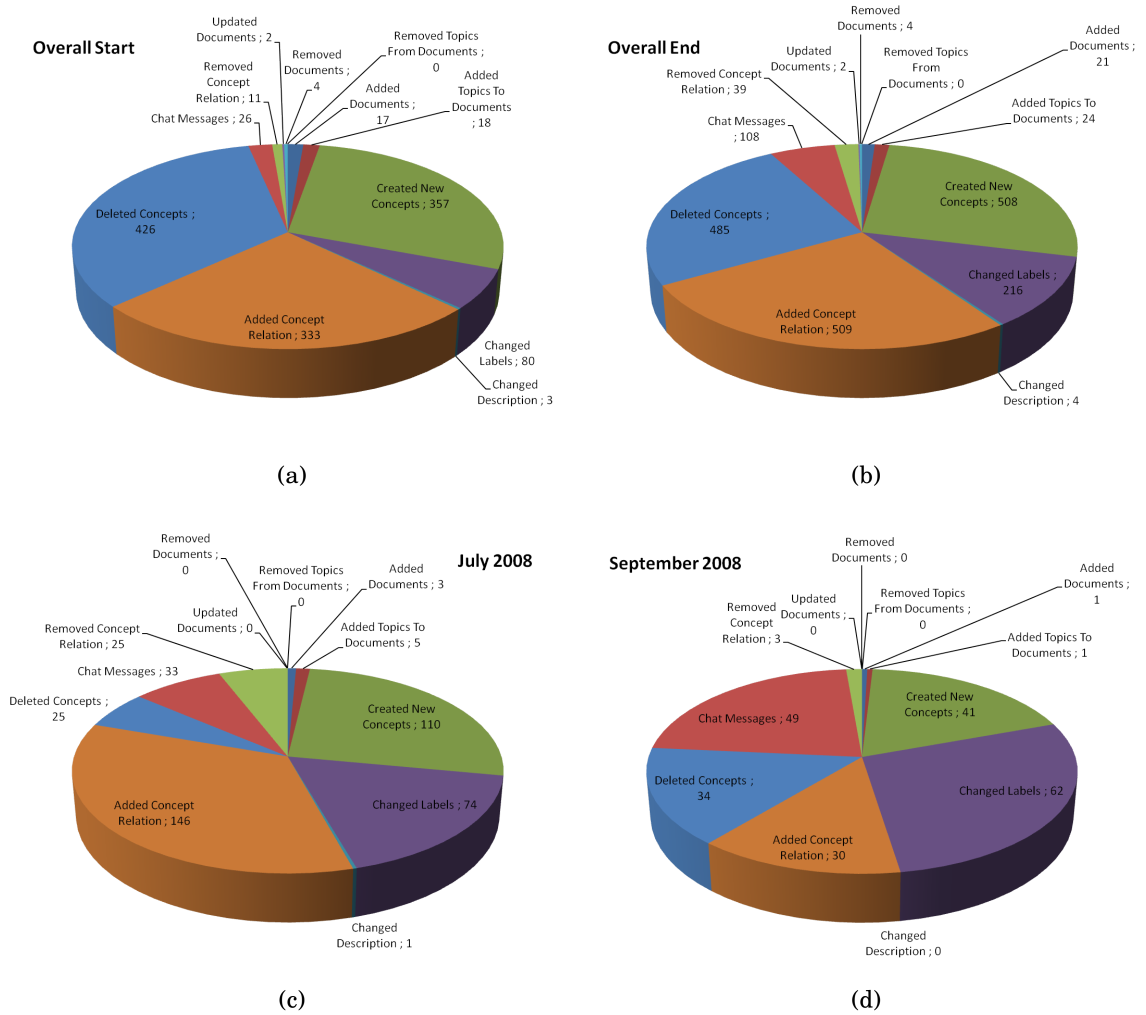


Figure 7.3.: Statistical results: (a) shows the overall activity by beginning of the oberservation and (b) the overall activity by the end of the oberservation. In comparison, (c) displys the activities that take place during July 2008 and (d) during September 2008.

Table 7.3 shows the activities of 16 users in total using the system in July and September 2008 together with the overall numbers by the beginning and end of the observation.

7.2.5. Discussion

Being the first evaluations with non-technical users, these evaluations could confirm the usefulness and validity of our approach and the SOBOLEO framework. The answers show that SOBOLEO is deemed to be suitable to structure a common terminology and to share and organize web resources with it. The results further helped us to identify fundamental weaknesses and next steps:

- **Seeding & gardening:** The need for seeding and gardening led us to the conceptual application of the Seeding, Evolutionary Growth and Reseeding theory by Fischer et al. (1994, 2001) and the conceptual and technical development of ontology maturing scaffolding methods and functions (q. v. Section 5.1).

- **Gardening the ontology:** We addressed the need for support in maintaining, pruning, and cleaning with the development of analysis services for ontology gardening activities and recommendations in the following years (q. v. section 6.4.2).

- **Information & support material:** To facilitate especially starting to work with the system, we produced manual and video materials.

- **Usability:** We addressed the issue of missing system feedback by integrating additional system dialogs, e. g., when adding a new relation would produce a cycle. To improve and support orientation and navigation in the ontology in the editor, we integrated a search functionality with additional auto-completion for ease of use (q. v. Section 6.4). Further, we addressed the difficulties of the drag'n'drop or the annotate popup hiding the text to annotate.

- **Discussion support:** The above observed development of community behaviors, i. e. align knowledge, asking for help, leadership and acceptance of guidance, and the importance of the chat as essential utility for collaborative work, emphasized the importance of the social and knowledge dimension (q. v. Section 4.3) as well as the need for a better discussion support. Thus, in a first step, we separated chat messages from automatically generated system messages in two tabs of "chat" and "log" (q. v. Section 6.4). That was followed by the integration of the concept of structured dialogs (q. v. Section 6.5).

- **Awareness & tracking changes:** In order to generate awareness, e. g., for community building, and to stay up-to-date on other people and their activities, we provided ATOM feeds for subscription. So users are automatically notified about new developments within the community (q. v. Section 6.6.3). Additionally, we integrated the display of similar people (related to activities) in the semantic browsing interface in order to make the user aware of and thus to foster communication and community building (q. v. Section 6.7.4).

- **Search for people:** This was the beginning of integrating people as "resources" and the development of the social semantic tagging framework not only limited to documents. So, the SOBOLEO framework used for the remote user tests provided not only the functionality to search for web resources but also for experts and contacts on a specific topic. The profile data are automatically gained from user activities like their annotation behavior. This search is use-integrated through is combination with the document search and makes also use of the ontology.

7.3. Case Study III: Workshop Reflections on Knowledge Modeling as a Maturing and Learning Process

In 2008, we organized the workshop "Reflections on Knowledge Modeling as a Maturing and Learning Process" as part of the PROLEARN/PALETTE/EATEL Summer School on Technology Enhanced Learning & Knowledge Management[7]. The workshop's goal was to widen the students perspective on learning; in particular to understand design processes as learning process and to experience a distributed knowledge modelling process as such a learning process. For that experience, we brought the SOBOLEO system into action.

[7] http://www.prolearn-academy.org/Events/Past%20Events/summer-school-2008

7.3.1. Procedure

The workshop took place on June 16 2008 as a half-day workshop with 24 participants of the Summer School being mostly PhD students with a mixed background in pedagogy, educational science, computer science and information systems. The workshop consisted of four sessions and began with an introduction of the topic, the participants and goals that were:

- Experience a distributed knowledge modeling process as a learning process
- Experience knowledge maturing in a collaborative tagging environment
- Get to know a tool for supporting knowledge maturing
- Identify opportunities and pitfalls of knowledge maturing
- Identify requirements for the design of tools, processes, and environments.

Then, we presented the SOBOLEO system together with its purpose as a tool for editing ontologies and collecting and annotating web resources as bookmarks that is collaboratively usable and web-based accessible. We gave a short introduction to the ontology editor, the collecting and annotating of web resources, search and navigational browsing and how to get access to the system.

This was followed by the first hands-on session of individual & collaborative tagging with SOBOLEO lasting 45 minutes. Before the trial we seeded the ontology with 15 cluster topics as seed concepts related to the domain of the participants and the resources to work with. The seed concepts were: 'Authoring', 'Collaborative Learning and CoPs', 'E-Portfolios', 'Game-based Learning', 'Information Visualization', 'Interoperability', 'Knowledge Management', 'Learning Modeling', 'LO Repositories & Infrastructure', 'Mobile Learning', 'Personalized Learning', 'Semantic Web', 'Usability', 'Web 2.0 and Social Software', and 'Workplace Learning'. We only provided the concepts with their preferred label but no further information or structure. We also kept the document space empty. We asked the participants to work with the resources from eLearning Papers[8] and the cluster topics in SOBOLEO we provided and to:

- Read and tag papers according to their interests in SOBOLEO
- Introduce new tags as they see fit
- Watch what others are doing.

During the first session we provided support with the system and arising questions.

Before the second hands-on session, we asked the participants to form groups of 3-5 persons according to thematic interests in order to cleanup their "mess", i. e. to consolidate it, in the next session. So, the participants split up into five groups.

In the second hands-on session we asked the participants to work for another 50 minutes in their groups with the previous tagged resources and concepts in SOBOLEO and a selected set of new papers and pages they had personally bookmarked but not yet in SOBOLEO. The specific tasks were to:

- Try to arrive at a shared and consolidated conceptualization (face to face or online) and input it into SOBOLEO
- Tag their selected new papers in SOBOLEO

[8] http://www.elearningpapers.eu

- Check how the shared conceptualization matches these resources
- Keep discussing and consolidating while adding new resources.

The participants were free to use additional materials such as flip chart, cards or whatever they saw to be useful. We concluded the workshop with an 50 minutes open discussion session to reflect the experience; concretely to report on the activities, opportunities and pitfalls and any suggestions for improvements and requirements for tools, services but also social processes. We collected the feedback on a flip chart.

7.3.2. Results

In total the participants created 187 concepts with 214 relations and 66 alternative labels. As the editor has not yet been multi-language supportive at this time, the alternative labels also included concept labels translations, e. g., into Spanish or Dutch. They collected 79 resources that were annotated with 3 concepts in average. Further, the participants made 92 search requests and sent 127 chat messages during the hands-on sessions. An overview of the individual activities gives Table 7.4.

During the second session, we could observe different work strategies the groups did apply. Whilst one group sat around one laptop discussing the individual steps and having one person making the modifications, another group split up the single tasks, i. e. one collecting new bookmarks, one extending the structure by creating new concept and the other making and cleaning the structure.

Similarly, one group completed the tasks step by step – e. g., first creating necessary concepts, then deleting unnecessary ones, and finally connecting the concepts – whilst the other one made modifications they just regarded necessary that moment. Another group mostly did gardening work because they felt an *"urge to bring in order"*.

Even though there was also one group mostly communicating via the integrated chat, there was not such much communication between the groups and some ignored the chat at all. So, one participant claimed that *"nobody was reading the chat!"*. Indeed, the individual groups were rather focused on consolidating "their" topic cluster. This brought up some conflicts. For instance, while one group was building up a new sub-structure, their concepts were deleted without discussion in the course of another group's gardening activities because that group simply deleted anything they found to be unrelated.

Thus, "missing awareness of discussion" and "missing rationale for concepts" were seen as the most prominent issues during the experience. It was also mentioned that it was not always easy to follow the activities in the editor. To overcome these issues, the participants suggested to introduce mechanisms for better awareness or mechanisms like locking concepts or marking them as "proposals" or "under construction", voting on the deletion of concepts as a more distributed and democratic approach or providing explanations why to add or delete concepts and relations.

Further appreciated features mentioned were a display of the history and the connection to persons; i. e. to see who added which bookmark/concept to decide if it's trust worthy. As additional observations the participants remarked different interpretations of what is "broader", "narrower", "related", and "alternative" and the use of Wikipedia as an external source for consolidation work by one group.

Despite the controversial editing activities, the participants regarded themselves to be successful in collecting and annotating resources and in collaboratively structuring the

Table 7.4.: Statistical data overview of the workshop "Reflections on Knowledge Modeling as a Maturing and Learning Process"

Target	Activity	Number
Concepts	created	187
	deleted	67
	renamed	56
	added alternative	66
	removed alternative	10
	changed description	5
Relations	added broader/narrower*	114
	removed broader/narrower	21
	added related	100
	removed related	16
Documents	added	79
	removed	9
	updated	34
Annotations	added	189
	removed	24
Misc	browsed concepts	160
	browsed key concepts	132
	performed searches	92
	sent chat messages	127

* not including automatically created relations with the container "prototypical concepts" when a new concept is used during the tagging process

domain based on the seed. Allover, the participants perceived SOBOLEO to be useful and easy to use even for those without any technical background.

7.3.3. Discussion

Concluding, the evaluation has shown that the SOBOLEO framework enables people with minimal training to collaborate in real time in the creation of an ontology and annotated web resources. The statistical data (see Table 7.4) show that all main functionalities by the system were widely used.

One major area of improvement that was identified, was that of edit conflicts and the invisibility of the rationale for particular concepts or relations. We have addressed this later through the introduction of structured dialogs (see Section 6.5). Also, we have integrated the visualization of connections between a person and the web resources s/he added on the profile pages (see Section 6.7.2). Another improvement was an enhanced history function to see all earlier changes.

The participants of the study also expressed interest in tools for voting or locking of concepts to help in resolving edit conflicts or prevent parallel changes to the same parts of the ontology. We did not create the tools to address this, since voting would slow down the edit process too much and locking risks blocking changes for too long. However, different ways to address these kinds of conflicts are included in the scaffolding methods and functions of the conceptual framework.

Another area for improvement identified by the participants was that of the visibility "who is currently editing what". While seemingly useful, this functionality could nevertheless not be integrated for time reasons.

The subgroup of overzealous gardeners, who deleted structures others were currently working out, further emphasized the importance of the social dimension (see Section 4.3) and e. g., rules governing a community's behavior (see Section 5.1.3.2).

7.4. Case Study IV: Course Experiments for the Collaborative Development of Knowledge Structures

This case study was conducted by our project partner University of Innsbruck in the context of the EU MATURE project. Its aim was to investigate requirements for the collaborative development of knowledge structures based on tool comparison and experiment and by applying the ontology maturing model. The experiment took place in the course of a seminar at the University of Innsbruck. To that end, we provided the SOBOLEO framework and technical support. The focus of use was on annotating and thus generating new concept ideas and structuring concepts with the editor. In the following, we will briefly summarize the work by our project partner. For the details of that work, we kindly refer to its publication in Gruber (2009a).

7.4.1. Procedure

The experiment was divided in three sub-experiments. The first one took place in June 2008 with 12 students within a seminar class and lasted about 30 minutes. The first sub-experiment needed to be interrupted due to technical issues. So, the second sub-experiment took place intentionally as a pretest with five university employees in October 2008. This test also lasted about 30 minutes. The last sub-experiment was conducted in November 2008 with 20 students of a seminar. Based on the experiences of the previous sub-experiments, the last experiment was adapted in its time frame and extended to two weeks. All participants were either students of or had a degree in information systems. Whilst the participants of the second sub-experiment already had experience in the area of knowledge structures, this was totally new for the participants in the first and third sub-experiment.

In every experiment, there was an introduction into the general subject and the SOBOLEO system. This was followed by the activity of collaboratively developing a shared knowledge structure. The activity was divided into three steps according to the first three phases of ontology maturing. Precisely, the test leader first presented the topic – *Web 2.0* in the first and second sub-experiment and *Event-driven Process Chains* in the last sub-experiment – for which to develop the knowledge structure. This simulated the first phase *Emergence of Ideas* of ontology maturing. In a second step, the participants started with the tagging activity; i. e. they were asked to collect bookmarks and assign tags as many as possible. This activity should simulate the second phase *Distribution in Communities* of ontology maturing. In the first two sub-experiments, the tagging activity lasted 10 minutes whilst in the third sub-experiment the participants tagged for 15 minutes during the class. The last step entailed collaboratively structuring the collected tags with the ontology editor according to the third phase *Formalization* of ontology maturing. Therefore, the participants of the first two sub-experiments had 20 minutes of time whilst the participants of the third sub-experiment were asked to work on the common structure individually from home during the following two weeks.

Subsequent to the activity, the test leader collected the participants feedback related to their experiences, usability of the tool and suggestions of improvement and general difficulties of collaboratively developing knowledge structures in an open discussion. The open discussion was audio-recorded. Additionally, the test leader analyzed the log data for the evaluation of the experiment. The log data of the first sub-experiment together with the evaluation results were afterwards made available to us.

7.4.2. Results

As mentioned above, the first and second sub-experiments needed to be interrupted due to technical problems. Nevertheless, the participants could quite easily create a shared vocabulary in every sub-experiment in a short time.

7.4.2.1. Results of Sub-Experiment 1

Table 7.5 shows an overview of the detailed activities performed in sub-experiment 1.

Summarizing, the experiment leaders came to the following conclusions for the first sub-experiment:

- SOBOLEO is easy to install and easy to use
- The drag'n'drop feature of the tree display reached its limit if there are too many new and unsorted tags
- Key concepts remained stable over time

Table 7.5.: Statistical data overview of sub-experiment 1

Target	Activity	Number
Concepts	created	100
	deleted	10
	renamed	12
	added alternative	1
	removed alternative	0
	changed description	0
Relations	added broader/narrower*	127
	removed broader/narrower	5
	added related	2
	removed related	0
Documents	added	59
	removed	2
	updated	15
Annotations	added	107
	removed	2
Misc	browsed concepts	31
	browsed key concepts	57
	performed searches	51
	sent chat messages	7

- The participants stated building a taxonomy to be easy however specifying detailed relationships to be difficult
- The participants stated building a shared knowledge structures might be easier within a community with a common interest or goal.

7.4.2.2. Results of Sub-Experiment 2

Regarding the second sub-experiment, they reported on the subsequent problems and suggestions of improvement (cf. Gruber, 2009a, pp.81).

Problems:

- The coordination was difficult when several users wanted to edit the ontology at the same time
- There was confusion when one user deleted a concept another user was just working on
- Mixing chat and log messages made it confusing
- Keeping an overview with many unstructured tags is difficult
- Copy & paste vs. cut & paste when dragging and dropping a concept

Suggestions for improvement:

- Provide a concept history for better comprehensibility
- Notifications when concepts are modified
- Display how often and by whom a concept is used

7.4.2.3. Results of Sub-Experiment 3

For the last sub-experiment the experiment leaders made the interesting observation that the participants did not hesitate to modify and organize concepts provided by others when structuring the concepts. The participants reported that existing concepts, on the one hand, were irritating when they were not created by themselves, on the other hand the existing concepts biased what they did next.

Even though the period for formalization was extended, most formalization activity took place during classes. Many concepts remained unorganized under the container concept “prototypical concepts”. The ontology itself was quite flat with a depth of two. To that end, the participants gave reasons that, besides the before mentioned usability difficulties with drag’n’drop and keeping an overview, they lacked basic knowledge on the very specific topic of event-driven process chains on the one hand, and that they could not see any personal advantage in using the system.

7.4.3. Discussion

This case study, unfortunately, was severely affected by technically problems, we could totally fix only afterwards. Nevertheless, the results provided us valuable insights. Interestingly, they confirmed results from other evaluations like the one in the rapid prototyping community taking place about the same time.

- **Towards our main principles:** (1) Especially the third sub-experiment shows, developing a shared understanding and knowledge structures are learning processes. You cannot formalize knowledge that is lacking among the participants. It similarly shows that collaboratively developing an ontology is inherently interwoven with the knowledge and social dimension. (2) Starting with tagging lowers barriers. The participants could easily start with and contribute to the shared vocabulary. (3) As stated by the participants they could fairly understand and deal with lower levels of formality and the given possibilities were seen as sufficient. (4) Tagging and developing a shared knowledge structure had no bearing with the participants actual daily work; however we see that as a prerequisite.
- **Motivation:** As a general big challenge stated by the experiment leaders for collaborative ontology development support, we conceptually and technically started to integrate explicit system feedback mechanisms as scaffolding methods to foster user contribution and motivation. In that course, mechanisms of gratitude, historical reminder and relative and social ranking are further investigated by A. Mazarakis, e. g., in Mazarakis et al. (2011).
- **Usability:** Regarding the usability issues emerged in the experiments, we improved the drag'n'drop behavior in the editor, introduced an automatic alphabetical ordering, separated chat and log messages into different panes and provided better system feedback messages.
- **Ontology overview & gardening:** Similar to the e. g., the results of evaluation in the rapid prototyping community, support in keeping an overview and gardening the ontology turned out to be an important issue. We addressed this need with the development of analysis services for ontology gardening activities and recommendations in the following years (q. v. section 6.4.2).
- **Tracking & comprehensibility of changes:** To overcome the invisibility of the rationale of changes, we introduced a history function for concepts to see earlier changes. We also addressed this later through the introduction of structured dialogs (see Section 6.5) that are attached to their subject of discourse such as concepts. Additionally, ATOM feeds enable the users to track changes within the collaboration space.

7.5. Case Study V: Course Experimental Study in Collaborative Tagging & Sensemaking

This case study, as the previous one, was conducted in the context of the EU MATURE project by the project partner Graz University of Technology. Its aim was to gain insights into basic level effects and associative tag activation in collaborative tagging in order to be able to later offer more effective tag recommendations. The investigation took place in an experimental study by dint of the SOBOLEO framework in the course of a seminar on cognitive models in technology enhanced learning at the university. To that end, we provided the SOBOLEO framework and technical support. In the following, we will briefly summarize the work by our project partner. For the details of that work, we kindly refer to its publication in Seitlinger (2009); Ley and Seitlinger (2010); Schoefegger et al. (2010); Braun et al. (2010a).

7.5.1. Procedure

The experimental study took place over a period of 10 weeks between October 2008 and January 2009 with 25 participants – 12 female and 13 male between the ages of 21 and 25. All of them were psychology students and participated for course credit. The participants were equally divided into four groups of 6 or 7 based on their knowledge on the given topic, computer literacy and attitude towards using computers for communication. Each group had to research together a topic related to their course subject and to that end collaboratively collect and tag web resources. The topics to research on were 1) "the use of Wikis in enterprises" and 2) the use of Weblogs in universities" – each to be worked on by two groups. Each group had their own SOBOLEO collaboration space not accessible by the others.

During first class, the participants were introduced into the SOBOLEO system by a guided walkthrough and a small exercise to be done by the participants in pairs of two. After class, group assignment took place and the participants received an email with the topic to work on and further task instructions. Thus, the participants were requested to collect and tag at least two relevant web resources per week with the SOBOLEO tools. They were also expected to collaboratively create an ontology that is, in the way of a collaborative sensemaking, to be discussed using the internal chat or external discussion forum to develop a shared understanding. They were also informed that their activities were regularly evaluated by using the SOBOLEO log data.

The course continued with weekly classes in which the participants received advice and an update about activities in SOBOLEO. After five weeks of study, on December 1 2008, two groups switched the topic to work on; i. e. one group from topic 1 to topic 2 and vice versa. The hypothesis was: *"that groups that had worked for the whole duration [..] would form a stronger representation in memory of the more specific tags and that they would rate their relevance higher than the groups that had worked for only half the duration on their topic."* (Ley and Seitlinger 2010) Therefore, their collaboration spaces were cleared so that they started again from scratch and the experiment leaders sent around another email with reinforcing the instructions.

Before the topic switch and at the end of the study, word-association-tests[9] and relevance rating tests[10] were conducted to the study's purpose.

The study was concluded by a post-questionnaire with five five-point Likert scale questions on the own understanding and satisfaction with SOBOLEO, the editor, the group communication, and the created ontology and an open discussion to collect the participants feedback and suggestions for improvement.

[9]Word-association-tests are used to collect the number of associations a certain tag induces. This gives insights about a the knowledge on certain concepts. Here, tags from the participants' own ontology were used.

[10]With that test the participants had to rate the importance of individual tags from their own ontology for fulfilling their task; i. e. describing and organizing bookmarks, on a five-point Likert scale ranging from strongly important to strongly unimportant.

7.5.2. Results

Unfortunately, the log data and questionnaire results could not be delivered to us for own analysis after finishing the study. So we have to draw back on the experiment leaders' statements and interpretations.

In total, the participants created 213 distinct tags within their groups and collected 238 web resources. Interestingly, the questionnaire and open discussion revealed that the groups that worked for the whole period of eleven weeks on the same topic (11w groups) rated their understanding for the topic lower than the two groups that worked only five weeks on the same topic (5w groups). Furthermore, against to the experiment leaders' expectations, the eleven-weeks-groups assessed the quality of their collaboratively developed ontology lower than the other two groups.

The participants of the 11w groups explained their lower rating of the achieved understanding by the complexity of the ontology that hampered retrieval of relevant information. So, they limited themselves to own resources. Similarly, they explained their lower rating of their ontology with that there were quite redundant tags in ontology that resulted in a loss of overview.

In total, the assessment of the SOBOLEO system's usability was neutral, however according to the free text answers they felt not sufficiently supported. Some had problems to access their collaboration space from home due to the server installation and network configuration provided by the Technical University of Graz. In detail, regarding the editor part and the level of formality, the participants mostly agreed that the provided structuring with broader, narrower and related relations are sufficient. Concerning group communication, the participants were dissatisfied with the communication mechanism. The SOBOLEO chat was unsatisfiable because it was not persistent and synchronous editing happened only sometimes. Thus, it could only rarely be used. The forum was perceived as being cumbersome because it was not work-integrated. Therefore, the groups were unsatisfied with the few discussions. This fact also resulted in a missing group identity. The participants indicated that they felt working alone on the topic and ontology and not being part of a group. At this point, the participants criticized that they were assigned to a group and could not form themselves.

Another interesting observation was that whilst the 5w groups were forced to actively deal with and reflect about the others' contributions because they had to develop a completely new ontology after five weeks, the 11w groups were less active in the second half of the study and only added new tags but modified others' tags. They indicated that the hitherto developed ontology seemed to be enough for the object of exercise and communication issues were hampering. Additionally, the 11w groups reported on motivational issues to be concerned with the same topic the whole time.

At the end, the experiment leaders concluded the following implications and suggestions for improvement. For consensus finding, chat functionality is necessary that is integrated into the development environment and persistent. Discussions and modifications need to be traceable and comprehensible to understand the rationale behind. Awareness features may further help, e. g., who is online and available for discussions or what is the status of a concept or also suggested by the participants was a rating feature for tags to indicate their importance for the ontology. Support activities are advisable. For instance, regular group discussions might help to overcome a cluttering ontology and decrease in motivation and stimulate gardening activities. Similarly, it is necessary that rules of collaboration are developed.

7.5.3. Discussion

This case study provided in some ways quite similar results as the previous ones. Identifying social dimension, discussion support, trackability & comprehensibility of changes, and ontology gardening as the main issues:

- **Social dimension:** Again, this evaluation emphasizes the importance of our model's social dimension. A certain degree of group identity is a prerequisite and it is necessary that the community itself can evolve and develop rules and norms. Therefore, this also plays an important role in our scaffolding methods and social semantic tagging design framework.
- **Discussion support:** Discussions form an integral part for consensus building. This study has shown that it is not sufficient to use any discussion tool but functionalities need to be integrated into the work environment.
- **Trackability & comprehensibility of changes:** We already detailed in the previous study how we addressed these issues.
- **Ontology gardening & reseeding:** Similarly, we addressed this in the previous study. Interesting to see, that "discarding everything and starting newly" might be used as very radical form of reseeding. Unfortunately, our resources were too limited to investigate this as scaffolding method in depth. But we could integrate the participants' suggestion to provide awareness on concept status as one type of ontology gardening recommendations.

7.6. Case Study VI: The Classic Roman Civil Engineering Community of Practice – Dialogs for Developing a Shared Understanding

The objective of this evaluation, which took place in the context of the MATURE EU project, was to support a community of practice (CoP) in collaboratively developing its understanding of a domain through interweaving the development of a shared information repository and vocabulary (ontology) with dialogs (see Section 6.5) about them.

Our community was a community of practice for Classic Roman Civil Engineering newly built up with alumni students of an eLearning course on this topic by the Spanish eLearning provider Structuralia. Structuralia offers its clients eLearning courses primarily in the construction sector. Apart from the courses themselves, they also offer individual learning solutions to their clients, including advice as to which of their courses may be relevant for them. Their past experience suggests that (virtual or non-virtual) courses can be a good platform for bringing together people within an organization who have common interests and may continue sharing knowledge afterwards, thus forming new communities.

The aim was to support the CoP in developing a shared understanding of their domain by interweaving the development of a shared information repository and vocabulary (ontology) and dialogs about them. That means the CoP collects and bookmarks web resources around their domain and builds up the common multilingual vocabulary (ontology) – as there is a lot of information in different languages relating to Classic Roman Civil Engineering – which is used to organize the web resources 'in action' through annotating them during the bookmarking process. Structured dialogs are used to: discuss and refine the ontology; critically discuss and assess the (controversial) resources (e. g., reflecting and debating the correctness and quality) – especially information in the Internet about Classic Roman Civil Engineering is often erroneous; and provide a record of dialog performance that is linked to the related content and thus accessible and searchable.

This illustrates how to more closely linking critical dialogs to their semantic implications for community knowledge management within a particularly controversial domain, and where these dialogs are themselves considered as a form of knowledge. Aligning dialog and semantics addresses a number of important research challenges. These include how to improve the understanding, maintenance and application of ontological knowledge structures to contextualized problems within work-based communities.

The implementation and evaluation with Structuralia and the CoP was divided in two main phases of formative evaluation and participatory design activities. In the next section we describe the first phase of formative evaluation and participatory design, where the emphasis was on working with the user community to refine designs, usage scenarios and design-context fits in general. Then we present the second phase of formative evaluation that trialed the approach within an authentic context of use to: establish its usability; suitability and value in the context of use; and, potential value in terms of supporting ontology maturing.

7.6.1. Procedure Phase 1 Formative Evaluation and Participatory Design Activities

The first phase of formative evaluation from July 2009 to January 2010 comprised four steps of participatory design activities:

1. **Design workshop with application partners:** The initial design workshop took place in a face-to-face meeting in Madrid on July 16-17 2009 with six user representatives participating (from Structuralia, and one of their specialist contacts). User feedback was collected about SOBOLEO, InterLoc and their proposed integration through exploratory discussions and structured walkthrough sessions. Notes were taken and details documented throughout. These sessions involved: context setting and brainstorming possibilities; demonstration of the software and explanations of their underlying rationale; follow-on brainstorming and proposing candidate technology-scenario setups; and, post-hoc consideration (after the meeting) and selection of the development-scenario proposition.

2. **Internal informal evaluation:** Regular informal discussions and walkthroughs involving the design team took place in November 2009, which culminated in the production and testing of video tutorials as preparatory material for evaluation for the initial evaluation with Structuralia members and their contacts.

3. **Initial evaluation with application partners:** For the initial evaluation of the envisioned design an integrated version of the tools InterLoc & SOBOLEO was provided together with video tutorials as preparatory material to test the mashup in an unsupervised manner. These were tested by two Structuralia personnel covering research and technical aspects, and the results of testing were discussed via email and in an additional evaluation session and a walkthrough which took place on December 17 2009 with the same members of Structuralia. The session took place online in a telephone conference using screen sharing support. Besides audio recording, we made notes to gather the users' feedback.

4. **Conceptual validation:** the conceptual verification was performed, which checked and aligned the approach against the Ontology Maturing Model – especially its phases (see also Section 6.8.1) – and in its broader context against the Knowledge Maturing Model. This has also been reported in Bradley et al. (2010) and published in Ravenscroft et al. (2010b).

7.6.2. Results Phase 1 Formative Evaluation and Participatory Design Activities

Here we summarize the findings from the phase 1 formative evaluation and participatory design activities. With the design workshop with application partners, in July 2009, we were successful in approving the tools' suitability, collaboratively developing the implementation setting for a community of practice of Classic Roman Civil Engineering and gathering additional requirements for the adaptation to the new setting.

At the end, we elaborated the scenario of implementing the approach for a specially initiated CoP of Classic Roman Civil Engineering. This was with the aim of giving support to the community that collaboratively develops a shared understanding of this domain by collecting and discussing (controversial) information and developing a common multilingual vocabulary. We identified four different cases for dialog support:

1. the ontology and its development and gardening process,
2. the resources (e. g., reflecting and debating the correctness and quality),
3. the resource classification (according to the ontology),
4. the topic in general (e. g., initiated by a teacher's question).

Based on that first design workshop, the following technical adaptations were necessary:

- Developing structured dialogs targeted to specifically support the above identified cases; i. e. an ontology maturing dialog (OM-D) to support discussions about ontology development and an critical discussion and reasoning dialog (CDR-D) aimed at reasoned discussion about a knowledge domain or particular perspectives and resources related to it
- Integrating our framework with InterLoc; i. e. providing webservice interface for remote applications
- Providing the users the opportunity to start dialogs from various points within SOBOLEO
- Providing user notifications for newly started dialogs
- Listing of dialogs and their content
- Transcript storage for providing dialogs as new resource
- Linkage of dialogs to related resources and concepts.

The feedback of the subsequent activities showed that the respective dialog and semantic technologies successfully combine to provide new and tangible informal learning and ontology maturing activities. The final prototype was mostly acceptable to users and addressed the requirements for greater and more 'critical' informal learning and knowledge management facilities. Structuralia users were confident of its applicability in the community of practice for the formative evaluation Phase 2. Finally, all the desired improvements were listed and prioritized.

For a successful implementation in the CoP and Phase 2 Formative Evaluation it was crucial to provide a system in the Spanish language so that it could be used by the participants based in Spain. Therefore the internationalization of the tools was introduced, which involved translating all linguistic features from English to Spanish.

Additional requested improvements were (with prioritization):

- Improved labeling and translation of user interface elements (mandatory)
- User guidelines (mandatory)
- Orchestrating the interaction design as a 'game' (mandatory)
- Overview of activities, especially changes to the ontology (high)
- Search for dialogs equivalent to annotated documents (high)
- Voting option on changes at the end of a dialog (medium)
- Automatic launch of InterLoc (medium)
- Running on Windows XP systems (there might be difficulties with Sun Java VM) (to be considered for bigger roll out)
- Restoring of previous ontology version (low; instead: use not-agreed changes as trigger to start maturing dialogs about).

Also important was a careful introduction as a learning game experience to the moderator and members of the community. Therefore, it was also necessary to elaborate the experience with a set of rules and guidelines for how to use and interact with the system, as a starting point to support the community building process. Guidelines such as, for example, that everybody can make changes but that it is also necessary to achieve an agreement about changes and therefore should be discussed. Or if the moderator (or someone else) detects faults or ambiguities etc. in changes, he/she might start a dialog to achieve a better understanding.

Regarding motivational aspects, we elaborated on the questions how to additionally motivate the community to use the tools and how to make them discuss, because the community was newly created and there haven't been any community rules and practices established, yet (which are indeed essential; see also the theory of CoPs by, e. g., Wenger (1999); Wenger et al. (2002)). Therefore we proposed:

- Regular mail on activities
- High quality newsletter by trainer/moderator
- Joint community events
- 'Meet the trainer' sessions
- Scheduling/inviting to dialog sessions.

7.6.3. Procedure Phase 2 Formative Evaluation

Based on the results from the first evaluation phase, we finished the development of the prototype and started with phase 2 of formative evaluation. The goal of the phase 2 was the investigation of the use in the user context and of ontology maturing with the key aims:

1. To user-test and refine the approach of collaboratively developing an understanding of a domain through interweaving the development of a shared information repository and vocabulary (ontology) with dialogs about them with the Structuralia team
2. To deploy the system within a relatively small-scale but authentic CoP of alumni students of Classic Roman Civil Engineering
3. To investigate the performance in terms of:
   a) potential value to the CoP
   b) acceptability and usability
   c) The degree to which it supported the collaborative development of understanding and ontology maturing
   d) Other emergent insights relevant to the design, development and deployment of socio-technical systems to support learning (informal and formal) and ontology maturing.

We addressed this with a pilot learning experience in a Spanish speaking community of practice of alumni elearning course students at Structuralia's site that is built up with this experience. Within the learning experiences, the students are provided specific topics once a week to be elaborated and discussed, i. e. collecting and critically discussing information (web pages) around the topic and structuring the topic.

The experience took place as an online-only experience from May 11 to June 8 2010. Structuralia contacted 15 people from two courses and invited them to participate in this experience. We could recruit 10 participants aged between 45 and 55, with an average age of 49, representing 65% of the programme. All of these were practising Industry Professionals, including Engineers, Architects and Industrial Engineers. The experience was moderated and guided by two representatives from Structuralia.

Before the official run, the group of participants were contacted to present the Guide of Use and Management of InterLoc and SOBOLEO, providing personalized support and access to demo and test sessions in both environments. Structuralia teams prepared several forms of input and four dialogs and discussions, with the idea of serving as incentives for students.

The first day of the experience a welcome message was sent to the students involved to remind them of experience, inviting them to connect and familiarize themselves with the software. To facilitate this access, the Structuralia virtual training platform provided a specific and direct access to the initiative. The virtual training platform of Structuralia is a familiar and trusted tool for students, hence the idea of using this route as a gateway to the experience. In this virtual environment presentations and support systems were set to better guide students. From the Structuralia platform, with a single click, the student could get into SOBOLEO, and from SOBOLEO to InterLoc.

Once a week, they provided selected topics to the vocabulary via the SOBOLEO editor and initiated dialogs related to these topics in order to animate and facilitate the experience. The experience started with six participants; four participants joined the experience later. Before the experience the participants received an invitation letter and platform access, the user manual and additional informative material. During the experience Structuralia provided support by email and phone. The developer team provided technical support that was mediated by Structuralia. There was no direct contact between the developers and the participants.

During the experience we used diary keeping about the activity, issues and user engaging process, and automatic system logging. At the end of the experience semi-structured interviews with the participants took place by phone based on an interview guideline to collect the feedback and jointly reflect on the positive and negative aspects identified. Additionally, there was a face-to-face debriefing with the application partner that lasted about 45 minutes that were audio recorded.

Previous to the experience, we conducted a case testing stage with the application partner team in order to further refine the approach. Feedback of the case testing was collected in the form of notes and an excel sheet and in an additional in-depth online discussions with additional note taking. The development team took these up and refined the tools and additional material. Together with Structuralia, last refinements of the experience planning and design were made.

The Road Map for this complete formative evaluation process covering translation, initial case-testing, and student experience and evaluation is given in Figure 7.4. During the deployment and evaluation process several complications arose because: the application partner site lost their two main contacts consecutively until a third successfully concluded the evaluation; initially the software was in English and had to be translated into Spanish depending on the application partner; additional clarifications, translations and evaluation steps were needed so that the English speaking members of the evaluation team understood and interpreted an experience and evaluation performed in Spanish.

Below we summarize the complete testing and user experience, based on the road map, and then present further details of the user experience in terms of the instantiations of SOBOLEO and the performed dialog through InterLoc. These user activities are then interpreted to explain how the new socio-technical activities that were supported represented informal learning and ontology maturing in connection with their CoP related to Classic Roman Civil Engineering.

7.6.4. Results Phase 2 Formative Evaluation

7.6.4.1. Results of the Translation and Case-testing

Prior to the experience, the work focused on preparing the translation of environments, InterLoc and SOBOLEO, into Spanish, while trying to preserve most features of both software tools. In parallel, unit testing was carried out to ensure the proper functioning of all the properties of the tools (functional and transactional).

The setting in the Spanish CoP brought up two additional requirements for a successful experience: (1) supporting a multilingual vocabulary because there are many web pages in different languages relating to Classic Roman Civil Engineering, that should be approached when studying some topics and (2) internationalization of the tools for Spanish users, i. e. translation of all interface features from English to Spanish. Thus we approached these two requirements in four steps:

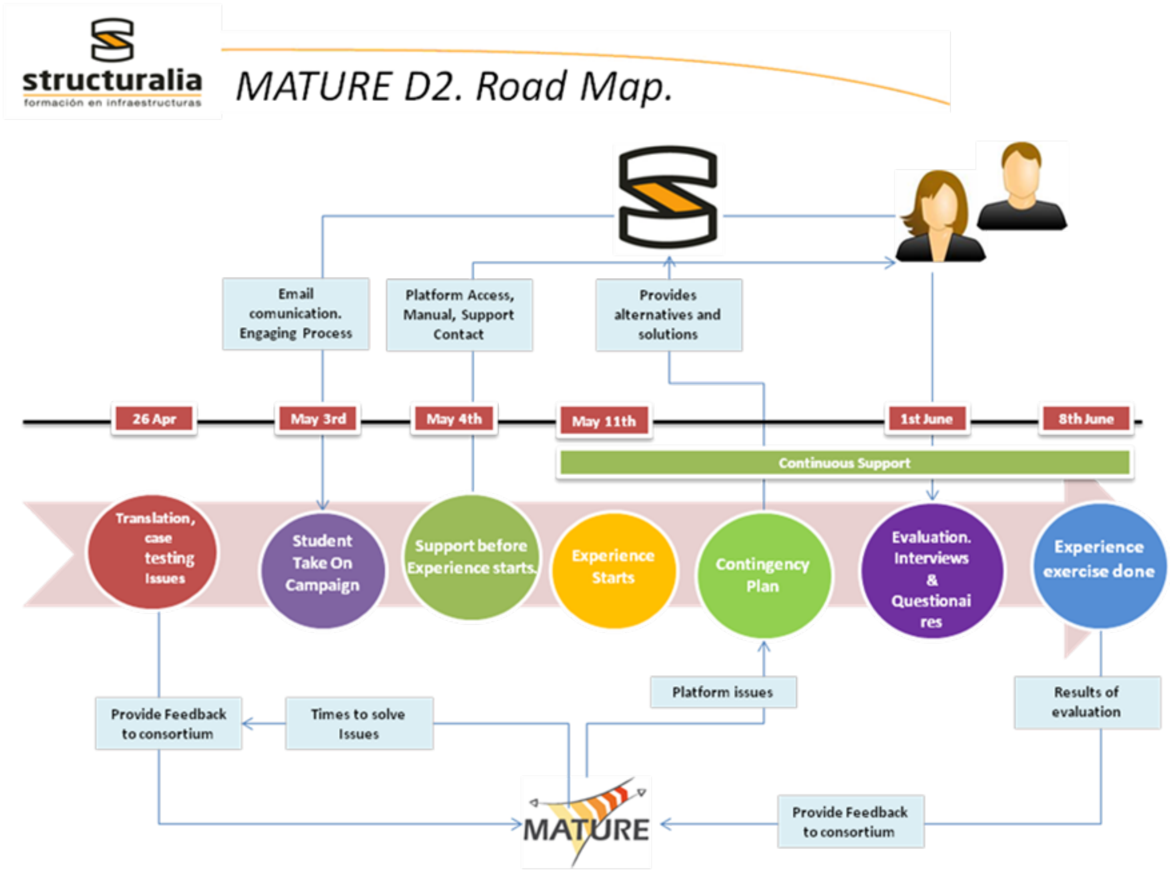


Figure 7.4.: Road map Phase 2 Formative Evaluation

1. Multilinguality of the vocabulary: the SOBOLEO back end as well as the front end needed to be extended and re-engineered to support multilinguality of the vocabulary.

2. Translation of the preparatory material: The User manual was also translated by the Structuralia's team and handed out to each participant prior to the experience.

3. Translation of the user interface: the user interface of SOBOLEO and InterLoc were translated (see Appendix 2).

4. Translation of the dialog features: the dialog features needed to be translated and adapted (see Appendix 2).

The Translation exercise was successfully completed, with the help of Structuralia staff and the development team working together to translate all the linguistic components of the interface from English to Spanish.

Technical issues were addressed relating to tagging documents, adding relationships and the display of the Moves and Openers. Due to technical and time constraints, it was not possible to provide a single sign-on for both tools, but access was facilitated by having consistent logins for both users. We amended the details of the experience design and agreed on having at least three terms provided by Structuralia each week as stimulus for the experience. And, user engagement issues were addressed related to motivation and reward for participation in the activity, through active promotion of the activity. It was agreed certificates would be distributed in recognition of people's participation and an additional gift in order to incentivize the students and increase their engagement.

7.6.4.2. Results of the User Experience

Between May 11 and June 8 2010, so for approximately one month, the users performed their live experience, in order to collaboratively develop their understanding of Classic Roman Civil Engineering. The detail of what they did has been recorded in a User Diary (for details of the user diary see Ravenscroft et al. (2010a)).

