

Linköping Studies in Science and Technology
Dissertations, No. 1891

**Fostering User Involvement in
Ontology Alignment
and Alignment Evaluation**

by

Valentina Ivanova



Linköping University
Department of Computer and Information Science
Division of Database and Information Techniques
SE-581 83 Linköping, Sweden

Linköping 2017

Edition 1:1

© Valentina Ivanova, 2017

ISBN 978-91-7685-403-7

ISSN 0345-7524

URL <http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-143034>

Published articles have been reprinted with permission from the respective
copyright holder.

Typeset using X_YTEX

Printed by LiU-Tryck, Linköping 2017

ABSTRACT

The abundance of data at our disposal empowers data-driven applications and decision making. The knowledge captured in the data, however, has not been utilized to full potential, as it is only accessible to human interpretation and data are distributed in heterogeneous repositories.

Ontologies are a key technology unlocking the knowledge in the data by providing means to model the world around us and infer knowledge implicitly captured in the data. As data are hosted by independent organizations we often need to use several ontologies and discover the relationships between them in order to support data and knowledge transfer. Broadly speaking, while ontologies provide formal representations and thus the basis, ontology alignment supplies integration techniques and thus the means to turn the data kept in distributed, heterogeneous repositories into valuable knowledge.

While many automatic approaches for creating alignments have already been developed, user input is still required for obtaining the highest-quality alignments. This thesis focuses on supporting users during the cognitively intensive alignment process and makes several contributions.

We have identified front- and back-end system features that foster user involvement during the alignment process and have investigated their support in existing systems by user interface evaluations and literature studies. We have further narrowed down our investigation to features in connection to the, arguably, most cognitively demanding task from the users' perspective—manual validation—and have also considered the level of user expertise by assessing the impact of user errors on alignments' quality. As developing and aligning ontologies is an error-prone task, we have focused on the benefits of the integration of ontology alignment and debugging.

We have enabled interactive comparative exploration and evaluation of multiple alignments at different levels of detail by developing a dedicated visual environment—*Alignment Cubes*—which allows for alignments' evaluation even in the absence of reference alignments.

Inspired by the latest technological advances we have investigated and identified three promising directions for the application of large, high-resolution displays in the field: improving the navigation in the ontologies and their alignments, supporting reasoning and collaboration between users.

The work has been supported by the Swedish Research Council (2010-4759), the Swedish Graduate School in Computer Science (CUGS), the Swedish e-Science Research Centre (SeRC) and the EU FP7 project VALCRI (FP7-IP-608142).

POPULÄRVETENSKAPLIG SAMMANFATTNING

Under de senaste 30 åren har webben fundamentalt och oåterkalleligen förändrat våra liv—hur vi arbetar, hur vi tar del av underhållning och hur vi kommunicerar. Webben innehåller miljarder av informationskällor i en uppsjö olika format—websidor, databaser, dokument, figurer, etc.—alla ihopkopplade genom en enorm mängd länkar. Webben är inte statisk, utan växer konstant och erbjuder ett överflöd av data, som kan användas för datadrivet beslutsfattande och andra tillämpningar.

Hur kan vi dra nytta av all denna data? Tänk dig att du planerar din sommarledighet. Skulle detta kunna göras av en automatisk reseplanerare? För att planera resan måste denna ta hänsyn till olika aspekter:

- Flightscheman—valda flighter måste passa dina semestertider och måste vara kompatibla med olika personliga preferenser och begränsningar—favoritresmål och väderförhållanden på dessa, bytestider på mellanliggande flygplatser, medlemskort för specifika flygbolag, undvikande av länder med visumkrav.
- Hotellvistelse—det skall vara en bra plats i en lugn stadsdel men inte för dyrt; hotellet skall ha bra användarrecensioner och bra allmänna kommunikationer till flygplats och sightseeingmål.
- Underhållning/sightseeing—hitta biljetter till kultur, sport eller andra evenemang under vistelsen, som inte kolliderar tidsmässigt och som ligger på bekvämt avstånd från hotellet.
- Mat—hitta högt rankade restauranger som uppfyller eventuella dietkrav.

Eftersom all denna data finns tillgänglig på webben, skulle en automatisk reseplanerare kunna behandla all tillgänglig data och planera din semester enligt ovanstående krav? I dag är svaret nej, detta är fortfarande inte möjligt. Eftersom innehåll på webben är kodat i ett format som lämpar sig för att läsas av människor är det svårt för maskiner att tolka innehållet. Maskiner ser endast strängar av symboler där människor ser en innebörd uttryckt i ord, fraser och meningar.

Att göra information maskinläsbar är därför en nyckelfråga idag. Fram tills alldeles nyligen har datorer endast lagrat, överfört och visat dokument, utan möjlighet att tolka informationen de innehåller, utan att kunna 'förstå' kunskapen representerad i dem. Detta håller dock på att förändras tack vare utvecklingen mot den Semantiska webben, där en av nyckelteknikerna är ontologier. Vi använder ontologier för att definiera betydelsen hos termer och relationerna dem emellan, samt för att automatiskt behandla information. Till exempel kan vi definiera vad en "bra plats" betyder, och vad "dyr" innebär. Genom att använda ontologier kan vi dra slutsatsen att en given restaurang serverar vegetarisk mat. Ibland använder dock olika ontologier olika men synonyma termer. Då måste vi hitta dessa relationer mellan termer i olika ontologier. Processen för att hitta sådana relationer kallas för justering av ontologier (*eng.* ontology alignment).

Den här avhandlingen handlar om problem inom justering av ontologier, alltså hur man finner korrekta relationer mellan termer i olika ontologier. Trots att flera automatiska metoder redan utvecklats krävs fortfarande att delar av processen utförs manuellt av människor för att uppnå högkvalitativa justeringar. De manuella delarna av processen kan dock vara svåra att utföra, särskilt för stora och komplexa ontologier. I den här avhandlingen undersöker vi olika sätt att hjälpa människor med justeringsprocessen och med att utvärdera kvaliteten på de upptäckta relationerna mellan termer.

Acknowledgments

When life brought me to Sweden and later to PhD studies I had never imagined the wonderful opportunities I would discover and the personal growth I would experience. They did not come for granted, though. Conducting research is challenging and even frustrating from time to time, but extremely rewarding. During these several years I had my hands full with challenges and experiences. These were constantly demanding to push my boundaries, something I started to enjoy and I will truly miss. Many people were part of my PhD endeavor and I wish to express my gratitude to them here.

I am sincerely and deeply thankful to my supervisor, Patrick Lambrix, who provided me with plentiful of research, and also management, opportunities and challenges and supported me while I was conquering them. These motivated and drove me forward! He provided an encouraging and relaxed working environment and has always been around for advice and a discussion. Honestly, I could not have imagined a better supervisor. Thank you, Patrick, for the challenges and opportunities you gave me!

I am especially grateful to Nahid Shahmehri, my co-supervisor, who is the main reason for me to start PhD studies. She is the one who first believed in my research talent and kindly advised me. I am also thankful to Lena Strömbäck and David Byers who introduced me to the wonderful world of research and showed me, early on, that I possess the strength to take this adventure.

On one of my conference trips I met Vania Dimitrova to whom I am very thankful for providing friendly discussions for the PhD life and life in general. During that week, in the course of several meetings she turned into my mentor and greatly helped me to address my insecurities and grow more confident.

During my time as a PhD student I was co-organizing several events—the VOILA! workshop series, the VISUAL workshop, two tracks at the OAEI and a special issue at the JWS. I am very grateful to all my co-organizers for the work we have done together and for sharing their rich experience. These events have spiced my PhD experience in a completely different way. My very special thanks go to Steffen Lohmann who, at the beginning, taught me basically everything for events' organization and gave me a good starting

point to build upon in the future. Thank you, Steffen, for this and for the wonderful experience while working together during the workshops!

I was happy to collaborate with several people for some of the papers in this thesis. Thank you, Emmanuel, Benjamin, Catia, Ernesto and Daniel for the inspiring discussions, the hard work we did together and for letting me learn from you.

The time here would not have been that enjoyable without my past and present ADIT colleagues who made the work environment fun and relaxing. Thank you, Vengat, Marcus, Dag, Jose and Olaf for the entertaining and sometimes strange discussions with you. Thanks go to Ulf for translating my popular Swedish abstract. Traveling around the world would not have been such a wonderful experience without sharing it with Zlatan with whom I had many trips. He was always around to answer my questions and help with various matters. We worked well together and I am really sorry we did not collaborate more.

I also thank the members of the IDA administrative group, who helped with various administrative matters and my conference trips. I especially thank Anne for her considerate and kind assistance and for making my life as a PhD student easier.

I am greatly thankful to my family and friends for their unquestioning support and constant encouragement. Their belief in the successful end of this adventure has always been helping me overcome difficult times. Early on, during my childhood, my family constantly taught me that I can achieve everything I wish for with hard work and persistence. This believe has played a huge role during my studies.

No words could express my appreciation to my life partner, Pavel, who was my continuous support but also, at times, my hardest critiquer. He shared the sunny and stormy weather with me. Thank you, Pavel, for your love and for being here!

Valentina Ivanova
November 2017
Linköping, Sweden

Contents

Abstract	iii
Acknowledgments	xi
Contents	xi
1 Introduction	3
1.1 Motivation	3
1.1.1 Ontologies	4
1.1.2 Ontology Alignment	5
1.1.3 Ontology Alignment Evaluation	6
1.2 Research Questions	7
1.3 Contributions	8
1.4 Research Methods	10
1.5 List of Publications	12
1.5.1 Included Papers	12
1.5.2 Other Publications	13
1.6 Thesis Outline	15
2 Background	17
2.1 Ontologies	17
2.1.1 Classification	18
2.1.2 Applications	19
2.2 Ontology Alignment	20
2.2.1 Ontology Alignment Framework	21
2.2.2 Ontology Alignment Evaluation	23
2.3 Ontology Debugging	24
2.3.1 Definitions	25
3 Summary of Papers	27
4 Related Work	31
4.1 Ontology Alignment	31
4.2 User Involvement in Ontology Alignment	33

4.3	Ontology Alignment Evaluation	37
4.3.1	Measures	37
4.3.2	Tools & Frameworks	39
4.4	Debugging of Ontology Networks	40
4.4.1	Detecting and Repairing Missing Structure in Ontology Networks	41
4.4.2	Semantic Defects in Ontology Alignments	43
4.5	Information Visualization	45
4.5.1	Visual Data Analysis and Exploration	45
4.5.2	Large and High-Resolution Displays	47
4.5.3	Navigation in Digital Environments	52
5	Conclusions and Future Work	55
5.1	Conclusions	55
5.1.1	User Involvement during Ontology Alignment	56
5.1.2	Ontology Alignment Evaluation	57
5.2	Future Work	58
	Bibliography	61
	Paper I	77
	Paper II	97
	Paper III	117
	Paper IV	135
	Paper V	145
	Paper VI	187



1

Introduction

1.1 Motivation

“*We are drowning in information but starved for knowledge.*” has been said more than thirty years ago¹. It is even more true today! For the past 30 years, the Web has, immensely and irreversibly, changed our lives—the way we work, we enjoy and communicate. The Web of Documents, as we call it, contains billions of information sources in a variety of formats—web pages, databases, documents, figures, etc.—interconnected through an enormous number of links. The Web is not static, it is constantly growing. Besides, we are living in the Big Data era which provides abundance of heterogeneous information sources² as highlighted by the popular 3Vs of Big Data³—volume, velocity and variety⁴.

The value of the knowledge captured in the Web of Documents has not been utilized to its full potential, as it is only accessible to human interpretation; today the knowledge is not available and cannot be utilized by machines

¹by John Naisbitt in *Megatrends: Ten New Directions Transforming Our Lives*, published in 1982.

²Both accessible and unaccessible on the Web.

³‘3Vs (of Big Data)’ refer to a widely adopted abbreviation which denotes the three initial challenges for management of large datasets as presented in [46]; *volume* refers to the amount of data, *velocity*—to the rate of (incoming) data and *variety*—to the diversity of formats for encoding data; at present, ‘5Vs of Big Data’ is in common use to reflect two other dimensions—*veracity* (uncertainty) and *value*.

⁴It is the challenge of *variety* that can be efficiently addressed by harnessing the expressive power of the Semantic Web techniques.

yet. Documents on the Web are mostly encoded in human-readable formats; extracting meaning from them is a task that only we, humans, can perform [15]. We, as humans, are capable of interpreting the variety of sources and formats, dealing with ambiguous, incomplete and overlapping information and to integrate them for the purpose of creating knowledge and fulfilling information needs. We, however, possess limited capabilities to process and comprehend the rapidly growing amounts of data, we cannot compete with the ever-increasing processing power of machines. In order to take advantage of the data at our disposal and to turn it into knowledge we need to enable machines to use it similarly to the way we do—interpret data in context, deal with various quality issues, integrate and draw conclusions from them in order to produce knowledge and fulfill tasks similarly to the way we do.

Making information machine-understandable is, therefore, a key problem nowadays. Until very recently, machines have only stored, transmitted and displayed documents without a means to make sense of the data they capture, and without ‘understanding’ the knowledge they convey. This is now starting to change with the evolution of the Web towards the Semantic Web, a concept introduced in 2001 in the seminal work of Tim Berners-Lee, James Hendler and Ora Lassila [15]. The Semantic Web encompasses a set of techniques for expressing and processing data in machine-readable formats. Owing to the Semantic Web progress we are now moving from the Web of Documents to the Web of Data where data captured in documents become accessible to machines for interpretation, integration and inference.

1.1.1 Ontologies

Ontologies are one of the key technologies in the Semantic Web, Linked Data being another. Ontologies provide means to describe and categorize real-world and abstract entities, the properties they could possess and the relationships in which they could participate. Thus, ontologies provide a shared vocabulary of a domain by formally representing the meaning of its concepts and relations and by defining rules for creating new concepts [96, 122]. They are a common language, machines can utilize to talk about the world. Ontologies serve as a basis for sharing, integrating and reusing knowledge and enable interoperability between systems. Tools, known as reasoning engines or reasoners, can infer new knowledge from ontologies.

Ontologies are complex information artifacts which can differ from each other in various aspects—by ontology definition language, by the capability to describe complex entities, by (number of) components. A recent survey [135] conducted among 118 respondents sheds light on ontology usage and illustrates such differences. According to this survey, there are three common languages with almost equal shares—OWL, RDF and RDFS—with OWL users employing different profiles, i.e., different capabilities to represent complex entities.

Ontologies have already been utilized in a variety of domains with significant interest in the Life Sciences—one of the early adopters which develops some of the largest ontologies. Almost a third of the survey’s respondents listed the Biomedical domain as a primary application area and the GeneOntology⁵ has been identified as one of the five most commonly used ontologies.

1.1.2 Ontology Alignment

There is no universal ontology and it is unlikely one will ever exist. Ontologies are developed by different people and organizations to fulfill different aims and reflect the views of their developers. Thus, there may exist several ontologies modeling the same domain which could differ in conceptual modeling, granularity level, vocabulary and domain coverage. Consequently, in order to enable data and knowledge transfer we need to use more than one ontology and to know how the concepts from the different ontologies are related to each other. Are two concepts equivalent, i.e., do they represent the same group of real-world or abstract entities? Does one subsume the other? Are two concepts incompatible?

These and related questions are the subject of investigation of the Ontology Alignment field^{6,7}. Broadly speaking, while ontologies provide formal representations and thus the basis, ontology alignment is the means to achieve knowledge sharing and reuse by providing techniques for integrating different information sources. Now, in the Big Data era, ontology alignment supplies techniques to turn the data kept in distributed, heterogeneous datasets into valuable knowledge.

One of the four user groups identified in the aforementioned survey [135] harnesses the expressive power of ontologies for the purposes of data integration at either schema or instance level. The interest in the ontology alignment area in the past 15 years has led to the development of many ontology alignment tools which, in most cases, apply fully automated approaches to compute *an alignment* (a set of relationships between the entities of a pair of input ontologies) without any human intervention. The progress in the field has been accelerated by a dedicated event—the Ontology Alignment Evaluation Initiative⁸ (OAEI) which has provided a discussion forum for developers and a platform for an annual evaluation of their tools.

User Involvement

Advancing the algorithms has not led to comparable improvements in the quality of the computed alignments [49], it has, therefore, been hypothesized

⁵It describes gene functions and how they are related—<http://www.geneontology.org/>.

⁶A broad and comprehensive introduction to the field can be found in [40].

⁷The terms Ontology Alignment, Ontology Mapping and Ontology Matching are often used interchangeably.

⁸<http://oaei.ontologymatching.org/>

that fully automated approaches are reaching a ceiling with regards to alignments' quality [102] and should only be considered as the first step in aligning ontologies [38]. Already ten years ago, practitioners have suggested that involving users will lead to a greater improvement in the alignments' quality than developing more accurate algorithms [16]. Similarly, an early survey dedicated to ontology mapping envisioned that better tools, rather than better algorithms, will lead to improvements in the quality [43].

Nearly half of the challenges identified in [118], and restated several years later in [117], are directly related to user involvement. These include *explanation of matching results to users*, fostering the *user involvement* in the matching process and *social and collaborative matching*. Another challenge aims at supporting users' collaboration by providing *infrastructure and support* during all phases of the alignment process.

Ontology alignment practitioners usually come from two backgrounds. Some practitioners—domain experts—possess an expertise in the domains which ontologies describe, as ontologies are usually domain specific and built to represent the knowledge in a particular area. These users are typically not trained in knowledge engineering. Other users possess technical expertise and formal training in the field of knowledge modeling and representation (knowledge engineers).

Users are most often involved in selecting and configuring matching strategies, validating automatically generated mappings, etc. The alignment process usually demands users exploring both (unfamiliar) ontologies in order to become familiar with them and their formal representations, and to understand their modelers' view of the domain. Further, users need to explore the mappings (*a mapping*⁹ represents a relationship between two concepts from different ontologies) computed by the tool's algorithms in order to determine their correctness and identify mappings missed by the system [44]. Thus, it is a cognitively demanding task that involves a high memory load and complex decision making. Furthermore, it is an inherently error-prone process as different users possess different levels of domain and knowledge representation expertise, due to human biases, experience and misinterpretations [45].

1.1.3 Ontology Alignment Evaluation

Once an alignment has been created, either fully- or semi-automatically, its quality is often evaluated by comparing it against a gold standard (a *reference alignment*—RA) and calculating measures such as precision, recall and F-measure¹⁰. These measures provide a good overall assessment of the quality of alignments in terms of the ratio of found mappings, missed mappings and wrongly suggested mappings. However, they do not allow comparison of alignments of specific parts of ontologies, or comparison of alignments to each

⁹Mapping and correspondence are interchangeable terms in this thesis.

¹⁰These measures are defined in Subsection 4.3.

other and to the RA at the detailed level of concepts and relations. Without means to compare the tools and algorithms at a detailed level, their strengths and weaknesses cannot easily be revealed and understood.

Furthermore, RAs are often not available, as their development is time and effort consuming and requires domain expertise. In the absence of RAs, the evaluation of alignments requires exploration and comparison of multiple alignments. This involves users performing tasks at different levels of granularity [5, 33, 113] such as determining regions with similar or different number of mappings between the alignments, determining common or rarely found mappings and characterizing mappings as correct or incorrect. These activities serve as a basis to decide how good the obtained alignment is and thus to compare alignment tools and algorithms. Currently, however, there is little support for performing these tasks in an interactive and flexible manner. Users and developers rely on custom scripts, which can be error-prone and time-consuming to develop and fine-tune. Additionally, the output of these scripts is cumbersome to explore with the growing size of the ontologies and number of alignments.

1.2 Research Questions

The discussion above makes a strong call for efficient support for the practitioners involved in the alignment process, as both aligning ontologies and evaluating the quality of the developed alignments are cognitively demanding tasks. Thus, the main subject of investigation of this thesis has been formulated in the following research question (RQ):

***RQ:** How to provide efficient user support during the process of ontology alignment and alignments' evaluation?*

This question consists of two parts which we have addressed separately. Since fully automated alignment approaches are only considered as the first step in creating alignments we have been interested in:

***RQ 1:** What features should an ontology alignment system provide in order to efficiently support users during the process of ontology alignment?*

As discussed earlier, users can be involved during different stages of the alignment process. One of the steps—validating candidate mappings—is likely the most effort- and time-demanding, especially when large and more complex ontologies are involved. It requires familiarity with both ontologies and exploring the mappings computed by the tools in order to determine their correctness and create mappings missed by the system. As explained, this is an inherently error-prone process due to different levels of users'

domain and knowledge representation expertise, experience, human biases, misinterpretations, etc. Thus, we have focused on the manual validation and have further investigated:

***RQ 1.1:** How to support users during the validation of the candidate mappings?*

Recently, with the development of technology and the associated cost reduction, large, high-resolution displays have become available at affordable prices. It has been pointed out that ‘when a display exceeds a certain size, it becomes qualitatively different’ [123]. A number of studies have shown improved performance and reduced cognitive load in an everyday office environment due to more peripheral awareness, glancing instead of windows switching to obtain additional information, flexibility in the organization of the space, etc. Environments where large displays are present are well-suited for activities involving several people where they can simultaneously work and discuss. Thus, we have looked into different means, beyond the traditional desktop and mouse interaction paradigms, and have formulated the following:

***RQ 1.2:** Are there benefits from applying large displays to ontology alignment for individual users and in a collaborative setting?*

The other part of the main research question is concerned with providing interactive means for evaluation of the alignments. As pointed out previously, users resort to writing custom scripts in order to evaluate their alignments at fine-grained level and thus to reveal the strengths and weaknesses of their tools and algorithms. This is also a laborious task as these scripts are crafted for every particular ‘question’ the user may have. Each new question would demand a new script. It implies that the user knows in advance what ‘questions’ to ask and does not support flexible exploration of several alignments and obtaining unexpected observations. Besides, comprehension of their results is cumbersome especially when the size and number of alignments grow. These considerations have led us to the following:

***RQ 2:** How to efficiently support users during the evaluation of ontology alignments?*

1.3 Contributions

The research questions above have been investigated in several of our works:

With respect to ***RQ 1:** What features should an ontology alignment system provide in order to efficiently support users during the process of ontology alignment?*

- In Paper I we conducted a literature review encompassing a number of promising works in the area and have identified a set of (both back-end and front-end) features that need to be supported by a semi-automated ontology alignment system. We then investigated if these features have been supported by the state-of-the-art tools. We paid special attention to the front-end features by further conducting two user interface evaluations.
- Due to the complexity of the alignment problem and in order to improve the alignments' quality, a debugging step is necessary and it has been identified as one of the desirable features of ontology alignment systems in Paper I. In Paper III we further investigated this issue in the context of taxonomies and taxonomy networks by integrating ontology alignment and debugging components. We showed that this integration leads to improved quality of both the alignments and the ontologies themselves.

With respect to **RQ 1.1:** *How to support users during the validation of the candidate mappings?*

- In Paper II we deepened our understanding of both back- and front-end issues by narrowing down the scope of literature sources to those focusing on user validation of *candidate mappings* (suggested by the tool but not yet checked by the user). Due to the differences in users' training, e.g., domain expert versus knowledge engineer, we have further considered the impact of users' expertise. We have also demonstrated that even if users make mistakes (up to 20%), user validation of candidate mappings leads to an improved alignments' quality.
- In Paper I we have identified several tasks users need to perform during manual validation and have further conducted a controlled experiment in order to reveal how efficiently these tasks have been supported by three systems.

With respect to **RQ 1.2:** *Are there benefits from applying large displays to ontology alignment for individual users and in a collaborative setting?*

- In Paper IV we conducted a literature review encompassing various fields—Cognitive Psychology, Navigation in Information Spaces, Human-Computer Interaction, Computer-Supported Cooperative Work and Software Engineering and have identified three promising directions for the application of large, high-resolution displays in the field of ontology alignment: improving ontologies' and alignments' navigation, supporting users' thinking process and collaboration between users.

