HAL open science

Towards ontology matching maturity: contributions to complex, holistic and foundational ontology matching

Cassia Trojahn dos Santos

► To cite this version:

Cassia Trojahn dos Santos. Towards ontology matching maturity: contributions to complex, holistic and foundational ontology matching. Artificial Intelligence [cs.AI]. Université Toulouse 2 Jean Jaurès, 2019. tel-03191561

HAL Id: tel-03191561

https://univ-tlse2.hal.science/tel-03191561

Submitted on 7 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MANUSCRIT

En vue de l'obtention du

HABILITATION À DIRIGER DES RECHERCHES DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *Université Toulouse 2 – Jean Jaurès (UT2J)*

Présentée et soutenue le *Date de défense (12/12/2019)* **par :**

Cassia TROJAHN DOS SANTOS

Towards ontology matching maturity: contributions to complex, holistic and foundational ontology matching

JURY

N. Aussenac-Gilles	DR CNRS	Garante
J. Euzenat	DR INRIA	Examinateur
G. Guizzardi	MCF HDR, Univ. Bolzano	Rapporteur
C. Faron-Zucker	MCF HDR, Univ. Nice	Rapportrice
A. Napoli	DR CNRS	Rapporteur
H. Stuckenschimdt	PR, Univ. Mannheim	Examinateur
N. Pernelle	MCF HDR, Univ. Paris-Sud	Examinatrice

École doctorale et spécialité :
MITT - Mathématiques, Informatique, Télécommunications de Toulouse
Unité de Recherche :
IRIT - Institut de Recherche en Informatique de Toulouse - UMR 5505

À Cécile, Clara et Gilles.
Et à ma maman, source inépuisable d'optimisme.

Acknowledgements

Research is inherently a collaborative activity. Without my colleagues the work I present here would never be possible! I will continue in Portuguese, in order to give a Brazilian accent to this manuscript, as a way of remembering from where I come.

Uma casa em obras, duas filhas pequenas (com seus sistemas imunitários em formação, pais entenderão!) e avós distantes alguns muitos quilômetros (Grenoble, França e Cachoeira do Sul, Rio Grande do Sul, Brazil). Um contexto não muito favoravél para escrever un HDR. Mas isso só foi possível graças à un trabalho de équipe: incansàvel papa Gilles, a fonte de otimistismo que se chama D. Iria (minha mãe), o sorriso das minhas filhas Clara e Cécile, e o apoio dos (muitos) anjos em forma de amigos. Obrigada!

Aos meus sogros Josiane et Pierre, que me socorrem e cuidam das minhas filhas e da minha casa à cada vez que me ausento para uma conferêncía ou reunião de trabalho. Obrigada!

Aos meus orientadores de doutorado e amigos para a vida, Renata Vieira e Paulo Quaresma. Renata, obrigada por compartilhar os alunos, a pesquisa. Isso foi fundamental para o marco científico. Obrigada por estar em Toulouse no momento que eu mais precisava. Obrigada pelas mamadeiras à Clara. Paulo, pelo teu otimismo de sempre, e pelas várias 'passadas' por Toulouse, pela oportunidade de estar do outro lado da sala, 10 anos depois. Isso se chama confiança. Obrigada!

Aos meus alunos de doutorado, Bernardo Severo, Daniela Schmidt, Élodie Thièblin, Jordane Dorne, Roger Granada, sem dúvidas aprendi mais com vocês do que vocês comigo. E aos posdocs, Helbert Arenas, Philippe Roussille, e Ba-Huy Tran. Obrigada!

Aos meus colegas de équipe, Nathalie Aussenac, Mouna Kamel e Laure Vieu, e de outras équipes, Imen Megdiche, Franck Ravat, Jiefu Song, e Olivier Teste. Obrigada por compartilharem as questões de pesquisa, pelas discussões interessantes! Mouna, obrigada pelas idas au Miral e pela tua visão estruturada e organizada da pesquisa. Um obrigada especial à Nathalie que aceitou me seguir nesta aventura (e pelos comentários construtivos). E aos pesquisadores de outros laboratórios, em particular Jérôme Euzenat, pela oportunidade de ter trabalhado contigo, e pelos comentários que serviram à melhorar este manuscrito. Obrigada ao Ondrej Zamazal e a Michelle Cheatham pelas discussões sobre os 'alinhamentos complexos'. E a todos que se investem nas campanhas OAEI. Obrigada!