Since the inception of the initiative, the Structuralia team provided terms to SOBOLEO and seeded debates in InterLoc. Initially 7 terms were added, which were "Acueducto, Calzada, Faro, Mina, Presa, Puente y Viaducto" and 4 debates were seeded in InterLoc, about "Red de Alcantarillado", "Faros desaparecidos", "Caminos a Roma" and "Presa". The Structuralia team, in order to monitor the progress of the users work and the depth and quality of the knowledge shared, intervened intensely in cases of low involvement by the students, posing suggestions for participation. Two daily connections were scheduled to control and streamline the work of students, providing personalized telephone support for students who needed help or guidance.

During the first week of their experience, although 5 students connected to the experience, only 4 of them made contributions. Participant I had 10 interventions on InterLoc´s dialogs, participant II 8 interventions, participant III 1 intervention, and participant IV created 2 new terms in SOBOLEO, "Puerto" and "Acueductos romanos". The Structuralia team added 4 new terms on SOBOLEO "Noria, Pozo, Galería, Oficio" and 2 dialogs in InterLoc "Oficios" and "Explotaciones mineras". The most interesting interaction that week was the involvement of students in dialogs with discussions between them on issues such as "Mina" or "Caminos a Roma".

During the second week of the experience 8 students were connected, although only 3 of them made contributions. Participant II with 6 InterLoc's dialogs; participant III with 2 InterLoc's dialogs and participant IV with 1 new term on SOBOLEO "Evolución histórica". Structuralia team had 4 new terms on SOBOLEO "Teatro, Anfiteatro, Minas de Río Tinto, Rueda de Cangilones"; 1 dialog "Teatro" and 1 web page (Tarraconensis). This week, there was a special student interest in advancing participant IV's ontologies, being the only student who entered SOBOLEO terms and descriptions.

During the third week the group experience continued with 8 students connected, 3 of them working and 2 not connected. The activity was led by one participant with 10 InterLoc's dialogs, a second participant with 3 InterLoc's interventions, and a third participant with 1 new term on SOBOLEO, "Distribución functional". The Structuralia team included 6 new terms on SOBOLEO "Velarium, Necrópolis, Mausoleo, Terma, Templo, Panteón" and 2 web pages (El tablero de piedra and Hipótesis de la exitencia de un teatro romano en Palma de Mallorca). In this week's rankings were the debate about the talks related to "Red de Alcantarillado" and "Explotaciones mineras". It is the week with more information on both tools. Unfortunately, there were no marks in the "People".

During the final week of their experience the group 8 students continued connected, only 2 of them working: one with 4 InterLoc's dialogs and the other with 1 InterLoc's dialog and 2 web pages. The Structuralia team added 3 web pages (Faros romanos en Hispania, Ingenie).

In total the key topics were accessed 242 times by the participants and moderator. During the experience, we lost two of the participants. According to the log data, all of the others entered SOBOLEO and browsed through the topics (e. g., Faro or Acueducto) and associated information. Three tags – Mina, Acueducto, Calzada – were accessed more than 20 times and six were accessed more than ten times – being Presa, Puerto,

Figure 7.5.: Case Study VI: Key topics and web documents created during the experience

Puente, Faro, Teatro, Antiteatro. The participants only added and annotated nine web documents. Figure 7.5 shows the key topics and added web documents. The editor was accessed 102 times by five participants but only three of them made any changes. In total 53 concepts were created and 17 deleted. The participants added 36 descriptions to these concepts of which nine were removed. They also added five alternative labels and changed four preferred labels (due to spelling issues). The participants structured the concepts with 58 broader, narrower and related relations of which 14 were removed again. A screenshot of the concepts created in the editor shows Figure 7.5.

In total, seven dialogs were performed. All of them were initiated by the moderator. The seven mostly asynchronous dialogs were performed about the following topics, that have been translated from Spanish: Dam, All Roads Lead to Rome, Missing Lighthouses, Sewerage Network, Mining, Crafts and Theatre. Three to five people participated in each dialog that was seeded and facilitated by the moderator, who was particularly active in animating all of the dialogs. An example dialog is given in detail in the Appendix C.1.

In the daily monitoring of the experience one of the objectives of the Structuralia team was to quickly detect any problem or malfunction of the systems, to minimize the impact on students. This work was intense over the initiative and dysfunctions contingencies arose that affected the smooth running of the experience. Through a scorecard these incidents were monitored and the resolution capability of the equipment involved in the tools was logged. The inability of InterLoc to be 100% available relegated the involvement of students, who found dialog more intuitive than ontology manipulations. In each tool, new entries or new contributions were recorded daily, identifying the author or authors, with the goal of a comprehensive control of the students' work. At all times the Structuralia team contributed to the system in an attempt to provide incentives and give some traction to the participation of students.

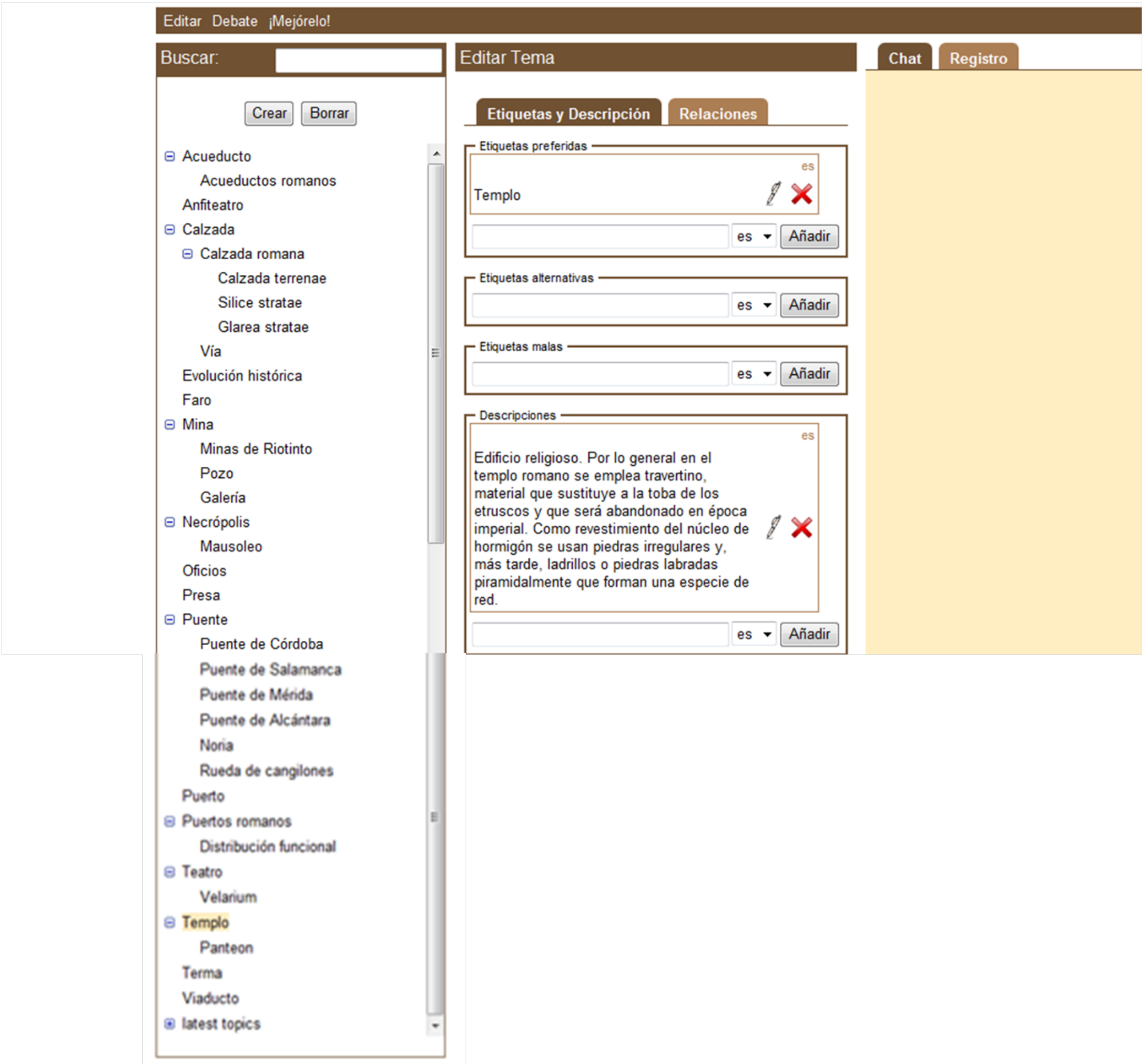


Figure 7.6.: Case Study VI: Concepts created during the experience with the SOBOLEO editor

7.6.4.3. Results from the Facilitators

In this Section we present an interpretation of the whole experience, from the initial inviting of the participants and their subsequent activities, that was provided by the two Structuralia staff facilitating the activity. According to observations that were recorded and reported by the Structuralia facilitators the experience as a whole was positive, especially for those who participated more actively with the tools. This was demonstrated by:

- The tools were key to the users in order to understand and develop their capabilities. Either alone or both combined (SOBOLEO & InterLoc).
- All active participation during the experience had been related to the actions done for setting up the tagging environment in SOBOLEO. Furthermore, none of the existing users have been able to create a dialog in InterLoc. The students worked with the dialogs initially created by Structuralia.
- Average age seems to be pretty high. As a technology barrier we need to consider the students did not have very developed computer skills. This kind of users found some technical problems which ended up in two of them withdrawing.
- Some of them experienced these technical issues and did not retry, did not dig into both application functionalities.
- The users' participation was low. Efforts at getting more work in SOBOLEO were made. Varying degrees of participation by teaching staff may have made a difference in this regard.
- The dialog moves & openers were very closely coupled with SOBOLEO functionalities. This also limited the possibilities of dialog and with this the experience of structured learning dialogs; except a few found that it worked well (create, question, affirm, challenge, Request for Information). This might be traced back to the fact that in the early phases of building a basic shared knowledge and understanding, more general moves & openers are needed; whereas the very specific moves & openers with their tight coupling are useful in later stages.
- An improvement that several students have pointed out is the ability to insert images. On the specific issue it would have been useful to represent many of the terms.
- It would have been interesting to have the opportunity to close some of the dialogs. Especially those that were completed.
- The functions around "People" were not used. However, this wasn't in the experience's focus and not explicitly introduced to the participants. As such, functions and information around people seem to be a new, not self-explanatory concept for the participants.

7.6.4.4. Results of Interviews

This Section describes interviews with six participants who completed the experience and also with one of the Structuralia facilitators. We conducted semi-structured interviews with six participants via telephone within a week of the experience. Interviews were based on a template. The template and answers can be found in the Appendix C.2.

Interviews with user participants The general assessment of the experience is positive: *"Positive. Working with tools has been very satisfying for me"*. The participants liked the collaborative work that the tools enable:

- *"The tools enable collaborative work that is very useful for gaining knowledge, applying to practice and comparing what you've learned. Let's keep alive curiosity about a subject."*
- *"Being immersed in this type of initiative makes you active. Knowing that your contributions will reach out to others, forces you to review, to deepen, to better prepare your way of speaking."*
- *"Surely you can increase knowledge. Sharing links, important information, saves much work and improves outcomes."*

On the downside, users mentioned the short duration of the experience and the technical difficulties that arose during the trial. One participant also indicated that he sometimes feared editing colleagues' work because s/he didn't want to upset them. This might be traced back to the users' unfamiliarity with the Web 2.0 culture where anybody can contribute and make changes. This was also emphasized by additional remarks on the importance of technical support.

Regarding the more specific feedback questions, the interviewees could only provide limited feedback on how the experience helped them and the group to improve knowledge and understanding. Regarding the tools usage, most of the participants used the SOBOLEO system as consumer; only one interviewee added new web pages and another one added information to the vocabulary. Four of the interviewees participated actively in discussions. As criteria for participation the confidence with the topic was mentioned: *"I've contributed topics and comments on those areas where I felt comfortable, with more knowledge."* The interviewees mostly participated after work; when they could find some free time.

Debriefing Interview with Facilitator A key application partner representative was interviewed, face-to-face, by a member of the LTRI Evaluation team approximately one week after the experience had finished. The representative was satisfied with the level of user engagement with the SOBOLEO-InterLoc ecosystem, but accepted that it would have been better to have more engagement than had occurred. There were daily posts and typically interaction of between 20-25 minutes/day for each student.

The level of recruitment was positive, as the 10 participants represented 65% of the Programme. The representative emphasised the challenge of explaining what MATURE project context and this approach of dialogs for collaboratively developing shared understanding 'was about' to these users, but once they started to engage he felt that their understanding got better, and they did mature their knowledge of the topic. But in particular, there was a barrier in explaining Web 2.0 concepts to those over 50.

The technical problems that were experienced in the early stages (i. e. mostly within the first week) related to the running of the adapted InterLoc application due to unanticipated crashing caused by 'special characters' in Spanish inevitably affected motivation and participation. Also, there was a brief period when the InterLoc server was down due to a 'freak' power failure in the area of London where the server was located. However, once the service was fixed and stabilized, most students regained their confidence with just 2 dropping out. So, as far as the total experience was concerned, there was a week with some problems and disruptions followed by smoother operation for about three weeks.

As mentioned above, there was less input from the students than expected, and it was recognized that more communication would have been advisable to contextualize and support the activities.

A significant barrier was the need to use three sign-ins, i. e. separate sign-ins to the Structuralia LMS, SOBOLEO and MATURE-InterLoc. There was no activity during weekends, but approximately 5 connections every day for 10 days. A common interaction would involve posting to the dialogs and tagging a word in SOBOLEO. Of particular importance to the experience was the animateur and facilitator role performed by the second Structuralia representative, who sometimes phoned participants to communicate with them directly. But unfortunately, due to resource limitations, it was not possible to organize and host scheduled 'Discussions with the Expert' as planned.

7.6.5. Discussion

Interpretation of the Dialogs Summarizing interpretations from all of the performed dialogs showed the following. Firstly, the translation of the structured dialog approach from English to Spanish seems to have performed quite well. Most of the Moves and Openers that were selected were used legitimately with the exception of a few of the maturing ones, such as 'Create' and 'Modify'. Accepting this is a small-scale study, this finding is particularly promising as it suggests that the deep pragmatic level structure of dialog should translate across different languages, from English to Spanish in this case.

Secondly, most of dialogs were of reasonable length, ranging from 8 to 17 considered contributions with just one unusually short dialog of 3 contributions. Also each contribution typically consisted of one or two sentences. Again this was promising, as the structured dialog approach promotes dialogs that involve contributions that approximate to a 'unit of thought', which typically maps to one or two sentences.

Thirdly, the Replies and Contributions were nicely linked, demonstrating the desired 'interthinking' on a topic. This dialog and thinking was shown to be of a collaborative knowledge building and critical inquiry in nature (e. g., Can you tell me, Let me explain, Is it the case), as is typical during early stages of InterLoc use. However, the users did not use the ontological classification dialog moves, or used them incorrectly. This is not unsurprising given the limited time of the trial (less than a month) and the somewhat less tangible nature of this form of activity, i. e. whereas critical discussion is a relatively natural activity, ontological classification is less so, even whilst adopting a semi-natural dialog interlingua. Although it was hoped that the use of the ontological classification moves would develop over time, towards the end of the trial, as users began to understand the correspondence with SOBOLEO operations, this was not found to be the case.

Fourthly, a legitimate and good range of Moves and Openers were used in performing these collaborative and critical inquiries. These included Moves such as Assertion, Info-Request and Question and Openers such as I think. . . , I've read that. . . , Let me explain. . . , Is it the case that. . . , And if. . . , This means. . . and Not so Sure. . . The sample of the dialogs showed that InterLoc supported coherent exploratory and inquiry dialogs about the topics, characterized by the offering of opinions (e. g., I think. . .), justifying of positions (e. g., I read that. . .), questioning (Is it the case that. . .) and the generation of implications (e. g., This means. . .).

Fifthly, in terms of knowledge maturing, these dialogs supported the features and phases of maturing that are specific to the designed dialog (Ravenscroft et al., 2010b). However, they did not demonstrate the use for maturing the ontologies that were related to it what might be traced back to the very close coupling with SOBOLEO functionalities. The performed dialogs implicitly showed Expression of Ideas about the domain (Phase 1), and Distributing in Communities (Phase 2), e. g., through collaborative discussing topics and then storing the dialogs in SOBOLEO, and clearly promoted the formalization (Phase 3) of dialog.

Usability and the User Experience The summary and implications arising from this formative evaluation can be broken down into: (a) usability and the user experience; (b) knowledge maturing and the collaborative development of understanding; and (c) implications for ongoing and future socio-technical developments.

The experience as a whole has been considered positive both by our application partner and the experience participants. The participants liked the collaborative work that the tools enable but also how they support to understand and develop capabilities. These findings also need to be interpreted in the context of the initial usability problems, as although these were relatively trivial in the technical sense, they were highly problematic in terms of the user experience. The lessons learned from this were that more extensive user testing with the instantiated experience should be performed prior to release with naïve users. During initial deployments the development teams needed to be available and respond rapidly to unanticipated problems that significantly affected the user experience. In this case, the fact that there were language differences and intermediary facilitators meant that the user's problems could not be detected and responded to as quickly as was desirable.

The students did not have highly developed computer skills. They were also unfamiliar with Web 2.0 / Social Software concepts where anybody can contribute and make changes. And the experience also seems to have been too short to overcome this unfamiliarity. Additionally, some technical problems limited the experience, which caused the drop out of two participants. Because of this, all active participation during the experience has been based on the moderator's initial input in SOBOLEO. Similarly to InterLoc, the students worked with the initially created dialogs from Structuralia. Furthermore, the close coupling of the dialog moves & openers to SOBOLEO semantic interactions seem to have limited the possibilities of dialog.

The narrative of the user experience, the instantiation of SOBOLEO and the performed InterLoc dialogs show that whilst the approach was usable in its basic sense, it was not used to stimulate clear examples of ontology maturing. This is not unsurprising, as by its nature, ontology maturing occurs over a prolonged period of time. The performed dialogs showed that the OM-D (containing critical and maturing features) supported maturing characteristics that are implicit in the designed dialogs and their storage within SOBOLEO (i. e. Phase 1, Phase 2), but no maturing of the ontologies were discussed or performed.

Summarizing, we have to be realistic about the 'fit' between socio-technical systems and their contexts of use. SOBOLEO in particular was anticipated as being appropriate as a tool to serve existing or developing community of practices. And in this case the user community was essentially assembled for the experience, rather than being 'tapped into' as such.

7.7. Conclusions

Case study IV and V have shown that the evaluation of our approach within an more or less artificial environment like university courses is quite limited. Therefore, we refrained from conducting any similar studies within a course context for semantic people tagging.

Nevertheless, both case studies emphasized the importance of our model's social dimension. That a certain degree of group identity is a prerequisite and it is necessary that the community itself can evolve and develop rules and norms of collaboration to which we may act supportively by providing an adaptable system framework and guidelines. This has also been confirmed by the evaluation with case study VI where the community has not yet been built up. Additionally, this has shown that it is not feasible to do this by just providing the users two tools.

8. Semantic People Tagging

By extending the group of people who can make competence and expertise assignments to encompass colleagues, semantic people tagging promises to achieve (1) a higher up-to-dateness and completeness of the employee profiles, (2) more realistic assessment of competencies and expertise than with self-assessment, and (3) additional awareness for the tagged person who can see his/her colleagues' perspective. At the same time, assignments by colleagues come with social risks, e. g., by the assignment of inappropriate tags. In order to explore the potential and risks of semantic people tagging prior to implementation and an evaluation in real enterprise setting, we conducted two field experiments with pen-and-paper prototypes. Additionally, we aimed for a conceptual validation with experts focus group. The first field experiment was targeted at the employee perspective, while the focus group was more targeted at the organizational perspective.

8.1. Case Study VII: People Tagging Field Experiments – Individual Acceptance of People Tagging

For evaluating the acceptance of people tagging from an employee perspective and to explore motivational and social aspects in particular, we have conducted two field experiments in two different environments. The field experiments took place in parallel with the system development and thus were based on pen-and-paper prototypes.

8.1.1. Procedure

Two research groups within the area of Computer Science were involved in the design study. Group I consisted of 50 people from two organizations, Group II of 63 people distributed over four organizations. Some of the people belonged to both research groups.

Both groups lacked awareness about the people's topics, interests and competencies within the groups. Both groups wanted to better exploit synergies and to know whom to ask for a problem at hand. Neither of the groups had competence management established, but it was considered to introduce this within Group I in order to improve/facilitate team staffing and career planning. Both groups were open for new technologies and familiar with tagging, Web 2.0 and semantic technologies (as they have been doing research and software development in these areas). They were, however, not familiar with competence management.

The first field experiment (FE I) took place with 39 participants of Group I in July 2008, the second field experiment (FE II) with 38 participants of Group II in September 2008. 17 people participated in both field experiments. Work atmosphere within both groups was frank and friendly. People worked together closely to very closely within their organizations and less closely across organization borders.

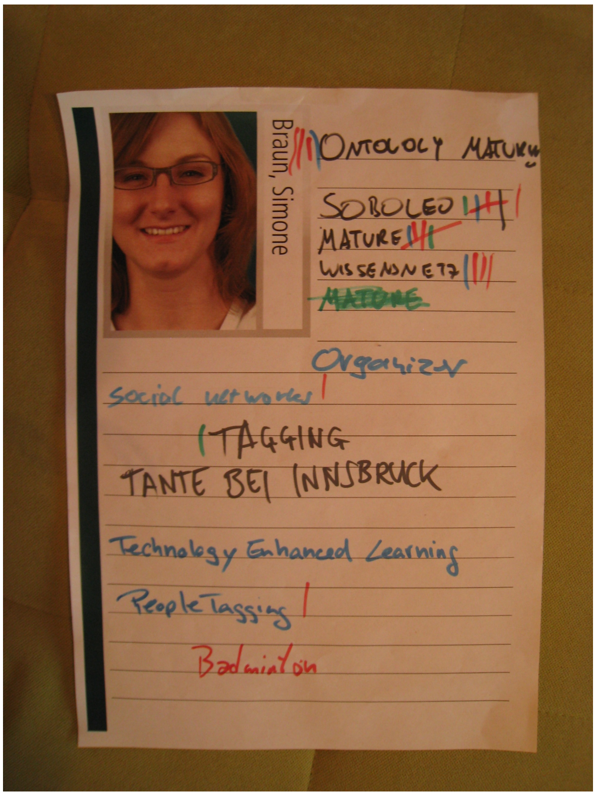


Figure 8.1.: Example of a paper-based poster used in field experiments to tag people.

The field experiments took place in the course of each research group's internal retreat. During these retreats, people tagging was an explicit item on the agenda and task during the three days of the retreat.

8.1.1.1. Materials and Instruments

We prepared paper-based posters for each group member (including not participating members). Each poster showed the name and photo of the person and blank lines to write down tags (see Figure 8.1). We prepared a seed list of tags. This seed list consisted of topics the people are dealing with in their daily work. This list was meant for inspiration and stimulation of the participants to start tagging. Further we prepared a presentation in order to introduce:

1. the topic of competence management and people tagging,
2. the motivation why to do people tagging in the group,
3. the task to do that was: (a) to tag the colleagues and yourself according to the interests associated with them (by writing the tag on the poster), (b) to use whatever tag found appropriate, (c) to use some from seed list, or ignore them completely, (d) to reuse tags of others, and (e) to indicate also if assigning the same tag as already there (by repeating the tag or by adding a multiplying factor).
4. the purpose of the experiment.

At the end of each field experiment, an extended discussion session together with the participants took place. The experiment leaders took notes. Afterwards, we digitalized

Figure 8.2.: People Tagging Setting of Field Experiment I showing the people's posters spread over the room walls and windows in the background. ©Valentin Zacharias

the posters for statistical analysis, i. e. a MS Access database was set up and each person with the assigned tags and number of application entered. We gave every group member her poster. After both experiments were conducted, we additionally sent an email with an overview of the assigned tags and their frequency to every group member. Afterwards, all members of the groups were asked to fill out a short online survey with seven questions.

The detailed execution of the activities, however, in FE I and FE II differed from each other. Whilst in FE I the prepared paper-based posters for each group member (including not participating members) were hung up spreading over the meeting room (see Figure 8.2), in FE II we had to hang up the posters on one wall of the meeting room due to space limitations. The seed list was hung up besides the presentation wall in both settings and both groups got the same task instructions.

In FE I we presented the introductory slides and explained the purpose of the experiment to be for: (a) individual reflection, (b) collective reflection (discussion session at the end of the experiment), and (c) moving towards competency profiles and competence development. As the question arose, we explicitly remarked that every tag is allowed including those related to hobbies or private issues, but that work related topics are in focus. However, in FE II due to unforeseen circumstances, the introductory slides were not presented. The introduction took place orally but without clearly communicating the background and purpose. Furthermore, it was explicitly remarked that every tag, also non-professional, is allowed. It was not said that work related topics are in focus. This difference in introducing people tagging led to some unforeseen and important results presented below.

In FE I, we asked the participants at the end of the introduction session to start walk around and tag and to continue with tagging in the following two days, whereupon the participants started to walk around and tag. After the session, the normal retreat program was continued. From time to time, we encouraged the participants to continue tagging during the breaks etc. At the end of the second day, we collected the posters and started to digitalize and analyze them. At the third day, we presented the participants a first overview of the results, mainly statistical numbers such as number of unique tags, number of tag applications, number of tags per person, most frequent tags, etc. and started then an open discussion. Some days later, we gave every person (including the

Table 8.1.: Statistical data overview of both field experiments in comparison

	FE I	FE II
# posters	50	63
# participants	39	38
# unique tags	585 (156 non-professional= 27%)	485 (226 non-professional = 47%)
# tag assignments	1807	1296
# tag reuse/tag	Ø 3,10 (median: 2)	Ø 2,67 (median: 2)
# tags/person	Ø 15 (median: 11)	Ø 11 (median: 9)
# tag applications/person	Ø 37 (median: 32)	Ø 21 (median: 15)

non-participants) her poster.

In FE II, the participants were asked to do people tagging for the following two days. In comparison to FE I, there was no explicit people tagging session in the agenda. From time to time, we encouraged the participants to continue tagging. At the third day, we started an open discussion. No result overview could be presented because we could only digitalize and analyze the collected posters after the event. Similarly, we gave every person (including the non-participants) their poster.

After both experiments were carried out, we sent every member of both groups an email with an overview of the assigned tags with their frequency and asked to fill out a short online survey of seven questions. People who are member of both field experiments were asked to fill out one survey for each.

8.1.2. Results

Overall people tagging has been regarded as positive and useful. People enjoyed the experiments and stated that *"it was fun"*. 8.1 shows an overview on the statistical data of both field experiments in comparison.

The participants appreciated reflecting about others' interests and competencies: *"tagging people forces you to think about what you actually know about others"*. They liked *"to learn about others"* and *"to get new insights"* in this way, in particular about people they are not so much in contact with. With the tags it was possible to get a quick overview and to see who works in the same area as oneself or has similar, also non-work related, interests (see also 8.2). The participants expressed the wish to have tool support that facilitates finding similar people or comparing people based on their tags. Concerning individual reflection, the participants enjoyed to see how others perceive them and what they associate with them.

It was stated that with single words a description is possible and that the tags *"converge to the right results"*. However, the participants also complained that tags are sometimes not expressive enough or misleading. They indicated that having more context and semantic information would be desirable. It happened that different tags were used for the same concept even on the same poster (e. g., use of both 'Personal Knowledge Management' and its acronym 'PKM'). The seed list was recognized only rarely as it was not integrated into the actual tagging process; i. e. the participants forgot about the list while tagging. The participants wished to have auto completion and suggestion support with more "semantics" during the tagging process.

Another issue the participants raised was the difficulty to start tagging from scratch with a blank sheet. Here again the participants asked for support functionalities or seeding, e. g., everyone tags oneself at first. On the other hand, another group of people stated that seeing the already assigned tags biased them towards confirming these tags instead of adding new ones.

In total the participants enjoyed people tagging as a social activity, i. e. walking around, meeting other participants in front of the posters and jointly reflecting about skills, competencies and (non-work related) interests.

At the same time, however, the joint reflection and discussion about other persons was also perceived as negative because it resulted in "talking about" instead of "talking with" people. This was particularly problematic in FE II where due to the missing introduction of people tagging some serious social issues arose. In FE II, a small number of participants saw people tagging as an intrusion into their privacy – they objected in particular to non-work related tags and to a small number of slightly offending tags. Interestingly even tags not seen as problematic by both tagger and taggee caused problems when read by people lacking the context needed to understand them in the playful way they were intended.

FE II has shown that it is very important to clearly communicate the purpose of people tagging, i. e. what it is intended for and why it is used and what happens with the data afterwards. It should be decided and communicated beforehand how to handle non-professional tags in general and it is necessary to create awareness of the sensitivity of (even slightly) offending tags.

Some participants of FE II also perceived the (partial) anonymity of tagging, i. e. that it was generally untraceable who tagged whom, as negative and as one reason for the high number of non-work related tags.

The general fear of transparency also arose as an important issue. The participants asked for more control over the tags assigned to them, i. e. that they should be able to decide which tags are publicly visible and which not. Some participants also asked for the possibility to opt out of people tagging altogether, to indicate that they don't want to be tagged, to display only self given tags (with only them being able to see tags by other people) or to disable tags from other people.

8.1.3. Discussion

This field experiments have shown that it is possible to retrieve competencies from tags and that people tagging supports reflection about individual and organizational competencies. However, the field experiments also identified important societal and privacy issues that have to be addressed. Addressing these issues needs to be done both with respect to the introduction process and tag visibility controls.

Table 8.2.: Survey data overview of both field studies in comparison

		FE I	FE II
# participants		29	29
Did you learn something new about your colleagues?	Yes	19	16
	Don't know	6	9
	No	2	2
	N.N.	2	2
Did you learn something new about how your colleagues see you?	Yes	13	14
	Don't know	2	3
	No	13	10
	N.N.	1	2
Concerning the expected number of tags: Were you tagged with	More tags	10	4
	Less tags	4	4
	As expected	10	11
	Don't know	5	9
	N.N.	0	1
Did you tag yourself?	Yes	11	6
	No	16	21
	N.N.	2	2

In particular the second field experiment has shown that the proper introduction and communication of purpose is one of the most important issues. Therefore we have developed with the design framework a method for implementing people tagging and continued research on organizational and social constraints related to culture and atmosphere as well as on implications of people tagging.

The other important question was the fear of transparency and the needs of people to control the tags that are visible on their profile. Therefore, we elaborated a fine granular design framework as an instrument to tackle this issue.

Some of the experiences related to technical tool support, like auto-completion support, were directly integrated into the software development process.

Further open questions arisen from the field experiments related to motivation and validity time; i. e. how people can be motivated to continuously tag each other (and not only once) in order to keep the profiles up to date and if the assigned tags can age, e. g., because they represent topics that are not practiced anymore. We incorporated these, for instance, into the scaffolding methods; i. e. explicit system feedback mechanisms (see e. g., 5.1.3.2), or into search heuristics with the indicator of "Time stamp of the tag assignment" (see e. g., 5.3.4).

To summarize, it was interesting to see that the arisen issues could be observed differently in both cases so that we conclude that we need to be careful with – from a technical perspective – rather minor aspects, like e. g., control of the individual over assigned tags, guidance through tag suggestion, visibility of tagging information (who has tagged whom).

8.2. Case Study VIII: HR Experts Focus Group – Organizational Dimension of Semantic People Tagging

For the conceptual validation from the organizational perspective, we have chosen a focus group with HR experts as a method. We presented the semantic people tagging approach in two separate sessions to a focus group of four experts – two German professors specialized in human resource, competence management and organizational development as well as two HR practitioners from large organizations.

8.2.1. Procedure

The first focus group session took place in April 2010 with the two German HR professors. The second focus group session was held in July 2010. Two practitioners participated along with the HR Professors.

For both meetings we met together in an informal discussion session. We introduced the semantic people tagging approach with a short presentation and system demonstration of the first prototype. This was followed by 2.5 hours of open discussion. We collected the focus group's feedback by note taking. Whilst the first session was intended to discuss the general approach, we focused the second session on relating design decisions to organizational characteristics.

8.2.2. Results

In the first session, the participants clearly recognized the novelty and potential of the approach, particularly for enhancing competence management through the integrated vocabulary development. The discussion identified the key issues of cultural awareness of system configuration as well as related potential conflicts of the participatory approach vs. the hierarchical structure of bigger organizations.

The second workshop suggested – against our expectations to relate design decisions to organizational characteristics – that it is a better approach to work on an introduction methodology that helps an organization to find the right configuration for itself. Furthermore, the organizational units were identified for which this would be most useful.

During the discussion the following key aspects emerged:

- People Tagging as a sub system in an organization has to be connected to other functions and systems (e. g., existing enterprise resource planning systems) – otherwise it may not yield full benefit.
- It should be integrated into everyday culture so that it lives; i. e. we need a system-culture-fit. For instance, the freedom of having everybody participating and transparency of changes across hierarchies might be alienating. This requires (1) flexibility and (2) a better understanding of culture and organization to which software configuration options have to be mapped. The identifications of these options might depend on the understanding of the cultural aspects. Thus the introduction of people tagging should involve an organizational assessment, including a detailed examination of the culture.
- Restrictions to the vocabulary might be needed to map to a strategy-oriented competence management approach, where you do not only collect, but make conscious reductions, i. e. prioritizations of competencies. Thus giving a basic structure as a form of guidance, e. g., by limiting the top most level of vocabulary and leaving the freedom for the detailing by tagging (also as a kind of brainpool). In this way, new developments can inform revisions in a controlled manner. Additionally, having employees participate in the corporate strategy can lead to a different form of identification and transparency.
- Flexibility in search strategies is required. Based on the information that is available and seen as a useful indicator; e. g., who is the tagger and what is his/her expertise? But not only to improve the usage experience of the individual, but also as an organizational constraint. For instance, the organization might not have everybody networked with everybody else. There are connections the organization wants to promote, e. g., sales and marketing or finance and sales, but sometimes there are unwanted connections: e. g., when sales representatives know someone in production – they might call the other and bypass the regular process to get something faster than others.
- Transparency of discussions and changes can interfere with organizational hierarchy, e. g., an employee always correcting and discussing his/her team leader' vocabulary changes might be interpreted as a weakness of the team leader by other team leaders.
- Flexibility might be needed on different levels:
  - on the individual (micro) level, e. g., one employee only wants self-assigned tags being displayed, the other any tags;

– on the department (meso) level, e. g., production department might get a more restricted search and vocabulary editing functionality than the R&D department;

– on the organization (macro) level, e. g., free vs. restricted vocabulary editing

Each can be different, but it could also lead to social effects if differences between departments are discovered – makes differences in culture and strategy transparent.

8.2.3. Discussion

The results from our HR expert focus group have further confirmed that there is no one-size-fits-all-system and that each target context of a people tagging system will require a different "configuration", which depends on cultural aspects as well as the actual goals that are associated with introducing people tagging. An analysis of the state of the art (see Section 9.2) has shown that there has been little research on identifying design options in a systematic way so that we have developed a framework for engineering people tagging systems as described in Section 5.3.

8.3. Case Study IX: British Career Guidance – A Real Enterprise Setting

We have implemented and evaluated ontology maturing in the application of social semantic tagging in the real enterprise setting of the British career guidance organization Connexions Northumberland[1] in the context of the MATURE EU project.

Connexions Northumberland is a local service providing help with decision making about study, job and career by offering impartial information, advice, guidance, and personal support to all 13-19 year olds, and to those up to 25 who have learning difficulties and disabilities, throughout the county of Northumberland (cf. . Figure 8.3). The Connexions Northumberland service is delivered by specially trained Personal Advisers (PAs) who are based in schools, colleges, training centers, and in a range of community settings distributed all over the county. Personal Advisers can help young people with all sorts of issues such as jobs, training, housing, money, relationships and health. Personal advisers come from a variety of backgrounds such as careers advice and guidance, youth work, health and social care, youth justice and education.

The knowledge and expertise required for the PAs' daily tasks is heavily context dependent and dynamic. The PAs build up a significant amount of expertise through experiencing concrete cases. However this knowledge-in-use and particularly knowledge on "who knows what" or "who has what expertise" is sparsely shared among the practitioners. Similarly, from the Human Resource development perspective, the organization is lacking sufficient information about needs and the current capabilities of PAs, i. e. what knowledge and expertise have they gained throughout handling the concrete cases.

Indeed, knowing-who is an essential element for efficient knowledge processes in career advising, e. g., for finding the right person to talk to. So, taking the scenario of where a novice PA needs to respond to a client query. The PA does not feel sufficiently confident

[1] http://www.connexions-northumberland.org.uk

Figure 8.3.: Location of the ceremonial county of Northumberland within England. ©Ordnance Survey

to respond adequately, so needs to contact a colleague who is more knowledgeable, for support. The key problems would be:

- How does the PA find the right person to contact
- How can the PA find people inside, and even outside, the employing organization?
- How can colleagues, who might be able to support the PA, be identified and contacted quickly and efficiently?

Currently, the PAs go through their personal notes to find the right person or they call several colleagues to ask if they can mediate a contact. There is no organization wide information pool on who knows what. After the PA might have found the right person, s/he updates his/her personal notes that identify the person to have expertise or interest in a certain topic. Occasionally, the PA may also share the newly gained knowledge from this concrete case, e. g., about others' expertise or collected and produced links and documents, in discussions and conversations with colleagues.

Typically, employee directories, which simply list staff and their areas of expertise, would be installed for such cases, however they were judged to be insufficient. One reason is that information contained in the directories is often outdated; or it is not described in an appropriate manner; or it focuses too much on 'experts'; and they often do not include external contacts (cf. Biesalski and Abecker, 2005; Schmidt and Kunzmann, 2007).

Additionally, Human Resource development wants to have sufficient information about the needs and current capabilities of current employees to make the right decisions; for instance to make a training plan for Connexions Northumberland's PAs. In service delivery contexts that must be responsive to the changing needs of clients, like Connexions services, it is necessary to establish precisely what additional skills and competencies are required to keep up with new developments. That means the Human Resource manager needs to know what additional skills and competencies are required and missing. Therefore, for instance, s/he needs to get an overview on what topics and requests the PAs demand to fulfill their daily work, i. e. what type of expertise is needed. On the other hand s/he needs to compare these needs with the current capabilities of the PAs, i. e. including the informal knowledge and expertise that the PAs gained throughout handling the concrete cases, in order to know how much of the requisite expertise already exists. Traditional top-down competence management approaches are similarly perceived as unsuitable as too rigid.

The implementation and evaluation with Connexions Northumberland was divided in two main phases of formative evaluation and participatory design activities. The goal was to answer the question if the system is useful for supporting maturing of a shared ontology and knowledge about others. To that end, we first introduced the system in a series of workshops (as detailed below). After passing a usability check point the system was integrated into normal operation.

8.3.1. Procedure Phase 1 Formative Evaluation and Participatory Design Activities

The first phase of the formative evaluation comprised two phases which consist itself of two variants. Phase (1a) of participatory design activities took place from October 2009 till January 2010 to the end of the iterative development and adaptation with Connexions Northumberland. Phase (1b) took placefrom December 2009 to January 2010 for the usability check that was a prerequisite for the larger scale test. The larger scale test was the second phase of formative evaluation that took place from February 2010 until June 2010 within an operational context with 15 people from Connexions Northumberland.

In phase (1a), two members from the HR development department at Connexions Northumberland were involved. In phase (1b), two managers and the head of service from Connexions Northumberland were involved in addition to the two HR development experts from (1a).

In phase (1a), we provided a functionally complete online version of the SOBOLEO system as well as a manual describing the usage of the tool. The HR development representatives tested the application on their own, without any involvement or presence from our side. The key representatives provided requirements, and suggestions for the development. This was a pre-test to see if the application was sufficiently self-explanatory and suitable for use. The results of testing (by Connexions Northumberland) and of adapting (by the development team) was discussed via email and in bi-weekly phone conferences.

In phase (1b), the participants of phase (1a) prepared a short training session based on their application, which was based on the manual for the application. Based on this, the participants approved passing the usability checkpoint.

8.3.2. Results Phase 1 Formative Evaluation and Participatory Design Activities

In phase 1, the participants found the system very easy to use from the very beginning and suitable for their organizational context. They did not need any additional instructions in addition to the manual that had been provided, and were confident introducing it to their colleagues without assistance from our team. During the testing, several suggestions for improvement were made:

- The original assumption of the system was that the application primarily is an overlay to existing internal or external web pages, e. g., an employee directory, a social networking site or similar. However, Connexions Northumberland wanted to focus on internal contacts first and did not have any such employee directory so that the "anchor" for bookmarking was missing. As a consequence, we have added an employee list (see 6.7.2) to the system framework, which in turn has been developing towards a full-blown replacement of employee directories or yellow pages.
- So far, the assumption had been that a user tags a person with certain topics. Then the topics can be consolidated if needed. However, it has turned out that users want to have the ability to add persons to concepts when they are in the ontology editor mode. This was implemented before phase 2 began (see also 6.7.4).
- A lot of comments and suggestions for improvement concerned the labeling of UI elements, which was in several cases too technical for the target context. The application partner made suggestions for improvements which have been implemented in several iterations.
- Likewise, several suggestions were intended to improve the introductory manual, which have been implemented, too[2].

8.3.3. Procedure Phase 2 Formative Evaluation

The following section describes the second phase of formative evaluation performed with career advisors at Connexions Northumberland. The piloting at Connexions Northumberland was prepared by the participatory design activities (phase 1 of the formative evaluation). This also included the specific system configuration and instantiation of the design framework (see below in Section 8.3.3.1).

The piloting was accompanied by three workshops. The first two workshops aimed at introducing the system to a selected group of employees of Connexions Northumberland and providing supervised access to the system for an initial trial for the duration of the workshop. This was intended as a kick-off event for the operational trial that followed. The third workshop was a user feedback workshop in order to collect the participants experiences after using the system over a longer period of four weeks during their daily work.

- **Workshop I:** The first workshop on March 26, 2010 in Blyth (UK) gave a general introduction of the system to a subset of 4 pilot users who attended the workshop. We developed a two part-questionnaire. The first part was to be answered at the beginning of the workshop, the second one at the end. So, we collected participants' expectations before a hands-on session in which they followed a task-oriented guide. At the end of the session, they were asked to fill in a post task questionnaire. Furthermore, any problems and suggestions arising during the hands-on session were collected. The workshop generally confirmed the approach, and provided several ideas and usability suggestions. The main issue of this workshop was a technical network configuration problem (proxy, network latency) that had

[2]This manual is available on request.

considerable impact on the user experience. Therefore, the operational pilot was postponed until these problems had been investigated and solved.

- **Workshop II:** A second workshop was conducted on May 11, 2010 with 2 participants from the first workshop and 6 new participants (8 in total). This workshop had the same approach as the first one. The measures taken to avoid the network-related problems were successful. The questionnaires and open discussion with this group have shown that these participants particularly liked the simplicity and the democratic philosophy of the approach.
- **Workshop III:** After the second workshop, a group of 15 users (some of whom did attend the workshop) started to use the system as part of their operational procedures. Some of the participants acted as multipliers for their teams, i. e. , they explained the tool to their team on their own. A third workshop on June 8, 2010 with a selection of 4 participants of the trial was used to gather semi-structured feedback about their general impressions, how they used the system, what they liked, what they disliked about the system, and how they considered it helpful for their work. The feedback on the helpfulness in particular was used to investigate the impact on maturing of the shared ontology and knowledge about others. The log data from the trial session was collected in order to analyze it.

8.3.3.1. Materials and Instruments

In preparation for the piloting and workshops at Connexions Northumberland, it was necessary to decide on the system configuration specific to the target context and to prepare the workshop materials; i. e. the scenario and tasks to guide the hands-on activity and the questionnaires.

System Configuration We elaborated the organization specific instantiation of the design framework together with a team of Human Resource, training, and knowledge management and team managers of the career advising organization in phase 1 of the formative evaluation in the iterative process of discussions and system demonstrations, tests and adaptations.

In the final concrete instantiation for piloting, even though the focus was on people tagging, it was decided to have people as well as web documents[3] as addable and taggable resources. However the introduction and guiding material was focused on people tagging and features for web documents were mentioned in passing. The extended dialog support was not enabled.

Employees were allowed to tag themselves and their colleagues without any further restrictions. Connexions Northumberland's focus is on its own employees, so tagging of external contacts was not envisioned but also not technically enforced.