With respect to **RQ 2:** *How to efficiently support users during the evaluation of ontology alignments?*

- In Paper V we have summarized experiences from organizing the OAEI Anatomy¹¹ and Interactive¹² tracks. While conducting analysis of the alignments submitted during the past 10 years we have identified several tasks that involve comparative assessment of multiple alignments.
- In Paper VI we have collected and analyzed several scenarios which demand comparative alignments' evaluation and have identified their shared tasks. Two of the scenarios are directly connected to our previous work—the comparative evaluation and exploration of several alignment for the purpose of comparing competing tools (Paper V) and for manually validating (Paper II) and debugging alignments (Paper III). We derived high-level interaction features to support these tasks and scenarios and implemented them in a prototype which has been built on top of a novel technique for interactive visual exploration of dynamic networks. We demonstrated the applicability of our approach with a walk-through scenario.

1.4 Research Methods

The research methods applied throughout the development of this thesis can be roughly grouped into two categories—(i) methods to study and organize existing approaches with respect to the user involvement perspective and (ii) various empirical methods to evaluate these approaches and their user interfaces with community benchmark datasets and human subjects. Our papers often combine methods from both categories in order to address different aspects of the studied questions.

We have addressed the former point by analyzing existing literature, conducting a case study and applying methods from grounded theory. In Papers I, II, IV and V we conducted several literature reviews with broad scope and aims for which we have used diverse literature sources. In Paper II we further employed an approach similar to grounded theory where we first extracted relevant features from existing literature, coded them and then grouped the codes into categories. Our work in Paper V—comparative assessment of several alignments—can be seen in the light of a case study used as a primary qualitative method¹³ [22]. Such case studies serve to obtain better understanding of users' tasks and working processes not necessary supported with tools and inform prototypes' design [22]. This case study together with our personal experience in the area were the basis for the functional requirements which informed the initial design of our tool presented in Paper VI.

¹¹<http://oaei.ontologymatching.org/2017/anatomy/>

¹²<http://oaei.ontologymatching.org/2017/interactive/>

¹³Such case studies are called exploratory case studies [35] and understanding work practices [61].

For our evaluations, the latter point above, we have combined qualitative and quantitative methods in order to obtain a richer understanding of the subject of inquiry [22, 50]. These methods differ in generalizability, precision and realism, and no study features all three [22]. Generalizability refers to the extent to which the results can be extended to other users and situations (than those considered in the study). Realism considers the similarity between the study context and the actual environment where the working process takes place. Precision refers to the extent to which the measurements can be considered certain and to the influence of other factors not intended in the design of the study.

While quantitative methods provide measurable data to study how changes in some factors impact others they are usually not conducted in a realistic setting and vary in generalizability. We distinguish between quantitative methods with human subjects, to assess usability of user interfaces, called controlled experiments¹⁴ and quantitative methods with community benchmark datasets, to study algorithms' features, called laboratory experiments. During controlled experiments participants follow a certain protocol and conduct tasks specified by the experimenter. This heavily impacts the realism and, depending on the selected tasks and participants, impacts the level of generalizability. Using benchmark datasets enables a comparison to other tools and approaches, however, due to the diversity of ontologies it does not allow precise generalization to other alignment cases. A higher degree of realism might be achieved when employing real-world test cases (as opposed to synthetic), however, due to the variable complexity of the ontologies they are unlikely to represent their diversity well.

On the other hand qualitative methods (think-aloud protocols, inspection methods, usage scenarios) are conducted in a more realistic setting and provide a richer understanding of the studied phenomenon [22]. Inspection methods, such as heuristic evaluation¹⁵, consist of evaluating a tool against a set of heuristics or guidelines developed by experts. Usage or walk-through scenarios¹⁶ are another qualitative evaluation method where (preferably) expert users are observed while analyzing their data in order to assess the extent to which the tool supports the analysis. Qualitative methods are also used with quantitative methods to provide additional insights and interpret results. These methods are called nested qualitative methods and include think-aloud protocols, experimenters' observations and collecting users' opinion during and after the experiments.

¹⁴The classification in [22] uses a laboratory experiment but we use a controlled experiment here to distinguish with the mentions of laboratory experiment in which algorithms are evaluated.

¹⁵A qualitative method called usability heuristics in [22].

¹⁶This type of evaluation is called confirmatory case study in [35]; both usage and walk-through scenarios falls into the scope of visual data analysis and reasoning (VDAR) in [61].

In Paper I we have employed quantitative, qualitative and nested qualitative methods respectively—a heuristic evaluation, a controlled experiment and collected observations and participants’ opinion, respectively, during and after the controlled experiment—to evaluate the usability of three state-of-the-art tools. As the evaluation of interactive visualization tools is challenging [22], methods such as case studies and usage scenario, are often employed [22, 61], as they are more likely to provide insightful observations than traditional controlled experiments. Thus in Paper VI we conducted a walk-through scenario in order to demonstrate the capability of our prototype to support comparative exploration and evaluation of several alignments.

We conducted laboratory experiments with several of the OAEI datasets in Papers I and II. In Paper II we studied the impact of erroneous validations on the alignments’ quality and in Paper III we studied the impact of the interaction between ontology alignment and debugging on the alignments’ quality.

1.5 List of Publications

1.5.1 Included Papers

Paper I V. Ivanova, P. Lambrix and J. Åberg. *Requirements for and Evaluation of User Support for Large-Scale Ontology Alignment*, In the Proceedings of the 12th Extended Semantic Web Conference - ESWC 2015, Lecture Notes in Computer Science, vol. 9088, pages 3-20.

Paper II Z. Dragisic, V. Ivanova, P. Lambrix, D. Faria, E. Jiménez-Ruiz and C. Pesquita. *User Validation in Ontology Alignment*, In the Proceedings of the 15th International Semantic Web Conference - ISWC 2016, Lecture Notes in Computer Science, vol. 9981, pages 200 - 217.

Paper III V. Ivanova and P. Lambrix. *A Unified Approach for Aligning Taxonomies and Debugging Taxonomies and Their Alignments*, In the Proceedings of the 10th Extended Semantic Web Conference - ESWC 2013, Lecture Notes in Computer Science, vol. 7882, pages 1-15.

Paper IV V. Ivanova. *Applications of Large Displays: Advancing User Support in Large Scale Ontology Alignment*, In the Proceedings of the Doctoral Consortium at the 15th International Semantic Web Conference - ISWC 2016, CEUR Workshop Proceedings, vol 1733, pages 50-57.

Paper V Z. Dragisic, V. Ivanova, H. Li, and P. Lambrix. *Experiences from the Anatomy track in the Ontology Alignment Evaluation Initiative*, Journal of Biomedical Semantics, vol.8, no.1, 2017.

Paper VI V. Ivanova, B. Bach, E. Pietriga and P. Lambrix. *Alignment Cubes: Towards Interactive Visual Exploration and Evaluation of Multiple Ontology Alignments*, In the Proceedings of the 16th International Semantic Web Conference - ISWC 2017, Lecture Notes in Computer Science, vol. 10587, pages 400 - 417.

1.5.2 Other Publications

P. Lambrix, V. Ivanova. **A unified approach for debugging is-a structure and mappings in networked taxonomies**, Journal of Biomedical Semantics, vol.4, no.1, 2013.

P. Lambrix, Dragisic Zlatan, and V. Ivanova. **Get My Pizza Right: Repairing Missing is-a Relations in ALC Ontologies**, JIST 2012, Lecture Notes in Computer Science, vol. 7774, pages 17-32.

V. Ivanova. **Integration of Ontology Alignment and Ontology Debugging for Taxonomy Networks**, Licentiate Thesis, Department of Computer and Information Science, Linköping University, Linköping, Sweden, 2014.

V. Ivanova, B. Bach, E. Pietriga, P. Lambrix. **Alignment Cubes: Interactive Visual Exploration and Evaluation of Multiple Ontology Alignments**, Posters & Demos @ ISWC 2017, CEUR Workshop Proceedings, vol. 1963. Demo.

P. Lambrix, Z. Dragisic, V. Ivanova, C. Anslow. **Visualization for Ontology Evolution**, VOILA 2016 @ ISWC 2016, CEUR Workshop Proceedings, vol. 1704, pages 54-67.

V. Ivanova, P. Lambrix. **User Involvement for Large-Scale Ontology Alignment**, VISUAL 2014 @ EKAW 2014, CEUR Workshop Proceedings, vol. 1299, pages 34-47.

P. Lambrix, F. Wei-Kleiner, Z. Dragisic, V. Ivanova. **Repairing missing is-a structure in ontologies is an abductive reasoning problem**, WoDOOM 2013 @ ESWC 2013, CEUR Workshop Proceedings, vol. 999, pages 33-44.

V. Ivanova, P. Lambrix. **A System for Aligning Taxonomies and Debugging Taxonomies and Their Alignments**, The Semantic Web: ESWC 2013 Satellite Events @ ESWC 2013, Lecture Notes in Computer Science, vol. 7955, pages 152-156. Demo.

V. Ivanova, P. Lambrix. **A System for Aligning Taxonomies and Debugging Taxonomies and Their Alignments**, Video Journal of Semantic Data Management Abstracts, Volume 2.

V. Ivanova, J. L. Bergman, U. Hammerling, P. Lambrix. **Debugging Taxonomies and their Alignments: the ToxOntology-MeSH Use Case**, WoDOOM 2012 @ EKAW 2012, pages 25-36.

V. Ivanova, P. Lambrix. **A System for Debugging Taxonomies and their Alignments**, WoDOOM 2012 @ EKAW 2012, pages 37-42. Demo.

B. Cuenca Grau, Z. Dragisic, K. Eckert, J. Euzenat, A. Ferrara, R. Granada, V. Ivanova, E. Jiménez-Ruiz, A. O. Kempf, P. Lambrix, A. Nikolov, H. Paulheim, D. Ritze, F. Scharffe, P. Shvaiko, C. Trojahn, O. Zamazal. **Results of the Ontology Alignment Evaluation Initiative 2013**, OM 2013 @ ISWC 2013, CEUR Workshop Proceedings, vol. 1111, pages 61-100.

Z. Dragisic, K. Eckert, J. Euzenat, D. Faria, A. Ferrara, R. Granada, V. Ivanova, E. Jiménez-Ruiz, A. O. Kempf, P. Lambrix, S. Montanelli, H. Paulheim, D. Ritze, P. Shvaiko, A. Solimando, C. Trojahn, O. Zamazal, B. Cuenca Grau. **Results of the Ontology Alignment Evaluation Initiative 2014**, OM 2014 @ ISWC 2014, CEUR Workshop Proceedings, vol. 1317, pages 61-104.

M. Cheatham, Z. Dragisic, J. Euzenat, D. Faria, A. Ferrara, G. Flouris, I. Fundulaki, R. Granada, V. Ivanova, E. Jiménez-Ruiz, P. Lambrix, S. Montanelli, C. Pesquita, T. Saveta, P. Shvaiko, A. Solimando, C. Trojahn, O. Zamazal. **Results of the Ontology Alignment Evaluation Initiative 2015**, OM 2015 @ ISWC 2015, CEUR Workshop Proceedings, vol. 1545, pages 60-115.

M. Achichi, M. Cheatham, Z. Dragisic, J. Euzenat, D. Faria, A. Ferrara, G. Flouris, I. Fundulaki, I. Harrow, V. Ivanova, E. Jiménez-Ruiz, E. Kuss, P. Lambrix, H. Leopold, H. Li, C. Meilicke, S. Montanelli, C. Pesquita, T. Saveta, P. Shvaiko, A. Splendiani, H. Stuckenschmidt, K. Todorov, C. Trojahn, O. Zamazal. **Results of the Ontology Alignment Evaluation Initiative 2016**, OM 2016 @ ISWC 2016, CEUR Workshop Proceedings, vol. 1766, pages 73-129.

1.6 Thesis Outline

The remaining of this thesis is organized as follows:

Chapter 2 provides background in the areas relevant to this dissertation. It discusses what an ontology is and presents its components. It further gives a brief overview of the ontology alignment and debugging areas.

Chapter 3 gives a short summary of the papers part of this dissertation.

Chapter 4 provides the context in which the work in this thesis has been carried out.

Chapter 5 concludes the thesis and discusses directions for future work.



2

Background

This chapter provides background in the areas relevant to this dissertation.

Section 2.1 discusses the term *ontology* and presents several definitions in the scientific literature. It then lists components of ontologies and briefly outlines several applications. Sections 2.2 and 2.3 give a brief overview of the areas of ontology alignment and debugging. Formal definitions relevant to Paper III are given in Section 2.3.1.

2.1 Ontologies

The term *ontology* originates from philosophy, where it denotes a branch dealing with the matters of being and existence. At the beginning of the 80's the term appeared in the Artificial Intelligence community and was later used to refer to a knowledge representation formalism [136] and spread in other Computer Science disciplines. There are different definitions of ontologies available in the scientific literature and some of the most popular are:

- An ontology defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary [96];
- An ontology is an explicit specification of a conceptualization [51];
- An ontology is a hierarchically structured set of terms for describing a domain that can be used as a skeletal foundation for a knowledge base [124];

- An ontology provides the means for explicitly describing the conceptualization behind the knowledge represented in a knowledge base [14];
- An ontology is a formal, explicit specification of a shared conceptualization [122];

All definitions share the view that ontologies explicitly describe a topic area. They model the world around us (or someone's view of the world) explicitly defining the meaning of its concepts, the existing relationships between them (for instance, *part-of*, *is-kind-of*, *is-located-in*, *is-not*) and rules for creating new concepts. The last definition supplies two additional important features of the ontologies—they provide a shared understanding of the area in question and are formally encoded in a machine readable language.

Ontologies consist of different components representing different aspects of our knowledge of the world. Lambrix [72] lists the following components from a knowledge representation point of view. Corcho et. al [25] define a similar list containing a minimal set of components:

- concepts (classes) represent a group of entities in a domain. They show what types of entities exist in a domain.
- instances (individuals) represent the actual entities, they may not be represented in ontologies.
- relations (roles, properties) represent different relationships between concepts in a domain, such as *part-of*, *is-kind-of* and *is-located-in*. Two main types of relationships have been distinguished in [121]: *taxonomic* and *associative*. As described in [121], *subsumption (is-a)* and *part-of* are taxonomic relationships; they organize the concepts in hierarchies. Subsumption relations (known also as *is-a*, *is-kind-of* or subclass relations) are the most often used in ontologies since they represent a common relationship that occurs in many domains. A subsumption relation shows that one set of entities is a subset of another set of entities. Associative relations show possible other relations that can exist between the concepts in a domain.
- axioms represent facts that are always true in the domain described by the ontology and are not represented by the other components [25, 72]. Axioms impose constraints on the values and relationships in which entities can participate.

2.1.1 Classification

The ontologies can be classified according to various criteria. Several one-dimensional classifications are shown in [111] in the context of a discussion regarding the usage of ontologies in software engineering and technology. Most of them consider how general the represented concepts are and the scope of the application of the ontologies—general, domain-, task- or application-specific. One of the classifications, given by [84] in a discussion regarding desirable and

required features for ontology languages, considers the complexity of the relationships that can be depicted in the domain of interest. This classification, referred to as ‘richness of the internal structure’, and the classification in [58], referred to as ‘subject of conceptualization’, are used as a foundation for the two-dimensional classification developed in [48]. Depending on the ‘richness of the internal structure’, i.e., the knowledge representation capabilities of an ontology, Gómez-Pérez et al. [48] defined eight categories of ontologies ranging from informally specified ontologies to ontologies precisely specified by formal languages. These eight categories can be further compacted to the four presented in [53, 128] and listed below. The classification developed by Lambrix [72] took into account the information represented by the components and arrives at a similar classification:

- glossaries and data dictionaries contain concepts with or without their definitions in a natural language;
- thesauri and taxonomies introduce, together with the concepts and their definitions, synonyms and relations such as *narrower* and *broader*;
- ontologies represented by metadata, XML schemas, data models, these models additionally provide properties and value restrictions;
- ontologies represented by logical languages. The ontologies represented by formal languages hold the most expressive knowledge representation capabilities. Their well-defined syntax and semantics allows for reasoning services such as consistency checking and classification.

The classification above encompasses the whole range of ontologies regarding their knowledge representation capabilities—from the so called lightweight to the heavyweight ontologies. The advantage of the former group is their simplicity at the price of reduced expressivity and high ambiguity. The advantage of the ontologies in the latter group is their powerful knowledge representation capabilities and inference mechanism at the price of complex development.

2.1.2 Applications

The ontologies have a wide range of applications to:

- provide mutual understanding of a domain and facilitate the communication between different agents in it [15] by enabling knowledge sharing and reuse [72];
- serve as a repository of information [72, 128];
- provide a query model for information sources explicitly structuring the domain knowledge [94, 128, 131];
- enable data integration of heterogeneous information sources [73, 94, 131].

Ontologies are a key technology for the Semantic Web and are intensively employed in other areas as well:

- Artificial Intelligence—knowledge representation and reasoning;
- Software Engineering—ontologies are used throughout all phases of the Software Engineering life cycle—starting from requirements specification, implementation, testing, deployment, maintenance and reuse [47, 54];
- Computer Security—for modelling software vulnerabilities, threats and counteractions [60] and in security requirements engineering [119];
- Bioinformatics and Systems Biology—many ontologies have already been developed in this domain—UMLS Metathesaurus¹, GeneOntology², NCI Thesaurus³, AMA⁴, FMA⁵, SNOMED-CT⁶ to name a few; they are used for specification, ontology-based search, data integration and exchange as discussed in [72, 81];
- E-commerce—such applications are discussed by Ding et al. [30]; one example is the GoodRelations ontology [59].

2.2 Ontology Alignment

Ontologies are developed by different people and organizations to fulfill different goals and reflect the views and needs of their developers. Thus, there may exist several ontologies modeling the same field which could differ in conceptual modeling, granularity level, vocabulary and domain coverage. Consequently, in order to enable data integration, knowledge sharing and reuse we need to use more than one ontology and to know how the components from the different ontologies are related to each other.

Finding these relationships is the subject of investigation of the *ontology alignment* field. A set of relationships between the components of two different ontologies is called an *alignment*. Each relation in the set is called a *mapping* or a *correspondence*. In Paper III we use the notion of *mapped concepts* which are the concepts that participate in mappings. Each (mapped) concept can participate in multiple mappings and alignments, i.e., a mapping can be of one of the following cardinalities: 1:1, 1:N (N:1) and N:M. While mappings can represent various relationships, in our work we consider *equivalence and subsumption mappings* between concepts. The equivalence mappings connect two concepts which represent the same set of entities. The subsumption mappings are relations between two concepts, where one of the concepts represents a set of entities that is a subset of the set of entities represented by the other concept. A set of ontologies connected through their alignments form a network—an *ontology network*.

¹https://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/

²<http://www.geneontology.org/>

³<https://ncit.nci.nih.gov/ncitbrowser/>

⁴http://www.informatics.jax.org/vocab/gxd/ma_ontology

⁵<http://si.washington.edu/projects/fma>

⁶<http://www.snomed.org/snomed-ct>

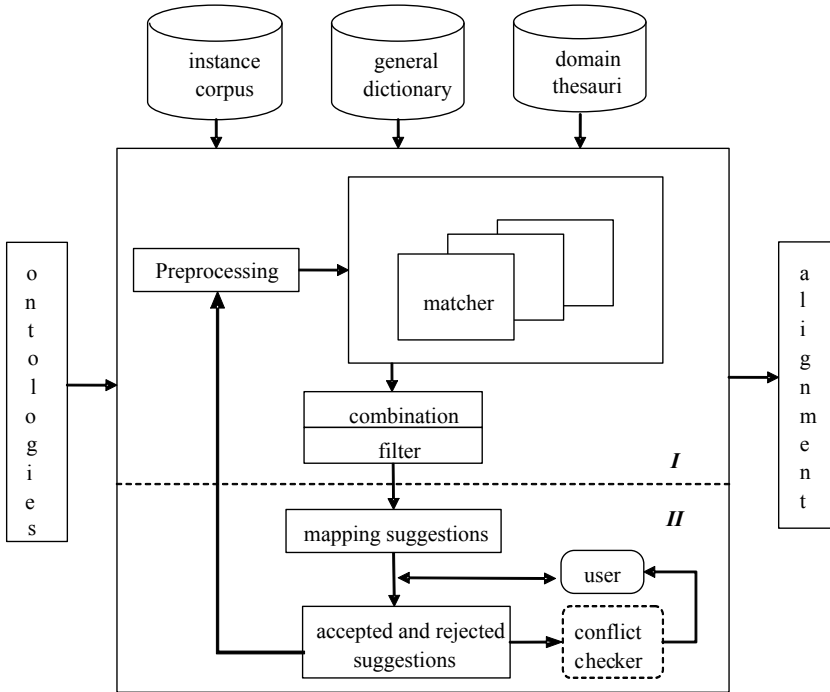


Figure 2.1: A general alignment framework [77].

2.2.1 Ontology Alignment Framework

Ontology alignment is an active research area where many (semi-)automatic tools have already been developed to address the challenges of discovering the relationships between two ontologies. The alignment process in most tools conforms to the general semi-automatic ontology alignment framework presented by Lambrix and Liu in [77] and depicted on figure 2.1. The input for a tool contains two ontologies and the output is an alignment. The alignment process presented in the framework goes through two phases. In phase I the tool generates possible mappings that are presented to the user for a manual validation in phase II. Phase I usually includes three steps:

Preprocessing step includes preliminary data processing, for instance, partitioning of the input ontologies or removing modifiers, such as definite and indefinite noun modifiers. The partitioning algorithms could employ partial alignments as shown in [77].

Running matchers to compute similarity values between pairs of components from the different ontologies. The similarity values represent an estimate that two components are connected. The matchers employ various strategies as described in [80] and listed below:

- linguistic strategies explore the linguistic similarity of the labels of the components. For instance, the labels are represented as sets of consecutive characters and then the similarity value between these components is calculated based on the number of characters in the intersection of these sets. Another strategy counts the number of insertions, deletions and modifications needed in order to make one of the labels identical to the other;
- structure-based strategies rely heavily on the structure of the ontologies. For example, it is more likely that two concepts are similar if there are established mappings between their siblings;
- constraint-based strategies consider different constraints, e.g., cardinalities, encoded in the ontologies. They are usually used to provide supplementary information, not as primary matchers;
- instance-based strategies assign similarity values based on the shared entities between the concepts in the different ontologies. The instances can be acquired from curated scientific resources (for instance, PubMed⁷ in the life sciences);
- strategies based on auxiliary sources use domain knowledge available from external sources, such as WordNet [93] and UMLS⁸, to find additional information for the concepts, e.g., synonyms, and the relationships between them.

Combining and filtering the similarity values obtained from the different matchers—most often the similarity values are combined using a weighted-sum approach in which each matcher is given a weight and the final similarity value is the weighted sum of the similarity values divided by the sum of the weights of the matchers. Another approach uses the maximal similarity value obtained from the matchers.

Those pairs with similarity values equal to or higher than a given threshold are retained in the final alignment or are presented to the user during phase II for manual validation. The latter are called *candidate mappings* or *mapping suggestions*; Some approaches, e.g., [23], use two thresholds—those pairs equal to or above the higher threshold are directly retained as (candidate) mappings while those between the two thresholds are further filtered with respect to the structure of the ontology and the pairs with similarity values above the higher threshold.

In phase II the candidate mappings are presented for validation to a user who can accept or reject them. Those validated as correct become part of the final alignment. Both the accepted and the rejected candidate mappings could be further used in the alignment process to avoid unnecessary computations and validations. A conflict checker may be used to detect possible conflicts.

⁷www.ncbi.nlm.nih.gov/pubmed/

⁸http://www.nlm.nih.gov/research/umls/about_umls.html

User Involvement

Phase II is becoming more important since after many years of experience the improvements of the fully-automatic approaches have not lead to comparable improvements in the alignments quality [49, 102]. While users may make mistakes, work by Jiménez-Ruiz et al. [68] has shown that manual validation improved the alignments' quality in the presented examples up to an error rate of 20%. Similarly, the evaluation campaigns in the OAEI Interactive track⁹ have shown that user interaction is still beneficial even if users make some mistakes. The exact threshold, however, would depend on the specific case and strategies implemented in the tool to employ the user's feedback.