Aos meus colegas du Mirail, Eric Jacoboni, Catherine Comparot, Caroline Thierry, Françoise Adreit, Pascal Sarda, Elsy Kaddoum, Romain Guillaume, Frederic Ferraty, e Brahim Hamid (que sempre me dizia 'escreve teu HDR'). Obrigada a todos por compartilhar a carga de aulas, pelos almoços e cafés, discussões pedagógicas e por me conceder esta pausa, mesmo quando o 'Pix arrive'! Cathy, obrigada por teu apoio de sempre e por ter encontrado a solução e me permitir 'curtir tranquila' o meu CRCT. Obrigada a todos os outros colegas do DMI (que de uma forma ou outra me ajudaram a crescer como humano e profissional.

Aos amigos de longa data, Sandra Collovini, Cristiane Pedron, Vivian Dadalt, em particular Carlos Ramisch (que me faz lembrar da pausa necessària das 16h) e que vai divertir e preencher de sol ainda mais os meus dias durante um ano em Toulouse. Obrigada!

Aos Trojahns e agregados do WhatApps, por me fazerem sentir próxima (de onde?), por compartilharem as fotos de churrasco dos domingos e dos chimarrões dos sábados, dos aniversários ... de tudo que os faz rir. E ao dindo e tio Augusto Klusener pelas consultas pediátricas via WhatApps, independente da hora e do dia da semana. Vocês são a parte mais forte de mim. Amo-vos todos!

Abstract

Interoperability between different ontologies and their instances is at the core of the Semantic Web. Ontology matching and instance matching, as distinct tasks, aim at facilitating the interoperability between different knowledge bases at their terminological and assertional levels, respectively. Both tasks are key ones in many applications as they are the basis for data exchange and integration. Ontology matching, in particular, is the task of generating a set of correspondences (i.e., an alignment) between the entities of different ontologies. This is an active research area that has fully developed over the last two decades, with approaches being systematically evaluated in the context of the Ontology Alignment Evaluation Initiative (OAEI) campaigns. A variety of matching approaches have been proposed so far, mainly dealing with the generation of simple alignments (i.e., alignments involving single entities from each of the ontologies to be matched) between pairs of ontologies (i.e., pairwise matching). Those approaches have as well been mainly applied on ontologies with the same level of abstraction, in particular domain ontologies (i.e., ontologies describing the entities related to a particular domain). However, simple correspondences are proven to be insufficient to fully cover the different types of heterogeneity between ontologies. More expressive (complex) correspondences are required instead. Moreover, in domains where several ontologies describing different but related aspects of the domain have to be linked together, matching multiple ontologies (i.e., holistic matching) simultaneously is necessary. Last but not least, while linking domain ontologies to foundational ontologies has proved to improve ontology quality and interoperability between domain ontologies, the problem of matching this kind of ontologies has been addressed to a lesser extent in the field. This is a challenging task, specially due to the different levels of abstraction of these ontologies. These matching contexts raise as well the problem of automatically evaluating the corresponding approaches, which requires dedicated evaluation strategies. This manuscript presents my contributions to the ontology matching field, in particular on the generation of expressive (complex) alignments, matching of multiple ontologies and foundational and domain ontology matching. It is the support of my habilitation to supervise research.