Every employee was seen as a participant. So we created for the second workshop (empty) profiles for all employees beforehand; i. e. it is allowed to tag any colleague without the taggee's explicit opt-in. The employees were supposed to develop and modify the vocabulary used for tagging on their own. Thus, the system automatically added new key words used during the tagging process to the vocabulary and changes to the vocabulary (e. g., by adding relations between tags) were immediately visible and effective (e. g., to the search). The extended ontology editor version was not enabled.

[3]Uploading local documents has not been available yet

The key representatives seeded the vocabulary with some topics in order to avoid cold start effects. There was no limitation or black list of terms. But the employees were not supposed to use negative or non-professional tags. During the tagging process, the system suggested existing tags from the vocabulary in order to foster reuse and thus the consolidation of the vocabulary. Tags were assigned without any further semantic differentiation. Assigned tags were immediately visible without previous approval by the taggee. The system did not block or delete specific tags automatically. At the beginning, it was decided that only the taggers can remove their tag assignments. Based on the participants feedback, this was revised so that also the taggee could remove tags assigned to him/her by others.

During the tagging process, the user could see the tags assigned by other users. We did not make a seeding of the profiles. Every user of the system could see the tags in an aggregated (by frequency) and anonymous representation without level differentiation. The search ranking was based on the tags of the taggee and their frequency. We additionally took into account the structure of the tags.

Task-Guide, Pre- and Post-Questionnaires In order to guide the hands-on session, we developed a scenario that represented a usual situation in the career advisers' daily work and in which we integrated four tasks how to use the new SOBOLEO system. These tasks were (see D.1):

- Tagging yourself
- Improving the vocabulary
- Searching for a contact person
- Tagging a colleague.

The subject matter of the pre-questionnaire (see Appendix D.2 and D.3) was usability and satisfaction centered and based on the German standard on user-centered design and DIN EN ISO 9241 "Ergonomic requirements for office work with visual display terminals". The content of this questionnaire was approved by the HR development representatives prior to its use.

The pre-questionnaire (see D.2) consisted of five questions on the participants' expectations of the system:

- How do you usually find useful people (without support from such a system)?
- What do you think of the system? What do you expect of the system specifically?
- What do you think are the system's advantages?
- What do you think are the system's disadvantages?
- Do you know of any other systems that you would use to find useful people? If so, which?

The post-questionnaire (see D.3) was centered on questions on user satisfaction, usability and motivation & barriers. It consisted of four parts with: (a) 20 five-point Likert scale questions on user satisfaction, (b) four free text questions on usability, (c) two five-point Likert scale questions and four free text questions on motivation & barriers, and (d) comment section.

8.3.3.2. Introduction Workshop I

Four employees of Connexions Northumberland participated in this first stage of this pilot. Before the presentation began, the Head of Service attended the meeting, to show her support and the commitment of the organization to participate in the pilot.

The team met together in an ICT training room. The system had already been loaded and a member of the ICT staff was available to answer questions. The group of participants included training managers and the information manager – no Connexions PAs were present. This group represented a range of confidence and competence in the use of ICT, with two less confident about using ICT than the others. They all understood that this system potentially offered their organization valuable support and had, in principle, agreed to convening a group of about 20 practitioners to test the system further. Details of this second phase of the pilot were also agreed.

The structure of the session (9:30am – 1pm) comprised a formal presentation on the MATURE project and the People Tagging approach and system with plenty of opportunity for discussion, comment and questions. Two of the group had tried the software three months previously in their organizational setting and had liked the software.

After the presentation, all four participants were asked to fill in a pre-questionnaire on user expectations (see D.2). Afterwards they were invited to log in to the system and carry out the set of tasks provided by us to guide the activity.

Our role was to provide support and to make notes on the participants' responses when using the system. The post-questionnaire on user satisfaction, usability, motivation and barriers concluded the initial trial session (see D.3). The questionnaires were based on DIN EN ISO 9241 "Ergonomic requirements for office work with visual display terminals". In a final open discussion panel, we collected the users' general feedback and experiences through taking notes of the conversations.

8.3.3.3. Introduction Workshop II

After improving the system with respect to the network environment (see results of workshop 1 below), a second workshop took place on May 11, 2010. The objective for the session was to introduce people tagging to a broader group of selected employees, especially Personal Advisers, of Connexions Northumberland and to provide supervised access to the system for an initial trial and motivate long term usage of the system.

Eight employees of Connexions Northumberland participated in the second stage of this pilot. These were three managers, who also attended the first workshop in March and five PAs in a mix of school based and community based PAs with varied expertise and experience. This group represented a spread of competence and confidence in the use of ICT.

The team met together in the ICT training room. The developer team had access to the room the evening before in order to make a system test. The system test was successful. Nevertheless a local system network was additionally prepared as a fallback solution in order to prevent known local network connection breakdowns, especially in the late morning hours, affecting the trial session again. In additional preparation, the system was adapted and improved (see below). The working space was seeded with basic personal data on organizational employees. A tag structure was entered and was available for participants on the system.

(a) (b)

Figure 8.4.: Hands-on trial with personal career advisers during the introduction Workshop II at Connexions Northumberland training center.

The structure of workshop 2 (9:30am – 2pm) corresponded to the workshop 1 structure and was only slightly adapted. There was a formal introduction and project presentation followed by a system demonstration to show the People Tagging approach. There was plenty of opportunity for discussion, comment and questions.

After the presentation, all new participants were asked to fill in a pre-questionnaire on user expectations. Afterwards all participants were invited to log in to the system and carry out a series of tasks provided by us. These tasks were the same as in the first workshop, and again related to tagging a colleague, self-tagging, improving the vocabulary, and searching for a contact person. The duration of the hands-on trial session lasted for approximately 50 minutes.

Our role was to provide support and to make notes on the participants' responses when using the system. A post-questionnaire on user satisfaction, usability, motivation and barriers concluded the initial trial session.

The provided questionnaires and tasks corresponded to the ones of the previous session. This means, three of the participants repeated the same tasks, but only the five new participants were asked to fill out the questionnaires. The comments and opinions of the other three were compared over both sessions in the final open discussion panel together with the participants' general feedback and experiences. All the responses to the questions are given in Appenix D.5.

8.3.3.4. User Feedback Workshop III

15 employees of Connexions Northumberland participated in the usage in an operational context phase after the previous introductory workshops. This phase took place from May 11 to June 8 2010. During this phase, the participants of the previous workshop acted as mediators and introduced the system to their own teams. To not disturb the longer term usage of the system, no developments had taken place between the second and the third workshop. At the end of the experience, we organized the third workshop with 4 participants of the trial in order to collect the participants' feedback.

The workshop lasted 3hrs in which we collected the participants' feedback by note taking. The three-parted session started with a round of personal statements (each participant 5-10 min) based on four central questions:

- How did you use the software (how frequently, for which purpose, individually or collaboratively, which part/feature)?
- What do you like about the software?
- What do you dislike? What are your suggestions for improvement?
- How did it help you or might it help you in the future?

This was followed by an open discussion about specific aspects that arose from the individual statements (30-60 min). Additionally, aspects of the developed design framework were used as guideline for discussion about the appropriateness of the design decisions. The third part of the session was targeted to future project planning.

8.3.4. Results Phase 2 Formative Evaluation

8.3.4.1. Results from Introduction Workshop I

The results from taking notes during the initial trial and the open discussion panel are divided into general remarks and those specifically related to technical issues.

General Feedback

- Staff development manager: would be useful to know what CPD people are doing, more on their skills profile and what they might be able to share with others;
- More generally, importance of encouraging reflective practice and this, in turn, leading to a sharing skills and knowledge;
- Work of personal advisers necessarily involves inter-agency learning (across six services) and this too means it is important to be able to tag people who possess particular types of knowledge and to share this knowledge within and beyond the service;
- Within the services there are also different 'hubs' of knowledge and this could be shared;
- There are concerns of sharing whole people tagging information with other services in general;

- People tagging could also play a role in leveraging specific types of staff development: for example, in relation to interview refresher training - having people who could discuss how this affects practice etc., would be useful;
- Once people can see the value of sharing their knowledge resources and making these accessible through people tagging, then this may lead to use of competency frameworks and link to more systematic forms of appraisal, review and training needs analysis (these already link to observation and development of professional practice, although at present are linked more to assessors and team managers rather than peers);
- Welcome the idea of having more support for identifying and using topic tags, with a collective review of skills and competence development;
- Finding the 'right person' to whom to direct a query would be useful, although it would be important that a person tag is time-bound, so people do not feel they are making a completely open-ended commitment;
- Visualization overviews would have a real benefit;
- Tagging of web pages and other resources could be part of a comprehensive package;
- The freedom to use the system in such a democratic way is appreciated very much;

Technical Feedback & Issues Initial problems arose when the four participants tried to log in to the system simultaneously, i. e. user information seems to be shared between the different user terminals. Additionally, the local Internet connection, which went down from time to time, affected the system use during the trial session. Nevertheless useful feedback was elicited, that is given in detail in Appendix D.4.3. This provided specific UI and design requirements related to Tagging, Editing, Searching and Browsing.

Pre-Questionnaire Results Two questionnaires were used to elicit further feedback, one administered before their experience with the system (pre-questionnaire) and one after their experience (post questionnaire).

All four participants answered all questions of the pre-questionnaire. The four users' statements showed surprisingly clear results: that they expected a bottom-up system which empowered them to shape it for their own needs. Furthermore, the system fills a gap in current infrastructure and organization, as they are often trying to 'find the right person to talk to', but don't know how. So the users were quite clear that they expected this approach to yield clear benefits.

Summing up, the participants usually find useful people by speaking and discussing with colleagues and personal contacts. The participants think the system will be useful and offers huge potential, especially for sharing knowledge. They expect finding people being easier and less time consuming. The biggest advantage, the participants see, is the collaborative and democratic nature of the system where everyone can contribute and amend. Another advantage is the wealth of information in one place.

Regarding the system's disadvantages, as it is a web-based system, it might be challenging for people with less confidence in using ICT. Additionally, it's dependent on people engaging with it, which was quite a new concept for them. Currently, they don't know any (internally available) system for finding useful people. Externally, they might use the web.

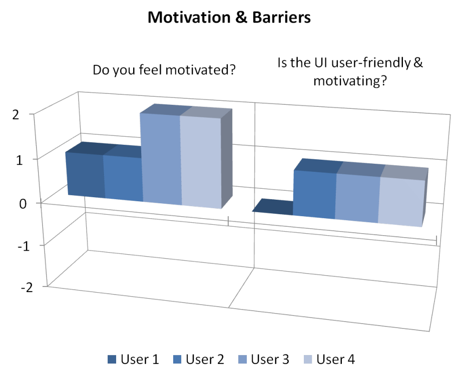


Figure 8.5.: Workshop I: Results on motivation and barriers (-2 = disagree; -1 = slightly disagree; 0 = neutral; 1 = slightly agree; 2 = agree)

Post-Questionnaire Results The answers of the post-questionnaire in terms of user satisfaction are collated in Table 8.3. In general, these results were promising. All responses were at least neutral, more than two thirds indicated "slightly agree" and above. The least user satisfaction can be observed with respect to quick system reaction (8), which could be clearly attributed to the network problems.

Regarding user satisfaction, the participants' feedback was positive. There were only two questions where participants answered with *slightly disagree* (*"the system is/reacts fast"*, *"The length of time spent learning about the system was appropriate."*) whereas the mean is *neutral* and respectively *neutral* to *slightly agree*. These can be traced back to the login and Internet connectivity problems that occurred during the trial. Above all, the participants could understand the system (questions 2-5, 13, 15) and were interested in many of the things the system offers (especially question 1 and 16).

Regarding usability, there were no fundamental problems stated in the post-questionnaire or things that needed to be changed apart from overcoming the technical problems related to the network and minor problems occurring during the trials. It was encouraging that the most complex part of the system (the taxonomy editor) was mentioned explicitly by half of the users as the most interesting one. The actual answers that were provided are given in Appendix D.4, and other interpretations are given below.

Comparing the system to their daily work, the participants mentioned the system's usefulness and ease of use. Besides a guiding document, there's nothing they missed or would like to be changed. They also stated that there's no function they would not use. The most interesting feature was different for every participant: editor, topic list (navigation), tagging, and ATOM feeds.

Similarly, they felt motivated by different aspects, e. g., simplicity, the support for communication, the participatory nature, and the support of CPD (Continuing Professional Development). The overall impression on motivation shows diagram 8.5. The barriers they mentioned relate to technical difficulties of integrating the SOBOLEO system into their IT environment. Additionally, they also stated that people might lack IT skills/confidence to use the system, especially the new culture that anyone can contribute, but also that they might lack time.

Table 8.3.: Workshop I: Results on user satisfaction (1 = disagree; 2 = slightly disagree; 3 = neutral; 4 = slightly agree; 5 = agree)

Questions on user satisfaction	# of answers	Min	Max	Mean
1) I am interested in many of the things the system offers.	4	4	5	4,75
2) I can easily find my way through the system.	4	4	5	4,25
3) I can quickly find what I'm looking for.	4	4	5	4,25
4) It's a logical system.	4	4	5	4,50
5) I can understand the system.	4	4	5	4,75
6) The user interface is attractive.	4	3	4	3,75
7) I had the feeling to have anything under control.	3	3	4	3,67
8) The system is / reacts fast.	3	2	4	3,00
9) The system supports me in finding what I'm looking for.	4	4	4	4,00
10) It's easy to navigate through the system.	4	4	4	4,00
11) I like the system.	4	4	5	4,50
12) With the system, I can easily find relevant people.	3	4	5	4,33
13) I can use the system efficiently.	4	4	5	4,25
14) The information provided is presented clearly.	4	4	5	4,50
15) The system is easy to use.	4	4	4	4,00
16) The system is free of annoying features and functions.	3	4	4	4,00
17) It's easy to see where I am exactly within the system.	4	3	4	3,75
18) The length of time spent learning about the system was appropriate.	4	2	4	3,50
19) Activating / Clicking on individual links resulted in exactly what I expected.	4	3	4	3,75
20) The system is easy to understand.	4	4	4	4,00

8.3.4.2. Design and Developments between Introduction Workshop 1 and Workshop 2

Along with the participatory design approach and iterative development process, we have taken the feedback from the first workshop to adapt and improve the system for the second trial:

- The previous problem of logging-in could be traced back to special local proxy caching mechanisms which caused a sharing of the login information and which could be eliminated by inhibiting the caching of the web application pages;
- We executed system performance and request-response-latency tests and could exclude these as the cause for the disturbance of the system use in the first trial. Thus, this could be traced back – together with the experience of the locals – to an unsteady local Internet connection under high load and the special situation in the ICT training room (i. e. the simultaneous access of all users from the same room). Therefore the previously mentioned local system network was provided for this trial. This might not affect the long term usage as the users access the system from different places and at different times;
- The Connexions Northumberland representatives provided a list of all employees that we included in the system's people directory;
- When tagging a person or adding a person to a topic, the user was now supported with an auto-completion functionality for entering the person's email address or name;
- System messages were revised;
- Query expansion when searching people was adapted to be more transparent;

8.3.4.3. Results from Introduction Workshop II

General Feedback After the formal introduction, project presentation & system demonstration – comments and questions from participants arose that are summarized in the following. Mostly, the comments originated from individuals, but did not cause any disagreement (unless otherwise stated) in the group so that basically these can be considered as group consensus. The responses are classified in terms of concerns expressed and counter arguments in support of using the system (see also D.5.3).

Concerns expressed included the following:

- that linking with these types of outside agencies would increase the workload of individual practitioners
- that some practitioners may abuse the system – 'lazy' colleagues may resist entering details about themselves and may tag others with expertise they may have (to deflect additional queries)
- that it could be a problem for some colleagues to identify areas of expertise with which they do not feel comfortable being identified

Counter arguments in support of using the system:

- the basic philosophy of the system is democratic – bottom up, rather than top down – and empowering the individual – often people feel out of control

- that management intervention might become needed, but the system is likely to work best where this is kept to a minimum
- that management solution might be needed as a complement
- that it's not about experts but about expertise – tagging does not mean expert, that's important
- that editing the own profile, i.e. removing tags that are inappropriate or in order to avid being contacted, is important
- that the system could increase the efficiency of working with outside agencies by allowing them search for PA expertise easily and quickly
- that the system has the potential to support increased levels of constructive interaction with external agencies (in line with policy expectation)
- that the system has the potential to enhance the sharing of resources likely to be increasingly rationed in times of economic stringency (e. g., learning from attendance at learning events can be shared quickly and easily with colleagues)
- that the system has the potential to support the learning of users through confidence building, related to peer perceptions of expertise
- that (eventually) the system would allow service users to select the expertise of the PAs to whom they wish to target a query

Technical Feedback & Issues This time, each participant had access to a PC and no problems were experienced logging into the system. Similarly, there were no difficulties with the local Internet connection that could have affected the system use during the trial session. The feedback we collected related to questions, comments and observations on the operation of the system. The responses that are recorded are detailed in Appendix D.5.4.

Pre-Questionnaire Results From the pre-questionnaire, all participants currently usually find useful people by asking colleagues, e. g., by phone or email. All participants considered the system in general as useful, but some made this dependent on people's participation and contribution. The biggest advantage, all the participants saw (see responses to question on advantages), is the system's ease to use and simplicity. Also mentioned was the opportunity to link up with partner agencies. Regarding the system's disadvantages, the participants mentioned that people might not contribute or be associated because they don't feel confident enough or an additional responsibility and with this an extra work load for a specific topic. Currently, except for one person, they did not know any other system to find useful people in a similar way.

Post-Questionnaire Results The comprehensive set of responses from the post-questionnaire related to user satisfaction are summarized in Table 8.4.

The responses reveal a similar pattern to that previously carried out. An interesting difference is that one of the participants had considerably more problems with the system (low rating in many of the questions, particularly on efficient usage) and attributes this to lack of time spent on the introduction (*disagree*). Despite that, there was on average higher agreement, which is clearly related to the overcoming of technical problems.

Table 8.4.: Workshop II: Results on user satisfaction (1 = disagree; 2 = slightly disagree; 3 = neutral; 4 = slightly agree; 5 = agree)

Questions on user satisfaction	# of answers	Min	Max	Mean
1) I am interested in many of the things the system offers.	5	4	5	4,2
2) I can easily find my way through the system.	5	2	5	3,8
3) I can quickly find what I'm looking for.	5	2	5	3,8
4) It's a logical system.	5	3	5	4,2
5) I can understand the system.	5	3	5	4,2
6) The user interface is attractive.	5	3	5	4,0
7) I had the feeling to have anything under control.	5	3	4	3,6
8) The system is / reacts fast.	5	2	4	3,2
9) The system supports me in finding what I'm looking for.	5	2	4	3,6
10) It's easy to navigate through the system.	5	2	5	4,0
11) I like the system.	5	4	5	4,4
12) With the system, I can easily find relevant people.	5	2	5	4,0
13) I can use the system efficiently.	5	1	4	3,4
14) The information provided is presented clearly.	5	3	5	4,4
15) The system is easy to use.	5	2	5	4,0
16) The system is free of annoying features and functions.	5	2	5	4,2
17) It's easy to see where I am exactly within the system.	5	3	5	4,0
18) The length of time spent learning about the system was appropriate.	5	1	4	3,0
19) Activating / Clicking on individual links resulted in exactly what I expected.	5	3	5	4,2
20) The system is easy to understand.	5	3	5	4,0

Regarding user satisfaction, the participants' feedback was positive, but in comparison to the first workshop, where the system was presented but not used, with more *neutral* to *slightly agree* answered question on average. There were also more questions answered with *disagree* or *slightly disagree* but only by one participant. Again, it's the system's reaction time and activating/clicking on individual links resulting in exactly what was expected that resulted in only an average of neutral satisfaction. Sticking out, the participants indicated that they like the system and the clear presentation of the information as well as that the system is free of annoying features and functions.

Given the better stability of the network, the users could use the system more intensely, which led to more detailed identification of usability improvements, e. g., better search functionality, case sensitivity. Users mentioned this time other features they liked best, which suggests that different individuals also put emphasis on different parts of the system.

Comparing the system to their daily work, the participants had different opinions. They find the system easier and especially useful when no colleague is nearby. But they still like to speak and discuss matters with people. They would be encouraged to use the system if the system was always up to date and had easier search functionality. Additionally, information such as more contact information or how who has been tagged could actually offer support. There are some minor issues they'd like to see changed regarding the search, case sensitivity and ordering of topics (each of them mentioned once). All stated that there was no function they would not use. Only one participant indicated that s/he would not use the tag webpage function. However, this function was not focussed on during the trial and only introduced on the side. The most interesting feature was different for every participant: monitoring of your own profile, browse people, possibility of defining broad topic areas, and bookmarking web pages. This suggests that actual usage will have a different focus by different individuals.

Similarly, they felt motivated by different aspects. The users again mentioned the simplicity of the system and quick access to more information than before. For instance, one participant stated: *"Like the lion! It's a very accessible and user-friendly site – best because of its simplicity. Even techno-phobes like myself feel very comfortable using it."* Diagram 8.6 shows the overall impression on motivation.

However, they mentioned at the second workshop particularly lack of time as a barrier to system usage as well as that this needs a culture that people agree on bringing knowledge (about them) to the system, which are classical knowledge management issues. Cultural aspects emerged again (two participants), i. e. that the system would have to be accompanied by a change of culture. Besides the aforementioned features that would encourage the participants, they added that links to outside agencies and seeing other colleagues using the system would make it more likely that they use the system.

8.3.4.4. Results User Feedback Workshop III

Personal Statements In the following, we present the participants personal statements in a collated form based on the four central questions together with general remarks. One participant (P3) had no opportunity to test the system and thus to provide his/her personal impressions.

Table 8.5.: Personal statements by the workshop participants

Participant	General	Use	Like	Dislike	Help
P1	•conducted training days with around 18 participants , positive feedback •had a meeting with information group •put a link on the intranet for easier access	•tidied up the topics, to make it presentable (10 minutes) •demonstrated the software	•easy to use •list of people •useful to tag the websites		•knowing what others in the offices are doing •useful for a central office •combination with web sites
P2		•curious to see how people have used it, which tags are used •expert notion discarded •used for tagging people (around 30 min) and gardening (25 min) (was on holiday) •personal perspective •multi hierarchy helps to overcome classification problems •also tagging persons not using the system •not tagged herself	•tagging, see how others are using the system •democratic approach •web application, looks good	•moving from the new topic list to the proper hierarchy •better: show who tagged •very plain, no pictures [might depend on the person]	•could be used for a time-limited classification •folder like in inbox •more web sites are accessed by recommendations than by searching, important functionality •headquarters context: communicating with them, who is interested •emails from PAs about experience, setting up a working group •inquiry or events organised would be triggers •Intranet start page is a problem, RSS should be used, but a cultural thing •fostering cross-team activities

P3	•feedback to two teams, explaining in an email how to use the site, team meeting, no IT resources in the meeting, was thus difficult to explain
	•positive feedback:
	  •relevance, potential use
	  •particularly link to web sites: too much information on the web, recommendations important, channeling
	•negative feedback:
	  •who keeps it up to date, responsibility for checking correctness ["symptomatic of how we run IT in the company"]
	  •nervous about tagging other people (embarassment, sometimes, but not dominantly)
	  •tagging battles feared →dialogues needed
	•has to go through team managers
	  •apathy to anything that is new
	  •if they don't have to use it, they won't
	  •permission to experiment
	  •combination of bottom up and top-down required
		•checked tags for team members [15 min]	•easy to use (used a lot of IT systems)		•identifying other staff members who could help
					•web sites
					•nervous about tagging, would not use tagging much (new to the organization), different to the existing
					  •putting pressure on people, potentially more work
					  •would not feel comfortable with being tagged
					  •explicitness is frightening
					•moving to another area

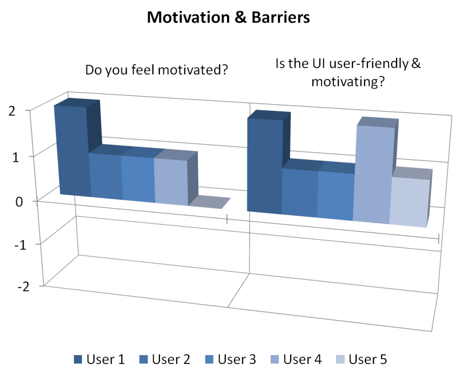


Figure 8.6.: Workshop II: Results on motivation and barriers (-2 = disagree; -1 = slightly disagree; 0 = neutral; 1 = slightly agree; 2 = agree)

Open Discussion Feedback In the open discussion, moderated by us, individuals came up with their impressions and responded to the questions from our side. These addressed impressions on the notion of experts, tagging and getting tagged, cultural and personality dependent aspects. The responses are given in Appendix D.6.1.

Aspects of Design framework We also discussed with the participants aspects of the design framework to reflect on the results of the participatory design approach so far. This included design options regarding transparency of user tagging activity (who can see what at which level of detail), opt-in and opt-out strategies for participants as well as different forms of control over the vocabulary. Options different from those that were selected at Connexions Northumberland were explicitly presented and the participants' feedback with respect to the best option for their context collected:

Who can be tagged?

- everyone, not just opt-in
- opt-out possible
- externals might be in the future

Who can tag?

- everyone can
- self-tagging would not yield enough information

Control over vocabulary

- like the balanced approach to control
- negative tags were not seen as a problem to avoid in a first step, only if there are problems

Visibility of the tagging

- ambivalence towards visibility of the tagger, should definitely not be visible to anyone else but the taggee

Search heuristics

- aging should be taken into account.

Summarized Feedback The evaluation has confirmed fundamental design assumptions:

- the democratic and participatory system philosophy in the context of making people's expertise more accessible is also appropriate for an organizational context and is considered to be more suitable than top-down approaches (this has been the consensus in the groups present at the workshop)
- although users with different types of usage preferences use the system in a very different way (tagger, gardener, searcher), there are sufficient active users both for tagging and gardening (this has been specifically subject of discussion at the final workshop)
- the simple tagging-based approach which does not require taggers to specify more than a weak [understanding] "associated with a topic or expertise" is appropriate, as semantically more precise statements (like "expert in") are rarely desired and would constitute a barrier to system usage (because of cognitive effort and hesitancy) (this could be observed particularly during the discussions about the notion of "expert", which appeared problematic and where participants were hesitant to judge who is an expert and who isn't).
- The weak semantics of the hierarchy (broader/narrower) as well as the possibility for multi-hierarchy (more than one broader term) appear to reduce the cognitive effort (this could be seen from (a) observations during the trials and (b) the absence of problems with vocabulary building compared to other experiences gathered with ontology-based approaches in the past)
- Practitioners agree that there is a pathway from tagging people's expertise towards human resource development and competency frameworks (this has been confirmed by both career advisors and the responsible for HR Development at the target site). The concrete connection between those activities, however, remains complex and not yet sufficiently understood.

During the later evaluation no major changes were suggested or requested. The additional functionality suggested seems to be more of an evolutionary enhancement.

8.3.5. Discussion

The evaluation has shown that introducing people tagging into an organization has to take serious account of the anticipated effects, which are related to the organizational culture, and current practices, but also on the individuals:

- People tagging is seen as intrusion into the personal sphere of each individual. It makes explicit what has so far only implicitly been known by a smaller group. This effect of transparency depends on the personality of the individual, but also the culture of the group.
- Although the tool does not prescribe any semantics of the tagging, the organizational context suggests certain semantics. In the case of Connexions Northumberland, this was "having expertise".

The workshops have shown that people tagging contributed to maturing shared ontology and knowledge about others' expertise particularly:

- The bundled presentation and handling of web resources and people's expertise makes it easier to share achievements in terms of expertise building with others in the organization, thus raising the expectation that this supports the learning of the individuals and that this increases the pace of maturing a shared ontology and knowledge about others' expertise.
- The system is considered to support reflective practice, which is interrelated with the vocabulary development.
- The expected influence on the social dimension is clearly seen by participants in the evaluation, although ambivalent. Particularly, the effect of making explicit personal expertise can have a varying impact based on the existing culture and rules. Connexions Northumberland associated "being tagged" and not objecting to it, committing to helping others on the topic.

Apart from suggestions for evolutionary development of the software, it has also become clear that there is a high potential in combining the people and the content dimension of maturing. Finding the right balance between contacting people and creating artefacts for sharing explicit knowledge helped to overcome the problem of being contacted too frequently.

In total we found out that most people accept the tool, view it as very user-friendly and easy to understand (e. g., all but one participant agreed or fully agreed on *The system is easy to understand* in the post-questionnaire). The simplicity was particularly seen as a specific strength. The usability questionnaires have also shown that there are no fundamental usability problems, although several suggestions for improvements came up during the workshops.

The participants also liked the way it can give them lots more information than they currently have and the basic philosophy of democracy which empowers the individual and where nobody is in charge but has all possibility to contribute (currently they often feel out of control because there is no possibility to easily contribute to a shared knowledge base like e. g., the intranet).

At the same time we identified some areas of concern. The participants stated that it might be difficult for some colleagues to identify areas of expertise with which they do not feel comfortable being identified; and it would be important that a person tag is time-bound, so people do not feel they are making a completely open-ended commitment. Another concern was that some practitioners may abuse the system – e. g., 'lazy' colleagues may resist entering details about themselves and may tag others with expertise they may have (to deflect additional queries). Thus editing the own profile, i. e. removing tags that are inappropriate or in order to avoid being contacted, is important.

In contrast to the organization's focus at the beginning, it was also mentioned that the system could increase the efficiency of working with outside agencies by allowing them search for personal advisers' expertise easily and quickly. On the other hand, the participants expressed concerns about sharing whole people tagging information with other services in general because they think that this might increase their workload too much.

8.4. Conclusions

The key lesson from the formative evaluation was the important interdependency between the technology introduced and the organizational and team culture: the people tagging tool can be a catalyst for (longer term) organizational development (and it has been viewed as such both in the focus groups and in the trials at Connexions Northumberland), but also the tool has to be adapted to the cultural characteristics of its target context.

While during the development, we had hoped that we could come up with a design framework that clearly links design decisions to cultural characteristics. However, it has turned out that this approach did not account for the complexities of cultural aspects and their development. From the evaluation results, we believe that a more appropriate approach is to develop guidelines for a system-culture fit, showing potentials and risks of certain design and configuration aspects that feed into the moderated introduction process. This also complies with organizational development best practices and promotes higher acceptance.

From a more global perspective, the evaluation has shown that users see great potential in combining people tagging with bookmarking. While originally the introduction at Connexions Northumberland mainly was supposed to focus on the tagging of people, the tagging of web resources, which was additionally technically available, was seen of equal importance – particularly reusing the same vocabulary for both. It was seen as a natural extension to have furthermore a content creation, maintenance, and quality rating tool, where it was considered important that those extensions shared "the same spirit", i. e. , a participatory, democratic – while still retaining some form of assurance about the quality of the collectively developed knowledge representations. Particularly, the evaluation has helped to solve an open question whether shared vocabularies across the different aspects of either organizing people or documents make sense from a user perspective – they actually want to have it.

Part IV.

Conclusions

9. Related Work

In the following we will discuss technical and application-oriented approaches related to our work, not yet discussed in the foundation part. We will first look at work related to social semantic software with a focus on social semantic bookmarking and semantic people tagging whilst we consider the latter one from the more general point of view of competence management. Finally, we look at work similar to our ontology maturing perspective.

9.1. Social Semantic Software

Social Semantic Software or Semantic Social Software combines social software with semantic technologies either to improve the creation of semantics by using the ease and collective intelligence of social software or to improve social software by the enrichment with machine-interpretable meta data.

Especially the former one harnesses collective intelligence for the creation of formalizations to tackle the Chicken-Egg problem of systems using semantic technologies that promise great functionality only once a large amount of knowledge is formalized and, however, stumble upon only few people willing to invest the time and money needed until this great functionality is visible (Hendler, 2008; Berners-Lee et al., 2006; Zacharias and Braun, 2008).

They strive to tackle this challenge by relying on three core properties (Breslin et al., 2009):

1. **Distribution** of effort over a group of people – so the additional effort for the individual user is small.
2. Reducing the effort needed for **formalizations** to build more powerful applications (compared to knowledge engineering tools) through the use of simple user interfaces and the use of knowledge representation languages of limited complexity.
3. **Voluntary** and unpaid contributions to reduce the cost of creating these formalizations.

Originally being two different perspectives and strands of development that Schaffert (2006b) describes as *Socially Enabled Semantic Web* and *Semantically Enabled Social Software*, they are actually used in both ways.

On closer inspection, we can identify seven broad classes of systems based on recent research (Mazarakis et al., 2011; Mika, 2007; Breslin et al., 2009):

- **Social Semantic Tagging Systems:** Based on the observation that a large number of people are successfully creating structured data with tagging applications, these approaches try to extend these systems with a bit more structure and formality (q. v. Section 5.2). The exact nature of the added formality differs between systems; it includes relations between tags (such as broader, narrower etc.), the use of concepts with attributes instead of tags (allowing e. g., alternative labels, descriptions, multi-word names) and even the introduction of arbitrary relations between tags or concepts. Examples for such systems are GroupMe (Abel et al., 2007) or Fuzzzy[1] (Lachica et al., 2008). We will discuss related social semantic tagging approaches for bookmarks in Section 9.1.2.

- **Social Semantic Wikis:** The second group of systems starts from the observation that people are spending large amounts of time creating semi-structured data in wikis. These systems then try to give people the tools and the support such that they can create and maintain structured information alongside wiki pages. The kind of information supported differs between systems; ranging from the typed links and attributes to full-fledged ontology editing support. The Semantic MediaWiki[2] (Krötzsch et al., 2007), IkeWiki (Schaffert, 2006a), Freebase[3] and MyOntology (Siorpaes and Hepp, 2007) are example for these kinds of systems. Because of their closeness to our approach, we will discuss them in more detail in Section 9.1.1.

- **Semantic Blogs & MicroBlogging:** Semantic Blogs and MicroBlogging tools aim for adding richer structure to blog posts, comments and topics and links between these and facilitating querying. semiBlog (Möller et al., 2005), semblog (Ohmukai et al., 2004), SMOB (Passant et al., 2008) are examples for these kind of applications.

- **Social Semantic Multimedia Sharing:** Besides textual content, it's also multimedia content such as images, audio, or video that users produce, annotate and share. Some examples are ImageNotion[4] (Walter, 2010) for images, ZemPod (Òscar Celma and Raimond, 2008), DBTunes[5] for audio, or SemaPlorer (Schenk et al., 2009).

- **Social Semantic Search:** Breslin et al. (2009) also consider search engines like Hakia[6], Swoogle[7] (Ding et al., 2004) and Powerset[8] as social semantic web applications. This classification is based on their aggregation and use of collective intelligence of content creators for retrieval. Woogle[9] (Happel, 2009) and TrueKnowledge[10] can be seen as more 'social' since they allow direct user changes. However, except for Woogle, these systems are rather concerned with the combination and use of knowledge and not so much with its creation.

- **Semantic Social Networks:** In semantic social networks structured content is created by the participants and semantic technologies (like FOAF[11]) are used to facilitate the exchange between systems. There are some popular social networking sites already providing FOAF export functionality, e. g., LiveJournal[12], Hi5[13] or friendfeed[14].

[1] http://fuzzzy.com
[2] http://semantic-mediawiki.org
[3] http://www.freebase.com
[4] http://imagenotion.com
[5] http://dbtune.org
[6] http://hakia.com
[7] http://swoogle.umbc.edu
[8] http://www.powerset.com
[9] http://www.teamweaver.org/wiki/index.php/Woogle
[10] http://www.trueknowledge.com
[11] http://www.foaf-project.org
[12] http://www.livejournal.com
[13] http://hi5.com
[14] http://friendfeed.com

- **Semantic Games with a Purpose:** The last group is inspired by the success of the gwap platform[15], based on the "Games with a Purpose" paradigm (von Ahn, 2006). This platform offers games that – as a side effect – also create structured data for the computer. OntoGame[16] is the approach that realized this for the Semantic Web (Siorpaes and Hepp, 2008). A recent initiative of the European INSEMTIVES project consortium[17] is the SemanticGames community portal[18] that aims to collect games for the creation of semantic content. From a motivational and feedback point of view, these applications stand apart from the other approaches since they aim mainly for providing fun, with semantic content creation as a side effect (Siorpaes and Simperl, 2010).

We will now look at related Social Semantic Wikis and Social Semantic Bookmarking Systems because of their closeness to our approach in a more detail. To our knowledge, there are not any related approaches using semantics for tagging people. General people tagging approaches will be discussed in Section 9.2.3.

9.1.1. Social Semantic Wikis

There has been and still is quite a lot of research and development on bringing together the Social and Semantic Web in (Social) Semantic Wikis, ranging from working prototypes to meanwhile established and large scale deployed applications like the Semantic Media Wiki. Indeed, we can again identify the two different strands of how wikis and semantic technologies are brought together. One strand aims for enriching wikis with semantic annotations, the other strand to use wikis and its collaborative principles for ontology development; being Platypus (Campanini et al., 2004) one of the first ones for the former strand or Wiki@nt (Bao and Honavar, 2004) for the latter one. Buffa et al. (2011) describe the two approaches with *"the use of wikis for ontologies"* and *"the use of ontologies for wikis"* and provide a broad overview and feature comparison of current approaches. Another overview focusing on the latter ones, i. e. using wikis and their principles for collaborative ontology development, provides Weller (2010).

- **Semantic Media Wiki:** As mentioned before, the Semantic Media may be the most established and popular one (Krötzsch et al., 2007). It is based on the MediaWiki engine (same as Wikipedia). The semantics to add are typed links between two wiki pages – to specify the relation between both pages – and attributes – data related to a page can directly be attached. The user can freely name the typed links and attributes when they edit a page, i. e. they are directly put into the wiki page code. For instance, *Berlin [[is capital of::Germany]]* inserts the typed link "is capital of" between the current page and the page Germany while *Berlin has a population of [[population:= 3,450,889]]* defines the attribute "population". The semantics are translated into RDF when saved and used for enhanced navigation and search. In this way, wiki categories become classes and pages their instances. However, this freedom also hampers its usage as one use may use instead of "population" the attribute "inhabitants", being different when querying. Similarly, there is no further support when creating the wiki code and the users have to remember each concept, property and attribute.

[15] http://www.gwap.com/gwap
[16] http://www.ontogame.org
[17] http://www.insemtives.eu
[18] http://www.semanticgames.org

- **Halo:** In order to tackle the difficulties in SMW, the Halo extension[19] has been developed as part of the Halo project[20]. Its main components are a semantic toolbar, ontology browser and graphical query interface. The semantic toolbar is for easier annotations so that users may not edit anymore the wiki code but can create and modify annotations and, for example, also new properties via input forms that are additionally equipped with auto-completion. With the ontology browser, the users can navigate through the category and property tree and related instances or property information and edit the elements. Recently, gardening support has been added that provides suggestions where to work on because one of eight issues, such as general consistency problems or missing annotations, has been identified. The Halo extension is meanwhile part of the semantic enterprise wiki bundle SMW+ distributed by ontoprise[21].

- **ACTIVE Ontology Editor:** Similarly recently, an ontology editor[22] has been developed as an extension to SMW at the University of Innsbruck as part of the European ACTIVE project (Bürger et al., 2010). It works with a meta model that consists of vocabularies, categories, properties and elements where a vocabulary comprises the others and elements are "normal" wiki articles. The editing of ontology elements is based on forms. For instance, when creating a category, the users have to provide a name and add it to a vocabulary. Further, super-categories, meta information like synonyms and a description may be provided and either existing or new properties may be added.

  On the ontology elements pages, the related vocabularies, categories, properties and elements are provided as links. Categories may be shown in a category tree that is drag'n'drop enabled to modify the hierarchical structure. In contrast to the Halo extension, it is not possible to create property hierarchies. There is a special functionality called "knowledge repair" that is intended to provide gardening support. It provides an overview on category statistics, identifies category cycles, redundant links between categories or categories with similar properties and elements with similar names.

  Whilst the Halo extension is quite mature, the ACTIVE Ontology Editor extension still is in July 2011 in its early prototypical status, as some features are buggy. For instance there is auto-completion support for existing ontology element, however during input the auto-completion box pops up somewhere on the page. Moreover, there is no evaluation information available, yet.

- **AceWiki:** AceWiki (Kuhn, 2010) pursues a different approach and aims to overcome the "subject-predicate-object" structures of the other wikis by using controlled natural language for ontology engineering. It uses Attempto Controlled English[23] (ACE), a subset of standard English, to formulate formal sentences that are translatable into OWL. The AceEditor[24] has been integrated into the wiki as a predictive authoring tool to create and modify ontological statements as such sentences. It works with menu boxes to select a particular type of word to build up a sentence. The type of words are "proper names" for instances, classes are presented as "nouns" and properties e. g., as "transitive verbs", "of-constructs" or "transitive adjective". They have to be created beforehand with the form-based word creator

[19] http://smwforum.ontoprise.com/smwforum/index.php/Halo_extension
[20] http://www.projecthalo.com
[21] http://smwforum.ontoprise.com/smwforum/index.php/SMW+
[22] http://smw-active.sti-innsbruck.at
[23] http://attempto.ifi.uzh.ch
[24] http://attempto.ifi.uzh.ch/webapps/aceeditor

tool. The type of word menu boxes are complemented by “function words” such as ’every’, ’no’, ’if’ etc. The individual boxes are only shown if possible for continuing the current sentence. Every ontology entity then is represented by a wiki page. The wiki page shows related sentences and collected information. Users may edit or delete sentences, retract or reassert them or leave a comment. Additionally, queries related to the entity may be used on the wiki page. Unfortunately, this approach focuses only on ontology building neglecting the aspects of work integration and application integration.

- **IkeWiki:** IkeWiki (Schaffert, 2006a) is a Java based web application that provides besides a wiki page content editor and an AJAX based editor for adding semantic annotations. With the AJAX editor users can add textual metadata to a page, associate types with a page and type incoming and outgoing links. Incoming and outgoing refences to related pages are provided in a side bar for further navigation. The editors are mainly intended for knowledge engineers and researchers; i. e. lay persons are supposed to fill the wiki with texts, maybe multimedia contents and simple relations which are then formalized by knowledge engineers. The developers have stopped working on IkeWiki and instead started with its successor KiWi Wiki as part of the KiWi platform that aims for providing a platform to build social semantic software upon (Schaffert et al., 2009).

- **SweetWiki:** SweetWiki (Semantic WEb Enabled Technology Wiki) has been a Java based semantic wiki research prototype from INRIA Sophia-Antipolis and uses the CORESE[25] Conceptual Resource Search Engine and the SeWeSe[26] Library (Buffa et al., 2011). SweetWiki is different from the other approaches as it supports social tagging; i. e. users can tag pages, images or attached files with freely chosen keywords that might be later organized with subClassOf and seeAlso relations. During the tagging the user is supported with suggestions of existing tags together with the information of their related category and how many other pages are tagged with. In order to organize and maintain the tags the integrated SeWeSe ontology editor is reused. Similar to IkeWiki, normal users are supposed to edit wiki pages and tag them with keywords and knowledge engineers subsequently check the keywords of the users and (re-)organize them by adding relationships.

There are further related semantic wiki projects we won’t discuss in more detail. To hide the wiki code and syntax and to avoid user mistakes, many approaches rely on form-based mechanisms. One example is the OntoWiki[27] developed by the AKSW research group at the University of Leipzig that aims to support the development of knowledge bases. It provides different views on instance data available with an information map like widget visualization and features like inline editing for RDF content. To support collaboration, user can keep track of changes and discuss any entity of the knowledge base. Another example is the MoKi Enterprise Modelling Wiki (Ghidini et al., 2009, 2010) developed as part of the European APOSDLE project that particularly aims for supporting the collaborative development of enterprise models and processes. The MoKi wiki is also based on Semantic Media Wiki. Ontologies may be imported and exported in OWL. An additional feature is a graphical browsing visualization that allows to browse through the isA, isPartOf hierarchies and processes and individuals and to edit these. Meanwhile, the focus has been set to process modeling. To that end, it provides a special graphical process editor.

[25] http://www-sop.inria.fr/edelweiss/software/corese/
[26] http://www-sop.inria.fr/teams/edelweiss/wiki/wakka.php?wiki=Sewese
[27] http://ontowiki.net

9.1.2. Social Semantic Bookmarking

Subsequently, we present similar approaches that enhance social bookmarking with semantics respectively let users create semantics in their usage. That means, the users may extend their folksonomy by adding semantic definitions or descriptions. The main distinctive, besides scope and degree of utilization of the semantics, is the degree of freedom the users are granted to contribute by themselves.