2.2.2 Ontology Alignment Evaluation

The increased interest in the topic of ontology alignment has led to the organization of annual events, such as the Ontology Matching workshop¹⁰ and the Ontology Alignment Evaluation Initiative, which provide discussion forums for developers and a platform for an annual evaluation of their tools. The OAEI consists of several tracks where the alignments computed by the tools (denoted below with *Align*) are evaluated by comparing them to a gold standard or a reference alignment (denoted below with *RA*) and computing general information retrieval measures such as *precision*, *recall* and *f-measure*. The precision measure reflects the ratio between the correct retrieved mappings and all mappings in the newly created alignment. The recall measure reflects the ratio between the correct retrieved mappings and all correct mappings (which are known to be correct according to, for instance, a reference alignment).

$$precision = \frac{|Align \cap RA|}{|Align|}$$

$$recall = \frac{|Align \cap RA|}{|RA|}$$

The f-measure combines precision and recall possibly with different weights controlled by the α parameter. When $\alpha = 1$ precision and recall have an equal weight.

$$f_measure_{\alpha} = (1 + \alpha) \frac{precision \cdot recall}{\alpha \cdot precision + recall}$$

Variations of precision and recall have been discussed in [39].

⁹<http://sws.ifl.uio.no/oaei/interactive/>

¹⁰<http://ontologymatching.org/>

2.3 Ontology Debugging

Developing ontologies and alignments is not a trivial task. As ontologies grow in size and complexity, the intended and unintended entailments become difficult to follow. Ontologies are usually developed by domain experts who often are not expert in knowledge representation and may not have experience with the capabilities of the knowledge representation languages. The same issues apply to alignments development. Concept discrepancies between the different ontologies, for instance, using one term for different real-world entities, are also sources of defects during the alignment. As a consequence, the ontologies, alignments and integrated ontology network may be incorrect, incomplete or inconsistent. Using them in semantically-enabled applications may lead to entailment of incorrect conclusions or valid conclusions may be missed.

To achieve highly reliable results from the semantically-enabled applications, it is necessary to have both high quality ontologies and high quality alignments. Debugging of the ontologies and alignments is a key step towards eliminating defects in them, which is essential for obtaining high-quality results in the semantically-enabled applications. The *ontology debugging* area deals with discovering and resolving defects in the structure of the ontologies and their alignments.

The defects differ [69] in nature and, consequently, in the complexity of their detection and repair.

- syntactic defects, such as incorrect format or missing tags, are easy to find and resolve using parsers;
- semantic defects have their origin in unintended inferences
 - unsatisfiable concepts are concepts that cannot have any instances, for instance a concept defined as the intersection of two disjoint concepts at the same time;
 - incoherent ontologies are ontologies that contain unsatisfiable concepts;
 - inconsistent ontologies contain contradictions, for example, an instance that belongs to two disjoint concepts at the same time.

The semantic defects can be found using reasoners, which are software programs that are able to derive logical consequences from a given set of asserted axioms, e.g., Jena¹¹, FaCT++¹², HermiT¹³.

- modeling defects are caused by modeling errors when encoding domain knowledge. Examples for such are missing and wrong relations. In order to detect and repair them domain knowledge is needed. The work pre-

¹¹jena.apache.org/

¹²<http://owl.man.ac.uk/factplusplus/>

¹³hermit-reasoner.com

sented in Paper III deals with missing and wrong subsumption relations and mappings.

2.3.1 Definitions

This subsection presents several extended definitions that are used in Paper III. This work focuses on taxonomies which are widely used since subsumption relationships are common in many domains.

Ontologies and Ontology Networks

The taxonomies consist of named concepts and subsumption (is-a) relations between the concepts. The following definition applies.

Definition 1 A *taxonomy* \mathcal{O} is represented by a tuple $(\mathcal{C}, \mathcal{I})$ where \mathcal{C} is its set of named concepts and $\mathcal{I} \subseteq \mathcal{C} \times \mathcal{C}$ is a set of asserted is-a relations, representing the is-a structure of the ontology.

The ontologies are connected into a network through alignments. We currently consider *equivalence* mappings (\equiv) and is-a mappings (*subsumed-by* (\rightarrow) and *subsumes* (\leftarrow)).

Definition 2 An *alignment* between ontologies \mathcal{O}_i and \mathcal{O}_j is represented by a set \mathcal{M}_{ij} of pairs representing the mappings, such that for concepts $c_i \in \mathcal{O}_i$ and $c_j \in \mathcal{O}_j$: $c_i \rightarrow c_j$ is represented by (c_i, c_j) ; $c_i \leftarrow c_j$ is represented by (c_j, c_i) ; and $c_i \equiv c_j$ is represented by both (c_i, c_j) and (c_j, c_i) .¹⁴

Definition 3 A *taxonomy network* \mathcal{N} is a tuple (\mathbb{O}, \mathbb{M}) with $\mathbb{O} = \{\mathcal{O}_k\}_{k=1}^n$ the set of the ontologies in the network and $\mathbb{M} = \{\mathcal{M}_{ij}\}_{i,j=1;i < j}^n$ the set of representations for the alignments between these ontologies.

Without loss of generality, we assume that the sets of named concepts for the different ontologies in the network are disjoint.

A significant part of our approach relies on knowledge *intrinsic* to the network, i.e., knowledge logically derivable from the network. The domain knowledge of an ontology network is represented by its induced ontology.

Definition 4 Let $\mathcal{N} = (\mathbb{O}, \mathbb{M})$ be an ontology network, with $\mathbb{O} = \{\mathcal{O}_k\}_{k=1}^n$, $\mathbb{M} = \{\mathcal{M}_{ij}\}_{i,j=1;i < j}^n$. Let $\mathcal{O}_k = (\mathcal{C}_k, \mathcal{I}_k)$. Then the *induced ontology* for network \mathcal{N} is the ontology $\mathcal{O}_N = (\mathcal{C}_N, \mathcal{I}_N)$ with $\mathcal{C}_N = \cup_{k=1}^n \mathcal{C}_k$ and $\mathcal{I}_N = \cup_{k=1}^n \mathcal{I}_k \cup_{i,j=1;i < j}^n \mathcal{M}_{ij}$.

¹⁴Observe that for every \mathcal{M}_{ij} there is a corresponding \mathcal{M}_{ji} such that $\mathcal{M}_{ij} = \mathcal{M}_{ji}$. Therefore, in Paper III we only consider the \mathcal{M}_{ij} where $i < j$.

Knowledge Bases

In the algorithms we use the notion of knowledge base (KB). The notion that we define here is a restricted¹⁵ variant of the notion as defined in description logics [6].

Definition 5 *Let \mathcal{C} be a set of named concepts. A **knowledge base** is then a set of axioms of the form $A \rightarrow B$ with $A \in \mathcal{C}$ and $B \in \mathcal{C}$. A model of the knowledge base satisfies all axioms of the knowledge base.*

In the algorithms we initialize KBs with an ontology. This means that for ontology $\mathcal{O} = (\mathcal{C}, \mathcal{I})$ we create a KB such that $(A, B) \in \mathcal{I}$ iff $A \rightarrow B$ is an axiom in the KB.

For the KBs, we assume that they are able to do deductive logical inference. Furthermore, we need the following reasoning services. For a given statement the KB should be able to answer whether the statement is entailed by the KB¹⁶. If a statement is entailed by the KB, it should be able to return the derivation paths (explanations) for that statement. The derivation paths, also called *justifications*, are used to show how a given statement has been entailed. For a given named concept, the KB should return the super-concepts and the sub-concepts.

The KBs can be implemented in several ways. For instance, any description logic system could be used. In our setting, where we deal with taxonomies, we have used an efficient graph-based implementation. We have represented the ontologies using graphs where the nodes are concepts and the directed edges represent the is-a relations. The entailment of statements of the form $a \rightarrow b$ can be checked by transitively following edges starting at a . If b is reached, then the statement is entailed, otherwise not. If $a \rightarrow b$ is entailed, then the *derivation paths* are all the different paths obtained by following directed edges that start at a and end at b . The *super-concepts* of a are all the concepts that can be reached by following directed edges starting at a . The *sub-concepts* of a are all the concepts for which there is a path of directed edges starting at the concept and ending in a .

¹⁵We only use concept names and no roles. The axioms in the TBox are of the form $A \dot{\subseteq} B$ or $A \dot{=} C$, and the ABox is empty.

¹⁶In our setting, entailment by ontology can be reformulated as entailment by KB.



3

Summary of Papers

This chapter summarizes the papers included in this thesis. Papers I-IV are connected to Research Question 1: *What features should an ontology alignment system provide in order to efficiently support users during the process of ontology alignment?* and dedicated to issues in relation to user involvement in the ontology alignment process. Papers I and II [65, 32] focus on arguably the most cognitively intensive part of the alignment process—the manual validation of the candidate mappings, Paper III [64] addresses the issue of debugging the ontologies and their alignments and Paper IV [62] investigates the benefits from the latest technological advances to the alignment process. Papers V and VI [33, 63] are connected to Research Question 2: *How to efficiently support users during the evaluation of ontology alignments?* and showcase the need of additional means, beyond precision and recall measures, for ontology alignment evaluation. Paper V summarizes our experiences from the organization of the OAEI Anatomy track and serves as a basis for the identification of common tasks during the assessment of several alignments. These tasks are further supported by our tool Alignment Cubes—described in Paper VI—which provides an interactive visual environment for comparative exploration and evaluation of multiple alignments.

Paper I: Requirements for and Evaluation of User Support for Large-Scale Ontology Alignment

Currently one of the challenges for the ontology alignment community is the user involvement in the alignment process. At the same time, the focus of the community has shifted towards large-scale matching which introduces an additional dimension to this issue. This paper aims to provide a set of requirements that foster the user involvement for large-scale ontology alignment tasks. Further, we present and discuss the results of a literature study for 7 ontology alignments systems as well as a heuristic evaluation and an observational user study for 3 ontology alignment systems to reveal the coverage of the requirements in the systems and the support for the requirements in the user interfaces.

Paper II: User Validation in Ontology Alignment

User validation is one of the challenges facing the ontology alignment community, as there are limits to the quality of automated alignment algorithms. In this paper we present a broad study on user validation of ontology alignments that encompasses three distinct but interrelated aspects: the profile of the user, the services of the alignment system, and its user interface. We discuss key issues pertaining to the alignment validation process under each of these aspects, and provide an overview of how current systems address them. Finally, we use experiments from the Interactive Matching track of the Ontology Alignment Evaluation Initiative (OAEI) 2015 to assess the impact of errors in alignment validation, and how systems cope with them as function of their services.

Paper III: A Unified Approach for Aligning Taxonomies and Debugging Taxonomies and Their Alignments

In this paper we address one of the desirable features for supporting user involvement in the ontology alignment process identified in Paper I—debugging of alignments—by proposing an integrated framework for aligning taxonomies and debugging taxonomies and their alignments and implementing a prototype system. By employing the knowledge inferred from the network (which consists of connected taxonomies), our approach allows for detecting and repairing defects in the alignments while finding and resolving defects in the taxonomies at the same time. The debugging process consists of three phases—detecting potentially missing relations/mappings, their validation and consecutively repairing by a user.

In our approach, if a subsumption relation can be derived in the network but not in a taxonomy itself a potential modeling defect¹ in the taxonomy is detected. An expert user needs to validate the defect which could be caused by an omitted relation in the taxonomy (which needs to be added) or an incorrect relation in the derivation path (consisting of taxonomies and alignments). In order to repair an incorrect relation existing relation(s) in the taxonomies or alignments need to be removed. To repair an omitted relation our framework provides different repairing actions (to a user to choose from) which may lead to adding new knowledge not previously derivable in the network. This new knowledge is in turn used to restart the process and potentially detect new defects. The repaired structure of the ontologies and alignments provides better input for (structure-based) alignment algorithms and (partial-alignment-based) preprocessing and filtering strategies.

As working with several taxonomies and alignments at the same time is likely infeasible, we provide user interface components for validating and repairing defects separately in the taxonomies and alignments. The user can switch between the different phases and components at any time which allows for an iterative and modular approach, where, for instance, some parts of the ontologies can be fully aligned and debugged before proceeding to other parts.

Paper IV: Applications of Large Displays: Advancing User Support in Large Scale Ontology Alignment

Producing alignments of the highest quality requires ‘humans in the loop’, however, user involvement is currently one of the challenges for the ontology alignment community. Ontology alignment is a cognitively intensive task and could be efficiently supported by user interfaces encompassing well-designed visualizations and interaction techniques. This work investigates the application of large, high-resolution displays to improve users’ cognitive support and identifies several promising directions for their application—improving ontologies’ and alignments’ navigation, supporting users’ thinking process and collaboration

Paper V: Experiences from the Anatomy track in the Ontology Alignment Evaluation Initiative

In this paper we focus on the Anatomy track of the Ontology Alignment Evaluation Initiative in the years 2007-2016 and the Anatomy part of the Interactive track in 2015-2016. We describe the data set and the changes it went through during the years as well as the challenges it poses for ontology alignment systems. Further, we give an overview of all systems that participated

¹We consider missing and wrong subsumptions.

in the track and the techniques they have used. We discuss the performance results of the systems and summarize the general trends.

Paper VI: Alignment Cubes: Towards Interactive Visual Exploration and Evaluation of Multiple Ontology Alignments

Ontology alignment is an area of active research where many algorithms and approaches are being developed. Their performance is usually evaluated by comparing the produced alignments to a reference alignment in terms of precision, recall and F-measure. These measures, however, only provide an overall assessment of the quality of the alignments, but do not reveal differences and commonalities between alignments at a finer-grained level such as, e.g., regions or individual mappings. Furthermore, reference alignments are often unavailable, which makes the comparative exploration of alignments at different levels of granularity even more important. Making such comparisons efficient calls for a ‘human-in-the-loop’ approach, best supported through interactive visual representations of alignments. Our approach extends a recent tool, Matrix Cubes, used for visualizing dense dynamic networks. We first identify use cases for ontology alignment evaluation that can benefit from interactive visualization, and then detail how our *Alignment Cubes* support interactive exploration of multiple ontology alignments. We demonstrate the usefulness of Alignment Cubes by describing visual exploration scenarios, showing how Alignment Cubes support common tasks identified in the use cases.



4

Related Work

This chapter provides related work in two main directions. In Section 4.1 we first discuss our work in the broad context of ontology alignment. The following Sections 4.2, 4.3 and 4.4 provide more details for each related area we outline in Section 4.1. We then touch upon (Section 4.5) selected works that fall in the broad scope of the Information Visualization field.

4.1 Ontology Alignment

Ontology alignment is an area of active research with a large community of researchers and practitioners and annual events attracting a significant number of participants. The progress in the field has been accelerated by events such as the Ontology Matching workshop and the Ontology Alignment Evaluation Initiative (OAEI) which have provided a discussion forum for developers and a platform for an annual evaluation of their tools. Both of them have now run for more than 10 years and have contributed to the advancement of the field. A recent comprehensive literature review of the field is presented in [99]. After years of research challenging and promising directions for future development have been outlined [118] and restated five years later [117]. Experiences and lessons learnt from the early OAEI campaigns have been summarized in [38] and Paper V, which focuses on the last 10 years of the OAEI Anatomy track, is in the same spirit.

The challenges identified in [117], although connected, can be roughly discussed in two general directions—challenges related to matching techniques

and approaches and their evaluation on the one side and to fostering user participation during the process on the other. Some aspects of these challenges have also appeared in a recent survey [99] which drew participants from the alignment community. Since many approaches have already been developed their evaluation is essential to reveal their strengths, weaknesses and directions for improvement. Existing evaluations, such as the OAEI, rely on manually curated alignments which require significant effort to develop and maintain. The first challenge—*large scale matching evaluation*—considers the demand for semi-automatic approaches for development of high-quality reference alignments to address the increasing size of the matching task. Another aspect is to introduce more accurate (in addition to precision and recall) as well as application specific evaluation measures. Section 4.3 further discusses this issue. As none of the developed algorithms is considerably better than the rest for a specific application, combining them usually provides more reliable results. Particular combinations could be tailored to application areas or datasets’ features, or both, which demands strategies for *matcher selection, combination and tuning*. Some matchers utilize *background knowledge* during the alignment process, for instance, curated resources such as WordNet and UMLS. Uncurated resources may be utilized as well. While including auxiliary resources often improves the number of relevant mappings, it also leads to decreasing the precision (i.e. more incorrect mappings). With two opposite tendencies in place—increasing the size of the matching task (demanding scalable techniques) and broadening the range of the applications performing it (including devices with limited resources)—the *efficiency of matching techniques* in terms of both computational time and memory consumption is becoming more and more important.

The other half of the challenges are related to supporting user participation in the process. So far most of the approaches have been focused on the first direction with little attention to fostering the *user involvement* in the process. However, advancement of the fully-automatic approaches and algorithms have not led to comparable improvements in the alignments’ quality [49, 102] and semi-automatic approaches are needed. Section 4.2 focuses on this issue. One important aspect is providing *explanations of matching results* to support different user groups in understanding of the results of the alignment algorithms and to thus support semi-automatic alignment. The manual curation of large alignments is a demanding task for a single user. It can be relaxed and encouraged through *social and collaborative matching* by involving several users who can discuss together problematic mappings. Such collaborative effort will demand metadata standards and proper *alignment management* frameworks providing *infrastructure and support* during all phases of the process—storage, version control, etc.

While our works are mostly in the direction of fostering user participation and alignment evaluation they also touch upon most of the challenges. Our visual evaluation environment—Alignment Cubes presented in Paper VI—

covers challenges from both directions—it extends the available means for evaluation beyond the availability of reference alignments and aggregated measures and can be used for simulating different combination strategies and parameters. Alignment Cubes promote user engagement and collaboration by facilitating alignments analysis and provide means for presenting explanations. As mentioned above, our most recent work in Paper V provides insights of the last ten editions of the OAEI Anatomy track evaluation campaign and surveys the matching algorithms used by the participating systems and their performance. The work in Paper I directly addresses the user involvement issues by identifying desirable back- and front-end features of alignment systems to promote user engagement. In Paper II we further narrow down our attention to the manual validation of candidate mappings computed by a tool, arguably the most cognitively intensive part of the alignment process. User engagement and collaboration is also the focus of Paper IV where we propose a way to benefit of recent technological developments. As the development of alignments is an error-prone process, our work in Paper III proposes methods and a tool to support discovering, explaining and resolving defects in the alignments.

4.2 User Involvement in Ontology Alignment

Many tools and algorithms for ontology alignment have already been developed by the community. In most cases ontology alignment systems apply fully automated approaches where an alignment is generated without any human intervention. However, after many years of experience in the OAEI, it has been observed by the community that there are limits to the performance (in terms of precision and recall of the alignments) of the automated systems since advancing the algorithms did not lead to comparable improvements in the alignments' quality [49, 102]. Higher quality alignments will be achieved with the help of better user interfaces exhibiting good scalability features, rather than more accurate matchers [16] where fully automated approaches should only be the first step in the alignments generation [38]. According to [108] user interaction is essential (in the context of large ontologies) for configuring the matching process, incremental matching and providing feedback to the system regarding the generated candidate mappings.

The demand for user involvement has been recognized by the alignment community and resulted in the introduction of the OAEI Interactive track¹ in 2013. Simulating user input by an (all-knowing) oracle led to the improvement of the alignments' quality in comparison to the fully automated approaches as shown during the OAEI Interactive track² evaluation campaigns and in [68].

¹The track evaluates the performance of interactive approaches by simulating user interaction. Currently, it does not evaluate the user experience or the user interfaces of the systems.

²<http://sws.ifi.uio.no/oaei/interactive/>

Further, these works have also demonstrated that even if users make some mistakes the interaction still leads to improved quality of the alignments.

Evaluation of interactive ontology alignment tools is considered by Paulheim et al. [102]. The authors discuss building a cost model for the different user interactions to assess the effort needed by a domain expert. Inspired by the evaluation of active learning tools, the authors also propose to plot the alignment after each user interaction and compare the tools according to the area under the learning curve. Work by Duchateau et al. [34] define a measure for assessing users' post matching effort on the basis of the number of candidate mappings and manual effort needed for the discovering of missed mappings.

Ontology alignment is a complex and challenging task imposing significant cognitive demands on the users [42, 45]. Users are most often involved in selecting matchers and configuring combination strategies, validating automatically generated mappings, etc. From the user perspective, manual validation of candidate mappings is arguably the most cognitively intensive part of the process as it usually involves exploration of both (unfamiliar) ontologies, in order to become familiar with them, and their formal representations in order to understand their modelers' view of the domain. Further, the user needs to explore the mappings computed by the tool's algorithms in order to determine their correctness and create mappings missed by the system. It is an inherently error-prone process due to different levels of users' domain and knowledge representation expertise, experience, human biases and misinterpretation.

Graphical interfaces are essential to support users during the manual validation of candidate mappings but also during the entire ontology alignment process. However, only about one third of the systems participating in the OAEI campaigns have any, while many systems do not provide such at all [117]. Tools' interfaces often resulted from the need to provide user input to matchers [42], and our works in Papers I and II reveal functionality and usability issues with them. They are rarely theoretically grounded and not based on advances in cognitive and behavioral sciences (except [45]).

As ontology alignment involves working simultaneously with at least two ontologies and numerous mappings interactive user interfaces have the potential to significantly facilitate their presentation and navigation and to reduce the cognitive load during the alignment process. The demand for comprehensive user interfaces is even more pressing given the trend towards growing size and complexity of the ontologies and the alignments.

Several works have explored these issues in depth [41, 42, 44, 45, 49, 74, 83]. Early on, Lambrix and Edberg [74] have conducted a comparative evaluation of two state-of-the-art (at that time) alignment systems—the PROMPT plugin for Protégé-2000 [98] and Chimaera [89]. The evaluation consisted of two parts—a literature study using predefined criteria not directly related to the user interfaces and a user interface evaluation with 8 novice users with

different backgrounds. The REAL methodology was selected for the interface evaluation. It consists of a questionnaire with questions grouped in four categories—relevance, efficiency, attitude and learnability. Time to complete a task and task success (in terms of precision and recall) were measured as well.

Work by Falconer et al. [44, 45] is probably one of the first accounts of cognitive support in the field of ontology alignment. Based on a small user study and review of relevant literature in cognition and decision making, the authors have developed a theoretical framework for cognitive support and a prototype implementation following the framework’s principles. The framework consists of four dimensions covering the interaction between a user and a semi-automatic alignment tool during the validation of candidate mappings—computing an initial alignment (Analysis and Generation), its presentation to the user (Representation Dimension), user analysis of the alignment (Analysis and Decision Making) and the interaction with the tool’s interface to change the initial alignment (Interaction Dimension).

Further, works by Euzenat et al. [41] and Granitzer et al. [49] collected and presented requirements for interactive ontology alignment systems in both individual and collaborative settings. These can be related to the four dimensions described above and include navigation and exploration capabilities at different levels of detail following the familiar visual information-seeking mantra [114]—overview, zoom, filter, detail on demand; explanation of matching results and highlighting promising regions worth user attention; interactions to modify the current alignment; assigning parts of the alignment to different users in a collaborative setting and providing communication means, such as annotating, commenting and voting, to facilitate collaboration.

Another work by Lanzenberger et al. [83] viewed ontology alignment in the light of exploratory data analysis and distinguishes three types of visualizations according to their purpose: exploration (to obtain insights and derive hypotheses), analysis (to confirm or reject hypotheses) and presentation (as means to communicate the findings). Based on these reflections the authors have based their prototype implementation on the following guidelines for exploratory data analysis—‘maximize insight into a data set’, ‘uncover underlying structure’, ‘extract important variables’, ‘detect outliers and anomalies’, ‘test underlying assumptions’. These guidelines can be considered higher-level tasks compared to those described in [41, 49].

The papers in this dissertation have contributed to the body of existing work in the following manner. In Paper I we have identified desirable back-end and front-end features for ontology alignment systems which contribute to fostering user engagement in (large-scale) ontology alignment problems. We have used the user interface features from [41, 44, 45, 49] to develop three categories according to the function they support—inspection, manipulation and explanation of alignments. The back-end features further include various complementary functionalities to assist the user in managing larger and more

complex ontologies and alignments (which may further pose additional visualization and interface requirements). They contribute to the development of responsive interfaces (in terms of reasonable response time [68]) and a complete infrastructure that supports the users during large-scale alignment tasks. We have further conducted a literature review to reveal how these features are supported by existing systems and two user interface evaluations of three state-of-the-art tools.