We analyzed and compared eight different systems: BibSonomy[28], Fuzzzy[29], GroupMe! [30], Twine[31], ZigTag[32], Faviki[33], gnirz[34] and Annotea[35]. BibSonomy, launched at the beginning of 2006, is the oldest system and Faviki, launched at the beginning of 2008 the youngest[36].

- **BibSonomy:** BibSonomy (Hotho et al., 2006b) is a system for the management of bookmarks of internet resources and publication entries. BibSonomy is a research project of the Knowledge and Data Engineering Group of the University of Kassel, Germany, that has launched the system at the beginning of 2006. BibSonomy offers functionality similar to that of well-known social bookmarking services but also functionality specifically tailored towards academics – e. g., sophisticated support for uploading and exporting bibliographic information in bibtex format. At its core, Bibsonomy differs from social bookmarking services by additionally offering users the possibility to create broader/narrower relations between tags. However, tag relationships are only local, i. e. , each user can (and has to) maintain his own relationships and cannot profit from others' contributions in that respect.

- **Fuzzzy:** Fuzzzy (Lachica et al., 2008) is a system for managing bookmarks of web pages and ISBN numbers. Fuzzzy is developed within the PhD project of Roy Lachica at the University of Oslo and its development started at the end of 2006. It is based on Topic Maps technology. Besides hierarchical and related tag relations, the users can choose of 22 specific predefined association types to link tags. These tag relations apply to the whole system and are editable by other users. Another main concept is voting for gardening and maintenance: the users can vote on bookmarks, tags a bookmark is annotated with, relations between tags, and users.

- **GroupMe!:** GroupMe (Abel et al., 2007) attempts to bridge the gap between the Semantic Web and Web 2.0 with an RDF based social bookmarking application. The Semantic Web Group at the University of Hannover in Germany has been developing GroupMe! since 2007[37]. The main unique functionality of GroupMe! is the extension of the tagging idea with the concept of 'groups' (collections): all annotated Internet resources (websites, music, videos, photos, and news feeds) can be organized into groups. These form another level of information that can be used for browsing and search.

[28] http://www.bibsonomy.de
[29] http://www.fuzzzy.com
[30] http://groupme.org
[31] http://twine.com
[32] http://zigtag.com
[33] http://faviki.com
[34] http://code.google.com/p/gnizr
[35] http://www.annotea.org
[36] Indeed the Annotea system is older but lacks the 'social' part of Social Semantic Bookmarking; q. v. Section 3.3
[37] Unfortunately, the service has not been available for some months.

- **Twine:** Twine has been a system that supported organizing and sharing bookmarks of web pages, images, videos, products, and books. Twine is a commercial product ran by Radar Networks has been publicly usable since October 2007 but has been stopped in May 2010. The main organizing principle of Twine was so called 'twines', a kind of user interest group. The users could join or create such interest groups in order to share bookmarks. Through bookmarks grouped in the same twine group, tags were set into relation. Another functionality of Twine was the faceted filtering of search results; i. e. the users could filter by seven categories like tags, people, places, or item type.

- **ZigTag:** ZigTag allows tagging and organizing bookmarks of Internet resources. ZigTag is run by ZigTag Inc., a small Canadian company, and had its beta release in April 2008. In contrast to the other approaches, the users can annotate their bookmarks with predefined tags, i. e. they choose from a given list of tags with a specific meaning. The users can additionally add their own tags for private usage. They cannot add further semantics. Tags are set into relation by the system.

- **Faviki:** The Faviki, launched in May 2008, is developed by the Serbian Web designer Vuc Milicic. The system distinguishes itself by relying on Wikipedia terms for the annotation of bookmarks, i. e. users annotate with tags that are titles of Wikipedia articles. Whilst Faviki users first have not been allowed to use tags any other than contained in Wikipedia[38], this has been relaxed for towards using websites as "definition" in general or adding their own label as synonym to existing tags. However adding semantic relations between tags is still not possible.

- **ginzr:** ginzr is a system for managing bookmarks of Internet resources using tags and folders. ginzr provides the users the possibility to define broader/narrower, related, and member-of relationships between tags according to the SKOS vocabulary. Additionally, the users can add geospatial information to the bookmarks. The system has been published as open source application under Mozilla Public License 1.1 by the Image Matters LLC company in March 2008. There is no publicly available installation. The system can be downloaded and hosted on one's own. However, despite several attempts, it was not possible for us to get the system to work properly. Therefore, we excluded this system from the overall comparison as we could not evaluate the system provider's information.

- **Annotea:** Annotea (Koivunen, 2006) is a metadata standard for semantic web annotations, it is implemented in a number of tagging tools and server applications. Annotea and its implementations have been developed by the W3C. Annotea differs from other approaches to social tagging in its emphasis on standards on decentrality, that it has sharing of bookmarks among services build in from ground up. However, Annotea has not been under active development for some years and for this reason was excluded from the overall comparison.

The first analysis was done in 2009 and published in Braun et al. (2009b). We have updated the analysis in June 2011 and applied our conceptual design framework and scaffolding methods for comparing the tools with SOBOLEO. Because Twine has been stopped in May 2010 and the GroupMe! service has not been available for some months, we made the comparison for both tools based on the data from 2009.

[38]meanwhile taking concepts from DBpedia

9.1.2.1. Comparison

The analyzed systems show a wide range of how social semantic bookmarking can be realized and supported. Table 9.1, 9.2, 9.3, 9.4 and 9.5 give a comparative overview of the systems and SOBOLEO regarding our design framework and scaffolding methods.

One major distinctive aspect is the degree of freedom of what users can contribute and what is done automatically. Thus, Fuzzzy, the most similar system to SOBOLEO, highly relies on the user community. It also aims at supporting the collaborative development of ontologies. To that end it provides its users a big latitude to freely define and edit tags and relations. Besides hierarchical and related tag relations, the users can choose of 22 specific predefined association types to link tags. However, mutual editing is nevertheless often limited so that it is the creator or special users with extra privileges who can modify knowledge elements. Unfortunately, there is also no information about evaluation available.

BibSononomy and GroupMe! limit their users to one specific type of relations they can add. The commercial systems ZigTag and Twine aim in offering highly automated support. The semantics in these systems are internally derived and predefined. The users have only little influence on the semantics itself. At the other end Faviki relies on standardization based on Wikipedia/DBpedia, targeted that users select from these for annotation.

Another characteristic is the scope and impact of the (user added) semantics. Within BibSonomy user added relations are only locally relevant and set on top of one user's tags. That means that only the creator can make use of the extra semantics e. g., for search. The other users can not benefit or even view the relations. They are only available in an aggregated overview. Within Fuzzzy changing tags and semantics have a system wide effect on every user. The additional semantics are mainly used to support better retrieval and tag-based navigation: e. g., on the retrieval of all bookmarked pages for one tag, the bookmarks associated with its sub-tags are also returned. All kinds of relations are used to facilitate navigation from one tag to another. In comparison, the SOBOLEO framework also makes use of the semantic information to augment the full text search of the contents of the bookmarked Internet resources.

Formal semantics do not seem to play an important role with the current generation of systems; to the author's best knowledge none of the systems employs an actual inference engine. This confirms the results from our evaluation that lightweight ontologies are very often sufficient for such use cases.

There is a big interest in particular to tackle the problem of how tagging data can be exchanged between systems. All the systems are offering (often multiple) ways to export the created resources and their annotations. However, the created semantics mostly are lost during the export. Besides SOBOLEO, it is only the GroupMe! system that allows to export the semantic data in RDF. Altogether there is considerable disagreement about the most important features and about suitable formats to exchange the tagging data.

Table 9.1.: Comparing Social Semantic Bookmarking Systems regarding Users

Users										
Design Aspect		General Options	BibSonomy	Fuzzzy	GroupMe!	Twine	ZigTag	Faviki	SOBOLEO
Types & Properties	Role	Tagger, creator, provider, owner, gardener, administrator	•Tagger/Creator •Administrator •Author	•Tagger/creator •Parliament member (elected group with extra privileges)	•Tagger/creator	•Tagger/creator •Twine member	•Tagger/creator	•Tagger/creator	•Tagger •Taggee •Administrator
	Social connectivity	(typed/directed) links, groups, none	•User groups •Directed links: friends, followers			•'twines' as user groups	•Groups •Directed links: friends	•Directed links: friends, followers	•Collaboration space as user group

Table 9.2.: Comparing Social Semantic Bookmarking Systems regarding Resources

Resources									
Design Aspect		General Options	BibSonomy	Fuzzzy	GroupMe!	Twine	ZigTag	Faviki	SOBOLEO
Types & Properties	Type of object	Bookmarks, people, multimedia objects, events etc.	•Bookmarks •Bibliographic information	•Bookmarks •Books (via ISBN) •Audio files (via URL) •Video files (via URL) •Events	•Bookmarks •RSS feeds •Video Files (via URL) •Images (via URL) •Audio files (via URL) •Groups	•Bookmarks •Documents	•bookmarks	•Bookmarks	•Bookmarks •Office Documents •Discussions •People
	Source of material	User-contributed, system-supplied, global	•Bookmarks: global •Bibliographic information: global + user-contributed	•Events: user-contributed •Others: global	•Global •Groups: user-contributed	•Bookmarks: global •Documents: user-contributed	•global	•global	•Bookmarks: global •Office Documents: user-contributed •People: user-contributed or system-supplied •People: internal contacts and/or external contacts
	Resource connectivity	Links, groups, none	•Through user groups	•'MyFavourites'	•groups	•'twines'			•Bookmarks and office documents may be linked with discussions
Rights & Control	Creation	Any user, restricted group, system	•Any user	•Any user	•Any user	•Any user	•Any user	•Any user	•People: any user, space member or system •Others: any user or space member
	Deletion & Modification	Any user, resource creator/provider, restricted group, system, none	•Creator/provider + system	•Creator/provider, parliament	•creator	•Creator	•creator	•creator	•People: administrator •Others: any user or space member
	Visibility	Any user, resource creator/provider, restricted group	•Any user or, resource creator/provider or restricted group (friends or specific user group)	•Any user	•Any user	•Any user or creator or twine members	•Any user or creator	•Any user or creator	•Any user or space member
Support & Tools	Content Control	File type filter, word filter, web filtering etc.	•Spam filtering a posteriori	•File type filter for audio and video files					•File type filter for office documents: PDF, MS Office 2004 and Office Open XML •"People" filter by email address of specific domain
	Mining techniques		•Title suggestion •Content description suggestion	•Title suggestion •Content description suggestion •Full text indexing	•Title suggestion •Content description suggestion	•Title suggestion •Content description suggestion •Full text indexing	•Title suggestion	•Title suggestion	•Title suggestion •Full text indexing

Table 9.3.: Comparing Social Semantic Bookmarking Systems regarding Ontology

Design Aspect			General Options	BibSonomy	Fuzzzy	GroupMe!	Twine	ZigTag	Faviki	SOBOLEO
			Ontology							
Types & Properties	Semantic Tags		Labels, description, multilingual, (multiple) connected	•Label •Multiple connected •System tags: •for:<groupname> •send:<username> •myown •sys:relevantFor:<groupname> •sys:author:<authorname> •sys:entryType:<bibtexEntryType> •sys:year:<year>	•Label •Description •Multilingual •Multiple connected	•Label •Multiple connected	•Label •Types of tags: people, place, organization, types of items	•Label •'zigtag' with description	•Label •Synonym •Description •Multilingual	•Preferred, alternative and hidden labels •Descriptions •additionally:competence types, instances, compositions •Multilingual •Multiple connected
	Relations		Types, reflexive, symmetric, transitive	•Broader/narrower •Related •Similar	•Hierarchical, related + 22 association types •Symmetric	•isInGroup •relatesToConcept •taggedInContext	•Relatedness via "twines"	•Synonymous •related	•Based on DBpedia relations •Relatedness	•Broader/narrower, symmetric + transitive •Related, symmetric •additionally: broader-/narrowerGeneric/-Instantive/-Partitive
Rights & Control	Creation	Who	Any user, gardener, restricted group, system	•Labels: any user •Broader/narrower: any user •Related, similar: system	•Labels: any user •Description: creator; any user can "hijack" the very •Relations: any user	•Labels: any user •isInGroup: any user • taggedInContext: system	•Labels: any user •Types of tags: system •Relations: system	•Labels: any user •Zigtags with description: system •Relations: system	•Labels + description: any user (for every new tag, a URL has to describe it) or system based on Dbpedia •Synonym: any user	•Broader/narrower/related: Any user or space member •Gerneric/Instantive/Partitive: automatically set by system
		Level of detail	Tags, tag meaning, relations, relation type							
	Deletion & Modification	Who	Any user, creator, gardener, restricted group, system, none	•Labels: creator •System tags: creator + system •Relations: creator •System relations: system	•Labels: creator, parliament •Description: creator, parliament •Relations: creator, parliament	•Labels: creator •isInGroup: creator •taggedInContext: system	•creator	•Labels: creator	•Synonym: creator	•Broader/narrower/related: Any user or space member •Gerneric/Instantive/Partitive: automatically set by system
		Level of detail	Tags, tag meaning, relations, relation type							
	Visibility	To whom	Any user, tag creator, restricted group	•Labels: any user •Broader/narrower: creator •Related/similar: system	•Labels: creator or any user •Description: any user •Relations: any user	•Labels: any user •Relations: any user	•Any user or creator or twine members	•any user or creator	•Synonym: creator •Others: Any user	•Any user or space member
		Level of detail	Tags, tag meaning, relations, relation type							
Support & Tools	Blacklist		Blocking , deleting; a priori , a posteriori checking	•Spam check a posteriori						
	Tagging Support		Blind tagging, viewable tagging, suggestive tagging; tags from vocabulary, automatic creation	•Viewable and suggestive tagging •Automatic creation of tags + broader/narrower relation	•Viewable and suggestive tagging	•Automatic creation	•Suggestive tagging •Automatic creation	•Suggestive tagging •Automatic creation	•Viewable and suggestive tagging	•Viewable and suggestive tagging •Automatic creation
	Ontology Enrichment		Exploiting lexical resources, statistical analysis	•Co-occurrence analysis to identify relatedness, cosine similarity for similar tags	•Identifies overlapping tags, empty tags and tags without relations for gardening activities		•Content analysis to identify type and relatedness of tags		•Automatic translation •Co-occurrence analysis to identify relatedness	•For gardening recommendations

Table 9.4.: Comparing Social Semantic Bookmarking Systems regarding Tag Assignments

				Tag Assignments						
Design Aspect			**General Options**	**BibSonomy**	**Fuzzzy**	**GroupMe!**	**Twine**	**ZigTag**	**Faviki**	**SOBOLEO**
Types & Properties	**Aggregation**		Multiple assignments, one time assignment	•Multiple assignments	•Multiple assignments	•Multiple assignments	•Multiple assignments	•Multiple assignments	•Multiple assignments	•Multiple or one time assignments
	Semantics of tagging		Without any further semantics, with specific system-/user-defined semantic relationship	•Without any further semantics: swrc::keywords	•'is about', 'has contributor', 'mentions' 'is a type of', 'is', 'has'	•Without any further semantics: keyword	•Without any further semantics	•Without any further semantics: commonTag:tagged	•Without any further semantics: commonTag:tagged	•Without any further semantics: commonTag:tagged •People: with or without weighting
Rights & Control	**Creation**	**Who**	Any user, resource owner/provider, restricted group, system based on mining techniques	•Resource owner/provider; others' resources may be copied	•Any user for whole semantics of tagging	•Group creator or any user	•Any user or twine members	•Any user	•Any user	•Any user or space members •People: any user or space members or system, with or without weighting •People: self-tagging
		Level of detail	Semantics of tagging							
	Deletion & Modification	**Who**	Any user, tagger, resource owner/provider, gardener, restricted group, system based on blacklist, none	•Tagger •System based on spam filtering	•Tagger for whole semantics of tagging	•none	•Tagger	•Tagger	•tagger	•Multiple assignments: any user or space members •One time assignments: tagger or taggee inc. weighting
		Level of detail	Semantics of tagging							
	Visibility	**Time**	Need for approval, blacklist blocking, immediate visibility, validity date	•Immediate visibility	•Immediate visibility	•Immediate visibility	•Immediate visibility	•Immediate visibility	•Immediate visibility	•Immediate visibility
		To whom	Any user, tagger, restricted group	•Any user or, resource creator/provider or restricted group (friends or specific user group)	•Any user or tagger for private tags	•Any user or creator or twine members	•Any user, creator or twine member	•Any user or creator	•Any user or resource creator/provider	•Any user or space members
		Level of detail	Aggregated and anonymous, tagger is visible, semantics are visible, details visible for restricted group	•Tagger is visible	•Tagger is visible •Semantics are visible	•Tagger is visible	•Tagger is visible	•Tagger is visible or anonymous	•Tagger is visible	•People: aggregated and anonymous; semantics visible •Others: tagger is visible
Support & Tools	**Blacklist**		Blocking , deleting; a priori , a posteriori checking	•Spam detection a posteriori						
	Tagging support		Blind tagging, viewable tagging, suggestive tagging; seeding from external sources	•Viewable and suggestive tagging •Auto-completion	•Viewable and suggestive tagging •Auto-completion		•Suggestive tagging	•Auto-completion •Suggestive Tagging	•Auto-completion •Viewable and suggestive tagging	•Auto-completion •Viewable and suggestive tagging
	Mining techniques		User behavior analysis, content mining	•FolkRank is based on user behavior analysis	•User behavior analysis for tagging support					•User behavior analysis and content mining for tagging support

Table 9.5.: Comparing Social Semantic Bookmarking Systems regarding Scaffolding Methods

			Scaffolding Methods						
Design Aspect		General Options	BibSonomy	Fuzzzy	GroupMe!	Twine	ZigTag	Faviki	SOBOLEO
Seeding	General	• Import functionality for ontology, resource and user information, and annotations • Candidate concepts extraction from texts	•Import: bookmarks from browser, from other websites: delicious and of Bibtex, RIS, EndNote files			•Import: bookmarks from browser and from other websites: delicious, digg	•Import: bookmarks from browser and from other websites: delicious	•Import: bookmarks from other websites: delicious •Import: uses DBpedia structure	•Import: of SKOS ontology from file
Evolutionary Growth	General	•Easy adding of new knowledge •Automatic concept creation •Annotation bookmarklet •Everywhere add/edit options •Tag Recommendations based on resource content •Make aware of new knowledge •Automatic publication •Search, Browse, Feed notification etc. •System feedback on own contributions •Make social use visible	•Automatic creation of tags + broader/narrower relation •Annotation bookmarklet •Everywhere add/edit options •Automatic publication (depending on visibility) •Limited effectiveness •BURST-feeds •Newest bookmarks on front page •Most recently used tags on front page •Notification based on system tags •Feedback on how often bookmarks and tags are used by others	•Annotation bookmarklet •Everywhere add/edit options •Automatic publication (depending on visibility) and effectiveness •RSS-feeds •Newly added tags and bookmarks •Bookmarks annotated with one's favorite tags •Feedback on who has bookmark or tag as favorite	•Automatic creation •Annotation bookmarklet •Everywhere add/edit options •Automatic publication (depending on visibility) and effectiveness •RSS-feeds •Latest group activities as Dashboard News	•Automatic creation •Annotation bookmarklet •Everywhere add/edit options •Content analysis to identify new tags and type for recommendations •Automatic publication (depending on visibility) •Limited effectiveness •RSS-feeds •Newest groups on front page •Latest twine group activities on front page	•Automatic creation •Annotation bookmarklet •Everywhere add/edit options •Automatic publication (depending on visibility) •Limited effectiveness •RSS-feeds •Newest bookmarks on front page •Feedback on who saved bookmark	•Annotation bookmarklet •Everywhere add/edit options •Automatic publication (depending on visibility) •Limited effectiveness •RSS-feeds •Newest bookmarks on front page •Popular tags on front page •Feedback on how often bookmarks and	•Automatic creation •Annotation bookmarklet •Everywhere add/edit options •Automatic publication and effectiveness •Tag recommendations based on content using external services •ATOM-feeds •Recent ontology changes in editor •System feedback
	Nudge	•Auto-completion •Disambiguation •Typing error correction •Blacklist •Tag recommendations from vocabulary; .e.g. own popular tags, others' tags •Social norms, e.g. naming convention •Training, Tutorials, Examples	•Auto-completion for tagging •Tag recommendations based on FolkRank, inc. own used tags and others' tags •Tutorials + help	•Auto-completion for tagging •Disambiguation for tagging + editing •Tag recommendations of often or recently used own tags •Help •Community guidelines	•Auto-completion •Help	•Auto-completion •Interactive help	•Auto-completion for tagging •Disambiguation for tagging + searc h • Tag recommendations of ,zigtags' from vocabulary based on content mining •FAQ	•Auto-completion for tagging + search •Disambiguation for tagging + search •Tag recommendations of recently used own tags •Help	•Auto-completion for tagging + search + editing •Disambiguation for tagging + search + editing •Tag recommendations from vocabulary •Manual
	Agree	•Discussions & Dialogues •Find the right discussion partners •Feedback mechanisms, e.g. rating, comments •Voting •Mutual editing/amending/ complementing •Revision management	•Rating as spam	•Comments on tags and resources •Voting on tags, relations , resources and tag assignments •Rating of resources •Limited mutual editing		•Forum •Rating of resources	•Discussing tags in forum •Comments on resources	•Rating as spam •Voting on the describing URL	•Chat in editor •Structured dialogs on resources, annotations and ontology •Rating of bookmarks and documents
Reseeding	General	•Easy vocabulary manipulation •History •Subscriptions to changes •Blacklist	•BURST-feeds	•Tag mapper •RSS-feeds	•RSS-feeds	•RSS-feeds	•RSS-feeds	•RSS-feeds	•Ontology editor •History of changes •ATOM-feeds
	Aggregate	•Make aware of (un-)used elements, discussion status •Identify inconsistencies, redundancies, or gaps •Recommendations for ontology enrichment •Knowledge completion based on usage data analysis	•Co-occurrence analysis to set relatedness, cosine similarity for similar tags •Title suggestion •Content description suggestion	•Identifies overlapping tags, empty tags and tags without relations for gardening activities •Title suggestion •Content description suggestion •Full text indexing	•Title suggestion •Content description suggestion	•Content analysis to set type and relatedness of tags •Title suggestion •Content description suggestion •Full text indexing	•Title suggestion	•Automatic translation •Co-occurrence analysis to set relatedness •Title suggestion	•Title suggestion •Full text indexing •Ontology gardening recommendations •User behavior analysis to fill profile •Expertise analytics for guidance

9.2. Competence Management

9.2.1. Competence Management in General

Traditionally, competence management approaches are conceived as top-down instruments (see e. g., Biesalski, 2006; Berio and Harzallah, 2005; Biesalski and Abecker, 2005). The basis are competency catalogs (Sicilia, 2005) modeled at irregular intervals by small expert groups and then to be used by the operational level in order to provide, update, and apply requirements and competency profiles.

However this method usually leads to communication and coordination problems between strategic and operational level. Schmidt and Kunzmann (2007) have proposed a closed-loop approach in which two-way communication between the different levels forms an integral (see Figure 9.1). This model is designed from a human resource development perspective. On the strategic level, the competence catalog and the requirement profiles for job roles are modeled in a continuous loop, taking into account corporate goals (in order to ensure that the catalog and the profiles are oriented toward the future) and feedback from the operational level. The operational level uses this vocabulary to describe the actual competency profiles of the individual employees. By comparing the actual competency profile with the requirements profile, it is possible to determine a competency gap, which can be addressed by development measures. Their outcomes should then improve work performance, which provides the indicators for setting up competency profiles, but also competency aspects which are not yet included in the competence catalog and thus have to be fed back to the strategic level.

Figure 9.1.: Integrated Model of Competence Management: A closed loop approach by (Schmidt and Kunzmann, 2007)

But even with a closed loop approach as outlined, there are still considerable problems when putting those approaches into practice. We will analyze in the following the competence modeling and diagnostics/assessment activities, which are in practice the most challenging ones.

9.2.2. Getting Competency Profiles

On the operational level, the most obvious problem is getting the competency profiles. One fundamental issue is that competencies cannot be measured, sensed, or observed directly. What we can observe is performance in various forms (Lau and Sure, 2002): assessment of learning outcomes, performance in every day job activities etc. All of these yield evidence from which a competency is usually deduced heuristically.

In practice, you can observe two approaches (Biesalski and Abecker, 2005): (1) self-assessment approaches in which employees themselves are asked to provide their competencies, sometimes mediated in a second step by their superior, and (2) external assessment approaches done by superiors or through formal assessment procedures.

While the latter approach is very expensive and cumbersome and thus can only be observed in limited areas, the first approach often fails because of missing motivation. This lack of motivation can be traced back to no immediate benefit for the employees. For instance, systems are hardly embedded into everyday work activities and have not proven their usefulness there. Or it can be even traced back to negative incentives; for instance, if you disclose your competencies, others will contact and perhaps disturb you or you will fear to appear not competent enough. As a result, employees might downplay or exaggerate their competencies as Becerra-Fernandez (2006) reports. Often, these competency profiles also do not contain information that is of high relevance to colleagues; for instance manually-updated repositories become particularly outdated (Mcdonald and Ackerman, 2000). Thus, recent and usually very specialized topics are not yet contained in the competency catalog because of the long update intervals.

Several studies address this problem by automatically extracting profile information from data the user generates in her daily work; e.g. from publications (Crowder et al., 2002), documents (Reichling et al., 2007) or community created contents in the Web (John and Seligmann, 2006; Breslin et al., 2007). Ley et al. (2010) propose a competence performance approach that derives competencies from executed tasks. In this approach, a task competency matrix is created together with domain experts. This matrix relates a set of tasks, e.g. required for a position, to a set of competencies needed to fulfill these tasks successfully. Based on this model, the system can infer a user's competency from her successful performance of a task in her daily work.

9.2.3. Social Tagging for Gathering Competence Information

Web 2.0 developments, mainly social networking approaches, have also brought forth solutions for describing and augmenting employee profiles from the purpose of those profiles for expert finding and community formation. These platforms are mainly based on the self-promotion paradigm: People can represent themselves with a profile and indicate their connections to other users. Further, in some of these approaches, the principle of social tagging and bookmarking is transferred to people; for instance Xing or theNTSH or Tagalag – the latter two meanwhile gone offline – were some of the first systems that allowed organizing your contacts with tags.

There also have been various people tagging applications on Facebook like Describe Me, Define Me or iDescribe (most of them have gone offline). These applications typically aim for entertaining rather than organizing and sharing knowledge on who knows what. The tagger stays anonymous and users can either choose a limited number of tags from a predefined list or create new tags to describe their friends.

An interesting approach focusing on the quality of tags and encouraging social connectedness has been Collabio. Collabio – short for Collaborative Biography – implemented ideas of Games with A Purpose (von Ahn, 2006) to tag people within the Facebook social network (Bernstein et al., 2009, 2010). Users can tag their friends in a game. Therefore, the users only see the tags assigned to a friend in an obscured tag cloud. When they start to describe the friend, guessed tags are uncovered and new tags are added to the tag cloud. For each tag, the users accumulate points equal to the number the tag is assigned to the friend. Only the friend him-/herself can see the whole uncovered tag cloud, who assigned which tag and delete tags if needed. However, self-tagging is not possible. To prevent the cold-start effect of a completely empty tag cloud, seed tags are used from a person's public profile.

In their evaluation, including a survey with 49 active users as respondents, the authors found that most tags capture affiliation, expertise, interests and hobbies. Uncommon tags describe miscellaneous and unusual information, nevertheless rated as fairly accurate descriptors. Whilst this evaluation shows that Collabio provides accurate and novel information about people, the authors mention three open design challenge: (a) motivate new users to join the game, (b) users stop using the system because of *tag exhaustion*; i. e. users do not know anymore what additional tags to add to their friends, (c) lack of semantics. In comparison with SOBOLEO, whilst it is not so much a problem of how to attract new users especially in an organizational context, we started to tackle the problem of how to keep users motivated, e. g., with feedback mechanisms. For (b) we provide tag suggestions based on the users' activities with the system and for (c) we provide tagging based on shared and continuous developed ontology.

For the enterprise context, IBM's Fringe Contacts (Farrell and Lau, 2006; Farrell et al., 2007b) was the first system that implemented people tagging. Within IBM's Fringe Contacts each employee can describe their colleagues, e. g., for contact management, or themselves, e. g., for self-presentation, by tagging them with arbitrary key words on their expertise and interests. Thus, step by step, a publicly visible tag cloud grows characterizing the individual employee. The authors indicate that this leverages network effects for setting up some sort of profile of the individual, and improves usefulness for the individual user of the system which, in turn, motivates to contribute. For instance, Farrell et al. (2007a) could state that tagging people was used to create communities.

Raban et al. (2011) from IBM Haifa studied self-tagging activity vs. tagging activity by and of others of a three-year-snapshot within their research enterprise employee directory. On the employees' profile two tag clouds display the tags the employee was tagged with and respectively used to tag others. Self-assigned tags are shown separately. The results show that users who tag themselves are the most productive contributers; both tagging themselves and others. Self-tagged users receive significantly more tags from other users. The study revealed that the more users tag themselves the more they get tagged by others or vice versa and the more people users tag, the more people tag them or vice versa[39]. However, when users tag themselves very extensively, they are tagged less by others and vice versa – there seem to be a saturation reached at 12-14 tags per user.

Razavi and Iverson (2009) extended OpnTag[40] (Iverson et al., 2008), an open source social bookmarking and note taking web tool, with a people tagging feature. The aim is to enhance relationship and personal privacy management for information sharing. By tagging people, users can categorize their contacts into different target groups to control

[39] Causality could not be inferred.

[40] http://sourceforge.net/projects/opntag

access to their personal information; e. g., to share a specific memo with all contacts tagged with 'java expert'. For each tag assignment, the user can specify its visibility; tagger only, taggee, people tagged with the same tag by the tagger, or anyone. Wang and Jin (2009) came up with a similar idea of using tags assigned to people within Fringe Contacts in order to selectively distribute messages in an automated way.

Thielen (2010) have recently analyzed to what extent the characteristics of social tagging systems are applicable for e-HRM tasks; especially to acquire information usable to augment and describe employees' competency profiles. The author presents a conceptual framework with profiles, tagger and tags as the identified dimensions and some basic characteristics. Whereas there are some overlaps with our design framework, our framework presents a more fine granular reflection of the design characteristics together with their impact. Thielen (2010) additionally provide an analysis of the reliability and validity of the competency related information. The author concludes that the absence of guidelines and rules and the lack of semantics that allow different interpretations of tags are the main disadvantages. Nevertheless, social tagging systems might be useful to gather more hidden or multi-perspective information.

Overall, there are different existing approaches that allow their users to tag each other, however the resulting employee profiles lack legitimation and commitment by the organization, especially with respect to the vocabulary used. The approaches do not provide support to overcome the gap and leverage the bottom-up topics to an organizational competences vocabulary. But that is a prerequisite for organizational competence management – ranging from team staffing, via human resource development to organizational competence portfolios.

To conclude, we can cover the technical and methodological aspects of the related approaches with semantic people tagging. Moreover, none of the approach supports semantic tagging that allow the development of an organizational competence vocabulary that is a pre-requesite for competence management. Here, semantic people tagging together with ontology maturing provides a holistic approach to close the gap.

9.3. Methods and Tools Related to Ontology Maturing

In this section, we look at related work from an ontology maturing perspective. 'ColBlend' and 'myOntology' are related in the aspect of collaborative & community-driven construction of ontologies, 'Seeding-Weeding-Fertilizing' is related regarding gardening as a scaffolding method. Finally, we will look at related collaborative editors for SKOS ontologies.

9.3.1. Collaborative and Community-driven Ontology Construction

ColBlend: Pereira and Soares (2008) also noted the insufficient support of a social and collaborative construction of a conceptualization in the state of the art. They therefore have directed their research towards the application of results from cognitives and the perspective similar to ours of meaning construction in and through collaboration and negotiation processes by groups of stakeholders in organizational contexts with a set of common objectives. ColBlend has been developed as supportive method based on the Conceptual Blending Theory (Fauconnier and Turner, 1998) where elements from different input spaces are mapped through common but more abstract representations in a generic space. An additional space, the blend, contains emergent structures derived from the input spaces but in neither to be found. Based on this, the authors propose a 7-step-method (Pereira et al., 2009): (1) a selected group of participants provides a preliminary conceptualization based on the defined goals and mission in the generic or “shared” space as starting point; (2) each participant or a group of representatives for each organization presents their proposal representation in their input space also called “conceptualization proposal space”, e. g., as a concept map, together with additional material for better understanding; (3) in the generic space,a common conceptual structure is manually or automatically generated without much negotiating; (4) the blend is run based on available background information (e. g., from input and generic spaces) to generate new conceptual structures; (5) the new structures are negotiated, agreed ones copied to the generic space and pushed back to the input spaces and the blend space is validated; (6) if input spaces need to be modified, the method reiterates at step 3; (7) the method is accomplished as soon as all participants agree with the conceptualization in the generic space.

To support the negotiations, the authors propose an additional consensus building workflow. In its preparation, the team and collaboration rules are defined. In the main phase, either “working with a single-text document” and “taking a visioning approach” is applied. In the former approach, the participants discuss and revise a working draft; in the latter approach focuses on “seeking agreement”, i. e. to identify what is wanted and how to get there. Finally, the negotiation process is evaluated and lessons learnt documented.

For technical support, the “Semantic System for Continuous Construction of Meaning” (SemSys4CCM) platform has recently been developed (Sá et al., 2010). It is based on Semantic Media Wiki and extensions to integrate CmapTools. Conceptual models are represented as concept maps whilst in the wiki concepts are automatically mapped to categories and connections to properties. The platform provides on a category page a concept map visualization that is re-generated every time the map is modified. A sidebar provides details, links for navigation and edit options that are add/delete/rename concept or connection and modify description. For negotiation, the users can discuss, comment and rate ontology elements. Third party applications may use the developed model through exports or the connected triple store. However, it is not obvious how the platform supports the generation and management of the different spaces, especially the blended space, that are proposed by the ColBlend method.

myOntology: Siorpaes et al. (2008) also propose a community-grounded approach to develop lightweight ontologies called *myOntology* that has been elaborated in the eponymous project[41]. The approach is based on six design principle that aim for similar purposes as ours: (1) *Tapping the Wisdom of Crowds* – claiming a diverse community being smarter and more agile and implying co-existence of conflicting ontology views that have to be brought to consensus; (2) *Openness and multimedia richness* – applying the wiki culture to lower entrance barriers and embedding multimedia objects illustrate intended meanings and enhance understanding; (3) *Lightweight ontologies* – to be easier understood by the community; (4) *Integration of linked data* – to reuse available data for ontology enrichment; (5) *Combination of human and computational intelligence* – to support the users in ontology building tasks and ensure quality; and (6) *Incentives* – transferring incentives from Web 2.0 applications to ontology building in order to foster user contributions.

The proposed LICONE (Lightweight and Community-Grounded Ontology Evolution) methodology adapts the methodology by Uschold and King (1995) in a five-step ontol-

[41] http://www.myontology.org

ogy lifecycle: (1) *Informal specification* – i. e. specifying the scope of the ontology to build; (2) *Collection of relevant named entities* – describing the ontology with labels, description, and lexical and multimedia resources; (3) *Typing of named entities* – by specifying concepts and properties according to the myOntology metamodel, which is a subset of OWL DL and SKOS element are additionally used; (4) *Adding taxonomic and non-taxonomic relations* – to create the subsumption hierarchy with subClassOf relations; (5) *Community-driven alignment*. In every step, the enrichment of specification with lexical and multimedia resources and maintenance of community consensus by checking the re-use of existing elements take place.

For technical support, the myOntology wiki system has been developed. The start page is the main navigational entry point. Three tag clouds provide an overview of available ontologies, concepts and properties; each having its own wiki page. A sidebar displays a typical hierarchy of concepts that can also be used for navigation or to edit the structure via drag'n'drop.

Editing and creating ontology entities is form-based and *"users can add a description, synonyms, images and videos, translations, tags and a seeAlso link"* (Siorpaes et al. 2008) for any entity. These are *"enhanced by gather data from the Web, such as Wikipedia, Flickr, YouTube, Wordnet, etc."*. For ontologies, the user may add concepts. Concepts may be created from the start page or the ontology page. For concepts, the users may add properties, create instances, SKOS mapping relations with other concepts being equivalent, partly overlapping, narrower and broader, or sub-/super-concepts. It is unclear what tags are and why SKOS broader/narrower relations and sub/superClassOf are used together. The sidebar structure is based on the latter one. For properties, the users may additionally specify equivalent properties and the allowed range and domain.

The menu bar provides a search box that also is used for navigation and creation; i. e. it provides auto-completion when typing ahead and additional suggest options to create a new ontology, concept or property. The advanced search provides filtering based on the types. Additional functionality are the creation of so called "freeze points" to capture a stable snapshot of an ontology that may be exported to OWL. There is a change log for versioning and some semi-automatic gardening mechanisms listing entries without description, multimedia content or not included in an ontology. Unfortunately, up to now there is no evaluation data available besides a small usability study with four students (Klotz, 2008).

To conclude, both approaches presented in this section are similar in its methods to ontology maturing and might be smoothly integrated into our approach. However, both approaches neglect the aspect of work-integration.

9.3.2. Gardening as Scaffolding Method

Seeding – Weeding – Fertilizing: This related work, in contrast to the previous approaches presented in this section, deals with tag gardening activities to enrich folksonomies (Weller and Peters, 2008; Peters and Weller, 2008) and is also quite similar to the SER model (q. v. 5.1.1) we applied for our scaffolding methods. The authors make use of the gardening metaphor where the folksonomy represents the garden and the individual tags the different plants that grow wildly and in a mess. The introduced gardening activities to enrich folksonomies with semantics are symbolized by weeding, seeding, landscape architecture and fertilizing.

Weeding is seen as the first activity of gardening and concerned with basic formatting and eliminating or revising unwanted tags because of classic folksonomy challenges like spelling mistakes, compound words, pluralisms (q. v. 2.8.2) but also because of spam. To that end, the authors propose (Peters, 2009; Weller, 2010):

- Simple editing functionalities to easily and quickly correct mistakes
- Automatic detection of typos, spelling variants, pluralisms and spam tags
- Auto-completion during the tagging process
- Suggestions and hints for potential mistakes with dialogs like "Did you mean..." during the tagging
- Blacklist for spam tags
- Advanced editing options for authorized users acting as gardeners
- Guidelines how to tag.

The second gardening activity is *seeding*. More infrequent but specific tags need to be circulated in order to overcome the suggestion of high-frequent but less discriminating tags, which more and more reduce expressivity of the folksonomy and precision in retrieval. This might be achieved by (a) an inverse tag cloud that no more shows the frequent tags but the infrequent ones in bigger font sizes, (b) by awareness displays for new or recently added tags, and (c) by providing infrequent or new tags as tag suggestions during the tagging process.

The third activity is the *landscape architecture* to be done by the gardeners. It aims for enhancing again the folksonomy's expressivity through adding semantic relations and thus is mainly concerned with semantic disambiguation, i. e. identifying homonyms and summarize synonyms and multilingual tags, and relating tags.

Fertilizing is the last gardening activity by which folksonomies are combined with other existing more complex ontologies. This might be done either by enriching the existing ontologies with folksonomy data or by using ontologies for complementing the folksonomy e. g., for query expansion or tag recommendations.

As tool support the tagCare tool has been developed. Even though it enables gardening activities across platforms – currently Flickr and Delicious, it is only for the personal folksonomy (Dittmann et al., 2009). With the tool, users can import and manage their personal tags. This includes basic editing actions like creating, renaming and deleting tags and summarizing synonymous tags under a preferred tag. Currently it provides auto-completion support and suggestions for editing the own tags based on user activities, spell checking mechanisms and general rules like for compound words or capitalization. Further advanced editing options and suggestions with semantic relations are intended to be implemented.

Weller (2010) suggested ontology gardening activities in analogy to tag gardening. *Weeding & Seeding* are seen quite similar the tag gardening activities and mainly concerned with deleting useless ontology concepts and creating new ones. Whilst *fertilizing* in tag gardening means the enhancement of the folksonomy with semantics from existing ontologies, ontologies might be fertilized with social information from folksonomies like usage or co-occurrence information. As an additional ontology gardening activity, the author introduces *harvesting* that is gathering information, i. e. terms or semantic relations, from external sources to enrich the ontology. For instance by analyzing tag distributions and frequencies to identify new or change of terms or co-occurrences for relatedness.

To conclude, the presented approach matches our scaffolding methods in Section 5.1 and we furthermore provide a technical implementation to support ontology gardening activities (q. v. 6.4.2).

9.3.3. Collaborative SKOS-Editors

With the official announcement as a recommendation by the W3C August 2009 and the emerging Linked Open Data community, SKOS has become quite popular and several editors from research but also commercial ones have recently shown up.

SKOSEd: One of the first editors and developed by members of the W3C working group is SKOSEd (Jupp et al., 2009). It is an open-source plugin for Protégé 4.0 to create and edit SKOS ontologies, however it does not provide all features of the SKOS recommendation. Furthermore, it only relies on Protégé collaboration support functionalities and it is unclear how both work together (q. v. 3.2.3.2).

iQvoc: Another very recent open source editor is iQvoc (Bandholtz et al., 2010). It is developed as a Ruby on Rails application and aimed for reference vocabulary management and especially SKOS XL editing in order to support specific complex lexicals like inflectional forms or term composition. The former one is also the reason for a quite restricted user right management. That means only specific members of a small editorial team are able to edit the vocabulary and eventually release new versions. For editing an entity, the editor creates a copy and with that also locks it for other editors until s/he or someone with more right unlocks it. After finishing the changes, the editor submits it and after a consistency check it might get released. Another focus of this IQvoc is the integration with the Linked Open Data world.

Enterprise Vocabulary Net: The web-based Enterprise Vocabulary Net[42] (EVN) is also very recent but commercial tool provided by TopQuadrant and provides a fine granular vocabulary and user role management. Thus, there are a “production vocabulary”, which is the vocabulary in use, and several “working copies”, created for editing purposes. It is also possible that some users with specific roles edit directly the production vocabulary (those changes are directly effective as soon as the users save them), however the workflow assumes to make modifications based on working copies. That means, users can create a working copy, make their changes and save these, which might then be published to the production vocabulary after approval. To that end, it also provides an extensive set of functionalities for version control and change history management. Reports and exports in various formats can be generated on the one hand for enterprise quality control and documentation purposes but they are also used to check working copies e. g., for SKOS constraint violations. There is no further application-oriented or work-integrated support nor functionality that helps the social dimension or development of a shared understanding.

[42] http://www.topquadrant.com/solutions/ent_vocab_net.html

PoolParty: The most similar editor to SOBOLEO and therefore maybe the most interesting one is PoolParty[43] (Schandl and Blumauer, 2010). PoolParty, which is a commercial product, is so close to SOBOLEO because the inventors, as stated by one of them, adopted our SOBOLEO ideas. Users can create web-based and collaboratively projects that can comprise multiple concept schemes with their concepts that are displayed in a hierarchy tree. Selecting a concept opens the concept schemes' or concepts' details. For concept documentation and notation purposes PoolParty additionally includes skos:notation, skos:scopeNote and skos:definition. Recently, the creation of own relation types as sub-properties of skos:related has been implemented. If such own relations are defined, the users are asked about the specific type of relation they want to use upon creation. Even though PoolParty is made for collaborative management, it does not provide real time multi-user support as we do.

PoolParty may also be used to upload and annotate documents either from a file or from a URL (however there is no browser plugin or bookmarklet available), thus providing a type of social semantic bookmarking. Similarly, the content is analyzed for references to existing concepts. Recently, new keywords are extracted as well and when saved as annotation, collected as "Free Concepts" – what we call "prototypical concepts". Free concepts are additionally subdivided in three folders – one for concepts that need approval, one for expert approval and one for 'idle' concepts; i. e. candidates for deletion. For integrity checks some quality queries are available that check the graph structure for completeness, cycles in the hierarchy and disjointness between related and broader/narrower concepts. This might be seen as ontology gardening recommendations.

Currently, the development focus is on Linked Open Data use cases. To that end, skos:exactMatch and skos:closeMatch have been introduced and users may link concepts to different data sources like DBpedia, WordNet, Sindice or Freebase (Schandl and Blumauer, 2010).

[43] http://www.poolparty.biz

10. Summary and Outlook

This chapter summarizes the main contributions of this thesis. It illustrates how the results of this thesis have already already been brought into practice and got a wider dissemination. It furthers shows a number of starting points for future research and concludes this thesis.

10.1. Contributions and Impact

The thesis provides a holistic approach to the question how to support a community of knowledge workers in the collaborative development of ontological knowledge structures used for organizing and sharing information resources. The solution consists of (a) Ontology Maturing as a new perspective and model for collaborative ontology development, (b) a conceptual design framework taking technical as well as non-technical aspects into account in order to derive real socio-technical systems, and (c) a technical framework as reference implementation. These results have been iteratively refined and proven in nine case studies with more than 250 participants involved.

A new Perspective and Novel Model for Collaborative Ontology Development: The **notion of Ontology Maturing** and the **Ontology Maturing Model** presented in Chapter 4 provide a new perspective on collaborative ontology development by knowledge workers that considers (1) the development of a shared understanding within social and collaborative learning processes; (2) the active participation of all stakeholders not only to distribute the effort but also to foster acceptance and up-to-dateness of the ontology; (3) variable levels of formality towards continuous transition between folksonomies and ontologies in order to encourage participation; (4) application-orientation and work integration to foster motivation but also acceptance and appropriateness of the ontology; and (5) usable evolving models to make shared understanding accessible.

The model provides the conceptual foundation and cross-domain description in order to help in understanding the process of such collaborative ontology development. It relates the different types of contributors with different levels of involvement and expected skills to the different phases of maturing. In Section 4.4.1 we have shown how to instantiate the model for a specific domain; i. e. in that case for competence ontologies. Furthermore, it describes ontology maturing in three different dimensions: the artifact dimension concerned with the created ontology elements, the knowledge dimension with the maturing and alignment of knowledge, and the social dimension with the development of competencies and social structures. In this way it addresses the requirements and shortcomings of current ontology development approaches.

A Conceptual Design Framework for Deriving Socio-technical Systems: The **conceptual design framework** presented in Chapter 5 complements the Ontology Maturing model and enables developers to derive and realize socio-technical systems for collaborative ontology development. It considers technical as well as non-technical aspects and provides guidelines to adapt to a given organizational setting. To that end we provide **methods and functions for scaffolding and guiding ontology maturing** presented in Section 5.1. They show for example how to support the consolidation process, how to achieve higher formalized knowledge units and structures and how to work against overgrowing. The **design framework for social semantic tagging** presented in Section 5.3 details it to the application and organizational context. To that end, we also came up with a general **definition and model of social semantic tagging** and the specializations **social semantic bookmarking** and **semantic people tagging** in Section 5.2. The design framework shows the different design aspects, their options and in particular their implications. This is done with a general perspective and then specialized to the cases of social semantic bookmarking and semantic people tagging. We further detailed the implications for usage of social semantic tags – concretely within search heuristics.

A Reference Implementation for Social Semantic Bookmarking & Semantic People Tagging: The **SOBOLEO framework** detailed in Chapter 6 presents a flexible culture-system-fit framework and reference implementation of the conceptual model and conceptual design framework. It is a framework that enables knowledge workers to construct and maintain ontologies, according to the principles of ontology maturing, in a collaborative and continuous way, integrated into the usage of these ontologies within their daily work. The overall technical framework that has outgrown the mere research prototype status over the years, provides a configurable and extensible architecture (q. v. Section 6.2.3) which others might reuse. It will be published as open source under GPLv3 by the end of the MATURE project in April 2012. It further provides reusable reference data models for social semantic bookmarking and semantic people tagging and competence ontology maturing (q. v. Section 6.3). These data models reuse and integrate with established semantic web standards.

The review of related work in Chapter 9 has shown that semantic people tagging is a novel notion for which SOBOLEO provides the first implementation ever (q. v. Section 9.2). We also applied our conceptual design framework for the comparison of social semantic bookmarking tools. The comparison shows that SOBOLEO is outstanding regarding design aspects of semantic tags/ontology and supporting scaffolding and guiding features (q. v. Section 9.1.2). For example, the majority of the tools still restricts the editing of relations between tags to only the private space and/or do not allow for a real community driven evolution of the semantic model. Regarding approaches related to ontology maturing presented in Section 9.3, we provide the hitherto most extensive evaluation. The review of related work on collaborative SKOS editors in Section 9.3.3 further shows that SOBOLEO has been a pioneer for collaborative SKOS editors and SKOS editors in general.

Empirical Evaluation of Concepts and Reference Implementation: We have proven the feasibility and usefulness of the overall approach and the usability of the SOBOLEO framework and its enabling of ontology maturing by nine **case studies** with more than 250 participants involved (q. v. Part III). This also represents the first and most extensive evaluation of social semantic bookmarking (q. v. Chapter 7) and semantic people tagging (q. v. Chapter 8).

The evaluations have shown that collaborative and work-integrated ontology development is more than maturing the mere artefact. To take this into account we have extended our ontology maturing model in order to take into account the artifact, knowledge and especially the social dimension. This differentiated view has also been integrated back into the knowledge maturing macro model and further extended towards a whole landscape in which our ontology maturing smoothly fits in (cf. Barnes et al., 2009, 2011).

Over the years, we have refined the SOBOLEO framework and our ideas got a wider dissemination; for instance by being adopted into the commercial PoolParty tool. Furthermore, SOBOLEO has already been and is used within other projects, e. g., the German BMWI project SABINE[1] or the European Collaborative Project MATRIX[2]. We have also introduced the approach of social semantic bookmarking into the tool environment for car construction at a large German automotive manufacturer – probably the first implementation of social semantic bookmarking in business use. Currently, we are integrating semantic people tagging at SAP to support their internal transfer processes.

10.2. Future Research

We provide semantic people tagging as a novel concept for, among others, bottom-up competence management. The SOBOLEO framework is a pioneer implementation for semantic people tagging with that brought semantic people tagging into use and provided first evaluations. The next challenging steps of research would be to further investigate the maturing of the (collective) knowledge about others' expertise by, e. g., starting to show the quality improvement of profile data or analyzing the social network and people's degree of "networkedness". Therefore, a long-term evaluation in practical and operational use needs to be conducted.

Motivation and incentives within social semantic (web) applications are a further promising field of research. We have started to leverage tagging motivations to ontology development and to integrate explicit system feedback mechanisms for user motivations. However, more extended user evaluations and research in general is necessary to investigate this area. For instance, it has also has not yet been explored in a systematic way how tagging motivations relate to actual user behavior. Explicit feedback mechanisms as incentives to increase user contributions and their dependency on personality traits are current research focus of Athanasios Mazarakis (Mazarakis et al., 2011; Mazarakis and van Dinther, 2011a,b).

A third interesting point for further research, is the deeper and systematic investigation of scaffolding methods in general and ontology gardening methods in particular. We presented a conceptual categorization and first implementations to that end. Additional existing approaches like the ones presented in Section 3.2.4.1 or new ones might be easily conceptually integrated into the categorization as well as technically integrated in to the SOBOLEO framework. Because the SOBOLEO framework is meant for extensions based on its generic interface for ontology gardening recommendations (q. v. Section 6.4.2).

We have shown in Section 4.4.2 conceptually and in Section 6.4.1 technically how we can support higher levels of formality for competence ontology purposes. Overall, our focus has been on the development of lightweight ontologies comprising the first three phases

[1] http://sabineprojekt.wordpress.com/
[2] http://matrix.gpi.kit.edu/

of ontology maturing. That was due to our case studies in evaluation that did not require heavyweight ontologies. Lacasta et al. (2010) stated: *"One of the most basic requirements to jump from terminological* [a/n: lightweight] *to formal* [a/n: heavyweight] *ontologies is to be able to classify the broader/narrower relationships* [..] *into specific categories such as is-a, instanceof, or is-part-of. Another important required transformation is the classification of the abstract related relationships into different specific subtypes".* Whilst we do not only conceptually but also technically support the former one, extensions towards the latter one still need to be implemented. In general, towards the transition to heavyweight ontologies, further research might be necessary on adequate use cases and supporting methods to enable a community of knowledge workers to develop such ontologies in a collaborative and work-integrated manner. In Zacharias and Braun (2008) we presented first attempts to tackle this challenge.

Over the years, the SOBOLEO framework has outgrown the mere research prototype status. To bring it to an eventual product release, however, it would still be necessary to integrate a more sophisticated user and data security management as this was not in focus of our research. At this point, we would like to refer, e. g., to the European Integrated Project TAS[3], which is an associated partner project of the MATURE project to provide answers to the issues of privacy and security[3]. Further performance tests might be necessary to support several hundreds of users with one server installation. To that end, we are currently working on a large scale roll out at igen Ltd. career and personal development services with up to 300 users by the end of 2011.

To conclude, this thesis presents a holistic approach towards collaborative and work-integrated ontology development by knowledge workers. Especially the aspect of work-integration had been neglected so far and makes this approach stand out from related work. It shows how an overall conceptual and socio-technical framework has been iteratively developed, which is feasible, useful and usable in practice.

[3] http://www.tas3.eu

Part V.

Appendix

A. Technical Details

A.1. Data Model Notation

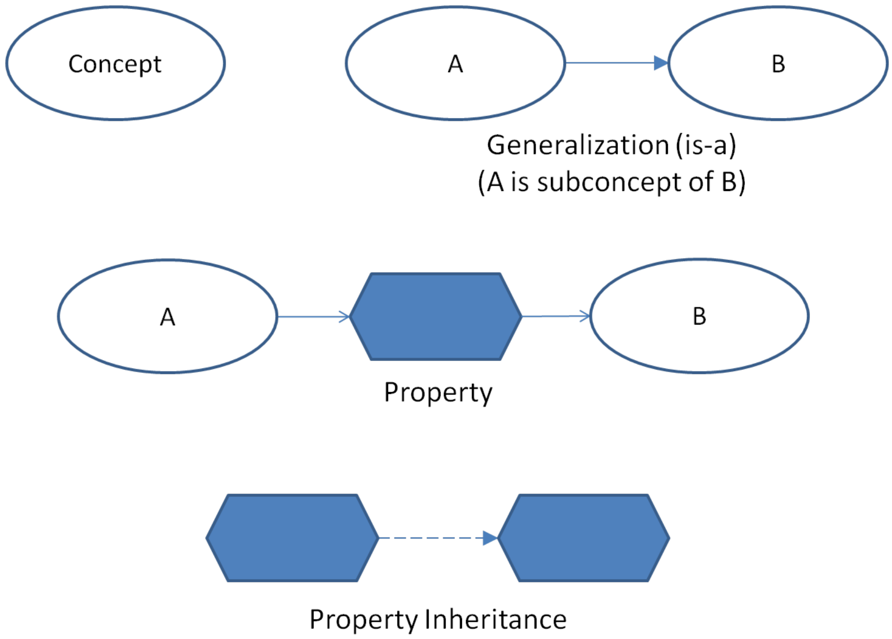


Figure A.1.: Data model notation.

B. Case Study II: The Rapid Prototyping Research Community

B.1. User Tests 1

Table B.1.: Results from observations and screen recording

Session	Task	Type	User 1	User 2	User 3	User 4
1	1	Notes	-Login editor -Creates concept for Rtejournal -Opens web pages and annotates them with Rtejournal -Switch to editor -Successful search for resources	-	-3 faults because popup was closed -2 questions related to annotations -training period needed	- No problems with annotating after having read the web documents
		Video analysis	-Login popup for annotation is confusing and cannot be closed -Annotation popup hides text -Save button without feedback	-	--	-
		Time	5 min	20 min	19 min	10 min
		Ask for advise	0	Several	3	0
		Faults	0	n.s.	5	0
		Success	100%	100%	100%	100%
1	4	Notes	-User opens editor -Creates new concepts -Does not look at existing concept	- Not visible that there is new term created with drag'n'drop	-n.a. [a/n: user did not perform the task due to session time out]	-problems with drag'n'drop
		Video analysis	-	-	-n.a.	-
		Time	5 min	5 min	--n.a.	10 min
		Ask for advise	0	n.s.	-n.a.	5
		Faults	0	n.s.	-n.a.	1
		Success	100%	100%	-n.a.	100%
1	5	Notes	-user enters his query, retrieves 4 links that he investigates	-	-n.a. . [a/n: user did not perform the task due to session time out]	-n.a. . [a/n: user did not perform the task due to session time out]
		Video analysis	-	-	-n.a.	-n.a.
		Time	5 min	5 min	--n.a.	--n.a.
		Ask for advise	0	n.s.	-n.a.	-n.a.
		Faults	0	n.s.	-n.a.	-n.a.
		Success	100%	100%	-n.a.	-n.a.
2	1	Notes	-User opens concept container 'collaborative concepts' -Starts using the chat and suggests himself as expert -User sorts concepts -Drag'n'drop overlay confusing because decoupled from cursor	- User is undecided	- Tree structure difficult to understand	- Very smooth task execution
		Video analysis	-	-	-	-
		Time	15 min	20 min	21 min	10 min
		Ask for advise	1	n.s.	1	0
		Faults	0	n.s.	2	0
		Success	100%	100%	100%	100%
2	2	Notes	-Supervisor gives hint for orientation - tries to search within ontology structure and switches to the browse area -User unsure if to add altLabel or broader concept and ask for advise -Needs time to find input field	-n.s.	n.s.	- Navagating the taxonomy tree not easy but solved via chat
		Video analysis	-	-	-	-
		Time	10 min	20 min	8 min	10 min
		Ask for advise	2	n.s.	0	4
		Faults	n.s.	n.s.	1	0
		Success	100%	50 %	100%	100%
2	4	Notes	-Searches for resources in the web -User does not know that PDFs are also annotable -Annotates 3 links with the given topics	-stops task because of annotation popup crashed	n.s.	-Works out smoothly -Learning effect visible
		Video analysis	--	-	-	-
		Time	5 min	10 min	8 min	15 min
		Ask for advise	n.s.	n.s.	0	3
		Faults	0	n.s.	1	0
		Success	100%	0%	100%	100%

C. Case Study VI: The Classic Roman Civil Engineering Community of Practice

C.1. User Experience

♦ Debatir acerca del siguiente tema: Red de alcantarillado

[redacted]
Dejame explicar... las aguas pluviales irian por la calzada hasta que en algun punto caerian a la red de alcantarillado

> [redacted]
> *Dejame explicar...* entiendo que irían por la calzada hasta un cauce, no a la alcantarilla porque entonces ésta tendría que tener una dimensión importante y no creo que, en general, esa sea la solución. En una ciudad extensa puede ser en alguna zona a lejada del rio, en la mayor parte de las poblaciones, no.

[redacted]
Crear... Por otro lado, me ha parecido muy interesante la aportación de Manuel que señalaba que ""terminaban desaguando sin ningún escrúpulo..."", contaminando ríos, por ejemplo. Creo que la preocupación por la contaminación y el medio ambiente son valores muy modernos que en la época romana no debían existir ¿o tenéis algún indicio de que si?

[redacted]
Pienso... que muy ligado al abastecimineto de agua potable de las ciudades, está el problema de la eliminación eficiente y limpia de residuos ¿Cuáles pensais que son las claves de la red de alcantarillado que consiguieron desarrollar los ingenieros romanos?

> [redacted]
> *Pienso...* que fue muy importante que separasen las aguas provenientes de la lluvia de las fecales que se producían en edificios, cuadras, etc.

> [redacted]
> *Pienso...* que también tuvo su importancia la gran capacidad con la que dimensionaban los depósitos en los que recogían las aguas de lluvia, lo que permitía retardar su incorporación a las puntos por los que se debía de evacuar. Cuando he dicho antes que separaban las aguas esto no es verdad en su totalidad porque si permitían que las pluviales circulasen por las alcantarillas, pero sólo parte de ellas
>
> > [redacted]
> > *Estas diciendo que...* podían retardar la incorporación y evacuación de aguas. Narciso ¿Podrías explicarme con un poco más de detalle cómo hacían esto?
> >
> > > [redacted]
> > > *He leido que...* las piscinas que construían en el interior de sus casas tenían en parte esa misión: todas las aguas que caían en el edificio se derivaban a la piscina, ésta tenía una salida pequeña que dosificaba el agua que se incorporaba al alcantarillado, mientras la piscina se iba llenando. Esto exige que la piscina esté ""poco llena"" cuando se inicia el chaparrón.

> [redacted]
> *Pienso...* que la redes de alcantarillado en general era de muy buena calidad en todo su recorrido, tanto urbano como su canalizacion fuera de las cuidades, pero que terminaban desaguando sin ningun escrupulo, contaminando, en rios u otro via de agua

[redacted]
Dejame explicar... hace 2000 años esta el mundo muy poco contaminado y no apreciaban el peligro que ahora nos acosa, ni tenian los romanos el potencial contaminante que ahora tenemos.

> [redacted]
> *Pienso...* que tienes toda la razón Manuel, lo has resumido perfectamente.

[redacted]
Dejame explicar... no dejaban acumular las aguas en las calles, evidente, circulaban por ellas no por los colectores, salvo en las grandes urbes (Cloaca Máxima). De todas las maneras creo que tienes razón, no he encontrado la información en la que se basaba mi idea de que el saneamiento era separativo. En algunos sitios(Baños, termas, etc9 si parece que salian separadas, en lo demás no. Perdón, estaba equivocado

> [redacted]
> *Pienso...* que gracias a vuestro debate, hemos aclarado bastante el tema. Os felicito por vuestra participación.

[redacted]
Dejame explicar... no creo que se preocuparan mucho de separar las aguas de fecales de las de lluvia. Almacenarian las aguas de lluvia en la piscina o en un algibe, que tendrian un aliviadero para expulsar el agua sobrante, pero las aguas fecales irian directamente al alcantarillado

> [redacted]
> *Estoy de acuerdo que...* no separaban del todo las aguas o que les preocupara mucho el tema, pero las pluviales las evacuabn por la superficie de las calzadas, mientras que las fecales las incorporaban al alcantarillado. Eso a mi manera de ver es separar las aguas
>
> > [redacted]
> > *Modificar...* la primera frase he querido decir: Estoy de acuerdo que no separaban mucho las aguas o que NO les procupara mucho el tema

[redacted]
Dejame explicar... piensa en las ciudades actuales, todas las aguas van al alcantarillado que tienen las dimenciones necesarias para el volumen de agua previsto avacuar, igual hacian los romanos. No creo que dejarar acumular las aguas de lluvia en las calzadas por que si fuera asi, en algun sitio inundarian las viviendas

Figure C.1.: Example Dialog in Detail

C.2. Participants Interviews

Table C.1.: General feedback on the experience

Participant	General opinion on the experience	Elicit any particularly positive points	Elicit any particularly negative points	Elicit any other points of interest
A1	The participant was the student who joined the initiative the latest. His assessment is positive but feels that having arrived late to the experience has limited their participation. Both SOBOLEO and Interloc found them interesting and useful tools.	Working in community with other professionals allows you to explore, go beyond what I would suggest individually. Surely you can increase knowledge. Share links, important information, it saves much work and improve outcomes. Knowing the concerns of other partners allows you to get involved in issues that perhaps you would not have proposed their own.	The experience has been very short. Incorporated later also into the debates made them and found them closed, with little opportunity to give continuity.	There has been a lot of participation by partners and this discouraged.
A2	In general he had a positive Review. Thought the process was able to indentify new ways for aquiring knowledge.	Being immersed in this type of initiative makes you active. Knowing that your contributions will reach out to others forces you to review, to deepen, to better prepare your speech.	He has not gained access to Interloc	Experiences like this help you break the ice with Web 2.0 tools. Technical support is very important. Whether it had not received this support have left the experience.
A3	Positive. He is the youngest student in the group. He was inexperienced and his work has been more to learn from others to contribute.	Working in collaboration with expert groups allows you to pose your questions and get immediate and reliable answers.	Continuous technical problems that greatly discouraged and prevented adequate paced experience.	-
A4	Positive. Working with tools has been very satisfying for him.	Enabling collaborative work is very useful tools for gaining knowledge, implement and contrast what they learned. Lets keep alive curiosity about a subject	Little participation by the other student. The experience has been somewhat short, "has known little."	Interloc failures have not really been a problem for work but have tarnished the discussions.
A5	Positive assessment but has been little involved	The possibility offered by the tools to deepen an issue with colleagues who can learn	Pick up the handling of tools, SOBOLEO especially, is rather complicated and sometimes you're afraid of upsetting the work they have been doing peer	I thought the course instructor would the experience from participating in
A6	It seemed interesting to know the tools we've provided. Thought to have many positive applications	Combining both tools can manage with great professionalism in a specific subject, but from their point of view we (STRUC) needed more support to clearly lead the project.	-	You have to dedicate time to see results, not serve to connect from time to time to see what peers are doing

Table C.2.: Specific feedback on the experience

Participant	How did/might the experience help you to improve your knowledge and understanding	How did/might the experience help the group to improve their knowledge and understanding	On SOBOLEO, What particular things did you do?	What others things did you do?
A1	-	-	Propose two Web pages. See proposals from other partners.	Participate in a discussion of Interloc
A2	-	-	Add terms. Complete details with descriptions	-
A3	-	-	Consult the terms and web pages that were being proposed.	Participate on InterLoc dialogues
A4	The experience helps broaden their knowledge, no doubt. Our topic in question would need more support chart. It misses to insert images that accompany and reinforce the arguments.	Participation has been lacking to assess this aspect.	Consult the terms and web pages that were being proposed.	Participate in all discussions of Interloc with my knowledge.
A5	He has been involved very little and was able to consult the terms that were being raised and the web pages that marked peers	He believes that the participants who have worked more experience have had a good opportunity to improve their knowledge	Consult the terms and web pages that were being proposed.	He has not gained access to Interloc
A6	It has helped me more to imagine the possibilities of the tools applied in professional work projects	-	Consult the terms and web pages that were being proposed.	Participate in several Interloc dialogues

Table C.3.: Specific feedback on the experience (continued)

Participant	What criteria did you use to participate in a particular subject?	What time of day did you do?	Is there anything else you would like to add about your experience with SOBOLEO and InterLoc?
A1	View the debate alive	At home after work	It takes a while operated in environments
A2	He has contributed words and comments on those areas where he felt comfortable with more knowledge	In free moments during the day.	It is important that the work of the backgrounds is gunned down by European universities.
A3	She wanted one in which she felt more informed.	Every day after work.	-
A4	He participates in all the issues, considering that in all of them he had something to say.	After work, daily.	He has not participate more in SOBOLEO because he considered with little fluent in technology. The proposed terms and discussions have seemed sufficient. In his view allowed a lot of game.
A5	-	On weekends, when it had free time.	-
A6	He participated in the discussions that were created at the beginning of the experiment	After work during the first week of experience	At first he thought it would have more time to participate much more than it really was possible. Apologize.

D. Case Study IX: British Career Guidance

D.1. Task-oriented Guide for Introduction Workshops

People Tagging: Trial Session

MATURE
www.mature-ip.eu

SCENARIO:

You are an employee of Connexions Northumberland Ltd. Yesterday you were assigned to a new client for a guidance interview. Because the client's case is very special, you need advice from a colleague. After a long search and asking around, you found a colleague, who can help. You just had a very helpful phone call with the colleague. You would like to remember that this colleague knows about the topic. You also like to share this information with other colleagues as this might be helpful for them as well when they need a person for advice about the same topic. Therefore you use the new **SOBOLEO People Tagging** (PT) system.

People Tagging: Tasks

www.mature-ip.eu

1) In order to facilitate later retrieval for you and your colleagues, you would like to retain the information about the person by using SOBOLEO PT.
   a. Please take a colleague at Connexions Northumberland of your choice and think about which topics you associate with him / her.
   b. Open the 'Tag People' area, add the person to the system, tag him / her with at least **2 topics** that previously came into your mind, and save the information.
   c. Make sure your information is correctly stored. Go to the 'Browse Topics' area and navigate to one of the topics you used and find the person you tagged.

2) You would also like to be found by your colleagues in areas where you too have expertise. Therefore you add information about yourself to the system.
   a. Go to the 'Browse People' area and find yourself in the list of people and open your profile.
   b. Tag yourselves with at least **2 topics**, save the information, and reload the page.

3) Now you would like to improve the topic list and add more information about the topics you used in 1) & 2) by using the SOBOLEO PT topic list editor.
   a. Open the 'Edit Topic List' area and select **one** of your previously used topics and add some more information (e.g. synonymous label or description).
   b. Take **two or three** of your topics and arrange them. Remove them from the 'latest topics' container. If necessary, add new topics.

4) In order to complete your client's case, you need another colleague's advice. Instead of asking around, you have a look at SOBOLEO PT.
   a. Go to the 'Home' area and search for people for a topic of your choice by entering your search terms into the text field. Include a topic of the topic list, you've seen in 3), in your search.
   b. Examine the results, select one person, and contact him / her.

D.2. Pre-Questionnaire

Pre-Questionnaire – SOBOLEO People Tagging

www.mature-ip.eu

I. User Expectations

You have received information about a new system which you will test today. How do you usually find useful people (without support from such a system)?

What do you think of the system? What do you expect of the system specifically?

What do you think are the system's advantages?

What do you think are the system's disadvantages?

Do you know of any other systems that you would use to find useful people? If so, which?

D.3. Post-Questionnaire

Post-Questionnaire – SOBOLEO People Tagging

MATURE
www.mature-ip.eu

I. User Satisfaction

#	Question	Disagree	-	Agree		
		--	-	-/+	+	++
1	I am interested in many of the things the system offers.	❍	❍	❍	❍	❍
2	I can easily find my way through the system.	❍	❍	❍	❍	❍
3	I can quickly find what I'm looking for.	❍	❍	❍	❍	❍
4	It's a logical system.	❍	❍	❍	❍	❍
5	I can understand the system.	❍	❍	❍	❍	❍
6	The user interface is attractive.	❍	❍	❍	❍	❍
7	I had the feeling to have anything under control.	❍	❍	❍	❍	❍
8	The system is / reacts fast.	❍	❍	❍	❍	❍
9	The system supports me in finding what I'm looking for.	❍	❍	❍	❍	❍
10	It's easy to navigate through the system.	❍	❍	❍	❍	❍
11	I like the system.	❍	❍	❍	❍	❍
12	With the system, I can easily find relevant people.	❍	❍	❍	❍	❍
13	I can use the system efficiently.	❍	❍	❍	❍	❍
14	The information provided is presented clearly.	❍	❍	❍	❍	❍
15	The system is easy to use.	❍	❍	❍	❍	❍

#	Question	Disagree	-	Agree		
		--	-	-/+	+	++
16	The system is free of annoying features and functions.	❍	❍	❍	❍	❍
17	It's easy to see where I am exactly within the system.	❍	❍	❍	❍	❍
18	The length of time spent learning about the system was appropriate.	❍	❍	❍	❍	❍
19	Activating / Clicking on individual links resulted in exactly what I expected.	❍	❍	❍	❍	❍
20	The system is easy to understand.	❍	❍	❍	❍	❍

II. Usability

How does this system compare to what you do on a daily basis?

Is something missing, that would encourage you to use the system on a regular basis?

Is there something you would like to be changed?

Are there any functions in the prototype that you would not use?

III. Motivation & Barriers

	Disagree		-		Agree
	--	-	-/+	+	++
Do you feel in general motivated in working with the system?	❍	❍	❍	❍	❍

Which aspects made you feel motivated?

Disagree		-		Agree
	--	-	-/+	+	++
Do you think the interface is user-friendly and motivating?	❍	❍	❍	❍	❍

Which features of the system do you find most interesting?

Do you see any barriers (e.g. usability, organizational, culture, lack of time, privacy issues) in using the system? If so, which ones?

What is missing and what would you like to be changed in order to make you more likely to use the system?

IV. Comment Section

If you have something to add, which was not covered by the questionnaire, please feel free to write it down in the comment box below.

Thank you for helping us to improve the demonstrator and being part of the evaluation.

Yours sincerely

The Mature Team

D.4. Results Introduction Workshop I

D.4.1. Pre-Questionnaire Results

How do you usually find useful people (without support from such a system)?

1. Conversations with others, organisations, Google, directories
2. By speaking to colleagues in my office, or enabling other colleagues for advice
3. By word of mouth. We have no staff directory in place that supports this. Personal recommendations, (often accidental) is really the only method.'oh, 'x' went on a course ... ask them'
4. Contact them via telephone / email. Speak face to face. Discuss with colleagues who else may be able to help me & then contact them as above. Use search on websites.

What do you think of the system? What do you expect of the system specifically?

1. It will speed up the process by being more specific, but will broaden the search
2. The system looks very useful. I expected it will make it easier to find people with useful knowledge & cut down the amount of time that it would take to find a useful person.
3. I think the system offers huge potential in terms of knowledge sharing and building up staff's e-confidence.
4. This will be a very useful system - appears fairly straight forward to use (I think!) Not sure what my expectations are - however - to usefully share knowledge / information more effectively to support staff with CPD.

What do you think are the system's advantages?

1. It will bring a wealth of useful material and make it possible to refine searches.
2. Having contacts, websites, tags and all the information in one place. Being a collaborative system which everyone can add to & amend will also be useful.
3. I like the democratic nature of the system and the fact it is self sustaining. It offers support to staff who, outside their teams, rarely meet or know each others' areas of expertise.
4. Immediacy / up to date. Users can edit themselves / not a management tool - user led. Can be adopted & refined. Creates dialogue between users.

What do you think are the system's disadvantages?

1. Is dependent on people engaging with it and editing information
2. It may be disadvantage for some people using a web-based system as some people have less well developed IT skills & will need extra help in using it.
3. As with any new system, explaining the method of using it will be a challenge and usage will take time to embed. Staff is used to quite restrictive IT practice. This programme will be quite revolutionary.
4. Some may feel that by being 'tagged' labels then an expert.

Do you know of any other systems that you would use to find useful people? If so, which?

1. None
2. None
3. I don't know any internally with Connexions Northumberland as our intranet has not been developed in a collaborative way. Externally I might use the web.
4. No other system my work environment. It would be something like Facebook outside work. There is a guidance forum (I think) through SSAT, but I haven't ever used it

D.4.2. Post-Questionnaire Results

How does this system compare to what you do on a daily basis?

1. Very well and can more time by narrowing search.
2. The system is easy to use with recognisable command functions. However there is no IT system here that fulfills this purpose, so SOBOLEO is very welcome.
3. I don't know any internally with Connexions Northumberland as our intranet has not been developed in a collaborative way. Externally I might use the web.
4. Quite different in the way I contact people / search for knowledge / update my information or understanding

Is something missing that would encourage you to use the system on a regular basis?

1. Hard to say as experience was limited
2. I think it would be useful to have a guide or help document somewhere on the system to be able to refer to
3. I can't think of anything - (other than the technical problems are encountered, but previously there were no problems)
4. Nothing I can think of

Is there something you would like to be changed?

1. Not really more
2. No
3. No, just minor points that Simone is aware of
4. No

Are there any functions in the prototype that you would not use?

1. Didn't use it enough
2. I would use all the functions
3. No, I would use all of the functions
4. None

Which features of the system do you find most interesting?

1. Editing
2. The topic list is a very interesting feature & the way in which it can be changed to make the list more useful
3. Being able to see how people 'tag' themselves & others. Adding users and RSS feeds.
4. All

Which aspects made you feel motivated?

1. Easy to use, and simple to understand
2. The ability for everyone to take part in using and amending the information
3. The potential for the system to involve staff in communicating their knowledge to others.
4. The potential for supporting staff with CPD. More effectively sharing of knowledge & best practice

Do you see any barriers (e. g., usability, organizational, culture, lack of time, privacy issues) in using the system? If so, which ones?

1. Our system proving to be inadequate in supporting this.
2. Lack of confidence with IT may be a issue with some members of staff. They may not feel confident in using something like this without bit of support. Our IT system might also be a barrier.
3. The culture of 'freedom' n use of IcT will take some time to embed. Time may also be an issue.
4. If our system doesn't adequately support Soboleo this could be a potential barrier. Perceived lack of time, staff worrying about being perceived as having 'specialist' knowledge.

What is missing and what would you like to be changed in order to make you more likely to use the system?

1. ?
2. I wouldn't change the system.
3. I think the system is very comprehensive, quite enough for a pil of *
4. Nothing I can think of.

D.4.3. Technical Feedback

Tagging:

- Auto-completion for email-addresses when tagging a new person would be helpful
- System message when asking for the tagged person's name is confusing
- After canceling email address input, the input form is broken

Editor:

- Editing of concept preferred, alternative and hidden labels by one click instead of double click
- D'n'D does not work properly with Internet Explorer

Search:

- Query expansion with broader topics might provide in confusing results, i. e. collecting container concept "latest topics" to be excluded;

Browse Topics/People:

- RSS feed for people would be useful
- Tag icon not obvious enough
- The directory of people should contain all employees of Connexions Northumberland
- When adding a person to topic, a list of existing people should be provided
- Adding and tagging links is great, uploading whole documents would be useful as well

D.5. Results Introduction Workshop II

D.5.1. Pre-Questionnaire Results

How do you usually find useful people (without support from such a system)?

1. Ask colleagues / manager. Ring colleagues.
2. Talking to colleagues. Connexions intranet. Internet. Resources kept at office e. g., leaflets about services that other agencies provide.
3. Asking colleagues; using the Intranet (Google search). Ringing organisations. E-Mailing.
4. Sending out an email query to PA's. Asking people I work with. Attending information / training events.
5. Ask people - who I ask is based on my own knowledge (often limited). The memex - it miss process.

What do you think of the system? What do you expect of the system specifically?

1. Good in theory - would be useful in-house. Looks very simple to use.
2. Could potentially be very useful
3. A useful tool - as long as everyone is prepared to use it. As long as the searches are wide ranging it should bring up information required.
4. Initial view - if people take up the opportunity & develop tagging, could be extremely useful resource. Needs to be developed quickly to maintain momentum
5. I like it - it wasn't what I expected but I really think it is a useful concept.

What do you think are the system's advantages?

1. Easy and quick access to the right person. More accurate Knowledge Building and cascading knowledge.
2. Appears to be easy to use. Potential to link up with other Connexions services and outside agencies.
3. Save time - get to people who know answers on who can tell you where to go. Help to share info: Help people who work in isolation / rural areas.
4. Ease of use, simple system, allows opportunity to individualise use by partner agencies.
5. Finding people with skills & knowledge to help my work. Develop training opportunities.

What do you think are the system's disadvantages?

1. People may not want to be contacted specifically in this field - they may feel an extra responsibility is not a good thing. It may add to work-load
2. Not sure that everyone within the organisation will use it. Could it be expanded to include documents? Although this is what the Intranet should be for.
3. Some people may not be keen to become involved, perhaps not have the confidence in themselves to identify with any particular topics.
4. If not developed quickly & tagging encouraged, it will not demonstrate individuals "expertise".
5. That the enthusiastics will use it & the workshy or cynical will not.

Do you know of any other systems that you would use to find useful people? If so, which?

1. Intranet - but this is often very slow & inaccurate
2. No
3. -
4. Sorry not great with IT / info system
5. 123people.com, Facebook, Bebo

D.5.2. Post-Questionnaire Results

How does this system compare to what you do on a daily basis?

1. It's easier to navigate and more logical.
2. The system would be useful if colleagues weren't around to ask for advice eg. when working alone in school.
3. Use of internet / computers is on a daily basis.
4. I still enjoy speaking to & discussing with people
5. Currently I don't do anything similar to this. So don't feel able to make a comparison.

Is something missing that would encourage you to use the system on a regular basis?

1. No
2. Ensuring the information is kept up to date. Other contact details for people eg. phone numbers.
3. Easier search facilities.
4. Being able to track who tagged who for what. Use of upper + lower case - get rid of case sensitivity.

5. No, but I would like it to be part of our own system rather than having to log onto a separate one.

Is there something you would like to be changed?

1. No
2. It would be easier to find relevant topics if they were in alphabetical order. Would be useful if it showed documents / people when you searched using part of a tag eg. searching 'Ponteland' doesn't bring up 'Ponteland HIgh School'
3. I like the browse people feature however I found the browse topics feature more difficult to navigate & search on.
4. See above.
5. No

Are there any functions in the prototype that you would not use?

1. No
2. No
3. Tag webpage - not really looked at this aspect of the system.
4. -
5. Not in principle. But I would have to add that I might find it hard to work into my own one

Which features of the system do you find most interesting?

1. The fact that you can update and monitor your own profile easily - adding to record your own CPD.
2. Clear text. Not too much information on the screen.
3. Browse people aspect.
4. Broad areas of expertise.
5. Ease of use. The potential to link websites.

Which aspects made you feel motivated?

1. How easy it is to use - edit - access
2. Prospect of being able to find out more information from colleagues.
3. Being able to identify someone quickly who may be able to provide help with a query.
4. Quickly develop useful information on other users.
5. Messing around with new IT & spying on my colleagues!

Do you see any barriers (e. g., usability, organizational, culture, lack of time, privacy issues) in using the system? If so, which ones?

1. Just when our system is going slow. Some colleagues may not like others 'tagging' subjects on without their permission - I would always ask. There would have to be a culture adopted & encouraged to keep profile information up-to-date.
2. Lack of time. Culture - some people can be very against new initiatives / anything technology-related.
3. People being prepared to identify areas of knowledge & agree to this being put on the system.
4. Time, speed of development, abuse of service.
5. Lack of time or maybe pure my own working culture (or maybe my pure lazy!)

What is missing and what would you like to be changed in order to make you more likely to use the system?

1. -
2. Links to outside agencies in particular Disabled Childrens' Team, Care Trust etc.
3. Within profiles a telephone contact would be useful.
4. Case sensitivity. Tracking of tagging.
5. Nothing to change. Looking at use I would want other colleagues to buy ? it as well. That would make me more likely to use it.

D.5.3. General Feedback

Concerns expressed included the following:

- that linking with these types of outside agencies would increase the workload of individual practitioners
- that some practitioners may abuse the system – 'lazy' colleagues may resist entering details about themselves and may tag others with expertise they may have (to deflect additional queries)
- that it could be a problem for some colleagues to identify areas of expertise with which they feel comfortable being identified

Counter arguments in support of using the system:

- the basic philosophy of the system is democratic – bottom up, rather than top down – and empowering the individual – often people feel out of control
- that management intervention might become needed, but the system is likely to work best where this is kept to a minimum
- that management solution might be needed as a complement
- that it's not about experts but about expertise – tagging does not mean expert, that's important

- that editing the own profile, i.e. removing tags that are inappropriate or in order to avid being contacted, is important
- that the system could increase the efficiency of working with outside agencies by allowing them search for PA expertise easily and quickly
- that the system has the potential to support increased levels of constructive interaction with external agencies (in line with policy expectation)
- that the system has the potential to enhance the sharing of resources likely to be increasingly rationed in times of economic stringency (e. g., learning from attendance at learning events can be shared quickly and easily with colleagues)
- that the system has the potential to support the learning of users through confidence building, related to peer perceptions of expertise
- that (eventually) the system would allow service users to select the expertise of the PAs to whom they wish to target a query

Comments in discussion after the trial session:

- I'm a technophobe – found it easy to use, accessible and friendly – lovely! I'll definitely use it! [individual statement]
- Like the way it can give us lots more information than we have now (not going through the whole process)
- Keep it nice and simple! – The current simplicity is fabulous!
- Opportunity to put web pages in – would find that really useful! I lose these now – but would be great to have a place to store.
- Can add a weblink as a PDF – can put deeper links to different types of documents.
- Can we link to the intranet – think we can!
- If it was easier to get into the system – more likely to use it regularly
- It's democratic, nobody is in charge but has the possibilities
- Need it to be used as much as possible – each participant to introduce to a colleague. Discuss with team managers.
- It was stressed that individuals should use as 'naturally' as possible.
- Need to address from the organizational perspective – what is needed, etc. (June workshop)

D.5.4. Technical Feedback

Unless otherwise indicated, the following observations relate to individual participants:

- How will we be able to manage the volume of data that potentially will be entered?
- Can we search for colleagues using a search facility (rather than scrolling down)
- It would be useful to have auto-completion support in the search field like Google; [mentioned by three participants]
- Search should not be case sensitive and recognize parts of multi-word topics; [mentioned by two participants]
- It would be useful to align telephone numbers to each individual on the system, as well as email addresses;
- It occurs that there are two topics with the same label because of case sensitivity
- D'n'D in the editor does not work properly
- When a relation is added to a topic under "latest topics", it should automatically be removed from the latest topics container;
- It's confusing that a person's profile page is not automatically updated after the user started tagging from there;
- The link to tag a person on his/her profile should be better visible
- It would be useful to directly add URLs on a person's profile
- When people are listed, it should be possible to navigate to their profile by clicking on their name
- Broader, narrower, and related topics for further navigation in the browse area should be in alphabetical order
- For three participants the difference between deleting a topic and removing a topic from a part of the tree was not obvious enough
- If the user removes a topic's last label, the system should provide the option to delete the topic
- System exceptions should be disabled

D.6. Results User Feedback Workshop III

D.6.1. Open Discussion

The responses – unless otherwise stated – can be considered a group consensus as they were discussed among the group members during the session.

Notion of experts:

- it is important that it is not about experts only, but slightly more experienced, which is much more helpful in practice
- seen similar to a rating system in a career guidance forum
- skills database experience; not as available as the people tagging tool

Tagging and getting tagged:

- would need a notification that you get tagged
- tagging others only with consent
- tagging could be seen as a compliment, opportunity for more conversation
- tagging better anonymous [individual opinion]

Culture:

- need to build trust, culture development
- mutuality
- otherwise tagged person might be afraid of getting approached too often etc.

Dependency on personality:

- networkers [like to get contacted and facilitate contact establishment to others]
- gardeners
- passive users
- non-users

D.6.2. Aspects of Design Framework

The participants' feedback with respect to the best option for their context collected:

Who can be tagged?

- everyone, not just opt-in
- opt-out possible
- externals might be in the future

Who can tag?