In Paper II we have deepened our understanding of the process of manual validation of candidate mappings and have analyzed it from two perspectives—user interface perspective and system services which contribute to reducing user intervention. We have further conducted another review narrowing down its scope to works focusing on these two perspectives. We have also touched upon the issue of the users’ (domain or knowledge representation) level of expertise and experience with the respective alignment system. Lastly, this work has also provided an overview of the results of the OAEI Interactive track in 2015³ to illustrate the impact of user error on the alignments’ quality and reveal differences between the implemented strategies to deal with erroneous input.

Furthermore, we have taken advantage of the latest technological developments and in Paper IV we have gone beyond regular visualization and interaction settings, e.g., desktop and mouse. There, we have investigated the application of large, high-resolution displays in the field of ontology alignment. Drawing from the characteristics of the matching task we have identified three promising directions for future research—improving the navigation in the ontologies and their alignments, space to support users’ thinking process and collaboration.

In Paper VI we have viewed the evaluation of ontology alignments as an exploratory task. We have identified several use-cases that would benefit from comparative assessment of several alignments at different level of detail and have discussed their shared analytical tasks and features that would benefit from visual support. To address them we have proposed an interactive visual environment for the simultaneous comparative exploration and evaluation of multiple alignments at different levels of granularity. Our work have provided new means for ontology alignment evaluation by enabling visual interactive assessment of ontology alignments.

Paper III has addressed one of the desirable features identified in Paper I—debugging of alignments—in more detail. As the development of alignments is an error-prone process, we have developed a visual tool for discovering, explaining and resolving defects in taxonomies and taxonomy networks. We focus on modeling defects which demand domain knowledge to discover and resolve. We take advantage of the knowledge inferred from the ontology net-

³<http://oaei.ontologymatching.org/2015/results/interactive/>

work to support the detection of modeling defects without the need of external domain knowledge.

4.3 Ontology Alignment Evaluation

We first discuss measures for ontology alignment evaluation in Subsection 4.3.1 and then frameworks and tools devoted to the evaluation task in Subsection 4.3.2.

4.3.1 Measures

Euzenat and Shvaiko [39] present a broad discussion on the evaluation of ontology matching systems. They first outline five general principles on which the evaluation should be based on: *systematic procedure* to allow for non-ambiguous and reproducible results, *continuity* to observe the progress in the field, clear evaluation rules and non-biased dataset (*quality and equity*), easily accessible datasets and results (*dissemination*) and *intelligibility* which concerns the analysis and explanation of the computed alignments. Our work in Papers V and VI addresses some of these principles. In Paper V we summarize and discuss the last 10 editions of the OAEI Anatomy track thus contributing to the continuity and dissemination principles. In Paper VI we address the last principle—intelligibility—by enabling interactive visual analysis and exploration of multiple alignments at different levels of detail in two of the three evaluation types mentioned further by the authors: *competence benchmarks* which allow for assessing tool’s performance on a particular task or input and *comparative evaluation* which considers comparing the performance of multiple tools on a common task.

After outlining these general evaluation principles, different evaluation measures and their application have been discussed and exemplified. The measures are divided into *compliance measures* which assess to what extent the evaluated alignment conforms to a reference alignment, *performance measures* which take into account the consumption of computational resources and *user-related measures* which aim to assess perceived user satisfaction and required interactions with the tool.

The compliance related measures include both classical information retrieval measures, such as precision, recall and F-measure, their complements (noise and silence) and overall, as well as measures which account for specific features of the alignments, such as weighted, relaxed and semantic precision and recall. *Semantic precision and recall* take into account if a mapping is entailed in the reference alignment. *Weighted precision and recall* take into account confidence values if provided in the alignment and may be used to suggest which mappings need human inspection, i.e. mappings with lower confidence may need manual validation while those with higher confidence can be directly added to the final alignment. *Relaxed precision and recall*,

discussed in detail in [36], take into account “near misses” and could be used to compute the user effort needed to correct the evaluated alignment. In the *effort-based measure* various weights can be assigned according to the type and difficulty to perform corrections in an alignment editor. The authors argue that the last measure is better than the *overall measure* which simply counts the edit distance between the evaluated and reference alignments.

Duchateau et al. [34] propose another compliance measure—*schema proximity* which assesses the quality of an integrated schema in comparison to a reference schema. This measure is an aggregation of three other measures—*completeness* which assesses the ratio of elements of the integrated schema present in the reference schema, *minimality* which measures the amount of additional elements in the integrated schema and *structurality* which assesses the structural similarity between the two schemas.

Further, Euzenat and Shvaiko [39] discuss performance measures. They consider the efficiency of algorithms in terms of speed, memory and scalability. These, together with the other measures listed above, may impact the perceived user satisfaction. User satisfaction is a highly subjective measure which may vary in a number of factors and their weights depending on the user and particular alignment task. Another user-related measure counts the number of times the tool presents a candidate mapping for user validation during the alignment process. It has been implemented in the OAEI Interactive track since its beginning in 2013. When a reference alignment is available it can be used to simulate an oracle user who uses the reference alignment to validate candidate mappings provided by the tool.

Being variations of precision and recall, the measures outlined above are aggregated measures expressing the “goodness” of alignment with single numbers, which may not always be possible to compute (i.e. confidence is rarely provided and a cost model needs to be built to compute the effort-based measures) and, importantly, they rely on the existence of reference alignments⁴. They summarize the alignment quality but do not allow for fine-grained differentiation between the approaches and algorithms to the level of single mappings. This is, however, highly demanded in order to reveal differences between approaches as tools often compute different mappings as showcased by Jiménez-Ruiz et al [66] and the evaluation in the OAEI Anatomy track in 2007⁵. In comparison, our work in Paper VI provides a different evaluation approach and means for evaluation even when no reference alignments are available. It allows flexible visual exploration of several alignments at gradually changing level of detail ranging from an overview of several alignments to the level of single mappings in an alignment. This allows for fine-grained

⁴When reference alignments are not available, the authors propose to create partial reference alignments by applying various sampling techniques. The samples are then evaluated by expert users and used to compute relative precision and recall. Thus, this method is also based on the existence of an alignment curated in advance.

⁵<http://oaei.ontologymatching.org/2007/results/anatomy/>

differentiation between the approaches and easily studying their impact in various (corner) cases. The measures above can be used to complement our work by highlighting promising parts of the alignments worth user attention.

4.3.2 Tools & Frameworks

Two alignment systems provide means for comparative evaluation of several alignments as means to facilitate alignment development. AgreementMaker [27] supports visualization of several alignments together by showing juxtaposed matrices computed by (a combination of) different matchers; an additional matrix can highlight disagreement in the computed mappings. If a reference alignment is available it can be overlaid on top of each matchers' matrix. Analytical tasks are further supported in its recent extensions [5, 85] by implementing parallel coordinate and pie chart views. The juxtaposed matrices resemble the most detailed views in our tool Alignment Cubes presented in Paper VI which also supplies coarse-grained views to support flexible exploration of the alignments. A recent tool, VOAR [113], favors another approach for visualizing multiple ontology alignments (including also alignments computed by different tools) at once. Alignments can be compared to a reference alignment and a tabular view with different icons for true/false positive/negatives presents the outcome. Visualizing multiple alignments is achieved by color-coded edges between the ontologies represented as indented trees. Due to this the display quickly becomes cluttered with the growing number of tools and concepts in both ontologies although filtering by threshold or mapping type is supported.

An early alignment evaluation framework, KitAMO [79], supports the study, evaluation and comparison of (combination of) alignment strategies based on their performance and compliance measures on test cases. The framework contains a database of evaluation cases (ontologies and reference alignments) and receives as input different alignment components (matchers, filters and combination algorithms). The user can decide via a user interface which evaluation case to run. KitAMO computes and stores the similarity values generated by the components and provides an analysis tool to compare run times and alignment quality in terms of precision, recall and F-measure for different combinations of components. Through the analysis tool the user can save the evaluation results into the database and produce an evaluation report. The evaluation results are presented in the form of sortable tables. The influence of (combinations of) matchers and thresholds on the alignment quality can be compared by sorting in different ways. To examine the alignments in more detail a table can be used to depict not found correct mappings and compare the mappings found by the different matchers.

The OAEI evaluation campaigns are conducted on the SEALS⁶ platform. It provides a scalable and extensible infrastructure for remote evaluation of semantic technologies. The alignments computed by the tools are compared to reference alignments and evaluated in terms of precision, recall, F-measure, run time, coherence and number of requests to an oracle. The results are in the form of text files which are further analyzed by custom scripts and presented on the OAEI web site in sortable tables.

Probably closest to our work w.r.t. the user interface is the framework for modeling and comparing matching systems—Auto Mapping Core (AMC)—described in [104]. It presents a general environment for building, executing and tuning matching processes/workflows (various matchers and combination strategies) and allows to plug in existing matchers. It provides a graphical editor for constructing matching workflows and several views to analyze the (intermediate) alignments. One of the views presents the current alignment as a cube where two of the dimensions present the source and target ontologies and the third dimension depicts the similarity values as bars—taller bars represent higher similarity values. While this work also employs the cube metaphor it only shows one alignment at a time, i.e., does not provide means for comparing several alignments with it. The cube is employed to easily differentiate between high and low similarity values; decomposing it to different projections is not considered.

While all three frameworks provide a rich back-end infrastructure for configuring and executing (combination of) alignment algorithms the tabular views are too limited to adequately address the simultaneous interactive comparison of several alignments and provide visual exploration at different granularity levels. The third framework devotes more attention to the visual presentation of the results but it only depicts one alignment at a time. In comparison our work focuses on the user interface and allows to visually explore multiple alignments together and thus it can be seen as a necessary complement to the back-end functionalities offered by these frameworks.

4.4 Debugging of Ontology Networks

This section presents related work in connection to Paper III and focuses on detection approaches relying on the knowledge inferred from the integrated network and thus independent of external domain knowledge. The detection of modeling defects, especially missing structure, is not trivial since it requires domain knowledge and may not lead to defects discovered by reasoners. Manual inspection, apart from being error-prone, is, of course, possible, however, tedious and even infeasible for very large ontologies. While discovering incorrect relations may be possible by examining all existing relations, finding missing ones requires significantly more effort as the user needs to

⁶<http://seals-project.eu/>—Semantic Evaluation At Large Scale

consider numerous possibilities. The process of debugging is similarly effort intensive as the alignment process and approaches facilitating detection and repairing of defects are highly needed to support the domain experts in the debugging task. In our approach we utilize the knowledge intrinsic to an ontology network and, additionally, ontology alignment algorithms for detecting missing mappings. As our approach for debugging alignments relies on detecting modeling defects in ontologies other complementary methods may be used including methods from ontology learning from text [19], linguistics [55, 132], machine learning and statistical methods [88]. Other approaches include presenting domain experts with multiple choice questions [1] and employing antipatterns to find common errors [26].

We first discuss several works carried out by Lambrix et al. [75, 76, 78] which served as a foundation of the work in Paper III. We then outline several approaches which employ knowledge in the network to detect possible modeling defects [7, 11, 18]. The work by Bada et al. [7] has been supported by a prototype implementation while the other two approaches are mostly manual, not supported by a dedicated tool. Bada et al. [7] detect defects in the ontologies and do not consider incorrect alignments while the other two approaches are focused on repairing the alignment (although Beisswanger et al. [11] discuss finding problems in the ontologies as well). At the end of this section we discuss debugging of semantic defects in ontology networks which are easily detected by logic-based reasoning. There the challenge lies in the efficient computation of repairing actions.

4.4.1 Detecting and Repairing Missing Structure in Ontology Networks

The approach for debugging modeling defects presented in Paper III is an extension of the approaches presented in [75, 76, 78]. The problem of repairing missing is-a relations in a single taxonomy was initially discussed in [78] with the assumption that the is-a structure of the taxonomy is correct. The authors present two algorithms for computing repairing actions for missing is-a relations. The first, which is similar to the one presented in Paper III, only computes solutions for a single missing is-a relation. The second extends it by taking into account the influence of the repairing actions of other missing is-a relations during the computation of the Source and Target sets. The results of the experimental evaluation of the extended algorithm show that such influences are not negligible and in some cases the repairing actions for different missing is-a relations influence each other. Missing is-a relations in the context of a taxonomy network with correct mappings are discussed in [76].

The work in [75] is continuation of [78] where the authors consider wrong is-a relations in the structure of the taxonomies. In contrast to [78] where the focus is on a single taxonomy, the context in [75] is a taxonomy network

with the assumption that the mappings in the network are correct. Paper III takes the approach further, considering wrong and missing subsumption and equivalence mappings in a taxonomy network. We see ontology alignment as a special kind of debugging which allows for completing the subsumption relations in a network. Therefore, we employ ontology alignment algorithms as an additional method for detecting missing mappings thus increasing the knowledge in the network.

After manual validation alternative repairing actions for the missing relations are computed by the repairing algorithms in Paper III. These are essentially also missing is-a relations which could lead to further detection of defects. Another work by Lambrix et al. [82] focusing on completing the subsumption hierarchy in ontologies presented in the *EL* family of description logics also computes repairing actions which add knowledge to the network.

Bada et al. [7] present the closest approach for detecting missing is-a relations to our approach discussed in Paper III. The authors describe a method for identifying nonalignments (essentially missing subsumptions) in the ontologies from the Open Biomedical Ontologies repository (OBOs). The nonalignments can be seen as the candidate missing is-a relations in our approach. They are detected based on properties, such as *is_metabolized_in*, while in our work we employ subsumption and equivalence relations. Similarly to the framework in our work, three phases can be distinguished in [7]—a phase for detecting nonalignments, an examination which resembles our validation phase and a repairing phase. During the examination, the nonalignments that should be aligned are identified (they are called discrepancies). The nonalignments that are not discrepancies are indicators of logical inconsistencies in the ontologies and should be resolved as well. The authors suggest two approaches for rectifying the discrepancies—either adding the missing subsumptions or removing the existing subsumptions. In contrast to our approach alternative repairing actions are not considered in the former case. This approach does not consider that nonalignments could be caused by incorrect alignments. In both approaches the search space during the detection phase is reduced—in our approach we only employ mapped concepts and in [7] only pairs of assertions with an already existing subsumption relation are checked.

Bodenreider et al. [18] present one of the early works considering debugging of alignments during their development. The authors compare two approaches for aligning earlier versions of AMA⁷ and NCI Thesaurus⁸—manual and lexical. During the lexical alignment matching pairs of concepts between the ontologies (i.e. mappings) are identified by lexically matching their labels. Lexically similar concepts are called anchors and other anchors are further identified using synonyms from the UMLS Metathesaurus⁹. The anchors are used to perform structural validation and resemble the mapped concepts in

⁷http://www.informatics.jax.org/vocab/gxd/ma_ontology

⁸<https://ncit.nci.nih.gov/ncitbrowser/>

⁹https://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/

our approach. The manual and lexical alignments are used to create a final alignment and a structural validation was performed for debugging purposes in order to remove pairs of concepts without structural similarity from it. In each ontology, the relations in which the anchors participate are examined and the existence of at least one common hierarchical relation among the concepts in the anchors across the ontologies is taken as positive structural evidence. That is, if there is a hierarchical relation between a pair of anchors in one of the ontologies so there should be a relation between the respective anchors in the other. While our approach would consider the relations between anchors which do not exist in one of the ontologies as candidate missing is-a relations, the authors use them for filtering out mappings from the final alignment.

Similarly to Bodenreider et al. [18], Arnold et al. [4] also uses the structure of the ontology to validate potential mappings.

Based on their experience, Beisswanger et al. [11] propose a set of ten requirements which could be seen as patterns for debugging alignments. The first half of the requirements covers mainly technical and versioning issues in the ontologies and alignments. The second part is focused on the content and completeness of the alignments taking into account subsumption and equivalence relations from structural and linguistic points of view. Closest to our approach presented in Paper III are two requirements that deal with structural completeness and resemble part of our detection phase. According to one of them, if there are equivalence mappings between a particular concept from one of the ontologies and more than one concept in the other, the concepts in the second ontology should be connected through equivalence relations. In the other requirement, a given equivalence mapping is examined to determine if the subclasses of one of the concepts are connected through subsumption mappings with the superclasses of the other and vice versa. Another pair of requirements can be seen as alignment algorithms where (part of) the labels (or local names) are compared. The checks for these requirements can be performed automatically to help with the identification of missing and wrong mappings and also missing and wrong subsumption relations in the ontologies.

4.4.2 Semantic Defects in Ontology Alignments

Merging the two input ontologies and their alignment may lead to unsatisfiable concepts rendering the entire network incoherent which does not allow for using standard reasoning services to make inferences from it. About half of the systems participating in the OAEI Anatomy track this year—2017¹⁰ produce incoherent alignments and many tools in the LargeBio track¹¹ produce alignments with large numbers of unsatisfiable classes. Meilicke [90] demonstrates the impact of the incoherent alignments on several applications.

¹⁰<http://oaei.ontologymatching.org/2017/results/anatomy/index.html>

¹¹<http://www.cs.ox.ac.uk/isg/projects/SEALS/oaei/2017/results/>

Detecting semantic defects is easily done with logic-based reasoners and the challenge is in repairing them. Jiménez-Ruiz et al. [67] defines three general principles which should not be violated by the integrated ontology. According to the *conservativity principle* and contrary to our approach, the authors propose that new subsumption relations in one of the ontologies should not be introduced by the alignment. The *consistency principle* states that the integrated ontology should be consistent and contain only satisfiable concepts. The *locality principle* assumes that mapped concepts share similar semantically related entities. Resolving semantic defects in the integrated ontology network can be done by modifying the ontologies or removing/modifying mappings from the alignment. Existing approaches prefer the latter starting with already coherent ontologies. Automatic repairing approaches follow the so called minimal impact paradigm and aim to minimize the number of removed mappings (called alignment repair or diagnosis) assuming that the alignment is largely correct [112]. This, however, may lead to removing correct mappings when a large number of incorrect mappings is in conflict with a correct one as discussed in [103, 112].

Approaches for computing alignment repair consist of computing minimal sets causing logical contradictions and applying heuristics to decide which mappings to remove, for instance by considering the confidence values. Some apply distributed description logics representing the network as a distributed ontology where subsumption and equivalence mappings are represented as a set of directional bridge rules [91] while others encode the network in Horn propositional representations [68, 112]. LogMap [68] and AML [112] apply different ontology modularization strategies in order to handle efficiently large ontologies. In contrast to AML, LogMap considers replacing equivalence mappings by the respective subsumptions which allows for reducing the number of removed mappings and thus minimizing the impact of the repair on the alignment. The approaches in [91, 112] compute a global solution and the one in [68] computes a local solution in shorter time. The approach in [91] has been implemented in ALCOMO [90] which presents the consequences of a given decision to the user. In interactive mode LogMap gives priority to user repairs while AML considers the repair as an offline process. Qi et al. [106] apply belief-based revision theory and propose a conflict-based operator for mapping revision. Similarly to Meilicke et al. [91] the authors base their work on the notion of minimal conflict sets. In another approach Wang et al. [133] exemplify four patterns of frequently occurring defects in mappings and propose repairing methods that are either automatic or user-driven. While most debugging approaches focus on the alignments given coherent ontologies Rodler et al. [110] introduced an interactive query approach based on reinforcement learning which considers the entire integrated ontology network.

4.5 Information Visualization

The field of Information Visualization (InfoViz) combines research from several disciplines—(cognitive) psychology, human-computer interaction, computer graphics and computer science [115]—with the aim to improve the understanding of complex abstract data [101]. As the main research question of this thesis is concerned with providing efficient user support during the ontology matching task, drawing from advances in InfoVis provides a theoretical foundation and a palette of techniques to apply in order to enhance user support in ontology alignment tools.

Card et al. [21] say that ‘the purpose of information visualization is to use perception to amplify cognition’. InfoVis leverages the humans’ most powerful perceptual channel—the visual system—to expand working memory capacity and ‘offload work from cognitive to perceptual system’ [21]. Visualizations serve as external memory [95], enable perceptual inference, group information together and support pattern and outliers recognition [21, 56, 95]. Visual representations support the exploration of unknown data, hypothesis generation [134] and acquisition of insight [95]. Interacting with visualizations allows for cognitive offloading [87] and is essential in order to facilitate understanding and to overcome limitations of the visual representations especially when complex and huge data have been analyzed [95]. Visualizations increase the accessibility of data [56] thus expanding the users of the data and are means for communication [56, 95].

4.5.1 Visual Data Analysis and Exploration

Keim et al. [70] distinguish three high-level goals of visualization—presentation, confirmatory and exploratory analyses. The purpose of the presentation is to ‘efficiently and effectively communicate the results of an analysis’, i.e., it considers already established facts. Confirmatory analysis aims to verify or falsify existing hypotheses. The last goal—exploratory analysis—involves developing of hypotheses by searching for structures and trends in the data with the support of interactive tools. Part of our work in Papers I and II falls into the scope of the confirmatory analysis where the user needs to confirm or reject a set of automatically obtained mappings while our work in Paper VI supports exploratory data analysis where the user develops hypotheses through the interactive exploration of several alignments.

Data analysis and exploration is undertaken in order to fulfill a particular high-level task, such as decision-making, and it is facilitated by interactive visualizations. Ware et al. [134]¹² distinguish the activities involved in interactive visualization in three broad feedback loops closely interacting with each other. The *data selection and manipulation loop* consists of basic low-level

¹²Chapter 10 Interacting with Visualizations

hand-eyes coordinated interactions such as selecting, hovering and moving objects. The *navigation and exploration loop* includes interactions ‘through which an analyst finds his or her way in a large visual data space’ and supports the development of mental models. This broad set of interactions include panning and zooming or walking and flying provided by the affordances of the implemented metaphor as well as focus, context and scale techniques, such as distortion and multiple coordinated views. The *problem-solving loop* involves high-level behaviors such as creating, refining and testing hypotheses.

Our work in Papers I, II and VI is in the context of the latter two loops—the navigation and exploration, and problem-solving loops. Revisiting Paper I in the light of this discussion, we have identified three categories of desirable tasks/activities to be supported in ontology alignment tools—manipulation, inspection and explanation. The inspection category resembles parts of the navigation and exploration loop—the user is actively moving through the information space acquiring information at each step while the explanation category contributes to the problem-solving loop by including activities supplying additional information in support of decision-making. After a decision has been taken, it is materialized by actions included in the manipulation category which includes domain-dependent tasks not mapped to a particular loop. Further, Paper II analyses existing alignment tools and identifies different interface features, connected to the three categories in Paper I, supporting the alignment process. While our work in Paper VI exhibits elements from the former loop, it mainly targets the problem-solving loop by identifying shared analytic tasks which contribute to hypotheses development and verification.

Comparing our work in Paper VI to the Heer’s et al. [57] taxonomy of ‘interactive dynamics’ highlights the steps of the analytic process supported by our tool and identifies areas for future improvement. The taxonomy presented by Heer et al. [57] captures ‘interactive dynamics’ during the visual analysis process and consists of three categories encompassing 12 critical types of tasks for supporting the analytic process. The lowest level—*data and view specification*—consists of low-level tasks that allow for data exploration starting by *visualize* to map data attributes to visual properties, *filter* and *sort* to focus on items of interest and discover patterns, and *derive* to obtain new data by transforming the existing data. The next level—*view manipulation*—consists of tasks for *selecting* items and *navigating* the dataset as well as providing coordinated views for linked exploration (*coordinate*), and flexible organization of the workspace (*organize*). The highest-level category—*process and provenance*—considers capturing and reproducing the analytic processes by *annotating* findings and *recording* the analytic steps. These are necessary not only to *share* the analysis with colleagues but also to refine, review and identify unexplored cues. Finally, the *guide* task considers supporting knowledge transfer in developing analytic processes and strategies. Our tool *Alignment Cubes* supports the tasks in the first two categories by flexibly changing visual encodings, interactive exploration by selecting, filtering and sorting by

different criteria and providing views at different levels of detail (e.g. small-multiples and aggregated views). The third category is yet to be supported in order to allow for sharing the results from the comparative analysis of multiple alignments.

While the analytic process is not explicitly considered in Papers I and II the discussion above is relevant to them as well. The explanation category mentioned earlier falls into the scope of the process and provenance category by capturing activities for explaining why and how mappings have been computed by the tool, validated in a certain way and the impact of their validation.