- everyone can
- self-tagging would not yield enough information

Control over vocabulary

- like the balanced approach to control
- negative tags were not seen as a problem to avoid in a first step, only if there are problems

Visibility of the tagging

- ambivalence towards visibility of the tagger, should definitely not be visible to anyone else but the taggee

Search heuristics

- aging should be taken into account

Bibliography

F. Abel, M. Frank, N. Henze, D. Krause, D. Plappert, and P. Siehndel. GroupMe! – Where Semantic Web Meets Web 2.0. In K. Aberer, K.-S. Choi, N. Noy, D. Allemang, K.-I. Lee, L. Nixon, J. Golbeck, P. Mika, D. Maynard, R. Mizoguchi, G. Schreiber, and P. Cudré-Mauroux, editors, *The Semantic Web*, volume 4825 of *Lecture Notes in Computer Science*, pages 871–878. Springer Berlin / Heidelberg, 2007. URL http://dx.doi.org/10.1007/978-3-540-76298-0_63. (Cited on pages 243 and 247)

K. Aberer, T. Catarci, P. Cudré-Mauroux, T. S. Dillon, S. Grimm, M.-S. Hacid, A. Illarramendi, M. Jarrar, V. Kashyap, M. Mecella, E. Mena, E. J. Neuhold, A. M. Ouksel, T. Risse, M. Scannapieco, F. Saltor, L. D. Santis, S. Spaccapietra, S. Staab, R. Studer, and O. D. Troyer. Emergent Semantics Systems. In M. Bouzeghoub, C. A. Goble, V. Kashyap, and S. Spaccapietra, editors, *Semantics of a Networked World*, volume 3226 of *Lecture Notes in Computer Science*, pages 14–43. Springer, 2004a. ISBN 3-540-23609-0. URL http://dx.doi.org/10.1007/978-3-540-30145-5_2. (Cited on pages 79 and 90)

K. Aberer, P. Cudré-Mauroux, A. M. Ouksel, T. Catarci, M.-S. Hacid, A. Illarramendi, V. Kashyap, M. Mecella, E. Mena, E. J. Neuhold, O. D. Troyer, T. Risse, M. Scannapieco, F. Saltor, L. D. Santis, S. Spaccapietra, S. Staab, and R. Studer. Emergent Semantics Principles and Issues. In Y.-J. Lee, J. Li, K.-Y. Whang, and D. Lee, editors, *Proceedings of the 9th International Conference on Database Systems for Advanced Applications (DASFAA'04)*, volume 2973 of *Lecture Notes in Computer Science*, pages 25–38. Springer, 2004b. ISBN 3-540-21047-4. (Cited on pages 79 and 90)

C. Allen. Tracing the evolution of social software, Oct. 2004. URL http://www.lifewithalacrity.com/2004/10/tracing_the_evo.html. (Cited on page 17)

H. Allert, H. Markannen, and C. Richter. Rethinking the Use of Ontologies in Learning. In M. Memmel and D. Burgos, editors, *Proceedings of the 2nd International Workshop on Learner-Oriented Knowledge Management and KM-Oriented Learning (LOKMOL 06), in conjunction with the First European Conference on Technology-Enhanced Learning (ECTEL 06)*, pages 115–125, Oct. 2006. URL http://sunsite.online.globule.org/Publications/CEUR-WS/Vol-213/paper21.pdf. (Cited on pages 3, 4, and 94)

M. Ames and M. Naaman. Why we tag: motivations for annotation in mobile and online media. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 971–980, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-593-9. doi: http://doi.acm.org/10.1145/1240624.1240772. URL http://dx.doi.org/10.1145/1240624.1240772. (Cited on pages xiii, 26, 27, and 30)

S. Angeletou, M. Sabou, and E. Motta. Improving Folksonomies Using Formal Knowledge: A Case Study on Search. In A. Gómez-Pérez, Y. Yu, and Y. Ding, editors, *Proceedings of the Asian Semantic Web Conference*, volume 5926 of *Lecture Notes in Computer Science*, pages 276–290. Springer, 2009a. ISBN 978-3-642-10870-9. URL http://dx.doi.org/10.1007/978-3-642-10871-6_19. (Cited on page 81)

S. Angeletou, M. Sabou, and E. Motta. Folksonomy Enrichment and Search. In L. Aroyo, P. Traverso, F. Ciravegna, P. Cimiano, T. Heath, E. Hyvönen, R. Mizoguchi, E. Oren, M. Sabou, and E. P. B. Simperl, editors, *Proceedings of the Extended Semantic Web Conference*, volume 5554 of *Lecture Notes in Computer Science*, pages 801–805. Springer, 2009b. ISBN 978-3-642-02120-6. URL http://dx.doi.org/10.1007/978-3-642-02121-3_59. (Cited on page 81)

E. Angus and M. Thelwall. Motivations for Image Publishing and Tagging on Flickr. In T. Hedlund and T. Yasar, editors, *Publishing in the Networked World: Transforming the Nature of Communication, 14th International Conference on Electronic Publishing (ELPUB 2010)*, Helsinki, Finland, May 2010. Hanken School of Economics. URL http://dhanken.shh.fi/dspace/bitstream/10227/599/14/13angus_thelwall.pdf. (Cited on pages 26 and 27)

R. Y. Arakji, R. Benbunan-Fich, and M. Koufaris. Exploring contributions of public resources in social bookmarking systems. *Decision Support Systems*, 47(3):245–253, 2009. (Cited on page 30)

J. L. S. Auer and T. Riechert. OntoWiki: A Tool for Social, Semantic Collaboration. In *Workshop on Social and Collaborative Construction of Structured Knowledge (CKC 2007) at WWW 2007*, Banff, Canada, 2007. URL http://www2007.org/workshops/paper_91.pdf. (Cited on page 172)

S. Auer. RapidOWL - an Agile Knowledge Engineering Methodology. In *Proceedings of 1st International Workshop on Semantic Technologies in Collaborative Applications (STICA 06), 26th-28th June, Manchester, UK*. IEEE Computer Society (P2623), June 2006. URL http://www.informatik.uni-leipzig.de/~auer/publication/RapidOWL.pdf. (Cited on page 74)

S. Auer and H. Herre. RapidOWL - An Agile Knowledge Engineering Methodology. In I. Virbitskaite and A. Voronkov, editors, *Perspectives of Systems Informatics*, volume 4378 of *Lecture Notes in Computer Science*, pages 424–430. Springer, 2006. ISBN 978-3-540-70880-3. URL http://dx.doi.org/10.1007/978-3-540-70881-0_36. (Cited on page 74)

M. Bächle. Social Software. *Informatik-Spektrum*, 29(2):121–124, Apr. 2006. (Cited on page 17)

T. Bandholtz, T. Schulte-Coerne, R. Glaser, J. Fock, and T. Keller. iQvoc – Open Source SKOS(XL) Maintenance and Publishing Tool. volume 699 of *CEUR Workshop Proceedings ISSN 1613-0073*, Feb. 2010. URL http://CEUR-WS.org/Vol-699/Paper2.pdf. (Cited on page 262)

J. Bao and V. Honavar. Collaborative ontology building with wiki@nt - a multi-agent based ontology building environment. In *Proceedings of the 3rd International Workshop on Evaluation of Ontology-based Tools (EON2004)*, pages 1–10, Oct. 2004. (Cited on page 244)

K. Barker, V. K. Chaudhri, S. Y. Chaw, P. Clark, J. Fan, D. Israel, S. Mishra, B. W. Porter, P. Romero, D. Tecuci, and P. Z. Yeh. A Question-Answering System for AP Chemistry: Assessing KR&R Technologies. In *Proceedings of the Ninth International Conference on Principles of Knowledge Representation and Reasoning*, pages 488–497, 2004. (Cited on pages 3 and 89)

M. Barla and M. Bieliková. On Deriving Tagsonomies: Keyword Relations Coming from Crowd. In N. T. Nguyen, R. Kowalczyk, and S.-M. Chen, editors, *Computational Collective Intelligence. Semantic Web, Social Networks and Multiagent System*, volume 5796 of *Lecture Notes in Computer Science*, pages 309–320. Springer, 2009. ISBN 978-3-642-04440-3. URL http://dx.doi.org/10.1007/978-3-642-04441-0_27. (Cited on page 82)

S.-A. Barnes, J. Bimrose, C. Bradley, A. Brown, D. Feldkamp, P. Franzolini, A. Kaschig, M. Kohlegger, C. Kunzmann, J. Magenheim, R. Maier, T. Nelkner, S. Nikles, U. Riss, A. Sandow, A. Schmidt, M. Shuttleworth, and S. Thalmann. Results of the Ethnographic Study and Conceptual Knowledge Maturing Model. MATURE Deliverable D1.1, April 2009. (Cited on page 266)

S.-A. Barnes, C. Bradley, A. Brown, J. Cook, A. Kaschig, C. Kunzmann, J. Magenheim, R. Maier, A. Mazarakis, A. Ravenscroft, U. Riss, A. Sandow, and A. Schmidt. Results of the representative study and refined conceptual knowledge maturing model. MATURE Deliverable D1.2, Apr. 2010. (Cited on page 112)

S.-A. Barnes, J. Bimrose, A. Brown, A. Kaschig, C. Kunzmann, T. Ley, R. Maier, J. Magenheim, A. Mazarakis, A. Sandow, A. Schmidt, and P. Seitlinger. Results of in-depth case studies, recommendations and final knowledge maturing model. MATURE Deliverable D1.3, April 2011. (Cited on page 266)

I. Becerra-Fernandez. Searching for experts on the Web: A review of contemporary expertise locator systems. *ACM Transactions on Internet Technologies*, 6(4):333–355, 2006. ISSN 1533-5399. doi: http://doi.acm.org/10.1145/1183463.1183464. URL http://dx.doi.org/10.1145/1183463.1183464. (Cited on pages 99 and 256)

S. Bechhofer and C. Goble. Towards Annotation using DAML+OIL. In *Workshop Knowledge Markup and Semantic Annotation*, Victoria B.C., Canada, Oct. 2001. (Cited on page 84)

S. Bechhofer, I. Horrocks, C. A. Goble, and R. Stevens. OilEd: a Reasonable Ontology Editor for the Semantic Web. In C. A. Goble, D. L. McGuinness, R. Möller, and P. F. Patel-Schneider, editors, *Description Logics*, volume 49 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2001. (Cited on page 64)

S. Bechhofer, Y. Yesilada, and B. Horan. COHSE: Knowledge-Driven Hyperlinks. In *Semantic Web Challenge at ISWC2006*, 2006. URL http://iswc2006.semanticweb.org/submissions/web_challenge.htm. (Cited on page 85)

S. Bechhofer, Y. Yesilada, R. Stevens, S. Jupp, and B. Horan. Using Ontologies and Vocabularies for Dynamic Linking. *IEEE Internet Computing*, 12(3):32–39, 2008. (Cited on page 85)

G. Begelman, P. Keller, and F. Smadja. Automated Tag Clustering: Improving search and exploration in the tag space. In *Collaborative Web Tagging Workshop at WWW2006, Edinburgh, Scotland*, 2006. URL http://www.pui.ch/phred/automated_tag_clustering/. (Cited on page 79)

G. Berio and M. Harzallah. Knowledge Management for Competence Management. *Journal of Universal Knowledge Management*, 0(1):21–38, 2005. (Cited on page 255)

A. Bernaras, I. Laresgoiti, and J. Corera. Building and Reusing Ontologies for Electrical Network Applications. In W. Wahlster, editor, *Proceedings of the 12th European Conference on Artificial Intelligence (ECAI'96)*, pages 298–302, Chichester, UK, 1996. John Wiley and Sons. ISBN 0471968099. (Cited on page 59)

T. Berners-Lee, Y. Chen, L. Chilton, D. Connolly, R. Dhanaraj, J. Hollenbach, A. Lerer, and D. Sheets. Tabulator: Exploring and analyzing linked data on the semantic web. In *In Procedings of the 3rd International Semantic Web User Interaction Workshop (SWUI06)*, 2006. URL http://serv2.ist.psu.edu:8080/viewdoc/summary?doi=10.1.1.97.950. (Cited on page 242)

M. Bernstein, D. S. Tan, G. Smith, M. Czerwinski, and E. Horvitz. Collabio: a game for annotating people within social networks. In A. D. Wilson and F. Guimbretière, editors, *Proceedings of the Annual ACM Symposium on User Interface Software and Technology (UIST'09)*, pages 97–100. ACM Press, 2009. ISBN 978-1-60558-745-5. (Cited on page 257)

M. S. Bernstein, D. S. Tan, G. Smith, M. Czerwinski, and E. Horvitz. Personalization via friendsourcing. *ACM Trans. Comput.-Hum. Interact.*, 17(2), 2010. (Cited on page 257)

E. Biesalski. *Unterstützung der Personalentwicklung mit ontologiebasiertem Kompetenzmanagement*. PhD thesis, Fakultät für Wirtschaftswissenschaften, Universität Karlsruhe, Karlsruhe, Germany, Aug. 2006. URL http://digbib.ubka.uni-karlsruhe.de/volltexte/documents/2755. (Cited on page 255)

E. Biesalski and A. Abecker. Human Resource Management with Ontologies. In K.-D. Althoff, A. Dengel, R. Bergmann, M. Nick, and T. Roth-Berghofer, editors, *Proceedings of WM2005, Workshop on IT Tools for Knowledge Management Systems: Applicability, Usability, and Benefits (KMTOOLS)*, volume 3782 of *Lecture Notes in Computer Science*, pages 499–507. Springer, 2005. ISBN 3-540-30465-7. (Cited on pages 2, 99, 220, 255, and 256)

N. Bikakis, G. Giannopoulos, T. Dalamagas, and T. Sellis. Integrating Keywords and Semantics on Document Annotation and Search. In *Proceedings of the 2010 international conference on On the move to meaningful internet systems: Part II*, OTM'10, pages 921–938, Berlin, Heidelberg, 2010. Springer-Verlag. ISBN 3-642-16948-1, 978-3-642-16948-9. (Cited on page 149)

K. Bischoff, C. S. Firan, W. Nejdl, and R. Paiu. Can All Tags Be Used for Search? In *CIKM '08: Proceeding of the 17th ACM conference on Information and knowledge management*, pages 193–202, New York, NY, USA, 2008. ACM. ISBN 978-1-59593-991-3. doi: 10.1145/1458082.1458112. URL http://dx.doi.org/10.1145/1458082.1458112. (Cited on pages 30 and 32)

S. Bloehdorn, K. Petridis, C. Saathoff, N. Simou, V. Tzouvaras, Y. Avrithis, S. Handschuh, Y. Kompatsiaris, S. Staab, and M. G. Strintzis. Semantic Annotation of Images and Videos for Multimedia Analysis. In A. Gómez-Pérez and J. Euzenat, editors, *The Semantic Web: Research and Applications: Proceedings of the Second European Semantic Web Conference, ESWC 2005, Heraklion, Crete, Greece, May 29-June 1, 2005*, volume 3532 of *Lecture Notes in Computer Science*, pages 592–607. Springer, Berlin–Heidelberg, Germany, 2005. URL http://dx.doi.org/10.1007/11431053_40. (Cited on page 84)

S. Bødker, P. Ehn, D. Sjögren, and Y. Sundblad. Co-operative Design – perspectives on 20 years with 'the Scandinavian IT Design Model. In *Proceedings of NordiCHI 2000, Stockholm, October 2000*, Stockholm, Sweden, 2000. URL http://cid.nada.kth.se/pdf/cid_104.pdf. (Cited on page 6)

T. Bogers. *Recommender Systems for Social Bookmarking*. PhD thesis, Tilburg University, Tilburg, The Netherlands, Dec. 2009. URL http://ilk.uvt.nl/~toine/phd-thesis/. (Cited on page 22)

W. N. Borst. *Construction of Engineering Ontologies for Knowledge Sharing and Reuse*. PhD thesis, Universiteit Twente, Enschede, Sept. 1997. URL http://doc.utwente.nl/17864/. (Cited on pages 42 and 43)

J. Bortz and N. Döring. *Forschungsmethoden und Evaluation*. Springer-Lehrbuch. Springer-Medizin-Verl., Heidelberg, 4., überarb. aufl. edition, 2006. ISBN 978-3-540-33305-0. (Cited on page 170)

E. Bozsak, M. Ehrig, S. Handschuh, A. Hotho, A. Maedche, B. Motik, D. Oberle, C. Schmitz, S. Staab, L. Stojanovic, N. Stojanovic, R. Studer, G. Stumme, Y. Sure, J. Tane, R. Volz, and V. Zacharias. KAON - Towards a Large Scale Semantic Web. In K. Bauknecht, A. M. Tjoa, and G. Quirchmayr, editors, *Proceedings of the Third International Conference on E-Commerce and Web Technologies EC-Web 2002*, volume 2455 of *Lecture Notes in Computer Science*, pages 304–313. Springer, 2002. ISBN 3-540-44137-9. URL http://www.aifb.uni-karlsruhe.de/WBS/ysu/publications/2002_ecweb_kaon.pdf. (Cited on page 64)

R. Brachman. What IS-A Is and Isn't: An Analysis of Taxonomic Links in Semantic Networks. *IEEE Computer*, 16:30–36, 1983. ISSN 0018-9162. doi: 10.1109/MC.1983.1654194. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1654194. (Cited on page 50)

C. Bradley, S. Braun, R. Brun, J. Cook, K. Hinkelmann, B. Hu, C. Kunzmann, T. Ley, A. Martin, A. Mazarakis, T. Nelkner, A. Ravenscroft, U. Riss, A. Schmidt, K. Schöfegger, B. Thönssen, N. Weber, and H. F. Witschel. Design and Delivery of Demonstrators of PLME / OLME and Tool Wrapper Infrastructure. MATURE Deliverable D2.2/3.2, Apr. 2010. (Cited on pages 9 and 198)

S. Braun and M. Hefke. Im WISSENSNETZ – Vernetzte Informationsprozesse in Forschungsverbünden. *Open Journal of Knowledge Management*, Feb. 2009. URL http://www.community-of-knowledge.de/fileadmin/user_upload/attachments/Im_Wissensnetz.pdf. (Cited on page 85)

S. Braun, M. Hefke, A. Schmidt, and N. Sevilmis. Im Wissensnetz: Linked Information Processes in Research Networks. In *Proceedings of the German e-Science Conference 2007 (GES 2007)*, May 2007. URL http://www.fzi.de/ipe/publikationen.php?id=1650. (Cited on page 85)

S. Braun, J. Busse, J. Franz, R. Schultz, and N. Sevilmis. E16: WISSENSNETZ Methodik Handbuch. deliverable, Dec. 2008. (Cited on page 7)

S. Braun, A. Brown, G. Attwell, P. Rustemeier, J. Magenheim, T. Nelkner, K. Schöfegger, and A. S. Tobias Ley. Pedagogical and usability foundations and concept for a PLME. MATURE Deliverable D2.1, Apr. 2009a. (Cited on page 7)

S. Braun, C. Schora, and V. Zacharias. Semantics to the Bookmarks: A Review of Social Semantic Bookmarking Systems. In A. Paschke, H. Weigand, W. Behrendt, K. Tochtermann, and T. Pellegrini, editors, *5th International Conference on Semantic Systems (I-SEMANTICS 2009), Proceedings of I-KNOW 09 and I-SEMANTICS 09*, pages 445–454, Graz, Austria, 2009b. Verlag der Technischen Universitt Graz. URL http://mature-ip.eu/files/papers/iknow09/semantics_to_the_bookmarks_review.pdf. (Cited on pages 110 and 248)

S. Braun, R. Brun, K. Hinkelmann, B. Hu, T. Ley, S. Lindstaedt, U. Riss, A. Schmidt, K. Schöfegger, P. Seitlinger, B. Thönssen, and N. Weber. Maturing Services Prototype V1. MATURE Deliverable D4.2, Apr. 2010a. (Cited on page 194)

S. Braun, C. Kunzmann, and A. Schmidt. People Tagging & Ontology Maturing: Towards Collaborative Competence Management. In D. Randall and P. Salembier, editors, *From CSCW to Web2.0: European Developments in Collaborative Design*, Computer Supported Cooperative Work, pages 133–154. Springer, Berlin/Heidelberg, 2010b. ISBN 978-1-84882-964-0. URL http://publications.andreas.schmidt.name/Braun_Kunzmann_Schmidt_PeopleTagging_book_2009.pdf. (Cited on page 101)

J. Breslin, A. Passant, and S. Decker. *The Social Semantic Web*. Springer-Verlag, Heidelberg, 2009. ISBN 978-3-642-01171-9. (Cited on pages 242 and 243)

J. G. Breslin, U. Bojars, B. Aleman-meza, H. Boley, L. J. Nixon, A. Polleres, and A. V. Zhdanova. Finding experts using internet-based discussions in online communities and associated social networks. In *First International ExpertFinder Workshop*, 2007. (Cited on page 256)

T. Bürger, M. Krötzsch, M. Luger, D. Vrandecic, and S. Wögler. Collaborative Articulation of Enterprise Knowledge (Demonstrator). ACTIVE Deliverable 1.3.2, 2010. (Cited on page 245)

D. Brickley and R. Guha. Rdf vocabulary description language 1.0: Rdf schema, Feb. 2004. URL http://www.w3.org/TR/rdf-schema/. (Cited on page 50)

D. Brickley and A. Miles. SKOS Core Vocabulary Specification. W3C Working Draft, Nov. 2005. URL http://www.w3.org/TR/2005/WD-swbp-skos-core-spec-20051102. (Cited on page 51)

M. Buffa, F. Gandon, P. Sander, C. Faron, and G. Ereteo. Sweetwiki: a semantic wiki. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(1), 2011. ISSN 1570-8268. URL http://www.websemanticsjournal.org/index.php/ps/article/view/138. (Cited on pages 244 and 246)

P. Buitelaar and P. Cimiano, editors. *Ontology Learning and Population: Bridging the Gap between Text and Knowledge*, volume 167 of *Frontiers in Artificial Intelligence and Applications*. IOS Press, Amsterdam, 2008. ISBN 978-1-58603-818-2. (Cited on page 77)

A. C. Bullinger. *Innovation and Ontologies: Structuring the Early Stages of Innovation Management*. Gabler, 1 edition, 2008. ISBN 3834912492. (Cited on pages 45 and 46)

V. Bush. As We May Think. *The Atlantic Monthly*, 176(1):101–108, 1945. (Cited on page 17)

S. E. Campanini, P. Castagna, and R. Tazzoli. Platypus Wiki: a Semantic Wiki Wiki Web. In *Semantic Web Applications and Perspectives, Proceedings of 1st Italian SemanticWeb Workshop*, Dec. 2004. URL http://semanticweb.deit.univpm.it/swap2004/cameraready/castagna.pdf. (Cited on page 244)

A. J. Cañas, G. Hill, R. Carff, N. Suri, J. Lott, T. Eskridge, G. Gómez, M. Arroyo, and R. Carvajal. CmapTools: A Knowledge Modeling and Sharing Environment. *Concept Maps: Theory, Methodology, Technology. Proceedings of the First International Conference on Concept Mapping*, 1:125–133, 2004. (Cited on page 76)

J. Cardoso. The semantic web vision: Where are we? *IEEE Intelligent Systems*, 22(5): 84–88, 2007. (Cited on page 3)

L. Carr, T. Miles-Board, G. Wills, A. Woukeu, and W. Hall. Towards a Knowledge-Aware Office Environment. In D. Karagiannis and U. Reimer, editors, *Practical Aspects of Knowledge Management*, volume 3336 of *Lecture Notes in Computer Science*, pages 129–140. Springer Berlin / Heidelberg, 2004. URL http://dx.doi.org/10.1007/978-3-540-30545-3_12. (Cited on page 85)

A. G. Castro, P. Rocca-Serra, R. Stevens, C. F. Taylor, K. Nashar, M. A. Ragan, and S.-A. Sansone. The use of concept maps during knowledge elicitation in ontology development processes - the nutrigenomics use case. *BMC Bioinformatics*, 7:267, 2006. (Cited on page 73)

C. Cattuto. Semiotic dynamics in online social communities. *The European Physical Journal C - Particles and Fields*, 46(/08/18/0):33–37, 2006. doi: 10.1140/epjcd/s2006-03-004-4. URL http://dx.doi.org/10.1140/epjcd/s2006-03-004-4. (Cited on page 35)

C. Cattuto, D. Benz, A. Hotho, and G. Stumme. Semantic Grounding of Tag Relatedness in Social Bookmarking Systems. In A. P. Sheth, S. Staab, M. Dean, M. Paolucci, D. Maynard, T. W. Finin, and K. Thirunarayan, editors, *The Semantic Web – ISWC 2008*, volume 5318 of *Lecture Notes in Computer Science*, pages 615–631, Berlin/Heidelberg, 2008. Springer. ISBN 978-3-540-88563-4. doi: 10.1007/978-3-540-88564-1_39. URL http://cxnets.googlepages.com/cattuto_iswc2008.pdf. (Cited on page 79)

B. Chandrasekaran, J. R. Josephson, and V. R. Benjamins. What are ontologies, and why do we need them? *IEEE Intelligent Systems and Their Applications*, 14(1):20–26, 1999. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=747902. (Cited on page 41)

C. Cheshire and J. Antin. The Social Psychological Effects of Feedback on the Production of Internet Information Pools. *Journal of Computer-Mediated Communication*, 13(3):705–727, 2008. (Cited on page 108)

P. Cimiano, A. Mädche, S. Staab, and J. Völker. Ontology Learning. In S. Staab and R. Studer, editors, *Handbook on Ontologies*, International Handbooks Information System, pages 245–267. Springer Berlin Heidelberg, 2009. ISBN 978-3-540-92673-3. URL http://dx.doi.org/10.1007/978-3-540-92673-3_11. (Cited on page 77)

F. Ciravegna and Y. Wilks. Designing Adaptive Information Extraction for the Semantic Web in Amilcare. In S. Handschuh and S. Staab, editors, *Annotation for the Semantic Web*. IOS Press, Amsterdam, 2003. (Cited on page 84)

T. Coates. An addendum to a definition of Social Software, 2005. URL http://www.plasticbag.org/archives/2005/01/an_addendum_to_a_definition_of_social_software/. (Cited on page 17)

N. Collier, K. Takeuchi, and A. Kawazoe. Open Ontology Forge: an environment for text mining in a Semantic Web world? In *Proc. International Workshop on Semantic Web Foundations and Application Technologies*, pages pp. 17–24, 2003. URL http://sites.google.com/site/nhcollier/publications-1. (Cited on page 85)

O. Corcho, M. Fernández-López, and A. Gómez-Pérez. Methodologies, tools and languages for building ontologies. Where is their meeting point? *Data & Knowledge Engineering*, 46(1):41–64, 2003. (Cited on pages 44 and 63)

F. Crestani. Application of spreading activation techniques in information retrieval. *Artificial Intelligence Review*, 11(6):453–482, 1997. (Cited on page 82)

R. Crowder, G. V. Hughes, and W. Hall. Approaches to locating expertise using corporate knowledge. *Int. Syst. in Accounting, Finance and Management*, 11(4):185–200, 2002. (Cited on page 256)

C. V. Damme, T. Coenen, and E. Vandijck. Deriving a Lightweight Corporate Ontology form a Folksonomy: a Methodology and its Possible Applications. *Scalable Computing: Practice and Experience - Int. J. for Parallel and Distributed Computing*, 9(4):293–301, 2008. URL http://www.scpe.org/?a=volume&v=38. (Cited on pages 22, 81, and 82)

danah boyd. issues of culture in ethnoclassification/folksonomy. Weblog, Jan. 2005. URL http://www.zephoria.org/thoughts/archives/2005/01/28/issues_of_culture_in_ethnoclassificationfolksonomy.html. (Cited on page 20)

DATech. Leitfaden usability, 2009. URL http://www.datech.de/share/files/Leitfaden-Usability.pdf. (Cited on page 175)

A. de Moor, P. D. Leenheer, and R. Meersman. DOGMA-MESS: A Meaning Evolution Support System for Interorganizational Ontology Engineering. In H. Schärfe, P. Hitzler, and P. Øhrstrøm, editors, *Proceedings of the 14th International Conference on Conceptual Structures (ICCS 2006)*, volume 4068 of *Lecture Notes in Computer Science*, pages 189–202. Springer, 2006. ISBN 3-540-35893-5. URL http://dx.doi.org/10.1007/11787181_14. (Cited on pages 72 and 90)

S. Decker and M. Frank. The Social Semantic Desktop. Technical Report DERI-TR-2004-05-02, DERI - Digital Enterprise Research Instiute Galway, Galway, Ireland, May 2004. URL http://www.deri.ie/fileadmin/documents/DERI-TR-2004-05-02.pdf. (Cited on page 85)

S. Decker, D. Fensel, F. van Harmelen, I. Horrocks, S. Melnik, M. C. A. Klein, and J. Broekstra. Knowledge Representation on the Web. In *Description Logics*, pages 89–97, 2000. (Cited on pages 3 and 90)

M. del Carmen Suárez-Figueroa. *NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse*. PhD thesis, Universidad Politécnica de Madrid, Madrid, Spain, June 2010. URL http://oa.upm.es/3879/2/MARIA_DEL-_CARMEN_SUAREZ_DE_FIGUEROA_BAONZA.pdf. (Cited on page 70)

M. del Carmen Suárez-Figueroa, G. A. de Cea, C. Buil, K. Dellschaft, M. Fernández-López, A. García, A. Gómez-Pérez, G. Herrero, E. Montiel-Ponsoda, M. Sabou, B. Villazon-Terrazas, and Z. Yufei. D5.4.1 NeOn Methodology for Building Contextualized Ontology Networks. NeOn Deliverable D5.4.2, Feb. 2008. (Cited on page 70)

M. del Carmen Suárez-Figueroa, E. Blomqvist, M. d'Aquin, M. Espinoza, A. Gómez-Pérez, H. Lewen, I. Mozetic, R. Palma, M. Poveda, M. Sini, B. Villazon, F. Zablith, and M. Dzbor. Revision and Extension of the NeOn Methodology for Building Contextualized Ontology Networks. NeOn Deliverable D5.4.2, Feb. 2009. (Cited on page 70)

M. del Carmen Suárez-Figueroa, A. Gómez-Pérez, E. Motta, and A. Gangemi, editors. *Ontology Engineering in a Networked World*. Springer, 2011. (Cited on page 70)

A. L. Delbecq and A. H. VandeVen. A Group Process Model for Problem Identification and Program Planning. *Journal Of Applied Behavioral Science VII*, pages 466–491, 1971. (Cited on page 68)

K. Dellschaft, H. Engelbrecht, J. M. Barreto, S. Rutenbeck, and S. Staab. Cicero: Tracking Design Rationale in Collaborative Ontology Engineering. In S. Bechhofer, M. Hauswirth, J. Hoffmann, and M. Koubarakis, editors, *Proceedings of the 5th European Semantic Web Conference*, volume 5021 of *Lecture Notes in Computer Science*, pages 782–786. Springer, 2008. ISBN 978-3-540-68233-2. URL http://dx.doi.org/10.1007/978-3-540-68234-9_58. (Cited on page 71)

M. Denny. Ontology Tools Survey, Revisited. Technical report, XML.com, July 2004. URL http://www.xml.com/lpt/a/2004/07/14/onto.html. (Cited on page 63)

A. Díaz, G. Baldo, and G. Canals. Co-Protégé: Collaborative Ontology Building with Divergences. In *DEXA Workshops*, pages 156–160. IEEE Computer Society, 2006. (Cited on page 75)

S. Dill, N. Eiron, D. Gibson, D. Gruhl, R. V. Guha, A. Jhingran, T. Kanungo, S. Rajagopalan, A. Tomkins, J. A. Tomlin, and J. Y. Zien. SemTag and Seeker: Bootstrapping the Semantic Web via Automated Semantic Annotation. In *Proc. of the Twelfth International World Wide Web Conference (WWW2003)*, pages 178–186. ACM Press, 2003. (Cited on page 86)

L. Ding, T. Finin, A. Joshi, R. Pan, R. S. Cost, Y. Peng, P. Reddivari, V. Doshi, and J. Sachs. Swoogle: a search and metadata engine for the semantic web. In *CIKM '04: Proceedings of the thirteenth ACM international conference on Information and knowledge management*, pages 652–659, New York, NY, USA, 2004. ACM. ISBN 1-58113-874-1. doi: http://doi.acm.org/10.1145/1031171.1031289. (Cited on page 243)

C. Dittmann, M. Dittmann, I. Peters, and K. Weller. Persönliches Tag Gardening mit tagCare. In M. Ockenfeld, editor, *Generation international - die Zukunft von Information, Wissenschaft und Profession. Proceedings der 31. Online-Tagung der Germany.*, pages 117–128. DGI Frankfurt a. M., 2009. URL http://wwwalt.phil-fak.uni-duesseldorf.de/infowiss/content/mitarbeiter/peters.php. Frankfurt am Main: DGI. (Cited on page 261)

J. Domingue. Tadzebao and WebOnto: Discussing, browsing, and editing ontologies on the web. In *In Proceedings of the 11th Knowledge Acquisition for Knowledge-Based Systems Workshop*, 1998. (Cited on pages 64 and 75)

M. Dubinko, R. Kumar, J. Magnani, J. Novak, P. Raghavan, and A. Tomkins. Visualizing Tags over Time. In *Proceedings of the 15th international conference on World Wide Web(WWW '06)*, pages 193–202, New York, NY, USA, 2006. ACM. ISBN 1-59593-323-9. URL http://dx.doi.org/10.1145/1135777.1135810. (Cited on page 32)

B. Dutta and F. Giunchiglia. Social tagging: Semantics are actually used. Technical report, University of Trento, Oct. 2008. URL http://eprints.biblio.unitn.it/archive/00001507/01/061.pdf. (Cited on page 31)

T. Economist. The data deluge: Businesses, governments and society are only starting to tap its vast potential, Feb. 2010. URL http://www.economist.com/node/15579717. (Cited on page 1)

T. Erickson and W. A. Kellogg. Social Translucence: An Approach to Designing Systems that Support Social Processes. *ACM Transactions on Computer-Human Interaction*, 7: 59–83, 2000. (Cited on page 23)

European Council. Lisbon European Council 23 and 24 March 2000: Presidency Conclusions, Mar. 2000. URL http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ec/00100-r1.en0.htm. (Cited on page 1)

J. Euzenat. Corporate Memory Through Cooperative Creation of Knowledge Bases and Hyper-Documents. In *In Proceedings of the 10th Knowledge Acquisition, Modeling and Management for Knowledge-based Systems Workshop (KAW'96*, pages 1–18, 1996. (Cited on page 75)

U. Farooq, T. G. Kannampallil, Y. Song, C. H. Ganoe, J. M. Carroll, and L. Giles. Evaluating tagging behavior in social bookmarking systems: metrics and design heuristics. In *GROUP '07: Proceedings of the 2007 international ACM conference on Conference on supporting group work*, pages 351–360, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-845-9. doi: http://doi.acm.org/10.1145/1316624.1316677. URL http://dx.doi.org/10.1145/1316624.1316677. (Cited on page 25)

A. Farquhar, R. Fikes, and J. Rice. The Ontolingua Server: A Tool for Collaborative Ontology Construction. *International Journal of Human-Computer Studies*, 46(6):707–727, 1997. (Cited on pages 63 and 75)

S. Farrell and T. Lau. Fringe Contacts: People-Tagging for the Enterprise. In F. Smadja, A. Tomkins, and S. Golder, editors, *Proc. of the Collaborative Web Tagging Workshop at WWW2006*, 2006. URL http://www.rawsugar.com/www2006/25.pdf. (Cited on page 257)

S. Farrell, T. Lau, and S. Nusser. Building Communities with People-Tags. In M. C. C. Baranauskas, P. A. Palanque, J. Abascal, and S. D. J. Barbosa, editors, *INTERACT (2)*, volume 4663 of *Lecture Notes in Computer Science*, pages 357–360. Springer, 2007a. ISBN 978-3-540-74799-4. (Cited on page 257)

S. Farrell, T. Lau, S. Nusser, E. Wilcox, and M. Muller. Socially augmenting employee profiles with people-tagging. In *UIST '07: Proceedings of the 20th annual ACM symposium on User interface software and technology*, pages 91–100, New York, NY, USA, 2007b. ACM. ISBN 9781595936792. doi: 10.1145/1294211.1294228. URL http://dx.doi.org/10.1145/1294211.1294228. (Cited on pages 101 and 257)

G. Fauconnier and M. Turner. Conceptual integration networks. *Cognitive Science*, 22:133–187, 1998. URL http://www.sciencedirect.com/science/article/B6W48-3Y2G1WN-H/2/98d434f1f5368ca8ad51ad2b0f259932. (Cited on page 258)

D. Fensel. *Ontologies A Silver Bullet for Knowledge Management and Electronic Commerce*. Springer, 2nd edition, 2004. (Cited on pages 44 and 63)

M. Fernández, A. Gómez-Pérez, and N. Juristo. METHONTOLOGY: from Ontological Art towards Ontological Engineering. In *Proceedings of the AAAI97 Spring Symposium Series on Ontological Engineering*, 1997. (Cited on page 60)

G. Fischer, R. McCall, J. Ostwald, B. Reeves, and F. Shipman. Seeding, Evolutionary Growth and Reseeding: Supporting the Incremental Development of Design Environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence*, CHI '94, pages 292–298, New York, NY, USA, 1994. ACM. ISBN 0-89791-650-6. doi: 10.1145/191666.191770. URL http://doi.acm.org/10.1145/191666.191770. (Cited on pages 103 and 184)

G. Fischer, J. Grudin, R. McCall, J. Ostwald, D. Redmiles, B. Reeves, and F. Shipman. Seeding, evolutionary growth and reseeding: The incremental development of collaborative design environments. In G. Olson, T. Malone, and J. Smith, editors, *Coordination theory and collaboration technology*, pages 447–472. Lawrence Erlbaum Associates Inc., Mahwah, N.J., May 2001. (Cited on pages 103, 104, 108, and 184)

C. Fluit, M. Sabou, and F. van Harmelen. Supporting User Tasks through Visualisation of Light-weight Ontologies. In S. Staab and R. Studer, editors, *Handbook on Ontologies*, International Handbooks on Information Systems, pages 415–434. Springer, 2004. ISBN 3-540-40834-7. (Cited on page 90)

G. W. Furnas, T. K. Landauer, L. M. Gomez, and S. T. Dumais. The vocabulary problem in human-system communication. *Communications of the ACM*, 30(11):964–971, 1987. ISSN 0001-0782. doi: 10.1145/32206.32212. URL http://dx.doi.org/10.1145/32206.32212. (Cited on page 38)

D. Gašević, D. Djurić, and V. Devedžić. *Model Driven Engineering and Ontology Development*. Springer, Berlin, 2. edition, 2009. ISBN 978-3-642-00281-6. doi: 10.1007/978-3-642-00282-3. (Cited on page 53)

J. H. Gennari, M. A. Musen, R. W. Fergerson, W. E. Grosso, M. Crubézy, H. Eriksson, N. F. Noy, and S. W. Tu. The evolution of Protégé: an environment for knowledge-based systems development. *Int. J. Hum.-Comput. Stud.*, 58:89–123, Jan. 2003. ISSN 1071-5819. doi: 10.1016/S1071-5819(02)00127-1. URL http://dx.doi.org/10.1016/S1071-5819(02)00127-1. (Cited on page 65)

C. Ghidini, B. Kump, S. Lindstaedt, N. Mahbub, V. Pammer, M. Rospocher, and L. Serafini. Moki: The enterprise modelling wiki. In L. Aroyo, P. Traverso, F. Ciravegna, P. Cimiano, T. Heath, E. Hyvönen, R. Mizoguchi, E. Oren, M. Sabou, and E. Simperl, editors, *The Semantic Web: Research and Applications*, volume 5554 of *Lecture Notes in Computer Science*, pages 831–835. Springer Berlin / Heidelberg, 2009. ISBN 978-3-642-02120-6. URL http://dx.doi.org/10.1007/978-3-642-02121-3_65. (Cited on page 246)

C. Ghidini, M. Rospocher, and L. Serafini. Moki: a wiki-based conceptual modeling tool. In *ISWC 2010 Posters & Demonstrations Track: Collected Abstracts*, volume 658 of *CEUR Workshop Proceedings (CEUR-WS.org)*, pages 77– 80, Shanghai, China, 2010. (Cited on page 246)

F. Giunchiglia, F. Giunchiglia, M. Marchese, M. Marchese, I. Zaihrayeu, and I. Zaihrayeu. Towards a theory of formal classification. In *Proceedings of the AAAI-05 Workshop on Contexts and Ontologies: Theory, Practice and Applications (C&O-2005)*, pages 1–8. AAAI Press. ISBN, 2005. (Cited on page 81)

S. Golder and B. A. Huberman. The Structure of Collaborative Tagging Systems. *Journal of Information Sciences*, 32(2):198–208, 2006. URL http://www.hpl.hp.com/research/idl/papers/tags/index.html. (Cited on pages 20, 25, 30, 31, 35, and 38)

A. Gómez-Pérez, M. Fernández-López, and O. Corcho. *Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. Advanced Information and Knowledge Processing. Springer, 1st edition, 2004. (Cited on pages 3, 45, 50, 54, 56, 57, 58, 61, and 63)

S. Grimm, A. Abecker, J. Völker, and R. Studer. Ontologies and the Semantic Web Foundations, Applications and Engineering. pages 1–75. 2011. (Cited on pages 41, 44, 50, 53, 54, and 56)

K. Groth. *On knowing who knows an alternative approach to knowledge management*. PhD thesis, Interaction and Presentation Laboratory (IPLab) Numerical Analysis and Computing Science (Nada) Royal Institute of Technology (KTH), Stockholm, Sweden, 2004. URL ftp://ftp.nada.kth.se/pub/documents/IPLab/TechReports/IPLab-220.pdf. (Cited on page 1)

B. Gruber. Kollaborative erstellung von wissensstrukturen: Experiment und vergleich zur ermittlung von anforderungen. Diploma thesis, Institut für Wirtschaftsinformatik, Fakultät für Betriebswirschaftslehre der Universität Innsbruck, Innsbruck, Austria, Jan. 2009a. (Cited on pages 189 and 192)

T. Gruber. Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human Computer Studies*, 43(5):907–928, 1995. (Cited on page 55)

T. Gruber. Ontology of Folksonomy: A Mash-up of Apples and Oranges. *International Journal on Semantic Web & Information Systems*, 3(2):1–11, 2007. (Cited on page 21)

T. Gruber. What is an Ontology?, 2009b. URL http://www-ksl.stanford.edu/kst/what-is-an-ontology.html. (Cited on pages 42 and 89)

T. R. Gruber. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, 5(2):199–221, 1993. URL http://tomgruber.org/writing/ontolingua-kaj-1993.pdf. (Cited on pages 42 and 43)

M. Gruninger. Designing and Evaluating Generic Ontologies. In *In Proceedings of Workshop on Ontological Engineering, 12th European Conference of Artificial Intelligence*, pages 53–64, 1996. (Cited on page 58)

M. Gruninger and M. S. Fox. Methodology for the Design and Evaluation of Ontologies. In *International Joint Conference on Artificial Inteligence (IJCAI95), Workshop on Basic Ontological Issues in Knowledge Sharing*, 1995. (Cited on page 58)

M. Gruninger and J. Lee. ONTOLOGY Applications and Design. *Communications of the ACM*, 45(2):39–41, Feb. 2002. (Cited on page 53)

M. Gruninger, O. Bodenreider, F. Olken, L. Obrst, and P. Yim. Ontology Summit 2007 – Ontology, taxonomy, folksonomy: Understanding the distinctions. *Applied Ontology*, 3 (3):191–200, 2008. (Cited on page 46)

N. Guarino. Formal Ontology and Information Systems. pages 3–15. IOS Press, 1998. (Cited on pages 42, 43, and 44)

N. Guarino and P. Giaretta. Ontologies and Knowledge Bases: Towards a Terminological Clarification. In *Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing*, pages 25–32, Amsterdam, 1995. IOS Press. (Cited on pages 41 and 42)

N. Guarino and C. A. Welty. Evaluating ontological decisions with OntoClean. *Communications of the ACM*, 45(2):61–65, 2002. URL http://citeseer.ist.psu.edu/guarino02evaluating.html. (Cited on page 62)

N. Guarino, D. Oberle, and S. Staab. What Is an Ontology? In S. Staab and R. Studer, editors, *Handbook on Ontologies*, International Handbooks on Information Systems, pages 1–17. Springer Berlin Heidelberg, 2009. ISBN 978-3-540-92673-3. URL http://dx.doi.org/10.1007/978-3-540-92673-3_0. 10.1007/978-3-540-92673-3_0. (Cited on pages xv, 45, 47, and 89)

J. R. Gutierrez-Pulido, M. A. G. Ruiz, R. Herrera, E. Cabello, S. Legrand, and D. Elliman. Ontology languages for the semantic web: A never completely updated review. *Knowledge-Based Systems*, 19(7):489–497, 2006. (Cited on page 50)

M. Guy and E. Tonkin. Folksonomies: Tidying Up Tags? *D-Lib Magazine*, 12(1), Jan. 2006. ISSN 1082-9873. URL `http://www.dlib.org/dlib/january06/guy/01guy.html`. (Cited on pages 24 and 38)

M. Hadzic, P. Wongthongtham, T. Dillon, and E. Chang. *Ontology-based Multi-Agent Systems*. Springer, Berlin Heidelberg, Germany, 1st edition, 2009. ISBN 978-3-642-01903-6. (Cited on page 53)

J. Hagelauer. Do Tags Die? Analyzing the Life Cycle of Tags in Social Bookmarking Systems. Master's thesis, Institut für Angewandte Informatik und Formale Beschreibungsverfahren des Karlsruher Instituts für Technologie, may 2011. (Cited on pages 142 and 143)

H. Halpin and H. Shepard. Evolving Ontologies from Folksonomies: Tagging as a Complex System. http://www.ibiblio.org/hhalpin/homepage/notes/taggingcss.html, 2006. URL `http://www.ibiblio.org/hhalpin/homepage/notes/taggingcss.html`. (Cited on pages 35 and 36)

H. Halpin, V. Robu, and H. Shepard. The Dynamics and Semantics of Collaborative Tagging. In K. Möller, A. de Waard, S. Cayzer, M.-R. Koivunen, M. Sintek, and S. Handschuh, editors, *Proceedings of the 1st Semantic Authoring and Annotation Workshop (SAAW'06)*, Athens (GA), USA, Nov. 2006. URL `http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-209/saaw06-full01-halpin.pdf`. issn=1613-0073. (Cited on page 36)

T. Hammond, T. Hannay, B. Lund, and J. Scott. Social Bookmarking Tools (I): A General Review. *D-Lib Magazine*, 11(4), Apr. 2005. URL `http://www.dlib.org/dlib/april05/hammond/04hammond.html`. (Cited on pages 20, 23, and 26)