4.5.2 Large and High-Resolution Displays

Recently, with the development of technology and the associated cost reduction, large, high-resolution displays have become available at affordable prices. For discussing large and high-resolution displays we adopt the view of Andrews et al. [3] who consider ‘...‘large, high-resolution display’ as being a display that is human scale ...the display’s size and resolution are closely matched to the sphere of perception and influence of the human body.’ and present several studies conducted in this broad setting ranging from several tiled displays to wall-sized displays.

Intuitively increasing the number of available pixels means providing more data as follows [3]: more data entities visible simultaneously, providing more dimensions (e.g., by for instance parallel coordinates), more details for a particular entity (without initial interaction with it by, e.g., hovering), data at different levels of detail (multiscale), higher data complexity and heterogeneity (revealing complex relationships), processing space (displaying history of views/transformations), space for sensemaking (employing spatial organization such as clusters and timelines [2, 17]) and enabling collaboration (by forming private and shared regions). Furthermore, more real estate enables designs which facilitate gaining insight from the data, such as providing multiple views of the same data (connected by linking and brushing techniques), embedded visualizations (embedding the detail in the overview) or combinations of both.

At the end of the 90’s Swaminathan et al. [123] pointed out that ‘when a display exceeds a certain size, it becomes qualitatively different: different design issues come into play and interaction design becomes full-blown environment design’. A number of studies have shown improved performance and reduced cognitive load in an everyday office environment due to more peripheral awareness, glancing instead of windows switching to obtain additional information, flexibility in the organization of the space, etc. A significant part of the benefits provided by large displays are in connection to humans’ spatial abilities. Large displays often replace virtual navigation by physical

‘thus allowing the user to exploit embodied human abilities such as spatial awareness, proprioception, and spatial memory’ [3].

Organization of Display Space

At the beginning of the millennium Grudin [52] pointed out that larger displays better utilize the peripheral view and reduce the cognitive load by glancing rapidly instead of switching windows to obtain supporting information. The benefits stem not from the possibility to make a single window larger but from the flexibility to organize and partition the available space. These benefits have been demonstrated in a sensemaking study by Andrews et al. [2].

Bi et al. [17] also observed different desktop partitioning and windows management strategies. All users partitioned the display into focal and peripheral regions for primary and secondary tasks and applications and adopted one of the following strategies: interactive applications closer to the center with passive applications on the periphery, and windows more relevant to the primary task closer to the center. Their users reported higher levels of engagement and relaxation when performing their daily tasks; they felt ‘surrounded’ by the task which contributed to improved attention. The authors observed that the benefits spread over tasks that require several windows (keeping them visible at the same time instead of switching between them), rich information tasks (by reducing the virtual navigation in, e.g., excel sheets and maps) and led to improved awareness of peripheral applications. Similar findings were reported by Ball et al. [8] where the primary task application was surrounded by supporting applications. The authors observed less window management and ‘decreased time switching between applications, allowing for decrease in cognitive load’ and their participants reported higher level of satisfaction. Similar workspace organization was reported by Knudsen et al. [71] who employed a 6-meters white board to study possible usage of the additional space during analytical tasks. The participants used the center area as a working/thinking area, which is configured by controls located on the vertical periphery, while special purpose views were fixed at the top or bottom of the white board.

Physical and Virtual Navigation

A major point of difference between regular- and large-size displays is in the navigation as physical navigation replaces virtual. While physical movement is more time consuming and tiring, coupling of physical motion and spacial memory would bring benefits in task performance [86]. Navigating large datasets in front of a regular monitor (by panning and zooming) could, however, be disorienting and distracting as discussed in [9]. Changing the display size, which replaces virtual navigation by physical, led to improved task performance while searching and navigating to a target on a map [10]. Ball et

al. [10] also observed that the increased display size (i.e., more information visualized) led users to change their strategy due to the better overview of the dataset they could obtain. The participants in another study [9] performed faster, felt less frustrated and stressed, and more confident in their replies while using high-resolution displays and physical navigation for the purposes of finding and comparing finely detailed data.

In contrast to previous studies which have not considered data manipulation tasks, Liu et al. [86] compared physical navigation in an ultra high-resolution display to virtual navigation on a regular display during a data classification task (e.g., reorganization). The authors attributed the improved performance in the hardest conditions to the use of more efficient strategies and better use of spatial memory when coupled with physical movement.

Shupp et al. [116] looked into different types of physical navigation by studying search and route tracing behavior during map navigation with 1, 12 and 24 displays in flat and curved display settings. Curving the display around the user replaces translational navigation with rotational. They found that the overall performance improved with the increased display size and it is best for the largest curved setting.

Navigation in Virtual Environments

Several studies in virtual environments demonstrated differences in navigation¹³ in regular and large displays setting. A study by Czerwinski et al. [28] on a large display with a wide field of view reduced the speed differences between males and females during a navigational task in a virtual environment. The wide field of view provided more environmental cues available for processing by the perceptual system instead of devoting resources for building a cognitive map of the environment. The authors suggested that wide field of view will benefit females with good spatial location memory for computational tasks that require navigation in complex information displays.

Similarly Ni et al. [97] showed improved performance for navigational tasks in large display conditions with more environmental cues (higher resolution or wider field of view). Varying the resolution on the large display impacted the performance less than changing the resolution in the small display condition. The authors suggested that this is due to ‘...the large displays improved navigation performance and hence reduced individual differences in spatial ability’. Providing a map during the training, as a wayfinding aid, to support spatial knowledge acquisition and cognitive map development, improved the performance in all conditions but had significant impact on the small display condition. The authors attributed these results to the ‘better optical flow

¹³Please observe the change in the meaning of the word *navigation* which until now has mostly referred to the act of moving in the space in front of a large display. In this subsection the usage of the word refers to exploring and finding how to reach different places in 3D virtual worlds.

cues' in the large display condition which might have supported these behaviors. The participants experienced more difficulties while navigating in the 3D world and relied more on the map in the small display conditions.

A series of experiments in a virtual world by Tan et al. [127, 125, 126] compared a large, projected-wall display and a standard desktop display (with equivalent content). In the wall condition the users adopted more efficient cognitive strategies for egocentric tasks and performed better in spatial orientation tasks, [127]; path integration, [125]; mental rotation, 3D navigational tasks, mental map formation and memory, [126]. The effect of the display size on performance seemed independent of other influencing factors, such as interactivity and mental aid (e.g., landmarks).

Sensemaking and Insight

Benefits from flexible space organization and coupling spatial memory and physical motion spread over complex tasks such as sensemaking.

A prominent study by Andrews et al. [2] showcased many of the already discussed benefits from the application of large displays during 'the cognitively demanding task of sensemaking'. Since sensemaking activities performed by the participants (students and professional analysts) 'are common for knowledge workers across many domains'—reading, identifying important points, arranging, etc.—the authors concluded that their findings can be generalized to other knowledge intensive tasks. The sensemaking process exhibits issues in relation to data overload and attention management which could be partially addressed by employing large, high-resolution displays. Using a small display requires intentional permanent effort in managing windows as also discussed in [109], e.g., constant context switching by hiding one window for the benefit from showing another. In such setting comparing information happens in the working memory instead of taking advantage of the humans' most powerful perceptual system—the visual system, which has been proven beneficial by, among others, Plumlee et al. [105].

The additional display space was employed by the study participants for two purposes: as external memory and a semantic layer. The persistent structures/representations created by the participants acted as external memory which allowed them to access part of the documents in their context and refresh their internal memory by rescanning, glancing, etc. without explicit context switching. The virtual navigation was replaced by physical, which although not quicker, has been claimed to be more efficient since spatial cues are available for the document. Having approximate cues facilitates recognition while at the same time reducing the recall from the memory. Arranging the documents in the space also conveyed information for the analysts—the documents were organized according to their relationships and those that were related ended up together. This is an additional cue that helps to identify their category and relationship between them by employing the perceptual

system instead of the memory. (One form of organization particularly stood out—an ordering where the documents were arranged according to their timeline horizontally and according to their topic vertically.)

Another study in the sensemaking domain by Reda et al. [109] compared exploration of crime data with multiple coordinated views in small and wall-sized displays (in curved settings) by general-public users. After the analysis the authors reported 74% more observations performed by the wall-display users. This may be due to the larger number of views simultaneously available for the wall-size display users in combination with ‘(hypothetically) lower cognitive cost of physical navigation may have induced participants to revisit information more frequently’. The scope of the observations performed on the large display were broader, integrating more dimensions of the dataset. This was attributed to reducing the distractions of view switching and staying longer in ‘insight-generating mental states’. The sessions with the wall display were longer and the authors grounded this finding with increased cognitive engagement during the analytic task. Increased engagement helps in keeping participants attention for longer and could have been fostered by replacing virtual navigation by physical which reduces the cognitive costs.

Shupp et al. [116] demonstrated that the type of insights is dependent on the displays configuration—flat vs curved setting. The type of initial insights obtained by their users differed—placing more pixels ‘within reach’ in the curved condition led to more detail-level insights (and was preferred by the users for search and comparison tasks) while the flat condition led to more overview-level initial insights (and it was preferred for insight tasks by the participants).

Knudsen et al. [71] conducted 11 workshops with data analysts from various domains—health care policy, website analysis, information retrieval, logistics—and observed them while walking through their recent data analysis tasks in front of a 6-meters white board. The analysis of the workshop activities revealed six common behaviors—four regarding the usage of display space and two regarding interaction strategies and gestures. The most prominent usage of the additional space during almost all of the workshops was to devote some areas for a certain purpose during the analysis with usually employing the central area for a working area. The vertical periphery was employed for configuration purposes while the horizontal was employed to permanently display aggregated data, (raw) data sets or particular views. The authors suggested that having additional space ‘one may use more of it to show data for longer periods of time.’ These persistent representations were used for both overview and detail views of the data. A common interaction was moving away from the display for obtaining an overview and moving closer to it for a detailed view. The participants also used the space to lay down multiple representations of data side-by-side and in an alternative approach (one-to-one) to interact with a single view which was transformed by applying, e.g., filters, for drilling down and comparing groups of data. It was also observed that

participants spread the data temporarily when they needed to choose from several options or modify it.

An important usage of the additional space was to externalize the thinking process by representing earlier steps both in the form of data processing flows and snapshots. These enable both backtracking and exploration of previously abandoned branches but also play a role of history of the manipulations that lead to the current state. Additionally, the participants found it useful to prepare a summary of their analysis and to mark the dead-ends reached during it. Such timeline would be beneficial for walking through collaborators during the different states of the dataset while explaining and verifying the different stages of the analysis. Using persistent views of the data to serve as an external memory was observed by Andrews et al. [2] and was beneficial during a learning task as discussed by Ragan et al. [107].

4.5.3 Navigation in Digital Environments

One of the complex processes supported by visual representations is navigation in digital spaces. Spatial navigation is a result of complex interaction of cognitive processes and large differences between individuals have been observed [137]. Since navigation in physical and digital environments involves similar behaviors, navigation in digital information spaces in connection to humans' spatial abilities has been a subject of a number of earlier studies (outside of the context of large displays). In their work on cognitive tests, Ekstrom et al. [37] distinguish two main spatial abilities—*spatial visualization* and *spatial orientation*. Spatial visualization is the 'ability to manipulate or transform the image of spatial patterns into other arrangements', while spatial orientation is the 'ability to perceive spatial patterns or to maintain orientation with respect to objects in space'.

Benyon et al. [12] posited that navigation related issues in information environments will have significant importance in the design of interactive systems and pointed out that such issues, especially in large information spaces, will lead to performance differences for different users in connection to retrieval time, confidence in the retrieved data, user satisfaction and learnability (as shown by Zhang [139]). Indeed, in a number of studies participants with higher spatial abilities were more efficient in information seeking tasks in archival search tools [31], hierarchical file systems [130, 129], online environment [92, 138], online shopping database system [13], hypermedia-based help system [29], hypertext system [20, 100] and a command line interface [13].

Furthermore, the studies also demonstrated that participants with lower spatial abilities got more often lost in hierarchical file systems [24, 31, 129] and visited more often the top-level table of contents [20]. Chen et al. [24] observed that these participants completed fewer tasks and 'were hesitant to explore large numbers of categories ...'. Campagnoni et al. [20] have suggested that participants with higher spatial visualization abilities are able to construct a

better mental model of the system which is further employed during search and navigation in it in the absence of visual cues. Providing navigational aids improved the performance for users with low spatial abilities [120] and for both users with low and high spatial abilities [100, 130, 138]. The navigational aids used in [120, 138] provided structural knowledge limiting the need to create a mental model. Pak et al. [100] further demonstrated that the aid could also impose demands on users' spatial abilities—users with higher abilities will perform better with aids demanding spatial abilities, such as maps (as opposite to a list with instructions).

**5**

Conclusions and Future Work

This chapter presents an outlook of the contributions of this thesis and draws directions for future work.

5.1 Conclusions

The Web contains an immense variety of structured, semi-structured and, most often, unstructured information sources in a variety of formats—web pages, databases, documents, figures, etc.—interconnected through an enormous number of links. Documents on the Web are mostly encoded in human-readable formats; extracting meaning from them is a task that only we, humans, can perform. Thus the knowledge captured in the Web is not utilized to its full potential, as it is only accessible to human interpretation and cannot be utilized by machines. In order to take advantage of the data at our disposal and to turn it into knowledge we need to enable machines to use it similarly to the way we do—interpret them in context, deal with various quality issues, integrate and draw conclusions from them in order to produce knowledge and fulfill tasks similarly to the way we do.

This has now started to change with the evolution of the Web towards the Semantic Web which encompasses a set of techniques for expressing and processing data in machine-readable formats. It provides means for automated agents to access the data captured in the Web of Documents, and to interpret, integrate and draw inferences from them in order to fulfill various tasks.

Ontologies are a key technology in the Semantic Web. They provide a shared, machine-readable vocabulary of a domain by formally representing the meaning of its concepts and relations and defining rules for creating new concepts. Ontologies serve as a basis for sharing, integrating and reusing knowledge and enable interoperability between systems.

Since ontologies are developed by different people and organizations to fulfill different goals, there may exist several ontologies modeling the same domain which could differ in conceptual modeling, granularity level, vocabulary and domain coverage. In order to employ several ontologies we need to find the relationships between their components. This is the subject of investigation of the ontology alignment field.

Many automatic approaches and algorithms have already been developed, however, for improving the quality of the alignments (in terms of precision and recall) user input is still necessary. Ontology alignment is a complex and challenging task imposing significant cognitive demands on the users. Involving users during the alignment process has not received enough focused attention and as a result the development of tools' interfaces has often been driven by the need to provide user input to matchers rather than by theoretical findings in the field of cognitive sciences and human-computer interaction.

5.1.1 User Involvement during Ontology Alignment

This dissertation investigates different issues related to involving users in the alignment process. Since fully automated alignment approaches are only considered the first step in creating alignments, we have been interested in the features that an ontology alignment system should provide in order to efficiently support users during the process of ontology alignment.

We have conducted a literature review encompassing a number of promising works in order to identify desirable back- and front-end features for ontology alignment systems. The front-end features consider the user interfaces of the tools; they are grouped in three categories according to their functionality—inspection, manipulation and explanation of alignments. In comparison to other works, we have also considered various back-end features which provide complementary functionalities and assist users in managing larger and more complex ontologies and alignments (and may further pose additional visualization and interface requirements). They contribute to the development of responsive interfaces and a complete infrastructure that supports the users during (large-scale) alignment tasks.

We have narrowed down the scope of our investigation to what we consider to be the most cognitively intensive part of the alignment process—manual validation of candidate mappings. Apart from discussing approaches for alignments visualization and necessary interactions we also discussed approaches for reducing user intervention and maximizing the value of user feedback. We also touched upon another little researched issue of the users' (domain or

knowledge representation) level of expertise and experience with the respective alignment system.

In order to reveal how the above aspects are supported by state-of-the-art tools we conducted two literature studies. We have further contributed one of the few user evaluations in the field by conducting a heuristic evaluation and an observational user study considering three alignment systems. Additionally, we provided an overview of the results of the OAEI Interactive track in 2015 to illustrate the impact of user error on the alignments' quality and reveal differences between the implemented strategies to deal with erroneous input.

As developing ontologies and alignments is an error-prone process, one desirable feature for alignment systems is to provide debugging capabilities. The detection of modeling defects, especially missing structure, is not trivial since it requires domain knowledge and may not lead to defects that could be discovered automatically. By employing the knowledge inferred in the integrated ontology network, we proposed methods for supporting domain experts in detecting and resolving missing and wrong relations in the ontologies and their alignments. We implemented one of the few dedicated tools where the debugging process is supported by interface components for validation and repairing which also supply explanations of the defects and recommendations for repairing them.

We have also considered the latest technological developments going beyond regular visualization and interaction settings, e.g., desktop and mouse, and have investigated the application of large, high-resolution displays in the field of ontology alignment. Drawing from the characteristics of the matching task we identified three promising directions for their application—improving the navigation in the ontologies and their alignments, space to support users' thinking process and collaboration.

5.1.2 Ontology Alignment Evaluation

Another important problem in the ontology alignment field is the evaluation of the alignments' quality. Alignments are compared to a reference alignment, if such is available, and general measures, such as precision and recall, are computed. However, most often, reference alignments are not available due to the cost to develop and maintain. Furthermore, even if reference alignments are available, these two numbers provide very limited understanding of the complex algorithms involved in the alignments' computation and other means for evaluation are necessary.

To address this issue we proposed a different evaluation approach and means for evaluation. We developed an interactive visual environment which allows for flexible visual exploration of several alignments at gradually changing levels of detail ranging from an overview of several alignments to the level of individual mappings in an alignment. This allows for fine-grained dif-

ferentiation between the approaches and facilitates studying their impact in various (corner) cases. We further provided insights of the last ten editions of the OAEI Anatomy track evaluation campaign and surveyed the matching algorithms used by the participating systems and their performance.

5.2 Future Work

We first discuss general research directions related to both developing alignments and their evaluation and then provide specific suggestions emerging from the papers included in this thesis.

General Directions

After studying various aspects relevant to fostering user involvement in the alignment process, deepening our understanding in the user expertise aspect appears to be one of the challenging, but promising, research directions. Most requirements investigations and user evaluations, including ours, have only considered novice users and predefined tasks. While recruiting actual users is infamously difficult, we need to understand the needs and requirements of actual users in order to improve the user support in the alignment tools. For this we need to go beyond task-based laboratory experiments and first conduct field experiments and interviews with practitioners which align ontologies on a regular basis. We need to consider users with different (levels and areas of) expertise as they likely have different requirements stemming from their different backgrounds. This will allow to develop tailored interfaces for users with varying needs.

Ontology alignment, as other knowledge-intensive areas, deals with an inherently intricate content. Development of alignments and their evaluation poses high demands on its users and involves a variety of cognitive processes—perception, attention, working memory, reasoning, etc. Thus, in order to improve the user support we should draw from findings in cognitive sciences and consider advances in decision and cognitive support achieved in Software Engineering with the help of InfoVis and HCI methods.

Ontology Alignment

Our user interface evaluations have revealed different usability issues, as the user interfaces of alignment systems have not been based on theoretical findings and have often resulted from the need to provide user input to matchers. Some seemingly trivial issues, such as visualizing multiple inheritance and search functionality, could become crucial in a large-scale setting. One issue that constantly appeared in our user study was the terminology adopted by the three systems. We have emphasized the importance of providing explanations for the matching results which is still insufficiently supported by the

current tools. As different visualizations are suitable for different tasks and could reveal different aspects of the data, it is interesting to study how different views support different tasks during the alignment. Additionally, more research into strategies for reducing the user intervention and maximizing the value from user feedback are also needed. Usability testing should be regularly performed to analyze the benefits of the proposed solutions and identify directions for improvement. Providing functional user interfaces with good usability features will ultimately support their adoption.

A fruitful direction for future research emerging from this thesis is the application of large, high resolution displays for the purpose of ontology alignment. Although we have discussed improving the navigation in the ontologies and their alignments, applying multiple views to support users' thinking processes and collaboration in the context of large displays, these problems are not specific for the large display setting and need further investigation in regular interaction settings as well.

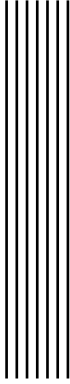
Ontology Alignment Evaluation

Alignment Cubes Our work on ontology alignment evaluation has just scratched the surface of a promising research direction. While we anticipate the benefits our tool brings to the process of alignments evaluation we are yet to quantify them through a comparative study including alternative tools (if any) or approaches (e.g., crafting scripts). Additionally, we need to evaluate the tool's usability in order to improve its ease of use, learnability, efficiency and understandability of the cube metaphor. Apart from the necessary evaluation, one interesting direction is to explore clustering and re-ordering algorithms to provide computational support for trends and patterns discovery. As already mentioned, introducing complementary views will be of help here as well. While interviews and field experiments cannot be replaced, one way to gather data and obtain insights for the actual user actions during the evaluation process is to log and analyze them.

As discussed in Subsection 4.5.1 *Alignment Cubes* support the tasks in the *data and view specification* and *data manipulation* categories from the task taxonomy for supporting interactive visual analysis developed by Heer et al. [57]. Adding support for the tasks in the third category—*process and provenance*—will enable sharing of the (intermediate) findings and recording of the entire analytic process as well as highlighting unexplored branches.

In the longer term, we are interested in investigating the integration of Alignment Cubes with the SEALS and HOBBIT platforms used in OAEI. This will also open the stage for investigating advantages, drawbacks and methodological issues around the two evaluation approaches: comparative visual exploration at a detailed level, and overall assessment of the quality of alignments.

Alignment Cubes Beyond Alignments Evaluation Another promising future direction is to investigate the adoption of the tool in different contexts such as workflows alignment and provenance as well as Linked Data exploration.



Bibliography

- [1] T. Alsubait, B. Parsia, and U. Sattler. “A Similarity Based Approach to Omission Finding in Ontologies”. In: *Ontology Engineering: 12th International Experiences and Directions Workshop on OWL, OWLED 2015, Revised Selected Papers*. Ed. by V. Tamma, M. Dragoni, R. Gonçalves, and A. Ławrynowicz. 2016, pp. 21–32.
- [2] C. Andrews, A. Endert, and C. North. “Space to Think: Large High-resolution Displays for Sensemaking”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. 2010, pp. 55–64.
- [3] C. Andrews, A. Endert, B. Yost, and C. North. “Information visualization on large, high-resolution displays: Issues, challenges, and opportunities”. In: *Information Visualization* 10.4 (2011), pp. 341–355.
- [4] P. Arnold and E. Rahm. “Semantic Enrichment of Ontology Mappings: A Linguistic-Based Approach”. In: *Proceedings of the 17th East European Conference on Advances in Databases and Information Systems (ADBIS '13)*. 2013, pp. 42–55.
- [5] J. Aurisano, A. Nanavaty, and I. Cruz. “Visual Analytics for Ontology Matching Using Multi-linked Views”. In: *Proceedings of the 1st International Workshop on Visualizations and User Interfaces for Ontologies and Linked Data (VOILA '15)*. Vol. 1456 of CEUR Workshop Proceedings. 2015, pp. 25–36.

- [6] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. Patel-Schneider, eds. *The description logic handbook*. Cambridge University Press, 2003.
- [7] M. Bada and L. Hunter. “Identification of OBO Nonalignments and Its Implications for OBO Enrichment.” In: *Bioinformatics* 24.12 (2008), pp. 1448–1455.
- [8] R. Ball and C. North. “Analysis of User Behavior on High-Resolution Tiled Displays”. In: *Proceedings of the 10th IFIP TC13 International Conference on Human-Computer Interaction (INTERACT '05)*. 2005, pp. 350–363.
- [9] R. Ball and C. North. “Effects of Tiled High-resolution Display on Basic Visualization and Navigation Tasks”. In: *CHI 2005 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05)*. 2005, pp. 1196–1199.
- [10] R. Ball, C. North, and D. A. Bowman. “Move to Improve: Promoting Physical Navigation to Increase User Performance with Large Displays”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. 2007, pp. 191–200.
- [11] E. Beisswanger and U. Hahn. “Towards valid and reusable reference alignments—ten basic quality checks for ontology alignments and their application to three different reference data sets”. In: *Biomedical Semantics* 3.1 (2012), S4.
- [12] D. Benyon and K. Höök. “Navigation in Information Spaces: supporting the individual”. In: *Proceedings of the 6th IFIP TC13 International Conference on Human-Computer Interaction (INTERACT '97)*. 1997, pp. 39–46.
- [13] D. Benyon and D. Murray. “Adaptive systems: from intelligent tutoring to autonomous agents”. In: *Knowledge-Based Systems* 6.4 (1993), pp. 197–219.
- [14] A. Bernaras, I. Laresgoiti, and J. Corera. “Building and Reusing Ontologies for Electrical Network Applications”. In: *Proceedings of the 12th European Conference on Artificial Intelligence (ECAI '96)*. (ECAI '96). 1996, pp. 298–302.
- [15] T. Berners-Lee, J. Hendler, and O. Lassila. “The Semantic Web”. In: *Scientific American* 284.5 (2001), pp. 29–37.
- [16] P. A. Bernstein and S. Melnik. “Model Management 2.0: Manipulating Richer Mappings”. In: *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data*. 2007, pp. 1–12.

-
- [17] X. Bi and R. Balakrishnan. “Comparing Usage of a Large High-resolution Display to Single or Dual Desktop Displays for Daily Work”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. 2009, pp. 1005–1014.
- [18] O. Bodenreider, T. Hayamizu, M. Ringwald, S. D. Coronado, and S. Zhang. “Of mice and men: aligning mouse and human anatomies.” In: *Proceedings of the American Medical Informatics Association (AIMA) Annual Symposium*. 2005, pp. 61–65.
- [19] P. Buitelaar, P. Cimiano, and B. Magnini, eds. *Ontology Learning from Text: Methods, Evaluation and Applications*. Vol. 123. Frontiers in Artificial Intelligence and Applications Series. 2005.
- [20] F. R. Campagnoni and K. Ehrlich. “Information Retrieval Using a Hypertext-based Help System”. In: *ACM Transactions on Information Systems (TOIS)* 7.3 (1989), pp. 271–291.
- [21] S. K. Card, J. D. Mackinlay, and B. Shneiderman, eds. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann, 1999.
- [22] S. Carpendale. “Evaluating Information Visualizations”. In: *Information Visualization*. Ed. by A. Kerren, J. T. Stasko, J-D. Fekete, and C. North. 2008, pp. 19–45.
- [23] B. Chen, H. Tan, and P. Lambrix. “Structure-Based Filtering for Ontology Alignment”. In: *Proceedings of 15th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE '06)*. 2006, pp. 364–369.
- [24] C. Chen and M. Czerwinski. “Spatial ability and visual navigation: an Empirical Study”. In: *The New Review of Hypermedia and Multimedia* 3 (1997), pp. 67–89.
- [25] O. Corcho, M. Fernández-López, and A. Gómez-Pérez. “Ontological Engineering: Principles, Methods, Tools and Languages”. In: *Ontologies for Software Engineering and Software Technology*. Ed. by C. Calero, F. Ruiz, and M. Piattini. 2006, pp. 1–48.
- [26] O. Corcho, C. Roussey, L. M. V. Blazquez, and I. Perez. “Pattern-based OWL Ontology Debugging Guidelines”. In: *Proceedings of 1st Workshop on Ontology Patterns (WOP 2009)*. Vol. 516 of CEUR Workshop Proceedings. 2009, pp. 68–82.
- [27] I. Cruz, C. Stroe, and M. Palmonari. “Interactive User Feedback in Ontology Matching Using Signature Vectors”. In: *Proceedings of the 28th IEEE International Conference on Data Engineering (ICDE '12)*. 2012, pp. 1321–1324.

- [28] M. Czerwinski, D. S. Tan, and G. G. Robertson. “Women Take a Wider View”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '02)*. 2002, pp. 195–202.
- [29] N. Dahlbäck, K. Höök, and M. Sjölander. “Spatial cognition in the mind and in the world: The case of hypermedia navigation”. In: *The 18th Meeting of the Cognitive Science Society*. 1996.
- [30] Y. Ding, D. Fensel, M. Klein, B. Omelayenko, and E. Schulten. “The Role of Ontologies in eCommerce”. In: *Handbook on Ontologies*. Ed. by S. Staab and R. Studer. 2004, pp. 593–615.
- [31] R. E. Downing, J. L. Moore, and S. W. Brown. “The effects and interaction of spatial visualization and domain expertise on information seeking”. In: *Computers in Human Behavior* 21.2 (2005), pp. 195–209.
- [32] Z. Dragisic, V. Ivanova, P. Lambrix, D. Faria, E. Jiménez-Ruiz, and C. Pesquita. “User Validation in Ontology Alignment”. In: *Proceedings of the 15th International Semantic Web Conference (ISWC '16)*. Vol. 9981 of Lecture Notes in Computer Science. 2016, pp. 200–217.
- [33] Z. Dragisic, V. Ivanova, H. Li, and P. Lambrix. “Experiences from the Anatomy track in the Ontology Alignment Evaluation Initiative”. In: *Biomedical Semantics* 8.1 (2017), p. 56.
- [34] F. Duchateau and Z. Bellahsene. “Designing a benchmark for the assessment of schema matching tools”. In: *Open Journal of Databases (OJDB)* 1.1 (2014), pp. 3–25.
- [35] S. Easterbrook, J. Singer, M-A. Storey, and D. Damian. “Selecting empirical methods for software engineering research”. In: *Guide to advanced empirical software engineering* (2008), pp. 285–311.
- [36] M Ehrig and J. Euzenat. “Relaxed precision and recall for ontology matching”. In: *Proceedings of the K-CAP 2005 Workshop on Integrating Ontologies*. Vol. 156 of CEUR Workshop Proceedings. 2005, pp. 25–32.
- [37] R. B. Ekstrom, J. W. French, H. H. Harman, and D. Dermen. *Manual for Kit of Factor-Referenced Cognitive Tests*. Tech. rep. Princeton, NJ: Educational Testing Service, 1976.
- [38] J. Euzenat, C. Meilicke, P. Shvaiko, H. Stuckenschmidt, and C. Trojahn dos Santos. “Ontology Alignment Evaluation Initiative: six years of experience”. In: *Data Semantics XV* (2011), pp. 158–192.
- [39] J. Euzenat and P. Shvaiko. “Evaluation of Matching Systems”. In: *Ontology Matching*. Ed. by J. Euzenat and P. Shvaiko. 2013, pp. 285–317.
- [40] J. Euzenat and P. Shvaiko, eds. *Ontology Matching*. Springer-Verlag New York, Inc., 2007.

-
- [41] J. Euzenat and P. Shvaiko. “User Involvement”. In: *Ontology Matching*. Ed. by J. Euzenat and P. Shvaiko. 2013, pp. 353–375.
- [42] S. Falconer and N. Noy. “Interactive Techniques to Support Ontology Matching”. In: *Schema Matching and Mapping*. Ed. by Z. Bellahsene, A. Bonifati, and E. Rahm. 2011, pp. 29–51.
- [43] S. Falconer, N. Noy, and M-A. Storey. “Ontology mapping - a user survey”. In: *Proceedings of the 2nd International Conference on Ontology Matching (OM '07)*. Vol. 304 of CEUR Workshop Proceedings. 2007, pp. 49–60.
- [44] S. Falconer, N. Noy, and M-A. Storey. “Towards Understanding the Needs of Cognitive Support for Ontology Mapping”. In: *Proceedings of the 1st International Conference on Ontology Matching (OM '06)*. Vol. 225 of CEUR Workshop Proceedings. 2006.
- [45] S. Falconer and M-A. Storey. “A Cognitive Support Framework for Ontology Mapping”. In: *Proceedings of the 6th International Semantic Web Conference and the 2nd Asian Semantic Web Conference (ISWC '06/ASWC '06)*. Vol. 4825 of Lecture Notes in Computer Science. 2007, pp. 114–127.
- [46] D. Laney META Group (today's Gartner). *3D Data Management: Controlling Data Volume, Velocity, and Variety*. 2001. URL: <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> (visited on 07/24/2017).
- [47] D. Gašević, N. Kaviani, and M. Milanović. “Ontologies and software engineering”. In: *Handbook on Ontologies*. Ed. by S. Staab and R. Studer. Second Edition. 2009, pp. 593–615.
- [48] A. Gómez-Pérez, M. Fernández-López, and O. Corcho, eds. *Ontological Engineering*. Springer-Verlag London, 2004.
- [49] M. Granitzer, V. Sabol, K. W. Onn, D. Luckose, and K. Tochtermann. “Ontology Alignment—A Survey with Focus on Visually Supported Semi-Automatic Techniques”. In: *Future Internet* (2010), pp. 238–258.
- [50] G. Grinstein, A. Kobsa, C. Plaisant, B. Shneiderman, and J. T. Stasko. “Which Comes First, Usability or Utility?” In: *Proceedings of the 14th IEEE Visualization 2003 (VIS'03)*. 2003, pp. 112–113.
- [51] T. R. Gruber. “A Translation Approach to Portable Ontology Specifications”. In: *Knowledge Acquisition* 5.2 (1993), pp. 199–220.
- [52] J. Grudin. “Partitioning digital worlds: focal and peripheral awareness in multiple monitor use”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '01)*. 2001, pp. 458–465.

- [53] N. Guarino, D. Oberle, and S. Staab. “What Is an Ontology?” In: *Handbook on Ontologies*. Ed. by S. Staab and R. Studer. Second Edition. 2009, pp. 1–17.
- [54] H. Happel and S. Seedorf. “Applications of Ontologies in Software Engineering”. In: *Proceedings of the 2nd International Workshop on Semantic Web Enabled Software Engineering (SWESE 2006)*. 2006, pp. 5–9.
- [55] M. A. Hearst. “Automatic acquisition of hyponyms from large text corpora”. In: *Proceedings of the 14th Conference on Computational Linguistics (COLING '92)*. Vol. 2. 1992, pp. 539–545.
- [56] J. Heer, M. Bostock, and V. Ogievetsky. “A tour through the visualization zoo”. In: *Queue* 8.5 (2010), pp. 1–22.
- [57] J. Heer and B. Shneiderman. “Interactive dynamics for visual analysis”. In: *Queue* 10.2 (2012), pp. 1–26.
- [58] G. van Heijst, A. T. Schreiber, and B. J. Wielinga. “Using Explicit Ontologies in KBS Development”. In: *Human-Computer Studies* 46.2–3 (1997), pp. 183–292.
- [59] M. Hepp. “GoodRelations: An Ontology for Describing Products and Services Offers on the Web”. In: *Proceedings of the 16th International Conference on Knowledge Engineering and Knowledge Management (EKAW '08)*. 2008, pp. 329–346.
- [60] A. Herzog, N. Shahmehri, and C. Duma. “An ontology of information security”. In: *Information Security and Privacy* 1.4 (2007), pp. 1–23.
- [61] T. Isenberg, P. Isenberg, J. Chen, M. Sedlmair, and T. Möller. “A Systematic Review on the Practice of Evaluating Visualization”. In: *IEEE Transactions on Visualization and Computer Graphics* 19.12 (2013), pp. 2818–2827.
- [62] V. Ivanova. “Applications of Large Displays: Advancing User Support in Large Scale Ontology Alignment”. In: *Proceedings of the Doctoral Consortium at the 15th International Semantic Web Conference (ISWC '16)*. Vol. 1733 of CEUR Workshop Proceedings. 2016, pp. 50–57.
- [63] V. Ivanova, B. Bach, E. Pietriga, and P. Lambrix. “Alignment Cubes: Towards Interactive Visual Exploration and Evaluation of Multiple Ontology Alignments”. In: *Proceedings of the 16th International Semantic Web Conference (ISWC '17)*. Vol. 10587 of Lecture Notes in Computer Science. 2017, pp. 400–417.
- [64] V. Ivanova and P. Lambrix. “A Unified Approach for Aligning Taxonomies and Debugging Taxonomies and Their Alignments”. In: *Proceedings of the 10th European Semantic Web Conference (ESWC '13)*. Vol. 7882 of Lecture Notes in Computer Science. 2013, pp. 1–15.

-
- [65] V. Ivanova, P. Lambrix, and J. Åberg. “Requirements for and Evaluation of User Support for Large-Scale Ontology Alignment”. In: *Proceedings of the 12th European Semantic Web Conference (ESWC '15)*. Vol. 9088 of Lecture Notes in Computer Science. 2015, pp. 3–20.
- [66] E. Jiménez-Ruiz, B. C. Grau, and I. Horrocks. “Is My Ontology Matching System Similar to Yours?” In: *Proceedings of the 8th International Conference on Ontology Matching (OM '13)*. Vol. 1111 of CEUR Workshop Proceedings. 2013, pp. 229–230.
- [67] E. Jiménez-Ruiz, B. C. Grau, I. Horrocks, and R. Berlanga. “Logic-based assessment of the compatibility of UMLS ontology sources”. In: *Biomedical Semantics 2.1* (2011), S2.
- [68] E. Jiménez-Ruiz, B. C. Grau, Y. Zhou, and I. Horrocks. “Large-scale Interactive Ontology Matching: Algorithms and Implementation”. In: *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI '12)*. 2012, pp. 444–449.
- [69] A Kalyanpur. “Debugging and Repair of OWL Ontologies”. PhD thesis. University of Maryland, 2006.
- [70] D. A. Keim, F. Mansmann, J. Schneidewind, and H. Ziegler. “Challenges in Visual Data Analysis”. In: *Proceedings of the 10th International Conference on Information Visualisation (IV '06)*. 2006, pp. 9–16.
- [71] S. Knudsen, M. R. Jakobsen, and K. Hornbæk. “An Exploratory Study of How Abundant Display Space May Support Data Analysis”. In: *Proceedings of the 7th Nordic Conference on Human-Computer Interaction (NordiCHI '12)*. 2012, pp. 558–567.
- [72] P. Lambrix. “Ontologies in Bioinformatics and Systems Biology”. In: *Artificial Intelligence Methods And Tools For Systems Biology*. Vol. 5. Computational Biology. 2004, pp. 129–145.
- [73] P. Lambrix. “Towards a semantic Web for bioinformatics using ontology-based annotation”. In: *Proceedings of the 14th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprise (WETICE '05)*. 2005, pp. 3–7.
- [74] P. Lambrix and A. Edberg. “Evaluation of ontology merging tools in bioinformatics”. In: *Proceedings of the Pacific Symposium on Biocomputing*. 2003, pp. 589–600.
- [75] P. Lambrix and Q. Liu. “Debugging Is-a Structure in Networked Taxonomies”. In: *Proceedings of the 4th International Workshop on Semantic Web Applications and Tools for the Life Sciences (SWAT4LS '11)*. 2011, pp. 58–65.

- [76] P. Lambrix and Q. Liu. “Debugging the missing is-a structure within taxonomies networked by partial reference alignments”. In: *Data & Knowledge Engineering* 86 (2013), pp. 179–205.
- [77] P. Lambrix and Q. Liu. “Using partial reference alignments to align ontologies”. In: *Proceedings of the 6th European Semantic Web Conference (ESWC '09)*. Vol. 5554 of Lecture Notes in Computer Science. 2009, pp. 188–202.
- [78] P. Lambrix, Q. Liu, and H. Tan. “Repairing the missing is-a structure of ontologies”. In: *Proceedings of the 4th Asian Semantic Web Conference (ASWC '09)*. Vol. 5926 of Lecture Notes in Computer Science. 2009.
- [79] P. Lambrix and H. Tan. “A tool for evaluating ontology alignment strategies”. In: *Data Semantics VIII* (2007), pp. 182–202.
- [80] P. Lambrix and H. Tan. “Ontology Alignment and Merging”. In: *Anatomy Ontologies for Bioinformatics*. Vol. 6. Computational Biology. 2008, pp. 133–149.
- [81] P. Lambrix, H. Tan, V. Jakoniene, and L. Strömbäck. “Biological Ontologies”. In: *Semantic Web*. 2007, pp. 85–99.
- [82] P. Lambrix, F. Wei-Kleiner, and Z. Dragisic. “Completing the is-a structure in light-weight ontologies”. In: *Biomedical Semantics* 6.1 (2015), p. 26.
- [83] M. Lanzemberger, J. Sampson, M. Rester, Y. Naudet, and T. Lattour. “Visual ontology alignment for knowledge sharing and reuse”. In: *Knowledge Management* 12.6 (2008), pp. 102–120.
- [84] O. Lassila and D. L. McGuinness. *The Role of Frame-Based Representation on the Semantic Web*. Tech. rep. 2001.
- [85] Y. Li, C. Stroe, and I. F. Cruz. “Interactive Visualization of Large Ontology Matching Results”. In: *Proceedings of the 1st International Workshop on Visualizations and User Interfaces for Ontologies and Linked Data (VOILA '15)*. Vol. 1456 of CEUR Workshop Proceedings. 2015, pp. 37–48.
- [86] C. Liu, O. Chapuis, M. Beaudouin-Lafon, E. Lecolinet, and W. E. Mackay. “Effects of Display Size and Navigation Type on a Classification Task”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. 2014, pp. 4147–4156.
- [87] Z. Liu and J. Stasko. “Mental Models, Visual Reasoning and Interaction in Information Visualization: A Top-down Perspective”. In: *IEEE Transactions on Visualization and Computer Graphics* 16 (2010), pp. 999–1008.
- [88] A. Maedche and S. Staab. “Discovering conceptual relations from text”. In: *Proceedings of the 14th European Conference on Artificial Intelligence (ECAI '00)*. 2000, pp. 321–325.

-
- [89] D. L. McGuinness, R. Fikes, J. Rice, and S. Wilder. “An environment for merging and testing large ontologies”. In: *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning (KR '00)*. 2000, pp. 483–493.
- [90] C. Meilicke. “Alignment Incoherence in Ontology Matching”. PhD thesis. Universität Mannheim, 2011.
- [91] C. Meilicke, H. Stuckenschmidt, and A. Tamilin. “Repairing Ontology Mappings”. In: *Proceedings of the 22nd Conference on Artificial Intelligence (AAAI '07)*. 2007, p. 6.
- [92] M. C. P. Melguizo, U. Vidya, and H. van Oostendorp. “Seeking information online: the influence of menu type, navigation path complexity and spatial ability on information gathering tasks”. In: *Behaviour & Information Technology* 31.1 (2012), pp. 59–70.
- [93] G. A. Miller. “WordNet: a lexical database for English”. In: *Communications of the ACM* 38.11 (1995), pp. 39–41.
- [94] S. Mukherjea, B. Bamba, and P. Kankar. “Information retrieval and knowledge discovery utilizing a biomedical Semantic Web”. In: *IEEE Transactions on Knowledge and Data Engineering* 17 (2005), pp. 1099–1110.
- [95] T. Munzner, ed. *Visualization analysis and design*. CRC press, 2014.
- [96] R. Neches, R. Fikes, T. Finin, T. Gruber, R. Patil, T. Senator, and W. P. Swartout. “Enabling technology for knowledge sharing”. In: *AI Magazine* 12.3 (1991), pp. 36–56.
- [97] T. Ni, D. A. Bowman, and J. Chen. “Increased Display Size and Resolution Improve Task Performance in Information-Rich Virtual Environments”. In: *Proceedings of the 32nd Graphics Interface Conference (GI '06)*. 2006, pp. 139–146.
- [98] N. Noy and M. Musen. “Algorithm and Tool for Automated Ontology Merging and Alignment”. In: *Proceedings of the 17th Conference on Artificial Intelligence (AAAI '00)*. 2000, pp. 450–455.
- [99] L. Otero-Cerdeira, F. J. Rodríguez-Martínez, and A. Gómez-Rodríguez. “Ontology matching: A literature review”. In: *Expert Systems with Applications* 42.2 (2015), pp. 949–971.
- [100] R. Pak, W. A. Rogers, and A. D. Fisk. “Spatial ability subfactors and their influences on a computer-based information search task”. In: *Human Factors* 48 (2006), pp. 154–165.
- [101] R. E. Patterson, L. M. Blaha, G. G. Grinstein, K. K. Liggett, D. E. Kaveney, K. C. Sheldon, P. R. Havig, and J. A. Moore. “A human cognition framework for information visualization”. In: *Computers & Graphics* 42 (2014), pp. 42–58.

- [102] H. Paulheim, S. Hertling, and D. Ritzke. “Towards Evaluating Interactive Ontology Matching Tools”. In: *Proceedings of the 10th European Semantic Web Conference (ESWC '13)*. Vol. 7882 of Lecture Notes in Computer Science. 2013, pp. 31–45.
- [103] C. Pesquita, D. Faria, E. Santos, and F. M. Couto. “To repair or not to repair: reconciling correctness and coherence in ontology reference alignments”. In: *Proceedings of the 8th International Conference on Ontology Matching (OM '13)*. Vol. 1111 of CEUR Workshop Proceedings. 2013, pp. 13–24.
- [104] E. Peukert, J. Eberius, and E. Rahm. “AMC-A framework for modeling and comparing matching systems as matching processes”. In: *Proceedings of the 27th IEEE International Conference on Data Engineering (ICDE '11)*. 2011, pp. 1304–1307.
- [105] M. D. Plumlee and C. Ware. “Zooming Versus Multiple Window Interfaces: Cognitive Costs of Visual Comparisons”. In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 13 (2006), pp. 179–209.
- [106] G. Qi, Q. Ji, and P. Haase. “A Conflict-Based Operator for Mapping Revision”. In: *Proceedings of the 8th International Semantic Web Conference (ISWC '09)*. Vol. 5823 of Lecture Notes in Computer Science. 2009, pp. 521–536.
- [107] E. D. Ragan, A. Endert, D. A. Bowman, and F. Quek. “How Spatial Layout, Interactivity, and Persistent Visibility Affect Learning with Large Displays”. In: *Proceedings of the 11th International Working Conference on Advanced Visual Interfaces (AVI '12)*. 2012, pp. 91–98.
- [108] E. Rahm. “Towards Large-Scale Schema and Ontology Matching”. In: *Schema Matching and Mapping*. Ed. by Z. Bellahsene, A. Bonifati, and E. Rahm. 2011, pp. 3–27.
- [109] K. Reda, A. E. Johnson, M. E. Papka, and J. Leigh. “Effects of Display Size and Resolution on User Behavior and Insight Acquisition in Visual Exploration”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. 2015, pp. 2759–2768.
- [110] P. Rodler, K. Shehekotykhin, P. Fleiss, and G. Friedrich. “Rio: Minimizing user interaction in debugging of aligned ontologies”. In: *Proceedings of the 7th International Conference on Ontology Matching (OM '12)*. Vol. 946 of CEUR Workshop Proceedings. CEUR-WS. org. 2012, pp. 49–60.
- [111] F. Ruiz and J. R. Hilera. “Using Ontologies in Software Engineering and Technology”. In: *Ontologies for Software Engineering and Software Technology*. Ed. by C. Calero, F. Ruiz, and M. Piattini. 2006, pp. 49–102.