S. Handschuh and S. Staab. Authoring and annotation of web pages in CREAM. In *WWW '02: Proceedings of the 11th international conference on World Wide Web*, pages 462–473, New York, NY, USA, 2002. ACM. ISBN 1-58113-449-5. doi: http://doi.acm.org/10.1145/511446.511506. (Cited on page 84)

S. Handschuh, S. Staab, and A. Maedche. CREAM: creating relational metadata with a component-based, ontology-driven annotation framework. In *K-CAP '01: Proceedings of the 1st international conference on Knowledge capture*, pages 76–83, New York, NY, USA, 2001. ACM. ISBN 1-58113-380-4. doi: http://doi.acm.org/10.1145/500737.500752. URL `http://dx.doi.org/10.1145/500737.500752`. (Cited on pages 83 and 84)

S. Handschuh, S. Staab, and F. Ciravegna. S-CREAM — Semi-automatic CREAtion of Metadata. In *13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02)*, pages 358–372, Siguenza, Spain, 2002. (Cited on page 84)

H.-J. Happel. Social search and need-driven knowledge sharing in Wikis with Woogle. In D. Riehle and A. Bruckman, editors, *5th International Symposium on Wikis and Open Collaboration*, pages 1–10. ACM, 2009. ISBN 978-1-60558-730-1. URL `http://doi.acm.org/10.1145/1641309.1641329`. (Cited on page 243)

C. Hayes, P. Avesani, and U. Bojars. An Analysis of Bloggers, Topics and Tags for a Blog Recommender System. In B. Berendt, A. Hotho, D. Mladenic, and G. Semeraro, editors, *From Web to Social Web: Discovering and Deploying User and Content Profiles*, volume 4737 of *Lecture Notes in Computer Science*, pages 1–20. Springer Berlin / Heidelberg,

2007. URL http://dx.doi.org/10.1007/978-3-540-74951-6_1. 10.1007/978-3-540-74951-6_1. (Cited on page 20)

P. Hayes, T. C. Eskridge, M. Mehrotra, D. Bobrovnikoff, , T. Reichherzer, and R. Saavedra. COE: Tools for Collaborative Ontology Development and Reuse. In *Proceedings of 3rd Knowledge Capture Conference - K-Cap 2005*, 2005a. (Cited on page 76)

P. Hayes, T. C. Eskridge, R. Saavedra, T. Reichherzer, M. Mehrotra, and D. Bobrovnikoff. Collaborative knowledge capture in ontologies. In *K-CAP '05: Proceedings of the 3rd international conference on Knowledge capture*, pages 99–106, New York, NY, USA, 2005b. ACM Press. ISBN 1-59593-163-5. doi: http://doi.acm.org/10.1145/1088622.1088641. (Cited on page 76)

S. Hayman. Folksonomies and Tagging: New developments in social bookmarking. In *Proceedings of Ark Group Conference: Developing and Improving Classification Schemes*, Sydney, June 2007. Ark Group. URL http://www.educationau.edu.au/jahia/Jahia/home/pid/482. (Cited on pages 20 and 38)

M. Heckner, S. Mühlbacher, and C. Wolff. Tagging tagging. Analysing user keywords in scientific bibliography management systems. *JODI: Journal of Digital Information*, 9(2), 2008a. ISSN 13687506. URL http://epub.uni-regensburg.de/6839/1/HeckneretalTaggingTaggingJoDI2008.pdf. (Cited on pages xv, 25, 30, 31, and 32)

M. Heckner, T. Neubauer, and C. Wolff. Tree, funny, to_read, google: what are tags supposed to achieve? a comparative analysis of user keywords for different digital resource types. In *SSM '08: Proceeding of the 2008 ACM workshop on Search in social media*, pages 3–10, New York, NY, USA, 2008b. ACM. ISBN 978-1-60558-258-0. doi: 10.1145/1458583.1458589. URL http://dx.doi.org/10.1145/1458583.1458589. (Cited on pages 25 and 31)

M. Heckner, M. Heilemann, and C. Wolff. Personal Information Management vs. Resource Sharing: Towards a Model of Information Behaviour in Social Tagging Systems. In *Int'l AAAI Conference on Weblogs and Social Media (ICWSM)*, pages 42–49, San Jose, CA, USA, May 2009. (Cited on pages xiii, 25, 26, 28, 29, 32, and 38)

M. Hefke, S. Braun, A. Abecker, R. Traphöner, J. Fran, N. Sevilmis, R. Schulz, G. Helferich, H.-J. Richter, R. Meyer, P. Kastl, H. Appel, and J. Busse. Abschlussbericht Im WISSENSNETZ - Vernetzte Informationsprozesse in Forschungsverbünden. deliverable, Apr. 2009. (Cited on page 7)

J. Heflin and J. Hendler. A Portrait of the Semantic Web in Action. *IEEE Intelligent Systems*, 16(2):54–59, Mar. 2001. ISSN 1541-1672. doi: http://dx.doi.org/10.1109/5254.920600. URL http://dx.doi.org/10.1109/5254.920600. (Cited on page 84)

J. Heil. *From an ontological point of view*. Oxford University Press, New York, USA, 2003. (Cited on page 41)

J. Hendler. Agents and the Semantic Web. *IEEE Intelligent Systems Journal*, 16(2): 30–37, 2001. (Cited on page 42)

J. Hendler. Web 3.0: Chicken Farms on the Semantic Web. *Computer*, 41:106–108, 2008. ISSN 0018-9162. doi: http://doi.ieeecomputersociety.org/10.1109/MC.2008.34. (Cited on page 242)

M. Hepp. Possible Ontologies: How Reality Constraints Building Relevant Ontologies. *IEEE Internet Computing*, 11(1):90–96, 2007. doi: http://doi.ieeecomputersociety.org/10.1109/MIC.2007.20. URL http://www.heppnetz.de/files/IEEE-IC-PossibleOntologies-published.pdf. (Cited on pages 3, 4, 87, 89, and 94)

M. Hepp, D. Bachlechner, and K. Siorpaes. OntoWiki: community-driven ontology engineering and ontology usage based on Wikis. In *WikiSym '06: Proceedings of the 2006 international symposium on Wikis*, pages 143–144, New York, NY, USA, 2006. ACM Press. ISBN 1-59593-413-8. doi: http://doi.acm.org/10.1145/1149453.1149487. (Cited on page 3)

A. R. Hevner, S. T. March, J. Park, and S. Ram. Design science in information systems research. *Management Information Systems Quarterly*, 28(1):75–106, 2004. URL http://www.hec.unil.ch/yp/HCI/articles/hevner04.pdf. (Cited on pages 5 and 6)

P. Heymann and H. Garcia-Molina. Collaborative Creation of Communal Hierarchical Taxonomies in Social Tagging Systems. Technical Report 2006-10, Computer Science Department, Apr. 2006. URL http://dbpubs.stanford.edu:8090/pub/2006-10. (Cited on page 81)

P. Heymann, A. Paepcke, and H. Garcia-Molina. Tagging Human Knowledge. In *Third ACM International Conference on Web Search and Data Mining (WSDM2010)*, pages 1–10. Stanford University, Feb. 2010. URL http://www.wsdm-conference.org/2010/proceedings/docs/p51.pdf. (Cited on pages 31 and 32)

D. Hinchcliffe. The state of Enterprise 2.0. Blog "Enterprise Web 2.0", Oct. 2007. URL http://blogs.zdnet.com/Hinchcliffe/?p=143. (Cited on pages xv, 19, and 20)

C. Holsapple and K. Joshi. A collaborative approach to ontology design. *Communications of the ACM*, 45(2):42–47, 2002. (Cited on pages 68 and 77)

I. Horrocks, P. F. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, and M. Dean. Swrl: A semantic web rule language combining owl and ruleml, May 2004. URL http://www.w3.org/Submission/SWRL/. (Cited on page 51)

A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. FolkRank: A Ranking Algorithm for Folksonomies. In *Proc. FGIR 2006*, 2006a. URL http://www.kde.cs.uni-kassel.de/stumme/papers/2006/hotho2006folkrank.pdf. (Cited on page 79)

A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. BibSonomy: A Social Bookmark and Publication Sharing System. In A. de Moor, S. Polovina, and H. Delugach, editors, *Proceedings of the Conceptual Structures Tool Interoperability Workshop at the 14th International Conference on Conceptual Structures*, Aalborg, Denmark, July 2006b. Aalborg University Press. (Cited on pages 172 and 247)

A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Trend Detection in Folksonomies. In Y. S. Avrithis, Y. Kompatsiaris, S. Staab, and N. E. O'Connor, editors, *Proc. First International Conference on Semantics And Digital Media Technology (SAMT)*, volume 4306 of *Lecture Notes in Computer Science*, pages 56–70, Berlin, Heidelberg, 2006c. Springer. ISBN 3-540-49335-2. (Cited on page 21)

H. Hu and X. Du. ConAnnotator: Ontology-Aided Collaborative Annotation System. In *CSCWD*, pages 850–855. IEEE, 2006. ISBN 1-4244-0165-8. (Cited on page 86)

H. Hu, Y. Zhao, Y. Wang, M. Li, D. Wang, W. Wu, J. He, X. Du, and S. Wang. Cooperative Ontology Development Environment CODE and a Demo Semantic Web on Economics. In Y. Zhang, K. Tanaka, J. X. Yu, S. Wang, and M. Li, editors, *7th Asia-Pacific Web Conference on Web Technologies Research and Development - APWeb 2005*, volume 3399 of *Lecture Notes in Computer Science*, pages 1049–1052. Springer, 2005. ISBN 3-540-25207-X. (Cited on page 86)

J. Hunter. Collaborative Semantic Tagging and Annotation Systems. *Annual Review of Information Science and Technology*, 43(1):187–239, 2009. (Cited on page 83)

J. Hunter, I. Khan, and A. Gerber. Harvana: harvesting community tags to enrich collection metadata. In *JCDL '08: Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries*, pages 147–156, New York, NY, USA, 2008. ACM. ISBN 978-1-59593-998-2. doi: http://doi.acm.org/10.1145/1378889.1378916. (Cited on page 84)

L. Iverson, M. N. Razavi, and V. Mirzaee. Personal and Social Information Management with OPNTAG. In J. Cordeiro and J. Filipe, editors, *ICEIS 2008 - Proceedings of the Tenth International Conference on Enterprise Information Systems*, pages 195–203, 2008. ISBN 978-989-8111-40-1. (Cited on page 257)

M. Jarrar and R. Meersman. Formal Ontology Engineering in the DOGMA Approach. In R. Meersman and Z. Tari, editors, *On the Move to Meaningful Internet Systems 2002: CoopIS, DOA, and ODBASE*, volume 2519 of *Lecture Notes in Computer Science*, pages 1238–1254. Springer, 2002. ISBN 3-540-00106-9. URL http://dx.doi.org/10.1007/3-540-36124-3_78. (Cited on page 72)

R. Jasper and M. Uschold. A Framework for Understanding and Classifying Ontology Applications. In *Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods (KRR5)*, 1999. (Cited on page 42)

A. John and D. Seligmann. Collaborative Tagging and Expertise in the Enterprise. In *Proceedings of the Workshop on Collaborative Web Tagging at WWW2006*, 2006. URL http://www.semanticmetadata.net/hosted/taggingws-www2006-files/26.pdf. (Cited on pages 21 and 256)

R. Jäschke, A. Hotho, C. Schmitz, B. Ganter, and G. Stumme. Discovering Shared Conceptualizations in Folksonomies. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(1):38–53, Feb. 2008. ISSN 1570-8268. doi: 10.1016/j.websem.2007.11.004. URL http://www.sciencedirect.com/science/article/B758F-4R53WD4-1/2/ae56bd6e7132074272ca2035be13781b. (Cited on page 81)

S. Jupp, S. Bechhofer, and R. Stevens. A flexible API and editor for SKOS. In *6th Annual European Semantic Web Conference (ESWC2009)*, pages 506–520, June 2009. URL http://data.semanticweb.org/conference/eswc/2009/paper/180. (Cited on page 262)

J. Kahan and M.-R. Koivunen. Annotea: an open RDF infrastructure for shared Web annotations. In *WWW '01: Proceedings of the 10th international conference on World Wide Web*, pages 623–632, New York, NY, USA, 2001. ACM Press. ISBN 1581133480. URL http://dx.doi.org/10.1145/371920.372166. (Cited on page 83)

Y. Kalfoglou. *Exploring Ontologies*, volume 1 of *Handbook of Software Engineering and Knowledg*. World Scientic Publishing Company, 2000. URL http://www.cs.toronto.edu/~nernst/papers/kalfoglou-seng-onto.pdf. (Cited on page 42)

A. Kalyanpur, J. Hendler, B. Parsia, and J. Golbeck. SMORE – Semantic Markup, Ontology, and RDF Editor, 2003. URL http://www.mindswap.org/papers/SMORE.pdf. (Cited on page 84)

A. Kalyanpur, B. Parsia, E. Sirin, B. C. Grau, and J. Hendler. Swoop: A Web Ontology Editing Browser. *Web Semantics: Science, Services and Agents on the World Wide Web*, 4(2):144 – 153, 2006. ISSN 1570-8268. doi: DOI:10.1016/j.websem.2005.10.001. URL http://www.sciencedirect.com/science/article/B758F-4HKMPW2-1/2/58f53d4884b0c1db1c53b8471a984f5e. (Cited on page 65)

S. Karapiperis and D. Apostolou. Consensus Building in Collaborative Ontology Engineering Processes. *j-jukm*, 1(3):199–216, Sept. 2006. (Cited on pages 68 and 77)

V. Kashyap, C. Bussler, and M. Moran. Applications of Metadata and Ontologies. In M. J. Carey, S. Ceri, P. Bernstein, U. Dayal, C. Faloutsos, J. C. Freytag, G. Gardarin, W. Jonker, V. Krishnamurthy, M. A. Neimat, P. Valduriez, G. Weikum, K. Y. Whang, and J. Widom, editors, *The Semantic Web*, Data-Centric Systems and Applications, pages 161–192. Springer Berlin Heidelberg, 2008. ISBN 978-3-540-76452-6. URL http://dx.doi.org/10.1007/978-3-540-76452-6_7. 10.1007/978-3-540-76452-6_7. (Cited on pages 83 and 87)

M. Kifer, L. G., and J. Wu. Logical foundations of object-oriented and frame-based languages. *Journal of the ACM*, 42(4):741–843, 1995. (Cited on page 51)

M. Kipp and D. G. Campbell. Patterns and Inconsistencies in Collaborative Tagging Systems: An Examination of Tagging Practices. *Proceedings Annual General Meeting of the American Society for Information Science and Technology*, 2006. URL http://www.citebase.org/abstract?id=oai:eprints.rclis.org:8315. (Cited on page 25)

M. E. Kipp. @toread and cool : Subjective, affective and associative factors in tagging. In *Proceedings Canadian Association for Information Science/L'Association canadienne des sciences de l'information (CAIS/ACSI)*, 2008. URL http://eprints.rclis.org/archive/00013788/. (Cited on page 31)

M. E. I. Kipp. *Information Organisation Practices on the Web: Tagging and the Social Organisation of Information*. Phd thesis, School of Graduate and Postdoctoral Studies, The University of Western Ontario, London, Ontario, Canada, July 2009. URL http://www.eskimo.com/~{}meik/papers/MEIK-Thesis-20090723-final.pdf. (Cited on pages 25 and 36)

A. Klotz. An Efficient User Interface for Community-Driven Ontology Engineering. Master's thesis, Faculty of Mathematics, Computer Science, and Physics of the University of Innsbruck, Innsbruck, Austria, Mar. 2008. (Cited on page 260)

G. Klyne and J. J. Carroll. Resource Description Framework (RDF): Concepts and Abstract Syntax, Feb. 2004. URL http://www.w3.org/TR/rdf-concepts/. (Cited on page 50)

H. Knublauch, R. W. Fergerson, N. F. Noy, and M. A. Musen. The Protégé OWL Plugin: An Open Development Environment for Semantic Web Applications. In S. A. McIlraith, D. Plexousakis, and F. van Harmelen, editors, *The Semantic Web – ISWC 2004*, volume 3298 of *Lecture Notes in Computer Science*, pages 229–243. Springer Berlin / Heidelberg, 2004. URL http://dx.doi.org/10.1007/978-3-540-30475-3_17. (Cited on page 65)

M. Koch and A. Richter. *Enterprise 2.0 – Planung, Einführung und erfolgreicher Einsatz von Social Software in Unternehmen*. Oldenbourg, München, 2008. (Cited on pages xv, 17, and 18)

M.-R. Koivunen. Semantic Authoring By Tagging with Annotea Social Bookmarks and Topics. In *Proc. of the 1st Semantic Authoring and Annotation Workshop (SAAW2006)*, 2006. URL http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-209/saaw06-full08-koivunen.pdf. (Cited on pages 83, 84, and 248)

C. Körner. Understanding the Motivation behind Tagging. ACM Student Research Competition - Hypertext 2009, July 2009. URL http://kmi.tugraz.at/staff/markus/documents/2009_ACM_HT09_Understanding_the_motivation_behind_tagging_POSTER.pdf. (Cited on page 25)

C. Körner, R. Kern, H.-P. Grahsl, and M. Strohmaier. Of categorizers and describers: an evaluation of quantitative measures for tagging motivation. In *HT '10: Proceedings of the 21st ACM conference on Hypertext and hypermedia*, pages 157–166, New York, NY, USA, 2010. ACM. ISBN 978-1-4503-0041-4. doi: 10.1145/1810617.1810645. URL http://dx.doi.org/10.1145/1810617.1810645. (Cited on page 25)

K. Kotis. On Supporting HCOME-3O Ontology Argumentation Using Semantic Wiki Technology. In R. Meersman, Z. Tari, and P. Herrero, editors, *On the Move to Meaningful Internet Systems: OTM 2008 Workshops*, volume 5333 of *Lecture Notes in Computer Science*, pages 193–199, Berlin / Heidelberg, Germany, 2008. Springer. ISBN 978-3-540-88874-1. URL http://dx.doi.org/10.1007/978-3-540-88875-8_39. (Cited on page 72)

K. Kotis and A. Papasalouros. Learning Useful Kick-off Ontologies from Query Logs: HCOME Revised. In L. Barolli, F. Xhafa, S. Vitabile, and H.-H. Hsu, editors, *Proceedings of the International Conference on Complex, Intelligent and Software Intensive Systems*, pages 345–351, Los Alamitos, CA, USA, 2010. IEEE Computer Society. ISBN 978-0-7695-3967-6. URL http://doi.ieeecomputersociety.org/10.1109/CISIS.2010.50. (Cited on page 72)

K. Kotis and A. Vouros. Human-centered ontology engineering: The HCOME methodology. *Knowledge and Information Systems*, 10(1):109–131, July 2006. ISSN 0219-1377. doi: http://dx.doi.org/10.1007/s10115-005-0227-4. URL http://www.icsd.aegean.gr/kotis/publications/HCOME-KAIS.pdf. (Cited on pages 71 and 72)

K. Kotis, G. A. Vouros, and J. P. Alonso. HCOME: A Tool-Supported Methodology for Engineering Living Ontologies. In C. Bussler, V. Tannen, and I. Fundulaki, editors, *Semantic Web and Databases. Second International Workshop - SWDB 2004*, volume 3372 of *LNCS*, pages 155–166, Berlin Heidelberg, Germany, Aug. 2004. Springer-Verlag. ISBN 978-3-540-24576-6. doi: http://dx.doi.org/10.1007/b106149. URL http://springerlink.metapress.com/content/pe7y548qc7brht6u/?p=3bf02a6fa0b1473381521a1e3c951659&pi=11. (Cited on page 71)

K. Kozaki, Y. Kitamura, M. Ikeda, and R. Mizoguchi. Hozo: An Environment for Building/Using Ontologies Based on a Fundamental Consideration of "Role" and "Relationship". In A. Gómez-Pérez and V. R. Benjamins, editors, *EKAW*, volume 2473 of *Lecture Notes in Computer Science*, pages 213–218. Springer, 2002. ISBN 3-540-44268-5. (Cited on pages 75 and 172)

K. Kozaki, E. Sunagawa, Y. Kitamura, and R. Mizoguchi. A Framework for Cooperative Ontology Construction Based on Dependency Management of Modules. In L. Chen, P. Cudré-Mauroux, P. Haase, A. Hotho, and E. Ong, editors, *ESOE*, volume 292 of *CEUR Workshop Proceedings*, pages 33–44. CEUR-WS.org, 2007. (Cited on page 75)

E. Kroski. The Hive Mind: Folksonomies and User-Based Tagging, Dec. 2005. URL http://infotangle.blogsome.com/2005/12/07/the-hive-mind-folksonomies-and-user-based-tagging/. http://infotangle.blogsome.com/2005/12/07/the-hive-mind-folksonomies-and-user-based- (Cited on page 38)

M. Krötzsch, D. Vrandecic, M. Völkel, H. Haller, and R. Studer. Semantic Wikipedia. *Journal of Web Semantics*, 5:251–261, Sept. 2007. (Cited on pages 243 and 244)

S. Krug. *Don't make me Think! A Common Sense Approach to Web Usability*. New Riders Publishing, Indianapolis, USA, 2000. (Cited on page 171)

T. Kuhn. *Controlled English for Knowledge Representation*. PhD thesis, Faculty of Economics, Business Administration and Information Technology of the University of Zurich, Zurich, Switzerland, 2010. (Cited on page 245)

B. Kump, T. Ley, K. Schöffegger, N. Weber, D. Theiler, A. Schmidt, S. Braun, M. Ramezani, J. Hagelauer, S. Brander, H.-F. Witschel, B. Hu, and T. Nelkner. Maturing Services Prototype V2. deliverable, Apr. 2011. (Cited on page 9)

W. Kunz and H. Rittel. Issues as elements of information systems. Working Paper 131, Institute of Urban and Regional Development, University of California, Berkeley, California, 1970. (Cited on page 70)

J. Lacasta, J. N. Iso, and F. J. Z. Soria. *Terminological Ontologies: Design, Management and Practical Applications*, volume 9 of *Semantic Web and Beyond*. Springer US, 2010. ISBN 978-1-4419-6981-1. URL http://dx.doi.org/10.1007/978-1-4419-6981-1_1. 10.1007/978-1-4419-6981-1_1. (Cited on pages 49, 54, 109, and 267)

R. Lachica, D. Karabeg, and S. Rudan. Quality, Relevance and Importance in Information Retrieval with Fuzzy Semantic Networks. *TMRA Germany*, 2008. URL http://home.ifi.uio.no/dino/KF/Lachica-Karabeg08.pdf. (Cited on pages 243 and 247)

R. Lambiotte and M. Ausloos. Collaborative Tagging as a Tripartite Network. In V. Alexandrov, G. van Albada, P. Sloot, and J. Dongarra, editors, *Proceedings of the International Conference on Computational Science - ICCS 2006*, volume 3993 of *Lecture Notes in Computer Science*, pages 1114–1117. Springer Berlin / Heidelberg, 2006. URL http://dx.doi.org/10.1007/11758532_152. 10.1007/11758532_152. (Cited on page 21)

V. Lanfranchi, F. Ciravegna, and D. Petrelli. Semantic Web-Based Document: Editing and Browsing in AktiveDoc. In A. Gómez-Pérez and J. Euzenat, editors, *ESWC*, volume 3532 of *Lecture Notes in Computer Science*, pages 623–632. Springer, 2005. ISBN 3-540-26124-9. (Cited on page 85)

T. Lau and Y. Sure. Introducing Ontology-based Skills Management at a large Insurance Company. In *Modellierung 2002, Modellierung in der Praxis - Modellierung für die Praxis, Tutzing, Deutschland, 25.-27. März 2002*, pages 123–134, 2002. URL http://www.aifb.uni-karlsruhe.de/WBS/ysu/publications/2002_modellierung_thlaysu.pdf. (Cited on page 256)

K. J. Lee. What goes around comes around: an analysis of del.icio.us as social space. In *CSCW '06: Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pages 191–194, New York, NY, USA, 2006. ACM. ISBN 1-59593-249-6. doi: 10.1145/1180875.1180905. URL http://dx.doi.org/10.1145/1180875.1180905. (Cited on pages 25 and 30)

P. D. Leenheer. *On Community-based Ontology Evolution*. PhD thesis, Vrije Universiteit Brussel, FACULTY OF SCIENCE, Department of Computer Science, Semantics Technology and Applications Research Lab, Belgium, May 2009. URL http://www.vub.ac.be/phd/verdedigingen2009/20090519a.pdf. (Cited on pages 3 and 72)

D. Lenat and R. Guha. *Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project*. Addison-Wesley, 1990. (Cited on page 58)

T. Ley and P. Seitlinger. A Cognitive Perspective on Emergent Semantics in Collaborative Tagging: The Basic Level Effect. In E. F. Cena, A. Dattolo, S. Kleanthous, C. Tasso, D. B. Vallejo, and J. Vassileva:, editors, *CEUR Workshop Proceedings of the International Workshop on Adaptation in Social and Semantic Web (SASWeb2010)*, volume 590, pages 13–18, 2010. URL http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-590/sasweb10_2.pdf. (Cited on page 194)

T. Ley, S. Lindstaedt, K. Schöfegger, P. Seitlinger, N. Weber, B. Hu, U. Riss, R. Brun, K. Hinkelmann, B. Thönssen, R. Maier, and A. Schmidt. Maturing Services Definition. deliverable, Apr. 2009. (Cited on page 79)

T. Ley, B. Kump, and D. Albert. A methodology for eliciting, modelling, and evaluating expert knowledge for an adaptive work-integrated learning system. *Int. J. Hum.-Comput. Stud.*, 68:185–208, April 2010. ISSN 1071-5819. doi: http://dx.doi.org/10.1016/j.ijhcs.2009.12.001. URL http://dx.doi.org/10.1016/j.ijhcs.2009.12.001. (Cited on page 256)

M. Li, D. Wang, X. Du, and S. Wang. Ontology Construction for Semantic Web: A Role-Based Collaborative Development Method. In Y. Zhang, K. Tanaka, J. X. Yu, S. Wang, and M. Li, editors, *Web Technologies Research and Development - APWeb 2005*, volume 3399 of *Lecture Notes in Computer Science*, pages 609–619. Springer Berlin / Heidelberg, 2005. URL http://dx.doi.org/10.1007/978-3-540-31849-1_60. (Cited on page 86)

X. Lin, J. E. Beaudoin, Y. Bul, and K. Desal. Exploring Characteristics of Social Classification. In *Proceedings 17th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research 17*, 2006. URL http://dlist.sir.arizona.edu/1790/. (Cited on pages 31 and 32)

H. Lindstone and M. Turoff. *The Delphi Method: Technology and Applications*. Addison-Wesley, Reading, MA, 1975. (Cited on page 68)

K. Ling, G. Beenen, P. Ludford, X. Wang, K. Chang, X. Li, D. Cosley, D. Frankowski, L. Terveen, A. M. Rashid, P. Resnick, and R. Kraut. Using Social Psychology to Motivate Contributions to Online Communities. *Journal of Computer-mediated Communication*, 10(4), 2005. URL http://jcmc.indiana.edu/vol10/issue4/ling.html. (Cited on page 108)

M. F. López, A. Gómez-Pérez, J. P. Sierra, and A. P. Sierra. Building a chemical ontology using Methontology and the Ontology Design Environment. *IEEE Intelligent Systems and Their Applications*, 14(1):37–46,

1999. URL http://www.aifb.uni-karlsruhe.de/Lehrangebot/Sommer2001/SemanticWeb/papers/chemical_ontology.pdf. (Cited on pages 60 and 64)

W. Maass, T. Kowatsch, and T. Münster. Vocabulary Patterns in Free-for-all Collaborative Indexing Systems. In P. Haase, A. Hotho, L. Chen, E. Ong, and P. C. Mauroux, editors, *Proceedings of the International Workshop on Emergent Semantics and Ontology Evolution (ESOE2007) at ISWC/ASWC2007, Busan, South Korea*, Nov. 2007. (Cited on page 21)

G. Macgregor and E. McCulloch. Collaborative tagging as a knowledge organisation and resource discovery tool. *Library Review*, 55(5):291–300, 2006a. (Cited on page 35)

G. Macgregor and E. McCulloch. Collaborative Tagging as a Knowledge Organisation and Resource Discovery Tool. *Library Review*, 55(5):291–300, 2006b. ISSN 0024-2535. doi: 10.1108/00242530610667558. URL http://www.emeraldinsight.com/10.1108/00242530610667558. (Cited on page 88)

A. Maedche. *Ontology Learning for the Semantic Web*. Kluwer Academic Publishing, Boston, 2002. (Cited on page 105)

R. Maier and S. Thalmann. Institutionalised collaborative tagging as an instrument for managing the maturing learning and knowledge resources. *International Journal for Technology Enhanced Learning (IJTEL)*, 1(1):70–84, 2008. URL http://www.inderscience.com/filter.php?aid=20231. (Cited on pages 25, 35, and 36)

I. Mainz. *Development and Implementation of Techniques for Ontology Engineering and an Ontology-based Search for Bioinformatics Tools and Methods*. PhD thesis, Universität Düsseldorf, Universitäts- und Landesbibliothek, Mathematisch- Naturwissenschaftliche Fakultät, WE Biologie, Physikalische Biologie, Düsseldorf, Germany, Jan. 2009. URL http://d-nb.info/99269776X/34. (Cited on page 75)

W. C. Mann and S. A. Thompson. Rhetorical Structure Theory: A Theory of Text Organization, ISI/RS-87-190. Technical report, ISI: Information Sciences Institute, Los Angeles, CA, 1987. (Cited on page 70)

L. B. Marinho, K. Buza, and L. Schmidt-Thieme. Folksonomy-Based Collabulary Learning. In A. P. Sheth, S. Staab, M. Dean, M. Paolucci, D. Maynard, T. W. Finin, and K. Thirunarayan, editors, *Proceedings of the International Semantic Web Conference*, volume 5318 of *Lecture Notes in Computer Science*, pages 261–276. Springer, 2008. ISBN 978-3-540-88563-4. URL http://dx.doi.org/10.1007/978-3-540-88564-1_17. (Cited on page 81)

B. Markines, C. Cattuto, F. Menczer, D. Benz, A. Hotho, and G. Stumme. Evaluating Similarity Measures for Emergent Semantics of Social Tagging. In *Proc. of the 18th Int. Conf. on WWW*, pages 641–650. ACM, 2009. ISBN 978-1-60558-487-4. doi: http://doi.acm.org/10.1145/1526709.1526796. (Cited on page 79)

J. A. Markowitz, J. T. Nutter, and M. W. Evens. Beyond is-a and part-whole: More semantic network links. *Computers & Mathematics with Applications*, 23(6–9):377–390, 1992. ISSN 0898-1221. doi: DOI:10.1016/0898-1221(92)90113-V. URL http://www.sciencedirect.com/science/article/B6TYJ-46NX0WK-175/2/317032ded09c41ee64ad46b2b64752a2. (Cited on page 114)

C. Marlow, M. Naaman, D. Boyd, and M. Davis. Position Paper, Tagging, Taxonomy, Flickr, Article, ToRead. In *Collaborative Web Tagging Workshop at WWW2006*, May 2006. URL http://www.rawsugar.com/www2006/29.pdf. (Cited on pages xiii, 21, 22, 26, 27, 28, 29, 35, 38, and 111)

A. Mathes. Folksonomies – Cooperative Classification and Communication Through Shared Metadata. Dec. 2004. URL http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html. (Cited on pages 20 and 38)

A. Mazarakis and C. van Dinther. Motivationssteigerung in Wikis durch systemneutrales Feedback. In *Mensch & Computer 2011*, Chemnitz, Germany, 2011a. (Cited on page 266)

A. Mazarakis and C. van Dinther. Motivation durch Feedbackmechanismen in Vorlesungswikis - Welche versprechen mehr Wirkung? In *DeLFI 2011*, Dresden, Germany, 2011b. (Cited on page 266)

A. Mazarakis, S. Braun, and V. Zacharias. Feedback in Social Semantic Applications. *International Journal of Knowledge Engineering and Data Mining (IJKEDM)*, 2011. (Cited on pages 108, 193, 242, and 266)

A. P. McAfee. Enterprise 2.0: The Dawn of Emergent Collaboration. *MITSloan Management Review*, 47(3):21–28, 2006a. URL http://sloanreview.mit.edu/the-magazine/articles/2006/spring/47306/enterprise-the-dawn-of-emergent-collaboration/. (Cited on pages 19 and 104)

A. P. McAfee. Enterprise 2.0, version 2.0. Blog "The Business Impact of IT", May 2006b. URL http://andrewmcafee.org/2006/05/enterprise_20_version_20/. (Cited on page 19)

A. P. McAfee. *Enterprise 2.0: new collaborative tools for your organization's toughest challenges*. Harvard Business Press, Boston, MA, 2009. ISBN 9781422125878. (Cited on page 19)

D. W. Mcdonald and M. S. Ackerman. Expertise recommender: a flexible recommendation system and architecture. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pages 231–240. ACM Press, 2000. ISBN 1581132220. doi: 10.1145/358916.358994. URL http://dx.doi.org/10.1145/358916.358994. (Cited on pages 99 and 256)

D. L. McGuinness. Ontologies Come of Age. In D. Fensel, J. Hendler, H. Lieberman, and W. Wahlster, editors, *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. MIT Press, 2003. URL http://www.ksl.stanford.edu/people/dlm/papers/ontologies-come-of-age-mit-press-(with-citation).htm. (Cited on pages xv, 44, 45, 46, and 47)

P. Merholz. Metadata for the Masses, Oct. 2004. URL http://adaptivepath.com/ideas/essays/archives/000361.php. (Cited on pages 20 and 38)

Merriam-Webster. Ontology - Definition and More from the Free Merriam-Webster Dictionary, 2011. URL http://www.merriam-webster.com/dictionary/ontology. http://www.m-w.com/. (Cited on page 41)

P. Mika. Ontologies are us: A unified model of social networks and semantics. In Y. Gil, E. Motta, V. R. Benjamins, and M. A. Musen, editors, *Proceedings of the International Semantic Web Conference (ISWC'05)*, volume 3729 of *LNCS*, pages 522–536, Berlin Heidelberg, Germany, 2005. Springer-Verlag. ISBN 3-540-29754-5. (Cited on pages 21, 44, 79, 81, and 90)

P. Mika. *Social Networks and the Semantic Web*, volume 5 of *Semantic Web and Beyond*. Springer, New York, 2007. ISBN 978-0-387-71000-6. doi: 10.1007/978-0-387-71001-3. (Cited on page 242)

A. Miles and S. Bechhofer. Skos simple knowledge organization system reference, Aug. 2009. URL http://www.w3.org/TR/2009/REC-skos-reference-20090818/. (Cited on page 51)

A. Miles and D. Brickley. SKOS Extensions Vocabulary Specification, Oct. 2004. URL http://www.w3.org/2004/02/skos/extensions/spec/2004-10-18.html. (Cited on pages 52 and 137)

A. Miles, B. Matthews, D. Beckett, D. Brickley, M. Wilson, and N. Rogers. SKOS: A language to describe simple knowledge structures for the web. In *Proceedings of XTech 2005*, 2005. (Cited on page 51)

D. R. Millen, J. Feinberg, and B. Kerr. Dogear: Social bookmarking in the enterprise. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 111–120, New York, NY, USA, 2006. ACM. ISBN 1-59593-372-7. doi: http://doi.acm.org/10.1145/1124772.1124792. (Cited on page 24)

K. Miller and B. Matthews. Having the Right Connections: the LIMBER Project. *Journal of Digital Information*, 1(8), 2001. ISSN 1368-7506. URL http://journals.tdl.org/jodi/article/viewArticle/34/35. (Cited on page 51)

V. Mirzaee and L. Iverson. Tagging: Behaviour and Motivations. In A. Grove, editor, *Proceedings of the Annual Meeting of the American Society for Information Science Technology (ASIS&T '09)*, volume 46, Silverspring, Maryland, USA, 2009. American Society for Information Science and Technology. (Cited on pages xiii and 29)

M. Missikoff, R. Navigli, and P. Velardi. The Usable Ontology: An Environment for Building and Assessing a Domain Ontology. In I. Horrocks and J. A. Hendler, editors, *International Semantic Web Conference*, volume 2342 of *Lecture Notes in Computer Science*, pages 39–53. Springer, 2002. ISBN 3-540-43760-6. URL http://link.springer.de/link/service/series/0558/bibs/2342/23420039.htm. (Cited on pages 55 and 105)

R. Mizoguchi and K. Kozaki. Ontology Engineering Environments. In S. Staab and R. Studer, editors, *Handbook on Ontologies*, International Handbooks Information System, pages 315–336. Springer Berlin Heidelberg, 2009. ISBN 978-3-540-92673-3. URL http://dx.doi.org/10.1007/978-3-540-92673-3_14. (Cited on page 63)

R. Mizoguchi, M. Ikeda, K. Seta, and J. Vanwelkenhuysen. Ontology for Modeling the World from Problem Solving Perspectives. In *Proc. of IJCAI-95 Workshop on Basic Ontological Issues in Knowledge Sharing*, pages 1–12, 1995. URL http://www.ei.sanken.osaka-u.ac.jp/pub/miz/IJCAI95WS.pdf. (Cited on page 43)

K. Möller, J. G. Breslin, and S. Decker. semiBlog - Semantic Publishing of Desktop Data. In *14th Conference on Information Systems Development (ISD2005)*, pages 855–866, Karlstad, Sweden, Aug. 2005. (Cited on page 243)

P. Monachesi and T. Markus. Using Social Media for Ontology Enrichment. In L. Aroyo, G. Antoniou, E. Hyvönen, A. ten Teije, H. Stuckenschmidt, L. Cabral, and T. Tudorache, editors, *Proceedings of the Extended Semantic Web Conference*, volume 6089 of *Lecture Notes in Computer Science*, pages 166–180. Springer, 2010. ISBN 978-3-642-13488-3. URL http://dx.doi.org/10.1007/978-3-642-13489-0_12. (Cited on page 82)

J. A. Moritz Weiten, Mika Meier-Collin. Metadata Annotation Tools. SEKT Deliverable D2.4.1, Jan. 2006. (Cited on page 85)

M. J. Muller, J. H. Haslwanter, and T. Dayton. Participatory Practices in the Software Lifecycle. In M. G. Helander, T. K. Landauer, and P. V. Prabhu, editors, *Handbook of human-computer interaction*, pages 256–300. North-Holland, 1997. ISBN 0 444 81876 6. (Cited on page 6)

G. Nagypál. Ontology Development Methodologies for Ontology Engineering. In R. Studer, A. Abecker, and S. Grimm, editors, *Semantic web services: concepts, technologies, and applications*, pages 108–134. Springer, 2007a. (Cited on pages 56 and 57)

G. Nagypál. *Possibly imperfect ontologies for effective information retrieval*. PhD thesis, Universität Karlsruhe, Fakultät für Informatik (Fak. f. Informatik) Institut für Programmstrukturen und Datenorganisation (IPD), Karlsruhe, 2007b. URL http://digbib.ubka.uni-karlsruhe.de/volltexte/1000007206. (Cited on pages 54 and 149)

R. Neches, R. Fikes, T. Finin, T. Gruber, R. Patil, T. Senator, and W. Swartout. Enabling Technology for Knowledge Sharing. *AI Magazine*, 12(3):36–56, Aug. 1991. (Cited on page 42)

T. Nelkner. An Infrastructure for Intercommunication between Widgets in Personal Learning Environments. In M. D. Lytras, P. O. de Pablos, E. Damiani, D. Avison, A. Naeve, and D. G. Horner, editors, *Best Practices for the Knowledge Society. Knowledge, Learning, Development and Technology for All*, volume 49 of *Communications in Computer and Information Science*, pages 41–48. Springer Berlin Heidelberg, 2009. ISBN 978-3-642-04757-2. URL http://dx.doi.org/10.1007/978-3-642-04757-2_5. (Cited on page 146)

T. Nelkner, B. HU, A. Martin, S. Brander, S. Braun, U. Riss, G. Attwell, K. Hinkelmann, and M. B. de Diego. Design and Delivery of Prototype Version V2 of PLME / OLME. deliverable, Apr. 2011. (Cited on page 9)

A. D. Nicola, M. Missikoff, and R. Navigli. A Proposal for a Unified Process for Ontology Building: UPON. In K. V. Andersen, J. K. Debenham, and R. Wagner, editors, *DEXA*, volume 3588 of *Lecture Notes in Computer Science*, pages 655–664. Springer, 2005. ISBN 3-540-28566-0. (Cited on page 62)

A. D. Nicola, M. Missikoff, and R. Navigli. A software engineering approach to ontology building. *Inf. Syst.*, 34(2):258–275, 2009. ISSN 0306-4379. doi: http://dx.doi.org/10.1016/j.is.2008.07.002. (Cited on page 62)

H. M. Niegemann, S. Hessel, M. Hupfer, S. Domagk, A. Hein, and A. Zobel. *Kompendium multimediales Lernen*. X.media.press. Springer-Verlag, Berlin, Heidelberg, 2008. ISBN 978-9-540-37225-7. doi: 10.1007/978-3-540-37226-4. (Cited on page 107)

J. Nielsen. *Usability Engineering*. Morgan Kaufmann, 1994. (Cited on page 171)

A. Norta, R. Yangarber, and L. Carlson. Utility Evaluation of Tools for Collaborative Development and Maintenance of Ontologies. *Enterprise Distributed Object Computing Conference Workshops, IEEE International*, 0:207–214, 2010. doi: http://doi.ieeecomputersociety.org/10.1109/EDOCW.2010.30. (Cited on page 75)

O. Nov and C. Ye. Why do people tag?: motivations for photo tagging. *Commun. ACM*, 53(7):128–131, 2010. ISSN 0001-0782. doi: 10.1145/1785414.1785450. URL http://dx.doi.org/10.1145/1785414.1785450. (Cited on page 30)

O. Nov, M. Naaman, and C. Ye. Motivational, Structural and Tenure Factors that Impact Online Community Photo Sharing. In *International AAAI Conference on Weblogs and Social Media*. Association for the Advancement of Artificial Intelligence, 2009. URL http://aaai.org/ocs/index.php/ICWSM/09/paper/view/206/426. (Cited on pages xiii, 28, and 29)

N. Noy, M. Sintek, S. Decker, M. Crubezy, R. Fergerson, and M. Musen. Creating Semantic Web contents with Protege-2000. *Intelligent Systems, IEEE*, 16(2):60–71, 2001. ISSN 1541-1672. doi: 10.1109/5254.920601. (Cited on page 65)

N. F. Noy and D. L. McGuinness. Ontology Development 101: A Guide to Creating Your First Ontology. Technical Report SMI-2001-0880, Stanford Knowledge Systems Laboratory and Stanford Medical Informatics, 2001. (Cited on pages 65 and 68)

N. F. Noy, A. Chugh, W. Liu, and M. A. Musen. A Framework for Ontology Evolution in Collaborative Environments. In I. F. Cruz, S. Decker, D. Allemang, C. Preist, D. Schwabe, P. Mika, M. Uschold, and L. Aroyo, editors, *International Semantic Web Conference*, volume 4273 of *Lecture Notes in Computer Science*, pages 544–558. Springer, 2006. ISBN 3-540-49029-9. (Cited on page 76)

N. F. Noy, A. Chugh, and H. Alani. The CKC Challenge: Exploring Tools for Collaborative Knowledge Construction. *IEEE Intelligent Systems*, 23(1):64–68, 2008. (Cited on page 172)

L. Obrst. Ontologies for semantically interoperable systems. In *CIKM '03: Proceedings of the twelfth international conference on Information and knowledge management*, pages 366–369, New York, NY, USA, 2003. ACM. ISBN 1-58113-723-0. doi: http://doi.acm.org/10.1145/956863.956932. URL http://dx.doi.org/10.1145/956863.956932. (Cited on pages xv, 45, 46, and 48)

Oertel and Schulz (eds.). User Tests: Usability Expertise. Im WISSENSNETZ Internal Report, Sept. 2008. (Cited on page 175)

I. Ohmukai, H. Takeda, and K. Numa. Personal Knowledge Publishing Suite with Weblog. In *13th International World Wide Web Conference (WWW'04), Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics*, May 2004. (Cited on page 243)

T. O'Keeffe. Organisational learning: a new perspective. *Journal of European Industrial Training*, 26:130–141, 2002. URL http://www.emeraldinsight.com/journals.htm?issn=0309-0590&volume=26&issue=2/3/4&articleid=837083&show=pdf. (Cited on page 3)