-
- [112] E. Santos, D. Faria, C. Pesquita, and F. M. Couto. “Ontology Alignment Repair through Modularization and Confidence-Based Heuristics”. In: *PLOS ONE* 10.12 (Dec. 2016), pp. 1–19.
- [113] B. Severo, C. Trojahn, and R. Vieira. “A GUI for Visualising and Manipulating Multiple Ontology Alignments”. In: *Proceedings of the Posters & Demonstrations Track at the 14th International Semantic Web Conference (ISWC '15)*. Vol. 1486 of CEUR Workshop Proceedings. 2015, pp. 37–48.
- [114] B. Shneiderman. “The eyes have it: A task by data type taxonomy for information visualizations”. In: *Proceedings of the 1996 IEEE Symposium on Visual Languages (VL '96)*. 1996, pp. 336–343.
- [115] B. Shneiderman and B. B. Bederson, eds. *The Craft of Information Visualization: Readings and Reflections*. Morgan Kaufmann Publishers Inc., 2003.
- [116] L. Shupp, C. Andrews, M. Dickey-Kurdziolek, B. Yost, and C. North. “Shaping the Display of the Future: The Effects of Display Size and Curvature on User Performance and Insights”. In: *Human-Computer Interaction* 24.1–2 (2009), pp. 230–272.
- [117] P. Shvaiko and J. Euzenat. “Ontology Matching: State of the Art and Future Challenges”. In: *Knowledge and Data Engineering* 25.1 (2013), pp. 158–176.
- [118] P. Shvaiko and J. Euzenat. “Ten Challenges for Ontology Matching”. In: *OTM* (2008), pp. 1164–1182.
- [119] A. Souag, C. Salinesi, and I. Comyn-Wattiau. “Ontologies for Security Requirements: A Literature Survey and Classification”. In: *Proceedings of the International Workshops on Advanced Information Systems Engineering (CAiSE 2012)*. Ed. by M. Bajec and J. Eder. 2012, pp. 61–69.
- [120] K. M. Stanney and G. Salvendy. “Information visualization; assisting low spatial individuals with information access tasks through the use of visual mediators”. In: *Ergonomics* 38.6 (1995), pp. 1184–1198.
- [121] R. Stevens, C. A. Goble, and S. Bechhofer. “Ontology-based knowledge representation for bioinformatics”. In: *Briefings in bioinformatics* 1.4 (2000), pp. 398–414.
- [122] R. Studer, V. R. Benjamins, and D. Fensel. “Knowledge Engineering: Principles and Methods”. In: *Data & Knowledge Engineering* 25.1–2 (1998), pp. 161–197.
- [123] K. Swaminathan and S. Sato. “Interaction Design for Large Displays”. In: *Interactions* 4 (1997), pp. 15–24.

- [124] B. Swartout, R. Patil, K. Knight, and T. Russ. “Toward Distributed Use of Large-Scale Ontologies”. In: *Proceedings of the 10th Workshop on Knowledge Acquisition for Knowledge-Based Systems*. 1997, pp. 138–148.
- [125] D. S. Tan, D. Gergle, P. G. Scupelli, and R. Pausch. “Physically Large Displays Improve Path Integration in 3D Virtual Navigation Tasks”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04)*. 2004, pp. 439–446.
- [126] D. S. Tan, D. Gergle, P. G. Scupelli, and R. Pausch. “Physically Large Displays Improve Performance on Spatial Tasks”. In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 13 (2006), pp. 71–99.
- [127] D. S. Tan, D. Gergle, P. G. Scupelli, and R. Pausch. “With Similar Visual Angles, Larger Displays Improve Spatial Performance”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. 2003, pp. 217–224.
- [128] M. Uschold and M. Gruninger. “Ontologies and Semantics for Seamless Connectivity”. In: *SIGMOD Record* 33.4 (2004), pp. 58–64.
- [129] K. J. Vicente, B. C. Hayes, and R. C. Williges. “Assaying and isolating individual differences in searching a hierarchical file system”. In: *Human Factors* 29.3 (1987), pp. 349–359.
- [130] K. J. Vicente and R. C. Williges. “Accommodating individual differences in searching a hierarchical file system”. In: *Man-Machine Studies* 29 (1988), pp. 647–668.
- [131] H. Wache, T. Vögele, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann, and S. Hübner. “Ontology-based integration of information—a survey of existing approaches”. In: *Proceedings of the International Joint Conference on Artificial Intelligence—01 Workshop: Ontologies and Information Sharing*. 2001, pp. 108–117.
- [132] T. Wachter, A. Wobst, M. Schroeder, H. Tan, and P. Lambrix. “A corpus-driven approach for design, evolution and alignment of ontologies”. In: *Proceedings of the Winter Simulation Conference (WSC '06)*. IEEE. 2006, pp. 1595–1602.
- [133] P. Wang and B. Xu. “Debugging ontology mappings: a static approach”. In: *Computing and Informatics* 27 (2008), pp. 21–36.
- [134] C. Ware, ed. *Information visualization: perception for design*. Third Edition. Elsevier, 2012.
- [135] P. Warren, P. Mulholland, T. Collins, and E. Motta. “Using Ontologies - Understanding the user experience”. In: *Proceedings of the 19th International Conference on Knowledge Engineering and Knowledge Management (EKAW '14)*. 2014, pp. 579–590.

-
- [136] C. Welty. “Ontology research”. In: *AI magazine* 24.3 (2003), pp. 11–12.
- [137] T. Wolbers and M. Hegarty. “What determines our navigational abilities?” In: *Trends in Cognitive Sciences* 14 (2010), pp. 138–146.
- [138] H. Zhang and G. Salvendy. “The implications of visualization ability and structure preview design for Web information search tasks”. In: *Human-Computer Interaction* 13 (2001), pp. 75–95.
- [139] Y. Zhang. “Searching for specific health-related information in MedlinePlus: Behavioral patterns and user experience”. In: *Association for Information Science and Technology* 65.1 (2014), pp. 53–68.

Department of Computer and Information Science
Linköpings universitet

Dissertations

Linköping Studies in Science and Technology

Linköping Studies in Arts and Science

Linköping Studies in Statistics

Linköping Studies in Information Science

Linköping Studies in Science and Technology

- No 14 **Anders Haraldsson:** A Program Manipulation System Based on Partial Evaluation, 1977, ISBN 91-7372-144-1.
- No 17 **Bengt Magnhagen:** Probability Based Verification of Time Margins in Digital Designs, 1977, ISBN 91-7372-157-3.
- No 18 **Mats Cedwall:** Semantisk analys av processbeskrivningar i naturligt språk, 1977, ISBN 91-7372-168-9.
- No 22 **Jaak Urmi:** A Machine Independent LISP Compiler and its Implications for Ideal Hardware, 1978, ISBN 91-7372-188-3.
- No 33 **Tore Risch:** Compilation of Multiple File Queries in a Meta-Database System, 1978, ISBN 91-7372-232-4.
- No 51 **Erland Jungert:** Synthesizing Database Structures from a User Oriented Data Model, 1980, ISBN 91-7372-387-8.
- No 54 **Sture Hägglund:** Contributions to the Development of Methods and Tools for Interactive Design of Applications Software, 1980, ISBN 91-7372-404-1.
- No 55 **Pär Emanuelson:** Performance Enhancement in a Well-Structured Pattern Matcher through Partial Evaluation, 1980, ISBN 91-7372-403-3.
- No 58 **Bengt Johnsson, Bertil Andersson:** The Human-Computer Interface in Commercial Systems, 1981, ISBN 91-7372-414-9.
- No 69 **H. Jan Komorowski:** A Specification of an Abstract Prolog Machine and its Application to Partial Evaluation, 1981, ISBN 91-7372-479-3.
- No 71 **René Reboh:** Knowledge Engineering Techniques and Tools for Expert Systems, 1981, ISBN 91-7372-489-0.
- No 77 **Östen Oskarsson:** Mechanisms of Modifiability in large Software Systems, 1982, ISBN 91-7372-527-7.
- No 94 **Hans Lunell:** Code Generator Writing Systems, 1983, ISBN 91-7372-652-4.
- No 97 **Andrzej Lingas:** Advances in Minimum Weight Triangulation, 1983, ISBN 91-7372-660-5.
- No 109 **Peter Fritzon:** Towards a Distributed Programming Environment based on Incremental Compilation, 1984, ISBN 91-7372-801-2.
- No 111 **Erik Tengvald:** The Design of Expert Planning Systems. An Experimental Operations Planning System for Turning, 1984, ISBN 91-7372-805-5.
- No 155 **Christos Levcopoulos:** Heuristics for Minimum Decompositions of Polygons, 1987, ISBN 91-7870-133-3.
- No 165 **James W. Goodwin:** A Theory and System for Non-Monotonic Reasoning, 1987, ISBN 91-7870-183-X.
- No 170 **Zebo Peng:** A Formal Methodology for Automated Synthesis of VLSI Systems, 1987, ISBN 91-7870-225-9.
- No 174 **Johan Fagerström:** A Paradigm and System for Design of Distributed Systems, 1988, ISBN 91-7870-301-8.
- No 192 **Dimitar Driankov:** Towards a Many Valued Logic of Quantified Belief, 1988, ISBN 91-7870-374-3.
- No 213 **Lin Padgham:** Non-Monotonic Inheritance for an Object Oriented Knowledge Base, 1989, ISBN 91-7870-485-5.
- No 214 **Tony Larsson:** A Formal Hardware Description and Verification Method, 1989, ISBN 91-7870-517-7.
- No 221 **Michael Reinfrank:** Fundamentals and Logical Foundations of Truth Maintenance, 1989, ISBN 91-7870-546-0.
- No 239 **Jonas Löwgren:** Knowledge-Based Design Support and Discourse Management in User Interface Management Systems, 1991, ISBN 91-7870-720-X.
- No 244 **Henrik Eriksson:** Meta-Tool Support for Knowledge Acquisition, 1991, ISBN 91-7870-746-3.
- No 252 **Peter Eklund:** An Epistemic Approach to Interactive Design in Multiple Inheritance Hierarchies, 1991, ISBN 91-7870-784-6.
- No 258 **Patrick Doherty:** NML3 - A Non-Monotonic Formalism with Explicit Defaults, 1991, ISBN 91-7870-816-8.
- No 260 **Nahid Shahmehri:** Generalized Algorithmic Debugging, 1991, ISBN 91-7870-828-1.
- No 264 **Nils Dahlbäck:** Representation of Discourse-Cognitive and Computational Aspects, 1992, ISBN 91-7870-850-8.
- No 265 **Ulf Nilsson:** Abstract Interpretations and Abstract Machines: Contributions to a Methodology for the Implementation of Logic Programs, 1992, ISBN 91-7870-858-3.
- No 270 **Ralph Rönnquist:** Theory and Practice of Tense-bound Object References, 1992, ISBN 91-7870-873-7.
- No 273 **Björn Fjellborg:** Pipeline Extraction for VLSI Data Path Synthesis, 1992, ISBN 91-7870-880-X.
- No 276 **Staffan Bonnier:** A Formal Basis for Horn Clause Logic with External Polymorphic Functions, 1992, ISBN 91-7870-896-6.
- No 277 **Kristian Sandahl:** Developing Knowledge Management Systems with an Active Expert Methodology, 1992, ISBN 91-7870-897-4.
- No 281 **Christer Bäckström:** Computational Complexity of Reasoning about Plans, 1992, ISBN 91-7870-979-2.
- No 292 **Mats Wirén:** Studies in Incremental Natural Language Analysis, 1992, ISBN 91-7871-027-8.
- No 297 **Mariam Kamkar:** Interprocedural Dynamic Slicing with Applications to Debugging and Testing, 1993, ISBN 91-7871-065-0.
- No 302 **Tingting Zhang:** A Study in Diagnosis Using Classification and Defaults, 1993, ISBN 91-7871-078-2.
- No 312 **Arne Jönsson:** Dialogue Management for Natural Language Interfaces - An Empirical Approach, 1993, ISBN 91-7871-110-X.
- No 338 **Simin Nadjm-Tehrani:** Reactive Systems in Physical Environments: Compositional Modelling and Framework for Verification, 1994, ISBN 91-7871-237-8.
- No 371 **Bengt Savén:** Business Models for Decision Support and Learning. A Study of Discrete-Event Manufacturing Simulation at Asea/ABB 1968-1993, 1995, ISBN 91-7871-494-X.

- No 375 **Ulf Söderman**: Conceptual Modelling of Mode Switching Physical Systems, 1995, ISBN 91-7871-516-4.
- No 383 **Andreas Kägedal**: Exploiting Groundness in Logic Programs, 1995, ISBN 91-7871-538-5.
- No 396 **George Fodor**: Ontological Control, Description, Identification and Recovery from Problematic Control Situations, 1995, ISBN 91-7871-603-9.
- No 413 **Mikael Pettersson**: Compiling Natural Semantics, 1995, ISBN 91-7871-641-1.
- No 414 **Xinli Gu**: RT Level Testability Improvement by Testability Analysis and Transformations, 1996, ISBN 91-7871-654-3.
- No 416 **Hua Shu**: Distributed Default Reasoning, 1996, ISBN 91-7871-665-9.
- No 429 **Jaime Villegas**: Simulation Supported Industrial Training from an Organisational Learning Perspective - Development and Evaluation of the SSIT Method, 1996, ISBN 91-7871-700-0.
- No 431 **Peter Jonsson**: Studies in Action Planning: Algorithms and Complexity, 1996, ISBN 91-7871-704-3.
- No 437 **Johan Boye**: Directional Types in Logic Programming, 1996, ISBN 91-7871-725-6.
- No 439 **Cecilia Sjöberg**: Activities, Voices and Arenas: Participatory Design in Practice, 1996, ISBN 91-7871-728-0.
- No 448 **Patrick Lambrix**: Part-Whole Reasoning in Description Logics, 1996, ISBN 91-7871-820-1.
- No 452 **Kjell Orsborn**: On Extensible and Object-Relational Database Technology for Finite Element Analysis Applications, 1996, ISBN 91-7871-827-9.
- No 459 **Olof Johansson**: Development Environments for Complex Product Models, 1996, ISBN 91-7871-855-4.
- No 461 **Lena Strömbäck**: User-Defined Constructions in Unification-Based Formalisms, 1997, ISBN 91-7871-857-0.
- No 462 **Lars Degerstedt**: Tabulation-based Logic Programming: A Multi-Level View of Query Answering, 1996, ISBN 91-7871-858-9.
- No 475 **Fredrik Nilsson**: Strategi och ekonomisk styrning - En studie av hur ekonomiska styrsystem utformas och används efter företagsförvärv, 1997, ISBN 91-7871-914-3.
- No 480 **Mikael Lindvall**: An Empirical Study of Requirements-Driven Impact Analysis in Object-Oriented Software Evolution, 1997, ISBN 91-7871-927-5.
- No 485 **Göran Forslund**: Opinion-Based Systems: The Cooperative Perspective on Knowledge-Based Decision Support, 1997, ISBN 91-7871-938-0.
- No 494 **Martin Sköld**: Active Database Management Systems for Monitoring and Control, 1997, ISBN 91-7219-002-7.
- No 495 **Hans Olsén**: Automatic Verification of Petri Nets in a CLP framework, 1997, ISBN 91-7219-011-6.
- No 498 **Thomas Drakengren**: Algorithms and Complexity for Temporal and Spatial Formalisms, 1997, ISBN 91-7219-019-1.
- No 502 **Jakob Axelsson**: Analysis and Synthesis of Heterogeneous Real-Time Systems, 1997, ISBN 91-7219-035-3.
- No 503 **Johan Ringström**: Compiler Generation for Data-Parallel Programming Languages from Two-Level Semantics Specifications, 1997, ISBN 91-7219-045-0.
- No 512 **Anna Moberg**: Närhet och distans - Studier av kommunikationsmönster i satellitkontor och flexibla kontor, 1997, ISBN 91-7219-119-8.
- No 520 **Mikael Ronström**: Design and Modelling of a Parallel Data Server for Telecom Applications, 1998, ISBN 91-7219-169-4.
- No 522 **Niclas Ohlsson**: Towards Effective Fault Prevention - An Empirical Study in Software Engineering, 1998, ISBN 91-7219-176-7.
- No 526 **Jochim Karlsson**: A Systematic Approach for Prioritizing Software Requirements, 1998, ISBN 91-7219-184-8.
- No 530 **Henrik Nilsson**: Declarative Debugging for Lazy Functional Languages, 1998, ISBN 91-7219-197-X.
- No 555 **Jonas Hallberg**: Timing Issues in High-Level Synthesis, 1998, ISBN 91-7219-369-7.
- No 561 **Ling Lin**: Management of 1-D Sequence Data - From Discrete to Continuous, 1999, ISBN 91-7219-402-2.
- No 563 **Eva L Ragnemalm**: Student Modelling based on Collaborative Dialogue with a Learning Companion, 1999, ISBN 91-7219-412-X.
- No 567 **Jörgen Lindström**: Does Distance matter? On geographical dispersion in organisations, 1999, ISBN 91-7219-439-1.
- No 582 **Vanja Josifovski**: Design, Implementation and Evaluation of a Distributed Mediator System for Data Integration, 1999, ISBN 91-7219-482-0.
- No 589 **Rita Kovordányi**: Modeling and Simulating Inhibitory Mechanisms in Mental Image Reinterpretation - Towards Cooperative Human-Computer Creativity, 1999, ISBN 91-7219-506-1.
- No 592 **Mikael Ericsson**: Supporting the Use of Design Knowledge - An Assessment of Commenting Agents, 1999, ISBN 91-7219-532-0.
- No 593 **Lars Karlsson**: Actions, Interactions and Narratives, 1999, ISBN 91-7219-534-7.
- No 594 **C. G. Mikael Johansson**: Social and Organizational Aspects of Requirements Engineering Methods - A practice-oriented approach, 1999, ISBN 91-7219-541-X.
- No 595 **Jörgen Hansson**: Value-Driven Multi-Class Overload Management in Real-Time Database Systems, 1999, ISBN 91-7219-542-8.
- No 596 **Niklas Hallberg**: Incorporating User Values in the Design of Information Systems and Services in the Public Sector: A Methods Approach, 1999, ISBN 91-7219-543-6.
- No 597 **Vivian Vimarlund**: An Economic Perspective on the Analysis of Impacts of Information Technology: From Case Studies in Health-Care towards General Models and Theories, 1999, ISBN 91-7219-544-4.
- No 598 **Johan Jenvald**: Methods and Tools in Computer-Supported Taskforce Training, 1999, ISBN 91-7219-547-9.
- No 607 **Magnus Merkel**: Understanding and enhancing translation by parallel text processing, 1999, ISBN 91-7219-614-9.
- No 611 **Silvia Coradeschi**: Anchoring symbols to sensory data, 1999, ISBN 91-7219-623-8.
- No 613 **Man Lin**: Analysis and Synthesis of Reactive Systems: A Generic Layered Architecture Perspective, 1999, ISBN 91-7219-630-0.
- No 618 **Jimmy Tjäder**: Systemimplementering i praktiken - En studie av logiker i fyra projekt, 1999, ISBN 91-7219-657-2.
- No 627 **Vadim Engelson**: Tools for Design, Interactive Simulation, and Visualization of Object-Oriented Models in Scientific Computing, 2000, ISBN 91-7219-709-9.

- No 637 **Esa Falkenroth:** Database Technology for Control and Simulation, 2000, ISBN 91-7219-766-8.
- No 639 **Per-Arne Persson:** Bringing Power and Knowledge Together: Information Systems Design for Autonomy and Control in Command Work, 2000, ISBN 91-7219-796-X.
- No 660 **Erik Larsson:** An Integrated System-Level Design for Testability Methodology, 2000, ISBN 91-7219-890-7.
- No 688 **Marcus Bjärelund:** Model-based Execution Monitoring, 2001, ISBN 91-7373-016-5.
- No 689 **Joakim Gustafsson:** Extending Temporal Action Logic, 2001, ISBN 91-7373-017-3.
- No 720 **Carl-Johan Petri:** Organizational Information Provision - Managing Mandatory and Discretionary Use of Information Technology, 2001, ISBN 91-7373-126-9.
- No 724 **Paul Scerri:** Designing Agents for Systems with Adjustable Autonomy, 2001, ISBN 91-7373-207-9.
- No 725 **Tim Heyer:** Semantic Inspection of Software Artifacts: From Theory to Practice, 2001, ISBN 91-7373-208-7.
- No 726 **Pär Carlshamre:** A Usability Perspective on Requirements Engineering - From Methodology to Product Development, 2001, ISBN 91-7373-212-5.
- No 732 **Juha Takkinen:** From Information Management to Task Management in Electronic Mail, 2002, ISBN 91-7373-258-3.
- No 745 **Johan Åberg:** Live Help Systems: An Approach to Intelligent Help for Web Information Systems, 2002, ISBN 91-7373-311-3.
- No 746 **Rego Granlund:** Monitoring Distributed Teamwork Training, 2002, ISBN 91-7373-312-1.
- No 757 **Henrik André-Jönsson:** Indexing Strategies for Time Series Data, 2002, ISBN 91-7373-346-6.
- No 747 **Anneli Hagdahl:** Development of IT-supported Interorganisational Collaboration - A Case Study in the Swedish Public Sector, 2002, ISBN 91-7373-314-8.
- No 749 **Sofie Pilemalm:** Information Technology for Non-Profit Organisations - Extended Participatory Design of an Information System for Trade Union Shop Stewards, 2002, ISBN 91-7373-318-0.
- No 765 **Stefan Holmlid:** Adapting users: Towards a theory of use quality, 2002, ISBN 91-7373-397-0.
- No 771 **Magnus Morin:** Multimedia Representations of Distributed Tactical Operations, 2002, ISBN 91-7373-421-7.
- No 772 **Pawel Pietrzak:** A Type-Based Framework for Locating Errors in Constraint Logic Programs, 2002, ISBN 91-7373-422-5.
- No 758 **Erik Berglund:** Library Communication Among Programmers Worldwide, 2002, ISBN 91-7373-349-0.
- No 774 **Choong-ho Yi:** Modelling Object-Oriented Dynamic Systems Using a Logic-Based Framework, 2002, ISBN 91-7373-424-1.
- No 779 **Mathias Broxvall:** A Study in the Computational Complexity of Temporal Reasoning, 2002, ISBN 91-7373-440-3.
- No 793 **Asmus Pandikow:** A Generic Principle for Enabling Interoperability of Structured and Object-Oriented Analysis and Design Tools, 2002, ISBN 91-7373-479-9.
- No 785 **Lars Hult:** Publika Informationstjänster. En studie av den Internetbaserade encyklopedins bruksegenskap, 2003, ISBN 91-7373-461-6.
- No 800 **Lars Taxén:** A Framework for the Coordination of Complex Systems' Development, 2003, ISBN 91-7373-604-X.
- No 808 **Klas Gäre:** Tre perspektiv på förväntningar och förändringar i samband med införande av informationssystem, 2003, ISBN 91-7373-618-X.
- No 821 **Mikael Kindborg:** Concurrent Comics - programming of social agents by children, 2003, ISBN 91-7373-651-1.
- No 823 **Christina Ölvingsson:** On Development of Information Systems with GIS Functionality in Public Health Informatics: A Requirements Engineering Approach, 2003, ISBN 91-7373-656-2.
- No 828 **Tobias Ritzau:** Memory Efficient Hard Real-Time Garbage Collection, 2003, ISBN 91-7373-666-X.
- No 833 **Paul Pop:** Analysis and Synthesis of Communication-Intensive Heterogeneous Real-Time Systems, 2003, ISBN 91-7373-683-X.
- No 852 **Johan Moe:** Observing the Dynamic Behaviour of Large Distributed Systems to Improve Development and Testing - An Empirical Study in Software Engineering, 2003, ISBN 91-7373-779-8.
- No 867 **Erik Herzog:** An Approach to Systems Engineering Tool Data Representation and Exchange, 2004, ISBN 91-7373-929-4.
- No 872 **Aseel Berglund:** Augmenting the Remote Control: Studies in Complex Information Navigation for Digital TV, 2004, ISBN 91-7373-940-5.
- No 869 **Jö Skåmedal:** Telecommuting's Implications on Travel and Travel Patterns, 2004, ISBN 91-7373-935-9.
- No 870 **Linda Askenäs:** The Roles of IT - Studies of Organising when Implementing and Using Enterprise Systems, 2004, ISBN 91-7373-936-7.
- No 874 **Annika Flycht-Eriksson:** Design and Use of Ontologies in Information-Providing Dialogue Systems, 2004, ISBN 91-7373-947-2.
- No 873 **Peter Bunus:** Debugging Techniques for Equation-Based Languages, 2004, ISBN 91-7373-941-3.
- No 876 **Jonas Mellin:** Resource-Predictable and Efficient Monitoring of Events, 2004, ISBN 91-7373-956-1.
- No 883 **Magnus Bång:** Computing at the Speed of Paper: Ubiquitous Computing Environments for Healthcare Professionals, 2004, ISBN 91-7373-971-5.
- No 882 **Robert Eklund:** Disfluency in Swedish human-human and human-machine travel booking dialogues, 2004, ISBN 91-7373-966-9.
- No 887 **Anders Lindström:** English and other Foreign Linguistic Elements in Spoken Swedish. Studies of Productive Processes and their Modelling using Finite-State Tools, 2004, ISBN 91-7373-981-2.
- No 889 **Zhiping Wang:** Capacity-Constrained Production-inventory systems - Modelling and Analysis in both a traditional and an e-business context, 2004, ISBN 91-85295-08-6.
- No 893 **Pernilla Qvarfordt:** Eyes on Multimodal Interaction, 2004, ISBN 91-85295-30-2.
- No 910 **Magnus Kald:** In the Borderland between Strategy and Management Control - Theoretical Framework and Empirical Evidence, 2004, ISBN 91-85295-82-5.
- No 918 **Jonas Lundberg:** Shaping Electronic News: Genre Perspectives on Interaction Design, 2004, ISBN 91-85297-14-3.
- No 900 **Mattias Arvola:** Shades of use: The dynamics of interaction design for sociable use, 2004, ISBN 91-85295-42-6.
- No 920 **Luis Alejandro Cortés:** Verification and Scheduling Techniques for Real-Time Embedded Systems, 2004, ISBN 91-85297-21-6.
- No 929 **Diana Szentivanyi:** Performance Studies of Fault-Tolerant Middleware, 2005, ISBN 91-85297-58-5.