T. O'Reilly. What Is Web 2.0. Design Patterns and Business Models for the Next Generation of Software. http://www.oreillynet.com/lpt/a/6228, Sept. 2005. URL http://www.oreillynet.com/lpt/a/6228. (Cited on page 17)

E. Oren, K. Möller, S. Scerri, S. Handschuh, and M. Sintek. What are Semantic Annotations? Technical report, DERI Galway, 2006. URL http://www.siegfried-handschuh.net/pub/2006/whatissemannot2006.pdf. (Cited on page 82)

R. Palma, P. Haase, Óscar Corcho, A. Gómez-Pérez, and Q. Ji. An Editorial Workflow Approach For Collaborative Ontology Development. In J. Domingue and C. Anutariya, editors, *ASWC*, volume 5367 of *Lecture Notes in Computer Science*, pages 227–241. Springer, 2008. ISBN 978-3-540-89703-3. (Cited on page 71)

S. Panke and B. Gaiser. “With My Head Up in the Clouds”. *Journal of Business and Technical Communication*, 23(3):318–349, 2009. doi: 10.1177/1050651909333275. URL http://jbt.sagepub.com/content/23/3/318.abstract. (Cited on pages xiii, 26, and 27)

A. Passant, T. Hastrup, U. Bojars, and J. Breslin. Microblogging: A Semantic and Distributed Approach. In C. Bizer, S. Auer, G. A. Grimnes, and T. Heath, editors, *Proceedings of the 4th Workshop on Scripting for the Semantic Web*, volume 368 of *CEUR Workshop Proceedings*, June 2008. URL http://CEUR-WS.org/Vol-368/paper11.pdf. (Cited on page 243)

P. F. Patel-Schneider, P. Hayes, and I. Horrocks. Owl web ontology language semantics and abstract syntax, Feb. 2004. URL http://www.w3.org/TR/2004/REC-owl-semantics-20040210/. (Cited on page 51)

I. Paulsen, D. Mainz, K. Weller, I. Mainz, J. Kohl, and A. von Haeseler. Ontoverse: Collaborative Knowledge Management in the Life Sciences Network. In *Proceedings of the German e-Science Conference 2007 (GES 2007)*, May 2007. URL http://edoc.mpg.de/316588. (Cited on page 75)

C. Pereira and A. L. Soares. Ontology Development in Collaborative Networks as a Process of Social Construction of Meaning. In R. Meersman, Z. Tari, and P. Herrero, editors, *OTM Workshops*, volume 5333 of *Lecture Notes in Computer Science*, pages 605–614. Springer, 2008. ISBN 978-3-540-88874-1. (Cited on page 258)

C. Pereira, C. Sousa, and A. L. Soares. A Socio-semantic Approach to Collaborative Domain Conceptualization. In R. Meersman, P. Herrero, and T. S. Dillon, editors, *OTM Workshops*, volume 5872 of *Lecture Notes in Computer Science*, pages 524–533. Springer, 2009. ISBN 978-3-642-05289-7. (Cited on page 259)

I. Peters. *Folksonomies : indexing and retrieval in Web 2.0*. Walter de Gruyter, Berlin, 2009. ISBN 9783598251795 3598251793. (Cited on pages xv, 3, 17, 18, 21, 25, 35, 38, and 261)

I. Peters and K. Weller. Tag gardening for folksonomy enrichment and maintenance. *Webology*, 5(3), Sept. 2008. URL http://www.webology.ir/2008/v5n3/a58.html. (Cited on page 260)

H. S. Pinto and J. P. Martins. Ontologies: How can They be Built? *Knowledge and Information Systems*, 6:442–464, July 2004. URL http://www.springerlink.com/content/0p5yqrdh5t5dvd06/fulltext.pdf. (Cited on page 56)

H. S. Pinto, C. Tempich, and S. Staab. DILIGENT: Towards a fine-grained methodology for DIstributed, Loosely-controlled and evolvInG Engingeering of oNTologies. In R. L. de Mantaras and L. Saitta, editors, *Proceedings of the 16th European Conference on Artificial Intelligence (ECAI 2004)*, pages 393–397. IOS Press, Aug. 2004. (Cited on page 69)

B. Popov, A. Kiryakov, A. Kirilov, D. Manov, D. Ognyanoff, and M. Goranov. KIM – Semantic Annotation Platform. In D. Fensel, K. P. Sycara, and J. Mylopoulos, editors, *International Semantic Web Conference*, volume 2870 of *Lecture Notes in Computer Science*, pages 834–849. Springer, 2003. ISBN 3-540-20362-1. (Cited on page 86)

E. Quintarelli. Folksonomies: power to the people. ISKO-Italy Uni-MIB meeting, June 2005. URL http://www-dimat.unipv.it/biblio/isko/doc/folksonomies.htm. (Cited on page 38)

D. R. Raban, I. Ronen, and I. Guy. Acting or reacting? Preferential attachment in a people-tagging system. *Journal of the American Society for Information Science and Technology*, 62(4):738–747, 2011. ISSN 1532-2890. doi: 10.1002/asi.21490. URL http://dx.doi.org/10.1002/asi.21490. (Cited on pages 121 and 257)

M. Ramezani. *Using Data Mining for Facilitating User Contributions in the Social Semantic Web* . Phd thesis, Karlsruher Institut für Technologie (KIT), Fakultät für Wirtschaftswissenschaften (Fak. f. Wirtschaftswiss.) Institut für Angewandte Informatik und Formale Beschreibungsverfahren (AIFB), Karlsruhe, Germany, Feb. 2011. URL http://digbib.ubka.uni-karlsruhe.de/volltexte/1000022794. (Cited on pages 79, 143, and 148)

M. Ramezani, H. F. Witschel, S. Braun, and V. Zacharias. Using machine learning to support continuous ontology development. In P. Cimiano and H. S. Pinto, editors, *EKAW*, volume 6317 of *Lecture Notes in Computer Science*, pages 381–390. Springer, 2010. ISBN 978-3-642-16437-8. (Cited on page 142)

A. Ravenscroft. Promoting thinking and conceptual change with digital dialogue games. *Journal of Computer Assisted Learning*, 23(6):453–465, 2007. (Cited on page 144)

A. Ravenscroft and S. McAllister. Investigating and promoting educational argumentation: towards new digital practices. *International Journal of Research and Method in Education (IJRME)*, 31(3):317–335, 2008. (Cited on page 144)

A. Ravenscroft, C. Bradley, J. Cock, A. Schmidt, S. Braun, S. Brander, R. Brun, J. Pueyo, T. Ley, P. Seitlinger, B. Thönssen, and F. Witschel. Formative Evaluation Report of 1st MATURE System Prototype and Requirements Method. deliverable, Apr. 2010a. (Cited on pages 9 and 203)

A. Ravenscroft, S. Braun, and T. Nelkner. Combining Dialogue and Semantics for Learning and Knowledge Maturing: Developing Collaborative Understanding in the 'Web 2.0 Workplace'. In *International Conference on Advanced Learning Technologies (ICALT) 2010*, pages 176–180. IEEE Computer Society, 2010b. (Cited on pages 144, 198, and 209)

M. N. Razavi and L. Iverson. Improving personal privacy in social systems with people-tagging. In *Proceedings of the ACM 2009 international conference on Supporting group work (Group'09)*, pages 11–20, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-500-0. doi: http://doi.acm.org/10.1145/1531674.1531677. URL http://doi.acm.org/10.1145/1531674.1531677. (Cited on page 257)

L. Reeve and H. Han. Survey of semantic annotation platforms. In *SAC '05: Proceedings of the 2005 ACM symposium on Applied computing*, pages 1634–1638, New York, NY, USA, 2005. ACM. ISBN 1-58113-964-0. doi: http://doi.acm.org/10.1145/1066677.1067049. (Cited on page 83)

T. Reichling, M. Veith, and V. Wulf. Expert Recommender: Designing for a Network Organization. *Computer Supported Cooperative Work*, 16(4-5):431–465, 2007. URL http://www.springerlink.com/index/10.1007/s10606-007-9055-2. (Cited on page 256)

V. Robu, H. Halpin, and H. Shepherd. Emergence of consensus and shared vocabularies in collaborative tagging systems. *ACM Transactions on the Web*, 3(4):1–34, 2009. ISSN 1559-1131. doi: http://doi.acm.org/10.1145/1594173.1594176. (Cited on pages 25 and 36)

C. Sá, C. Pereira, and A. L. Soares. Supporting Collaborative Conceptualization Tasks through a Semantic Wiki Based Platform. In R. Meersman, T. S. Dillon, and P. Herrero, editors, *OTM Workshops*, volume 6428 of *Lecture Notes in Computer Science*, pages 394–403. Springer, 2010. ISBN 978-3-642-16960-1. (Cited on page 259)

C. Saathoff, S. Schenk, and A. Scherp. KAT: The K-Space Annotation Tool. In *Proccedings of the SAMT 2008 Demo and Poster Session*, 2008. URL http://www.uni-koblenz.de/~saathoff/publications/samt08-kat-demo.pdf. (Cited on page 84)

M. Sabou, M. d'Aquin, and E. Motta. Exploring the Semantic Web as Background Knowledge for Ontology Matching. *Journal on Data Semantics*, 11:156–190, 2008. (Cited on page 81)

E. Santos-Neto, M. Ripeanu, and A. Iamnitchi. Tracking Usage in Collaborative Tagging Communities. In *Workshop on Contextualized Attention Metadata (CAMA'07)*, June 2007. URL http://arxiv.org/pdf/0705.1013. masses may not only support folksonomies but may also hinder their usage for individuals, efficiency of tagging systems decreases with growing population,. (Cited on page 25)

L. Sauermann, A. Bernardi, and A. Dengel. Overview and Outlook on the Semantic Desktop. In S. Decker, J. Park, D. Quan, and L. Sauermann, editors, *Proceedings of the 1st Workshop on The Semantic Desktop at the ISWC 2005 Conference*, volume 175 of *CEUR Workshop Proceedings*, pages 1–19. CEUR-WS, Nov. 2005. URL http://www.dfki.uni-kl.de/~sauermann/papers/Sauermann+2005d.pdf. (Cited on page 85)

Òscar Celma and Y. Raimond. Zempod: A semantic web approach to podcasting. *Journal of Web Semantics*, 6(2):162–169, 2008. URL http://dx.doi.org/10.1016/j.websem.2008.01.003. (Cited on page 243)

S. Schaffert. IkeWiki: A Semantic Wiki for Collaborative Knowledge Management. In *1st International Workshop on Semantic Technologies in Collaborative Applications (STICA'06)*, Manchester, UK, June 2006a. URL http://www.wastl.net/download/paper/schaffert06_ikewiki.pdf. (Cited on pages 243 and 246)

S. Schaffert. Semantic Social Software: Semantically enabled Social Soft-ware or Socially enabled Semantic Web. In S. Schaffert and Y. Sure, editors, *Semantic Systems. From Visions to Applications*. OCG Verlag, 2006b. (Cited on page 242)

S. Schaffert, J. Eder, S. Grünwald, T. Kurz, and M. Radulescu. KiWi - A Platform for Semantic Social Software (Demonstration). In *ESWC'09: The Semantic Web: Research and Applications, Proceedings of the 6th European Semantic Web Conference*, pages 888–892, Heraklion, Greece, June 2009. (Cited on page 246)

B. Schandl. *An Infrastructure for the Development of Semantic Desktop Applications*. PhD thesis, University of Vienna, Sept. 2009. URL http://eprints.cs.univie.ac.at/132/. (Cited on page 85)

T. Schandl and A. Blumauer. PoolParty: SKOS Thesaurus Management Utilizing Linked Data. In L. Aroyo, G. Antoniou, E. Hyvönen, A. ten Teije, H. Stuckenschmidt, L. Cabral, and T. Tudorache, editors, *The Semantic Web: Research and Applications*, volume 6089 of *Lecture Notes in Computer Science*, pages 421–425. Springer, 2010. ISBN 978-3-642-13488-3. URL http://dx.doi.org/10.1007/978-3-642-13489-0_36. (Cited on page 263)

S. Schenk, C. Saathoff, S. Staab, and A. Scherp. SemaPlorer – Interactive Semantic Exploration of Data and Media based on a Federated Cloud Infrastructure. *Journal of Web Semantics*, 7(4):298–304, 2009. (Cited on page 243)

A. Schmidt. Knowledge Maturing and the Continuity of Context as a Unifying Concept for Knowledge Management and E-Learning. In *Proceedings of I-KNOW '05, Special Track on Integrating Working and Learning*, 2005. (Cited on page 95)

A. Schmidt. Enabling Learning on Demand in Semantic Work Environments: The Learningin Process Approach. In J. Rech, B. Decker, and E. Ras, editors, *Emerging Technologies for Semantic Work Environments: Techniques, Methods, and Applications*. IGI Publishing, 2008. URL http://publications.andreas.schmidt.name/SWE_Schmidt_EnablingLearningOnDemand.pdf. (Cited on page 102)

A. Schmidt. *Situationsbewusste Informationsdienste für das arbeitsbegleitende Lernen*. PhD thesis, Karlsruher Institut für Technologie (KIT), Fakultät für Informatik, Institut für Programmstrukturen und Datenorganisation (IPD), Karlsruhe, Germany, Feb. 2009. URL http://digbib.ubka.uni-karlsruhe.de/volltexte/1000012939. (Cited on page 140)

A. Schmidt and C. Kunzmann. Towards a Human Resource Development Ontology for Combining Competence Management and Technology-Enhanced Workplace Learning. In R. Meersman, Z. Tahiri, and P. Herero, editors, *On The Move to Meaningful Internet Systems 2006: OTM 2006 Workshops. Part I. 1st Workshop on Ontology Content and Evaluation in Enterprise(OntoContent 2006)*, volume 4278 of *Lecture Notes in Computer Science*, pages 1078–1087. Springer, 2006. URL http://www.andreas-p-schmidt.de/publications/schmidt_kunzmann_OntoContent06.pdf. (Cited on pages xv, 101, and 102)

A. Schmidt and C. Kunzmann. Sustainable Competency-Oriented Human Resource Development with Ontology-Based Competency Catalogs. In M. Cunningham and P. Cunningham, editors, *Expanding the Knowledge Economy: Issues, Applications, Case Studies. Proceedings of E-Challenges 2007*, Amsterdam, 2007. IOS Press. URL http://publications.professional-learning.eu/schmidt_kunzmann_sustainable-competence-management_eChallenges07.pdf. (Cited on pages 2, 99, 101, 220, and 255)

A. Schmidt, K. Hinkelmann, T. Ley, S. Lindstaedt, R. Maier, and U. Riss. Conceptual Foundations for a Service-oriented Knowledge and Learning Architecture: Supporting Content, Process and Ontology Maturing. In S. Schaffert, K. Tochtermann, and T. Pellegrini, editors, *Networked Knowledge - Networked Media: Integrating Knowledge Management, New Media Technologies and Semantic Systems*, pages 79–94. Springer, 2009. URL http://publications.andreas.schmidt.name/schmidt_et_al_learning_knowledge_architecture_conceptual_foundations_2009.pdf. (Cited on page 103)

J. Schmidt. Social Software: Onlinegestütztes Informations-, Identitäts- und Beziehungsmanagement. *Forschungsjournal Neue Soziale Bewegungen*, 19(2):37 – 47, June 2006. ISSN 09339361. URL http://www.bamberg-gewinnt.de/wordpress/wp-content/pdf/SocialSoftwareFJNSB_preprint.pdf. (Cited on page 17)

C. Schmitz. *Self-Organized Collaborative Knowledge Management*. PhD thesis, Universität Kassel, Sept. 2007. URL http://www.uni-kassel.de/hrz/db4/extern/dbupress/publik/abstract.php?978-3-89958-325-0. (Cited on pages 21 and 38)

C. Schmitz, A. Hotho, R. Jäschke, and G. Stumme. Mining Association Rules in Folksonomies. In V. Batagelj, H.-H. Bock, A. Ferligoj, and A. Žiberna, editors, *Data Science and Classification: Proc. of the 10th IFCS Conf.*, Studies in Classification, Data Analysis, and Knowledge Organization, pages 261–270, Berlin, Heidelberg, 2006. Springer. (Cited on page 81)

P. Schmitz. Inducing Ontology from Flickr Tags. In *Proceedings of the Workshop on Collaborative Web Tagging at WWW2006*, Edinburgh, Scotland, May 2006. URL http://.citeulike.org/user/ryanshaw/article/740688. (Cited on pages 30 and 81)

K. Schoefegger, P. Seitlinger, and T. Ley. Towards a user model for personalized recommendations in work-integrated learning: A report on an experimental study with a collaborative tagging system. In *Proceedings of the 1st Workshop on Recommender Systems for Technology Enhanced Learning (RecSysTEL 2010)*, volume 1, pages 2829–2838. Elsevier, 2010. URL http://dx.doi.org/10.1016/j.procs.2010.08.008. (Cited on page 194)

G. Schreiber, H. Akkermans, A. Anjewierden, R. de Hoog, N. Shadbolt, W. V. de Velde, and B. J. Wielinga. *Knowledge Engineering and Management: The CommonKADS Methodology*. MIT Press, Cambridge, Mass., 2nd ed. edition, 1999. URL http://books.google.com/books?id=H1XOW_1fsIEC&hl=de. (Cited on page 61)

R. Schroeter, J. Hunter, and A. Newman. Annotating Relationships Between Multiple Mixed-Media Digital Objects by Extending Annotea. In *The Semantic Web: Research and Applications, Proceeedings of the 4th European Semantic Web Conference (ESWC2007), June 3-7 , 2007, Innsbruck, Austria*, pages 533–548, 2007. doi: 10.1007/978-3-540-72667-8_38. (Cited on page 84)

Schulz (ed.). User Tests: Usability Expertise. Im WISSENSNETZ Internal Report, Oct. 2008. (Cited on page 182)

P. C. Seitlinger. Kognitionspsychologische Aspekte von Collaborative Tagging: Assoziative Aktivierung und Basiskategorie-Effekte bei der gemeinschaftlichen Verschlagwortung von Internetressourcen. Master's thesis, Karl-Franzens-Universität Graz, Graz, Austria, 2009. (Cited on page 194)

S. Sen, S. K. Lam, A. M. Rashid, D. Cosley, D. Frankowski, J. Osterhouse, F. M. Harper, and J. Riedl. tagging, communities, vocabulary, evolution. In *CSCW '06: Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pages 181–190, New York, NY, USA, 2006. ACM Press. (Cited on pages 21, 23, 25, 30, 31, 32, 35, 36, 108, and 115)

S. W. Sen. *Nurturing Tagging Communities*. Phd thesis, UNIVERSITY OF MINNESOTA, Minnesota, USA, Mar. 2009. URL http://conservancy.umn.edu/bitstream/49984/1/Sen_umn_0130E_10255.pdf. (Cited on pages xv, 35, and 37)

L. Seremeti and A. Kameas. Tools for Ontology Engineering and Management. In R. Poli, M. J. Healy, and A. D. Kameas, editors, *Theory and Applications of Ontology: Computer Applications Theory and Applications of Ontology: Computer Applications*, pages 131–154. Springer, 2010. ISBN 978-90-481-8846-8. URL https://springerlink3.metapress.com/content/n72x345854g268h4/resource-secured/?target=fulltext.pdf&sid=fjfjkybj0taf5045n5ksflyk&sh=www.springerlink.com. (Cited on page 63)

C. Shirky. Ontology is Overrated: Categories, Links, and Tags, 2005. URL http://www.shirky.com/writings/ontology_overrated.html. (Cited on page 38)

M. Sicilia. *Ontology-based competency management: infrastructures for the knowledge-intensive learning organization*, pages 302–324. Idea Group, Hershey, 2005. (Cited on page 255)

E. Simperl and M. Luczak-Rösch. Collaborative Ontology Engineering: A Survey. *(to appear)*, 2011. (Cited on page 68)

E. Simperl, C. Tempich, and T. Bürger. Methodologies for the creation of semantic data. In *Handbook of Metadata, Semantics and Ontologies*. World Scientific Publishing Co, 2011. (Cited on page 83)

E. P. B. Simperl and C. Tempich. Ontology Engineering: A Reality Check. In R. Meersman and Z. Tari, editors, *OTM Conferences (1)*, volume 4275 of *Lecture Notes in Computer Science*, pages 836–854. Springer, 2006. ISBN 3-540-48287-3. (Cited on pages xv, 56, and 57)

R. Sinha. A cognitive analysis of tagging, 2005. URL http://www.rashmisinha.com/archives/05_09/tagging-cognitive.html. (Cited on pages 38 and 117)

K. Siorpaes and M. Hepp. myOntology: The Marriage of Ontology Engineering and Collective Intelligence. In *Bridging the Gep between Semantic Web and Web 2.0 (SemNet 2007)*, pages 127–138, 2007. URL http://www.kde.cs.uni-kassel.de/ws/eswc2007/proc/TheMarriage.pdf. (Cited on page 243)

K. Siorpaes and M. Hepp. Games with a Purpose for the Semantic Web. *IEEE Intelligent Systems*, 23(11):50–60, May 2008. ISSN 1541-1672. doi: 10.1109/MIS.2008.45. URL http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=4525142. (Cited on page 244)

K. Siorpaes and E. Simperl. Human Intelligence in the Process of Semantic Content Creation. *World Wide Web*, 13:33–59, 2010. ISSN 1386-145X. URL http://dx.doi.org/10.1007/s11280-009-0078-0. 10.1007/s11280-009-0078-0. (Cited on pages 82, 83, 87, and 244)

K. Siorpaes, M. Hepp, A. Klotz, and M. Hackl. myOntology: Tapping the Wisdom of Crowds for Building Ontologies. Technical report, STI Innsbruck, Innsbruck, Austria, Oct. 2008. URL http://www.sti-innsbruck.at/fileadmin/documents/technical_report/myontology-techreport.pdf. (Cited on pages 259 and 260)

B. Smith and H. Burkhardt. *Handbook of Metaphysics and Ontology*. Philosophia, 1991. (Cited on page 41)

B. Smith and C. Welty. FOIS introduction: Ontology—towards a new synthesis. In *FOIS '01: Proceedings of the international conference on Formal Ontology in Information Systems*, New York, NY, USA, 2001. ACM Press. ISBN 1581133774. doi: http://dx.doi.org/10.1145/505168.505201. URL http://dx.doi.org/10.1145/505168.505201. (Cited on pages xv, 45, 46, and 47)

G. Smith. Folksonomy: Social Classification. http://atomiq.org/archives/2004/08/-folksonomy_social_classification.html, Aug. 2004. (Cited on page 20)

M. K. Smith. VIEWER TAGGING IN ART MUSEUMS: COMPARISONS TO CONCEPTS AND VOCABULARIES OF ART MUSEUM VISITORS. In J. Furner and J. T. Tennis, editors, *Proceedings of the 17th ASIS&T SIG/CR Classification Research Workshop*, volume 17 of *Advances in classification research*, 2006. URL http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.85.3019. (Cited on page 38)

D. Sánchez, J. Cavero, and E. Martínez. The Road Toward Ontologies. In R. Sharman, R. Kishore, and R. Ramesh, editors, *Ontologies*, volume 14 of *Integrated Series in Information Systems*, pages 3–20. Springer US, 2007. ISBN 978-0-387-37022-4. URL http://dx.doi.org/10.1007/978-0-387-37022-4_1. 10.1007/978-0-387-37022-4_1. (Cited on page 41)

J. F. Sowa. Ontologies for Knowledge Sharing. In *n Manuscript of the invited talk at Terminology and Knowledge Engineering Congress (TKE '96)*, 1996. URL http://www.ei.sanken.osaka-u.ac.jp/pub/miz/IJCAI95WS.pdf. (Cited on page 44)

L. Specia and E. Motta. Integrating Folksonomies with the Semantic Web. In *Proc. of the European Semantic Web Conference (ESWC2007)*, volume 4519 of *LNCS*, pages 624–639, Berlin Heidelberg, Germany, July 2007. Springer-Verlag. URL http://people.kmi.open.ac.uk/motta/papers/SpeciaMotta_ESWC-2007_Final.pdf. (Cited on page 79)

L. F. Spiteri. Structure and form of folksonomy tags: The road to the public library catalogue. *Webology*, 4(2), June 2007. URL http://www.webology.ir/2007/v4n2/a41.html. (Cited on pages 25 and 38)

P. Spyns, R. Meersman, and M. Jarrar. Data Modelling versus Ontology Engineering. *SIGMOD Record*, 31(4):12–17, 2002. (Cited on page 72)

P. Spyns, Y. Tang, and R. Meersman. An ontology engineering methodology for DOGMA. *Applied Ontology*, 3(1-2):13–39, 2008. (Cited on page 72)

S. Staab and R. Studer, editors. *Handbook on Ontologies*, volume 2nd ed. of *International Handbooks on Information Systems*. Springer, Berlin–Heidelberg, Germany, 2009. (Cited on pages 50 and 54)

S. Staab, R. Studer, H.-P. Schnurr, and Y. Sure. Knowledge Processes and Ontologies. *IEEE Intelligent Systems*, 16:26–34, Jan. 2001. URL http://www.google.de/url?sa=t&source=web&cd=1&ved=0CBsQFjAA&url=http3Drep1B0kZqQ&sig2=ej0rCzNpi9xpYsXKrn4How. (Cited on page 61)

L. Stojanovic. *Methods and Tools for Ontology Evolution*. PhD thesis, University of Karlsruhe (TH), Germany, 2004. (Cited on pages 141 and 143)

R. Studer, R. Benjamins, and D. Fensel. Knowledge Engineering: Principles and Methods. *Data & Knowledge Engineering*, 25(1-2):161–198, Mar. 1998. (Cited on pages 42 and 44)

Y. Sure, M. Erdmann, J. Angele, S. Staab, R. Studer, and D. Wenke. OntoEdit: Collaborative Ontology Engineering for the Semantic Web. In I. Horrocks and J. Hendler, editors, *Proceedings of the First International Semantic Web Conference 2002 (ISWC 2002), June 9-12 2002, Sardinia, Italia*, volume 2342 of *LNCS*, pages 221–235. Springer, 2002. URL http://www.aifb.uni-karlsruhe.de/WBS/ysu/publications/2002_iswc_ontoedit.pdf. (Cited on pages 62, 70, and 75)

Y. Sure, S. Staab, and R. Studer. On-To-Knowledge Methodology (OTKM). In S. Staab and R. Studer, editors, *Handbook on Ontologies: International Handbook on Information Systems*, pages 117–132. Springer, 2004. ISBN 3540408347. (Cited on page 61)

B. Swartout, R. Patil, K. Knight, and T. Russ. Toward Distributed Use of Large-Scale Ontologies. In *AAAI Symposium on Ontological Engineering*, pages 138–148, 1997. (Cited on pages 59, 64, 66, and 75)

W. Swartout and A. Tate. Guest Editors' Introduction: Ontologies. *IEEE Intelligent Systems*, 14:18–19, 1999. ISSN 1094-7167. doi: http://doi.ieeecomputersociety.org/10.1109/MIS.1999.747901. (Cited on page 42)

H. Takeda, P. Veerkamp, T. Tomiyama, and H. Yoshikawam. Modeling design processes. *AI Magazine*, 11(4):37–48, 1990. (Cited on pages xv and 5)

J. W. Tanaka and M. Taylor. Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitve Psychology*, 23(3):457–482, July 1991. ISSN 0010-0285. (Cited on page 39)

J. Tang, H. fung Leung, Q. Luo, D. Chen, and J. Gong. Towards ontology learning from folksonomies. In *IJCAI'09: Proceedings of the 21st international jont conference on Artifical intelligence*, pages 2089–2094, San Francisco, CA, USA, 2009. Morgan Kaufmann Publishers Inc. URL http://ijcai.org/papers09/Papers/IJCAI09-344.pdf. (Cited on page 82)

F. Tekin. Konzeption zur Konsolidierungsunterstützung in der kollaborativen Ontologieentwicklung. master thesis, Karlsruher Institut für Technologie (KIT), Fakultät für Wirtschaftswissenschaften (Fak. f. Wirtschaftswiss.) Institut für Angewandte Informatik und Formale Beschreibungsverfahren (AIFB), Karlsruhe, Germany, Jan. 2009. (Cited on page 105)

C. Tempich. *Ontology Engineering and Routing in Distributed Knowledge Management Applications*. PhD thesis, Universität Karlsruhe (TH), Fakultät für Wirtschaftswissenschaften, Karlsruhe, Germany, Aug. 2006. URL http://digbib.ubka.uni-karlsruhe.de/volltexte/documents/1705. (Cited on pages 69 and 71)

C. Tempich, E. P. B. Simperl, M. Luczak, R. Studer, and H. S. Pinto. Argumentation-Based Ontology Engineering. *IEEE Intelligent Systems*, 22(6):52–59, 2007. URL http://doi.ieeecomputersociety.org/10.1109/MIS.2007.103. (Cited on page 70)

J. Tennison and N. R. Shadbolt. APECKS: a Tool to Support Living Ontologies. In B. Gaines and M. Musen, editors, *11th Knowledge Acquisition for Knowledge-Bases Systems Workshop (KAW98)*, pages 1–20, 1998. (Cited on page 75)

S. Teufel, C. Sauter, T. Mühlherr, and K. Bauknecht. *Computerunterstützung für die Gruppenarbeit*. Addison- Wesley, Bonn, Paris, 1995. (Cited on page 17)

T. D. Thanh. *Process-oriented Semantic Web Search*. PhD thesis, Karlsruher Institut für Technologie (KIT), Fakultät für Wirtschaftswissenschaften (Fak. f. Wirtschaftswiss.) Institut für Angewandte Informatik und Formale Beschreibungsverfahren (AIFB), Amsterdam, The Netherlands, Feb. 2011. (Cited on page 149)

P. I. Thielen. Social Tagging Systems – Shall we Use the Collaborative and Collective Approach to Gather Competency Related Information? In S. Strohmeier and A. Diederichsen, editors, *Proceedings of the 3rd European Academic Workshop on Electronic Human Resource Management*, volume 570, pages 186–205. CEUR-WS, May 2010. URL http://ceur-ws.org/Vol-570/paper012.pdf. (Cited on page 258)

J. Thom-Santelli, M. J. Muller, and D. R. Millen. Social tagging roles: publishers, evangelists, leaders. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1041–1044, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-011-1. doi: http://doi.acm.org/10.1145/1357054.1357215. URL http://dx.doi.org/10.1145/1357054.1357215. (Cited on page 25)

E. Tonkin, E. M. Corrado, H. L. Moulaison, M. E. I. Kipp, A. Resmini, H. Pfeiffer, and Q. Zhang. Collaborative and Social Tagging Networks. In *Ariadne*, number 54, Jan. 2008. ISBN 1361-3200. (Cited on page 35)

J. Trant. Studying Social Tagging and Folksonomy: A Review and Framework. *Journal of Digital Information*, 10(1), 2009. URL http://journals.tdl.org/jodi/article/view/269/278. (Cited on pages 20 and 35)

T. Tudorache, N. F. Noy, and M. A. Musen. Collaborative Protege: Enabling Community-based Authoring of Ontologies. In C. Bizer and A. Joshi, editors, *International Semantic Web Conference*, volume 401 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2008a. (Cited on pages 76 and 172)

T. Tudorache, N. F. Noy, S. Tu, and M. A. Musen. Supporting Collaborative Ontology Development in Protégé. In A. P. Sheth, S. Staab, M. Dean, M. Paolucci, D. Maynard, T. W. Finin, and K. Thirunarayan, editors, *International Semantic Web Conference*, volume 5318 of *Lecture Notes in Computer Science*, pages 17–32. Springer, 2008b. ISBN 978-3-540-88563-4. (Cited on page 76)

T. Tudorache, J. Vendetti, and N. F. Noy. Web-Protege: A Lightweight OWL Ontology Editor for the Web. In C. Dolbear, A. Ruttenberg, and U. Sattler, editors, *OWLED*, volume 432 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2008c. (Cited on page 76)

G. Tummarello, C. Morbidoni, and M. Nucci. Enabling Semantic Web Communities with DBin: An Overview. In I. F. Cruz, S. Decker, D. Allemang, C. Preist, D. Schwabe, P. Mika, M. Uschold, and L. Aroyo, editors, *International Semantic Web Conference*, volume 4273 of *Lecture Notes in Computer Science*, pages 943–950. Springer, 2006. ISBN 3-540-49029-9. (Cited on page 172)

J. D. Ullman. *Principles of Database and Knowledge-Base Systems: Volume II: The New Technologies*. W. H. Freeman & Co., New York, NY, USA, 1990. ISBN 071678162X. (Cited on page 51)

V. Uren, P. Cimiano, J. Iria, S. Handschuh, M. Vargas-Vera, E. Motta, and F. Ciravegna. Semantic annotation for knowledge management: Requirements and a survey of the state of the art. *Journal of Web Semantics*, 4(1):14–28, 2006. ISSN 1570-8268. doi: {DOI}10.1016/j.websem.2005.10.002. (Cited on pages 82, 83, 86, and 87)

M. Uschold. Building Ontologies: Towards a Unified Methodology. In *Proceedings of Expert Systems '96, the 16th Annual Conference of the British Computer Society Specialist Group on Expert Systems*, Cambridge, UK, Dec. 1996. (Cited on page 43)

M. Uschold and M. Grüninger. Ontologies: Principles, Methods and Applications. *Knowledge Sharing and Review*, 11(2):93–136, June 1996. (Cited on pages 42, 44, 45, 53, 55, 56, 58, and 68)

M. Uschold and M. Gruninger. Ontologies and Semantics for Seamless Connectivity. *SIGMOD Record*, 33(4):58–64, 2004. (Cited on pages 44, 45, 46, 47, 90, and 94)

M. Uschold and M. King. Towards a Methodology for Building Ontologies. In *Workshop on Basic Ontological Issues in Knowledge Sharing, held in conjunction with IJCAI-95*, Montreal, Canada, 1995. (Cited on pages 58 and 259)

V. Vaishnavi and W. Kuechler. Design Research in Information Systems. last updated August 16, 2009, Jan. 2004. URL http://www.isworld.org/Researchdesign/drisISworld.htm. (Cited on pages xiii, 5, and 6)

C. Van Damme, M. Hepp, and K. Siorpaes. FolksOntology: An Integrated Approach for Turning Folksonomies into Ontologies. In *In Proceedings of the ESWC Workshop "Bridging the Gap between Semantic Web and Web 2.0" (SemNet 2007)*, pages 57–70, 2007. (Cited on pages 79 and 80)

G. van Heijst, A. T. Schreiber, and B. J. Wielinga. Using explicit ontologies in KBS development. *International Journal Human-Computer Studies*, 46(2):183–292, 1997. URL http://dx.doi.org/10.1006/ijhc.1996.0090. (Cited on page 43)

M. Vargas-Vera, E. Motta, J. Domingue, M. Lanzoni, A. Stutt, and F. Ciravegna. MnM ontology driven semi-automatic or automatic support for semantic markup. In A. Gómez-Pérez and V. R. Benjamins, editors, *Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web, Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management, EKAW02*, volume 2473 of *LNAI*, pages 379–391, Berlin, 2002. Springer Verlag. (Cited on page 86)

J. C. A. Vega, O. Corcho, M. Fernández-López, and A. Gómez-Pérez. WebODE: a scalable workbench for ontological engineering. In *K-CAP*, pages 6–13. ACM, 2001. ISBN 1-58113-380-4. (Cited on page 64)

M. Völkel. *Personal Knowledge Models with Semantic Technologies* . PhD thesis, Karlsruher Institut für Technologie (KIT), Fakultät für Wirtschaftswissenschaften (Fak. f. Wirtschaftswiss.) Institut für Angewandte Informatik und Formale Beschreibungsverfahren (AIFB), Karlsruhe, Germany, July 2010. URL http://digbib.ubka.uni-karlsruhe.de/volltexte/documents/1453712. (Cited on page 27)

L. von Ahn. Games with a Purpose. *Computer*, 39(6):92–94, June 2006. ISSN 0018-9162. doi: 10.1109/MC.2006.196. URL http://www.cs.cmu.edu/~biglou/ieee-gwap.pdf. (Cited on pages 244 and 257)

G. A. Vouros, K. Kotis, C. Chalkiopoulos, and N. Lelli. The HCOME-3O Framework for Supporting the Collaborative Engineering of Evolving Ontologies. In L. Chen, P. Cudré-Mauroux, P. Haase, A. Hotho, and E. Ong, editors, *Proceedings of the International Workshop on Emergent Semantics and Ontology Evolution (ESOE2007) at ISWC/ASWC2007*, volume 292 of *CEUR Workshop Proceedings*, pages 95–107. CEUR-WS.org, 2007. URL http://ceur-ws.org/Vol-292/paper10.pdf. (Cited on pages 71 and 72)

Z. Vrandecic. *Ontology Evaluation*. PhD thesis, Karlsruher Institut für Technologie (KIT), Fakultät für Wirtschaftswissenschaften (Fak. f. Wirtschaftswiss.) Institut für Angewandte Informatik und Formale Beschreibungsverfahren (AIFB), Karlsruhe, Germany, 2010. (Cited on page 55)

T. V. Wal. Explaining and Showing Broad and Narrow Folksonomies, Feb. 2005. URL http://www.personalinfocloud.com/2005/02/explaining_and_.html. (Cited on page 21)

T. V. Wal. Folksonomy Coinage and Definition, Feb. 2007. URL http://vanderwal.net/folksonomy.html. (Cited on page 20)

A. Walter. *Hochwertige Bildsuche mittels empirisch fundierter semantischer Verfahren*. PhD thesis, Karlsruher Institut für Technologie (KIT), Fakultät für Informatik, Oct. 2010. URL http://digbib.ubka.uni-karlsruhe.de/volltexte/documents/1518757. (Cited on pages 83, 149, and 243)

F. Wang and M. J. Hannafin. Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, 53(4):5–23, 2005. URL http://lopezlearning.net/files/19511441FenWangArticle-2.pdf. (Cited on page 6)

Q. Wang and H. Jin. Selective message distribution with people-tagging in user-collaborative environments. In *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems (CHI'09)*, pages 4549–4554, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-247-4. doi: 10.1145/1520340.1520698. URL http://doi.acm.org/10.1145/1520340.1520698. (Cited on page 258)

R. Wash and E. Rader. Public bookmarks and private benefits: An analysis of incentives in social computing. *Proceedings of the American Society for Information Science and Technology*, 44(1):1–13, 2007. doi: 10.1002/meet.1450440240. URL http://dx.doi.org/10.1002/meet.1450440240. (Cited on pages xiii and 28)

W. Waterfeld, M. Weiten, and P. Haase. Ontology Management Infrastructures. In M. Hepp, P. D. Leenheer, A. de Moor, and Y. Sure, editors, *Ontology Management*, volume 7 of *Semantic Web And Beyond Computing for Human Experience*, pages 59–87. Springer, 2008. ISBN 978-0-387-69900-4. (Cited on page 63)

K. E. Weick, K. M. Sutcliffe, and D. Obstfeld. Organizing and the Process of Sensemaking. *ORGANIZATION SCIENCE*, 16(4):409–421, 2005. doi: 10.1287/orsc.1050.0133. URL http://orgsci.journal.informs.org/cgi/content/abstract/16/4/409. (Cited on page 35)

D. Weinberger. Tagging and why it matters. Berkman Center Research Publication No. 2005-07, May 2005. URL http://ssrn.com/abstract=870594. (Cited on page 20)

M. Weiten. OntoSTUDIO® as a Ontology Engineering Environment. In J. Davies, M. Grobelnik, and D. Mladenic', editors, *Semantic Knowledge Management*, pages 51–60. Springer Berlin Heidelberg, 2009. ISBN 978-3-540-88845-1. URL http://dx.doi.org/10.1007/978-3-540-88845-1_5. (Cited on page 66)

K. Weller. Folksonomies and Ontologies: Two New Players in Indexing and Knowledge Representation. In *Applying Web 2.0. Innovation, Impact and Implementation: Online Information 2007 Conference Proceedings*, pages 108–115, London, 2007. URL http://wwwalt.phil-fak.uni-duesseldorf.de/infowiss/admin/public_dateien/files/35/1197280560weller009p.pdf. (Cited on pages xv, 46, and 48)

K. Weller. *Knowledge Representation in the Social Semantic Web*, volume 3 of *Knowledge and Information*. De Gruyter Saur, Berlin, Germany, 2010. (Cited on pages 63, 75, 114, 244, and 261)

K. Weller and I. Peters. Seeding, Weeding, Fertilizing. Different Tag Gardening Activities for Folksonomy Maintenance and Enrichment. In S. Auer, S. Schaffert, and T. Pellegrini, editors, *Proceedings of I-Semantics '08, International Conference on Semantic Systems*, pages 100–117, 2008. URL http://wwwalt.phil-fak.uni-duesseldorf.de/infowiss/admin/public_dateien/files/35/1221222331triple-i_t.pdf. (Cited on page 260)

C. Welty, F. Lehmann, G. Gruninger, and M. Uschold. Ontology: Expert Systems All Over Again? In *Invited panel at AAAI-99: The National Conference on Artificial Intelligence*, Austin, Texas, USA, 1999. (Cited on page 45)

E. Wenger. *Communities of Practice. Learning, Meaning, and Indentity*. Cambridge University Press, Cambridge, UK, 1999. (Cited on page 199)

E. Wenger, R. Mcdermott, and W. M. Snyder. *Cultivating Communities of Practice*. Harvard Business School Press, Boston, USA, 1 edition, Mar. 2002. (Cited on page 199)

R. Wetzker. *Graph-Based Recommendation in Broad Folksonomies*. PhD thesis, Von der Fakultät IV – Elektrotechnik und Informatik der Technischen Universität Berlin, Berlin, Germany, May 2010. (Cited on page 148)

H. F. Witschel, B. Hu, U. Riss, B. Thönssen, K. Hinkelmann, R. Brun, A. Martin, K. Schöfegger, T. Ley, and S. Braun. Model of organizational requirements and of supporting services of the OLME. deliverable, Apr. 2009. (Cited on page 9)

J. Wu and A. Pinsonneault. Facilitating Sensemaking in Organizations Through Social Navigation Systems. In *Proceedings of the Fourteenth Americas Conference on Information Systems (AMCIS 2008)*. Association for Information Systems, 2008. (Cited on page 35)

W. Wu, X. Du, H. Hu, and N. Ma. An Ontology-Based and Cooperative Annotation System. In Z. Shi, K. Shimohara, and D. D. Feng, editors, *Intelligent Information Processing*, volume 228 of *IFIP*, pages 537–542. Springer, 2006. ISBN 978-0-387-44639-4. (Cited on page 86)

K. Xu, Y. Chen, Y. Jiang, R. Tang, Y. Liu, and J. Gong. A Comparative Study of Correlation Measurements for Searching Similar Tags. In C. Tang, C. Ling, X. Zhou, N. Cercone, and X. Li, editors, *Advanced Data Mining and Applications*, volume 5139 of *Lecture Notes in Computer Science*, pages 709–716. Springer Berlin / Heidelberg, 2008. URL http://dx.doi.org/10.1007/978-3-540-88192-6_75. (Cited on page 79)

Z. Xu, Y. Fu, J. Mao, and D. Su. Towards the Semantic Web: Collaborative Tag Suggestions. In *Proceedings of Collaborative Web Tagging Workshop at 15th International World Wide Web Conference*, 2006. URL http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.97.4194. (Cited on pages 25 and 30)

V. Zacharias and S. Braun. Tackling the Curse of Prepayment – Collaborative Knowledge Formalization Beyond Lightweight. In *1st Workshop on Incentives for the Semantic Web (INSEMTIVE), 7st International Semantic Web Conference ISWC 2008*, CEUR Workshop Proceedings, 2008. (Cited on pages 242 and 267)

L. Zhou. Ontology learning: state of the art and open issues. *Inf. Technol. and Management*, 8:241–252, Sept. 2007. ISSN 1385-951X. doi: 10.1007/s10799-007-0019-5. URL http://dx.doi.org/10.1007/s10799-007-0019-5. (Cited on page 77)

A. Zollers. Emerging Motivations for Tagging: Expression, Performance and Activism. In *Proceedings of the Workshop on Tagging and Metadata for Social Information Organization at WWW2006*, Banff, Alberta, Canada, 2007. URL http://www2007.org/workshops/paper_55.pdf. (Cited on pages xiii and 28)

G. L. Zuniga. Ontology: its transformation from philosophy to information systems. In *Proceedings of the International Conference on Formal Ontology in Information Systems 2001*, pages 187–197. ACM Press, 2001. (Cited on page 41)