- No 933 **Mikael Cäker:** Management Accounting as Constructing and Opposing Customer Focus: Three Case Studies on Management Accounting and Customer Relations, 2005, ISBN 91-85297-64-X.
- No 937 **Jonas Kvarnström:** TALplanner and Other Extensions to Temporal Action Logic, 2005, ISBN 91-85297-75-5.
- No 938 **Bourhane Kadmiry:** Fuzzy Gain-Scheduled Visual Servoing for Unmanned Helicopter, 2005, ISBN 91-85297-76-3.
- No 945 **Gert Jervan:** Hybrid Built-In Self-Test and Test Generation Techniques for Digital Systems, 2005, ISBN 91-85297-97-6.
- No 946 **Anders Arpteg:** Intelligent Semi-Structured Information Extraction, 2005, ISBN 91-85297-98-4.
- No 947 **Ola Angelmark:** Constructing Algorithms for Constraint Satisfaction and Related Problems - Methods and Applications, 2005, ISBN 91-85297-99-2.
- No 963 **Calin Curescu:** Utility-based Optimisation of Resource Allocation for Wireless Networks, 2005, ISBN 91-85457-07-8.
- No 972 **Björn Johansson:** Joint Control in Dynamic Situations, 2005, ISBN 91-85457-31-0.
- No 974 **Dan Lawesson:** An Approach to Diagnosability Analysis for Interacting Finite State Systems, 2005, ISBN 91-85457-39-6.
- No 979 **Claudiu Duma:** Security and Trust Mechanisms for Groups in Distributed Services, 2005, ISBN 91-85457-54-X.
- No 983 **Sorin Manolache:** Analysis and Optimisation of Real-Time Systems with Stochastic Behaviour, 2005, ISBN 91-85457-60-4.
- No 986 **Yuxiao Zhao:** Standards-Based Application Integration for Business-to-Business Communications, 2005, ISBN 91-85457-66-3.
- No 1004 **Patrik Haslum:** Admissible Heuristics for Automated Planning, 2006, ISBN 91-85497-28-2.
- No 1005 **Aleksandra Tešanovic:** Developing Reusable and Reconfigurable Real-Time Software using Aspects and Components, 2006, ISBN 91-85497-29-0.
- No 1008 **David Dinka:** Role, Identity and Work: Extending the design and development agenda, 2006, ISBN 91-85497-42-8.
- No 1009 **Iakov Nakhimovski:** Contributions to the Modeling and Simulation of Mechanical Systems with Detailed Contact Analysis, 2006, ISBN 91-85497-43-X.
- No 1013 **Wilhelm Dahllöf:** Exact Algorithms for Exact Satisfiability Problems, 2006, ISBN 91-85523-97-6.
- No 1016 **Levon Saldamli:** PDEModelica - A High-Level Language for Modeling with Partial Differential Equations, 2006, ISBN 91-85523-84-4.
- No 1017 **Daniel Karlsson:** Verification of Component-based Embedded System Designs, 2006, ISBN 91-85523-79-8.
- No 1018 **Ioan Chisalița:** Communication and Networking Techniques for Traffic Safety Systems, 2006, ISBN 91-85523-77-1.
- No 1019 **Tarja Susi:** The Puzzle of Social Activity - The Significance of Tools in Cognition and Cooperation, 2006, ISBN 91-85523-71-2.
- No 1021 **Andrzej Bednarski:** Integrated Optimal Code Generation for Digital Signal Processors, 2006, ISBN 91-85523-69-0.
- No 1022 **Peter Aronsson:** Automatic Parallelization of Equation-Based Simulation Programs, 2006, ISBN 91-85523-68-2.
- No 1030 **Robert Nilsson:** A Mutation-based Framework for Automated Testing of Timeliness, 2006, ISBN 91-85523-35-6.
- No 1034 **Jon Edvardsson:** Techniques for Automatic Generation of Tests from Programs and Specifications, 2006, ISBN 91-85523-31-3.
- No 1035 **Vaida Jakoniene:** Integration of Biological Data, 2006, ISBN 91-85523-28-3.
- No 1045 **Genevieve Gorrell:** Generalized Hebbian Algorithms for Dimensionality Reduction in Natural Language Processing, 2006, ISBN 91-85643-88-2.
- No 1051 **Yu-Hsing Huang:** Having a New Pair of Glasses - Applying Systemic Accident Models on Road Safety, 2006, ISBN 91-85643-64-5.
- No 1054 **Åsa Hedenskog:** Perceive those things which cannot be seen - A Cognitive Systems Engineering perspective on requirements management, 2006, ISBN 91-85643-57-2.
- No 1061 **Cécile Åberg:** An Evaluation Platform for Semantic Web Technology, 2007, ISBN 91-85643-31-9.
- No 1073 **Mats Grindal:** Handling Combinatorial Explosion in Software Testing, 2007, ISBN 978-91-85715-74-9.
- No 1075 **Almut Herzog:** Usable Security Policies for Runtime Environments, 2007, ISBN 978-91-85715-65-7.
- No 1079 **Magnus Wahlström:** Algorithms, measures, and upper bounds for Satisfiability and related problems, 2007, ISBN 978-91-85715-55-8.
- No 1083 **Jesper Andersson:** Dynamic Software Architectures, 2007, ISBN 978-91-85715-46-6.
- No 1086 **Ulf Johansson:** Obtaining Accurate and Comprehensive Data Mining Models - An Evolutionary Approach, 2007, ISBN 978-91-85715-34-3.
- No 1089 **Traian Pop:** Analysis and Optimisation of Distributed Embedded Systems with Heterogeneous Scheduling Policies, 2007, ISBN 978-91-85715-27-5.
- No 1091 **Gustav Nordh:** Complexity Dichotomies for CSP-related Problems, 2007, ISBN 978-91-85715-20-6.
- No 1106 **Per Ola Kristensson:** Discrete and Continuous Shape Writing for Text Entry and Control, 2007, ISBN 978-91-85831-77-7.
- No 1110 **He Tan:** Aligning Biomedical Ontologies, 2007, ISBN 978-91-85831-56-2.
- No 1112 **Jessica Lindblom:** Minding the body - Interacting socially through embodied action, 2007, ISBN 978-91-85831-48-7.
- No 1113 **Pontus Wärnestål:** Dialogue Behavior Management in Conversational Recommender Systems, 2007, ISBN 978-91-85831-47-0.
- No 1120 **Thomas Gustafsson:** Management of Real-Time Data Consistency and Transient Overloads in Embedded Systems, 2007, ISBN 978-91-85831-33-3.
- No 1127 **Alexandru Andrei:** Energy Efficient and Predictable Design of Real-time Embedded Systems, 2007, ISBN 978-91-85831-06-7.
- No 1139 **Per Wikberg:** Eliciting Knowledge from Experts in Modeling of Complex Systems: Managing Variation and Interactions, 2007, ISBN 978-91-85895-66-3.
- No 1143 **Mehdi Amirjoo:** QoS Control of Real-Time Data Services under Uncertain Workload, 2007, ISBN 978-91-85895-49-6.
- No 1150 **Sanny Syberfeldt:** Optimistic Replication with Forward Conflict Resolution in Distributed Real-Time Databases, 2007, ISBN 978-91-85895-27-4.
- No 1155 **Beatrice Alenljung:** Envisioning a Future Decision Support System for Requirements Engineering - A Holistic and Human-centred Perspective, 2008, ISBN 978-91-85895-11-3.

- No 1156 **Artur Wilk:** Types for XML with Application to Xcerpt, 2008, ISBN 978-91-85895-08-3.
- No 1183 **Adrian Pop:** Integrated Model-Driven Development Environments for Equation-Based Object-Oriented Languages, 2008, ISBN 978-91-7393-895-2.
- No 1185 **Jörgen Skågeby:** Gifting Technologies - Ethnographic Studies of End-users and Social Media Sharing, 2008, ISBN 978-91-7393-892-1.
- No 1187 **Imad-Eldin Ali Abugessaisa:** Analytical tools and information-sharing methods supporting road safety organizations, 2008, ISBN 978-91-7393-887-7.
- No 1204 **H. Joe Steinhauer:** A Representation Scheme for Description and Reconstruction of Object Configurations Based on Qualitative Relations, 2008, ISBN 978-91-7393-823-5.
- No 1222 **Anders Larsson:** Test Optimization for Core-based System-on-Chip, 2008, ISBN 978-91-7393-768-9.
- No 1238 **Andreas Borg:** Processes and Models for Capacity Requirements in Telecommunication Systems, 2009, ISBN 978-91-7393-700-9.
- No 1240 **Fredrik Heintz:** DyKnow: A Stream-Based Knowledge Processing Middleware Framework, 2009, ISBN 978-91-7393-696-5.
- No 1241 **Birgitta Lindström:** Testability of Dynamic Real-Time Systems, 2009, ISBN 978-91-7393-695-8.
- No 1244 **Eva Blomqvist:** Semi-automatic Ontology Construction based on Patterns, 2009, ISBN 978-91-7393-683-5.
- No 1249 **Rogier Woltjer:** Functional Modeling of Constraint Management in Aviation Safety and Command and Control, 2009, ISBN 978-91-7393-659-0.
- No 1260 **Gianpaolo Conte:** Vision-Based Localization and Guidance for Unmanned Aerial Vehicles, 2009, ISBN 978-91-7393-603-3.
- No 1262 **AnnMarie Ericsson:** Enabling Tool Support for Formal Analysis of ECA Rules, 2009, ISBN 978-91-7393-598-2.
- No 1266 **Jiri Trnka:** Exploring Tactical Command and Control: A Role-Playing Simulation Approach, 2009, ISBN 978-91-7393-571-5.
- No 1268 **Bahlol Rahimi:** Supporting Collaborative Work through ICT - How End-users Think of and Adopt Integrated Health Information Systems, 2009, ISBN 978-91-7393-550-0.
- No 1274 **Fredrik Kuivinen:** Algorithms and Hardness Results for Some Valued CSPs, 2009, ISBN 978-91-7393-525-8.
- No 1281 **Gunnar Mathiason:** Virtual Full Replication for Scalable Distributed Real-Time Databases, 2009, ISBN 978-91-7393-503-6.
- No 1290 **Viacheslav Izosimov:** Scheduling and Optimization of Fault-Tolerant Distributed Embedded Systems, 2009, ISBN 978-91-7393-482-4.
- No 1294 **Johan Thapper:** Aspects of a Constraint Optimisation Problem, 2010, ISBN 978-91-7393-464-0.
- No 1306 **Susanna Nilsson:** Augmentation in the Wild: User Centered Development and Evaluation of Augmented Reality Applications, 2010, ISBN 978-91-7393-416-9.
- No 1313 **Christer Thörn:** On the Quality of Feature Models, 2010, ISBN 978-91-7393-394-0.
- No 1321 **Zhiyuan He:** Temperature Aware and Defect-Probability Driven Test Scheduling for System-on-Chip, 2010, ISBN 978-91-7393-378-0.
- No 1333 **David Broman:** Meta-Languages and Semantics for Equation-Based Modeling and Simulation, 2010, ISBN 978-91-7393-335-3.
- No 1337 **Alexander Siemers:** Contributions to Modelling and Visualisation of Multibody Systems Simulations with Detailed Contact Analysis, 2010, ISBN 978-91-7393-317-9.
- No 1354 **Mikael Asplund:** Disconnected Discoveries: Availability Studies in Partitioned Networks, 2010, ISBN 978-91-7393-278-3.
- No 1359 **Jana Rambusch:** Mind Games Extended: Understanding Gameplay as Situated Activity, 2010, ISBN 978-91-7393-252-3.
- No 1373 **Sonia Sangari:** Head Movement Correlates to Focus Assignment in Swedish, 2011, ISBN 978-91-7393-154-0.
- No 1374 **Jan-Erik Källhammer:** Using False Alarms when Developing Automotive Active Safety Systems, 2011, ISBN 978-91-7393-153-3.
- No 1375 **Mattias Eriksson:** Integrated Code Generation, 2011, ISBN 978-91-7393-147-2.
- No 1381 **Ola Leifler:** Affordances and Constraints of Intelligent Decision Support for Military Command and Control - Three Case Studies of Support Systems, 2011, ISBN 978-91-7393-133-5.
- No 1386 **Soheil Samii:** Quality-Driven Synthesis and Optimization of Embedded Control Systems, 2011, ISBN 978-91-7393-102-1.
- No 1419 **Erik Kuiper:** Geographic Routing in Intermittently-connected Mobile Ad Hoc Networks: Algorithms and Performance Models, 2012, ISBN 978-91-7519-981-8.
- No 1451 **Sara Stymne:** Text Harmonization Strategies for Phrase-Based Statistical Machine Translation, 2012, ISBN 978-91-7519-887-3.
- No 1455 **Alberto Montebelli:** Modeling the Role of Energy Management in Embodied Cognition, 2012, ISBN 978-91-7519-882-8.
- No 1465 **Mohammad Saifullah:** Biologically-Based Interactive Neural Network Models for Visual Attention and Object Recognition, 2012, ISBN 978-91-7519-838-5.
- No 1490 **Tomas Bengtsson:** Testing and Logic Optimization Techniques for Systems on Chip, 2012, ISBN 978-91-7519-742-5.
- No 1481 **David Byers:** Improving Software Security by Preventing Known Vulnerabilities, 2012, ISBN 978-91-7519-784-5.
- No 1496 **Tommy Färnqvist:** Exploiting Structure in CSP-related Problems, 2013, ISBN 978-91-7519-711-1.
- No 1503 **John Wilander:** Contributions to Specification, Implementation, and Execution of Secure Software, 2013, ISBN 978-91-7519-681-7.
- No 1506 **Magnus Ingmarsson:** Creating and Enabling the Useful Service Discovery Experience, 2013, ISBN 978-91-7519-662-6.
- No 1547 **Wladimir Schamai:** Model-Based Verification of Dynamic System Behavior against Requirements: Method, Language, and Tool, 2013, ISBN 978-91-7519-505-6.
- No 1551 **Henrik Svensson:** Simulations, 2013, ISBN 978-91-7519-491-2.
- No 1559 **Sergiu Rafilciu:** Stability of Adaptive Distributed Real-Time Systems with Dynamic Resource Management, 2013, ISBN 978-91-7519-471-4.
- No 1581 **Usman Dastgeer:** Performance-aware Component Composition for GPU-based Systems, 2014, ISBN 978-91-7519-383-0.
- No 1602 **Cai Li:** Reinforcement Learning of Locomotion based on Central Pattern Generators, 2014, ISBN 978-91-7519-313-7.
- No 1652 **Roland Samlaus:** An Integrated Development Environment with Enhanced Domain-Specific

- Interactive Model Validation, 2015, ISBN 978-91-7519-090-7.
- No 1663 **Hannes Uppman**: On Some Combinatorial Optimization Problems: Algorithms and Complexity, 2015, ISBN 978-91-7519-072-3.
- No 1664 **Martin Sjölund**: Tools and Methods for Analysis, Debugging, and Performance Improvement of Equation-Based Models, 2015, ISBN 978-91-7519-071-6.
- No 1666 **Kristian Stavåker**: Contributions to Simulation of Modelica Models on Data-Parallel Multi-Core Architectures, 2015, ISBN 978-91-7519-068-6.
- No 1680 **Adrian Lifa**: Hardware/Software Codesign of Embedded Systems with Reconfigurable and Heterogeneous Platforms, 2015, ISBN 978-91-7519-040-2.
- No 1685 **Bogdan Tanasa**: Timing Analysis of Distributed Embedded Systems with Stochastic Workload and Reliability Constraints, 2015, ISBN 978-91-7519-022-8.
- No 1691 **Håkan Warnquist**: Troubleshooting Trucks – Automated Planning and Diagnosis, 2015, ISBN 978-91-7685-993-3.
- No 1702 **Nima Aghaee**: Thermal Issues in Testing of Advanced Systems on Chip, 2015, ISBN 978-91-7685-949-0.
- No 1715 **Maria Vasilevskaya**: Security in Embedded Systems: A Model-Based Approach with Risk Metrics, 2015, ISBN 978-91-7685-917-9.
- No 1729 **Ke Jiang**: Security-Driven Design of Real-Time Embedded System, 2016, ISBN 978-91-7685-884-4.
- No 1733 **Victor Lagerkvist**: Strong Partial Clones and the Complexity of Constraint Satisfaction Problems: Limitations and Applications, 2016, ISBN 978-91-7685-856-1.
- No 1734 **Chandan Roy**: An Informed System Development Approach to Tropical Cyclone Track and Intensity Forecasting, 2016, ISBN 978-91-7685-854-7.
- No 1746 **Amir Aminifar**: Analysis, Design, and Optimization of Embedded Control Systems, 2016, ISBN 978-91-7685-826-4.
- No 1747 **Ekhlotz Vergara**: Energy Modelling and Fairness for Efficient Mobile Communication, 2016, ISBN 978-91-7685-822-6.
- No 1748 **Dag Sonntag**: Chain Graphs – Interpretations, Expressiveness and Learning Algorithms, 2016, ISBN 978-91-7685-818-9.
- No 1768 **Anna Vapen**: Web Authentication using Third-Parties in Untrusted Environments, 2016, ISBN 978-91-7685-753-3.
- No 1778 **Magnus Jandinger**: On a Need to Know Basis: A Conceptual and Methodological Framework for Modelling and Analysis of Information Demand in an Enterprise Context, 2016, ISBN 978-91-7685-713-7.
- No 1798 **Rahul Hiran**: Collaborative Network Security: Targeting Wide-area Routing and Edgework Attacks, 2016, ISBN 978-91-7685-662-8.
- No 1813 **Nicolas Melot**: Algorithms and Framework for Energy Efficient Parallel Stream Computing on Many-Core Architectures, 2016, ISBN 978-91-7685-623-9.
- No 1823 **Amy Rankin**: Making Sense of Adaptations: Resilience in High-Risk Work, 2017, ISBN 978-91-7685-596-6.
- No 1831 **Lisa Malmberg**: Building Design Capability in the Public Sector: Expanding the Horizons of Development, 2017, ISBN 978-91-7685-585-0.
- No 1851 **Marcus Bendtsen**: Gated Bayesian Networks, 2017, ISBN 978-91-7685-525-6.
- No 1852 **Zlatan Dragisic**: Completion of Ontologies and Ontology Networks, 2017, ISBN 978-91-7685-522-5.
- No 1854 **Meysam Aghighi**: Computational Complexity of some Optimization Problems in Planning, 2017, ISBN 978-91-7685-519-5.
- No 1863 **Simon Ståhlberg**: Methods for Detecting Unsolvable Planning Instances using Variable Projection, 2017, ISBN 978-91-7685-498-3.
- No 1879 **Karl Hammar**: Content Ontology Design Patterns: Qualities, Methods, and Tools, 2017, ISBN 978-91-7685-454-9.
- No 1887 **Ivan Ukhov**: System-Level Analysis and Design under Uncertainty, 2017, ISBN 978-91-7685-426-6.
- No 1891 **Valentina Ivanova**: Fostering User Involvement in Ontology Alignment and Alignment Evaluation, 2017, ISBN 978-91-7685-403-7.
- Linköping Studies in Arts and Science**
- No 504 **Ing-Marie Jonsson**: Social and Emotional Characteristics of Speech-based In-Vehicle Information Systems: Impact on Attitude and Driving Behaviour, 2009, ISBN 978-91-7393-478-7.
- No 586 **Fabian Segelström**: Stakeholder Engagement for Service Design: How service designers identify and communicate insights, 2013, ISBN 978-91-7519-554-4.
- No 618 **Johan Blomkvist**: Representing Future Situations of Service: Prototyping in Service Design, 2014, ISBN 978-91-7519-343-4.
- No 620 **Marcus Mast**: Human-Robot Interaction for Semi-Autonomous Assistive Robots, 2014, ISBN 978-91-7519-319-9.
- No 677 **Peter Berggren**: Assessing Shared Strategic Understanding, 2016, ISBN 978-91-7685-786-1.
- No 695 **Mattias Forsblad**: Distributed cognition in home environments: The prospective memory and cognitive practices of older adults, 2016, ISBN 978-91-7685-686-4.
- Linköping Studies in Statistics**
- No 9 **Davood Shahsavani**: Computer Experiments Designed to Explore and Approximate Complex Deterministic Models, 2008, ISBN 978-91-7393-976-8.
- No 10 **Karl Wahlin**: Roadmap for Trend Detection and Assessment of Data Quality, 2008, ISBN 978-91-7393-792-4.
- No 11 **Oleg Sysoev**: Monotonic regression for large multivariate datasets, 2010, ISBN 978-91-7393-412-1.
- No 13 **Agné Burauškaite-Harju**: Characterizing Temporal Change and Inter-Site Correlations in Daily and Sub-daily Precipitation Extremes, 2011, ISBN 978-91-7393-110-6.
- Linköping Studies in Information Science**
- No 1 **Karin Axelsson**: Metodisk systemstrukturerings- och skapa samstämmighet mellan informationssystem-arkitektur och verksamhet, 1998, ISBN 9172-19-296-8.
- No 2 **Stefan Cronholm**: Metodverktyg och användbarhet - en studie av datorstödd metodbaserad systemutveckling, 1998, ISBN 9172-19-299-2.

- No 3 **Anders Avdic:** Användare och utvecklare - om anveckling med kalkylprogram, 1999, ISBN 91-7219-606-8.
- No 4 **Owen Eriksson:** Kommunikationskvalitet hos informationssystem och affärsprocesser, 2000, ISBN 91-7219-811-7.
- No 5 **Mikael Lind:** Från system till process - kriterier för processbestämning vid verksamhetsanalys, 2001, ISBN 91-7373-067-X.
- No 6 **Ulf Melin:** Koordination och informationssystem i företag och nätverk, 2002, ISBN 91-7373-278-8.
- No 7 **Pär J. Ågerfalk:** Information Systems Actability - Understanding Information Technology as a Tool for Business Action and Communication, 2003, ISBN 91-7373-628-7.
- No 8 **Ulf Seigerroth:** Att förstå och förändra systemutvecklingsverksamheter - en taxonomi för metautveckling, 2003, ISBN 91-7373-736-4.
- No 9 **Karin Hedström:** Spår av datoriseringens värden - Effekter av IT i äldreomsorg, 2004, ISBN 91-7373-963-4.
- No 10 **Ewa Braf:** Knowledge Demanded for Action - Studies on Knowledge Mediation in Organisations, 2004, ISBN 91-85295-47-7.
- No 11 **Fredrik Karlsson:** Method Configuration method and computerized tool support, 2005, ISBN 91-85297-48-8.
- No 12 **Malin Nordström:** Styrbar systemförvaltning - Att organisera systemförvaltningsverksamhet med hjälp av effektiva förvaltningsobjekt, 2005, ISBN 91-85297-60-7.
- No 13 **Stefan Holgersson:** Yrke: POLIS - Yrkeskunskap, motivation, IT-system och andra förutsättningar för polisarbete, 2005, ISBN 91-85299-43-X.
- No 14 **Benneth Christiansson, Marie-Therese Christiansson:** Mötet mellan process och komponent - mot ett ramverk för en verksamhetsnära kravspecifikation vid anskaffning av komponentbaserade informationssystem, 2006, ISBN 91-85643-22-X.