Contents

Chapter 1

Introduction

This manuscript is the support to defend my *Habilitation à Diriger les Recherches*. In my PhD thesis, defended at University of Évora (Portugal) in 2009, I have proposed an approach relying on negotiation and argumentation models for combining ontology alignments from different matching systems. That proposal has shown how different points of view can be reconciled under different criteria and strategies (voting, strength, kind of matching approach). During my postdoctoral work, at INRIA Grenoble Rhône-Alpes (France), I have continued working on the field, but with the perspective of evaluating matchers and their alignments. I have been working on the evaluation of semantic technologies, in particular automating ontology matching evaluation. From 2012, as assistant professor at University of Toulouse 2 – Jean Jaurès and researcher at Toulouse Research Institute of Computer Science (IRIT), I have been working on topics such as complex ontology matching, holistic ontology matching, alignment of foundational and domain ontologies, alignment evaluation, and alignment visualization. I have also been working on knowledge extraction from text (with a particular focus on hierarchical relation extraction), semantic integration of Earth observations and contextual data, and multidimensional analysis of RDF statistical data. This chapter gives an overview of my main contributions in the topics I have been working on so far. However, as the larger part of my research activity and my major contributions so far are related to ontology matching, the remaining chapters of this document will be dedicated to the contributions to this field.

1.1 Context

The past decades have seen impressive development in the area of semantic technologies, mostly driven by the idea of creating a *semantic web* [Berners-Lee et al., 2001], as a source of knowledge accessible by machines. The vision of a semantic web relies on the idea of having the data on the Web exposed with annotations in a way that it can be used by machines for processing, integration and re-use across various applications. This idea is essentially based upon the notion of *ontology*, a knowledge representation structure that allows for organizing the concepts of a domain of interest, and relations between these concepts, and for expressing their semantics. The vision of a semantic web has been materialized with the *Linked Open Data* (LOD) initiative, where structured data are (ideally) exposed as instances of ontologies and linked across knowledge bases. The LOD has become a rich source of knowledge for several domains (such as Geography, Linguistics, Life Sciences, Social Networking) with about 1,239 datasets with 16,147 links[1].

Interoperability between different ontologies (and their instances) is at the core of the semantic web. *Ontology matching* and *data interlinking*, as distinct tasks, aim at facilitating the interoperability between different knowledge bases at their terminological (conceptual schemes) and assertional (data) levels, respectively. Both tasks are key ones in many applications as they are the basis for data exchange, integration and linking. Ontology matching, in particular, can be seen as the task of generating a set of correspondences (i.e., an alignment) between the entities of different ontologies, usually one source and one target ontologies [Euzenat and Shvaiko, 2013]. Correspondences express relationships between ontology entities, for instance, that an `Author` in one source ontology is equivalent to `Writer` in one target ontology, or that `Writer` in the source is subclass of `Person` in the target. A set of correspondences between two ontologies is called an alignment. An alignment may be used, for instance, to generate query expressions that automatically translate instances of these ontologies under an integrated ontology or to translate queries bearing on one ontology into queries with respect to the other.

A variety of matching approaches has been proposed so far in the literature. A review of them can be found in diverse surveys focusing on different aspects of the matching task in general, as schema matching [Rahm and Bernstein, 2001, Shvaiko and Euzenat, 2005], ontology matching [Kalfoglou and Schorlemmer, 2003, Noy, 2004, De Bruijn et al., 2006, Euzenat and Shvaiko, 2013] or with a focus on databases [Doan and Halevy, 2005]. While several classifications of approaches have been proposed in those works, the main distinction between each approach is based on the type of knowledge encoded within each ontology, and the way it is utilized when identifying correspondences between features or structures within the ontologies. They can be classified along the many features that can be found in ontologies (annotations, structures, instances, semantics), or with regard to the kind of disciplines they belong to (e.g., statistics, combinatorics, semantics, linguistics, machine learning, or data analysis). The different strategies have been combined in a variety of ways, including the exploitation of background knowledge from external resources, such as WordNet [Lin and Sandkuhl, 2008, Yatskevich and Giunchiglia, 2004], exploring existing alignments within an indirect matching [Zhang and Bodenreider, 2005, Jung et al., 2009, Hecht et al., 2015], involving the user in the loop [Jiménez-Ruiz et al., 2012, Dragisic et al., 2016], matching tuning [Lee et al., 2007, Ritze and Paulheim, 2011], reasoning with alignments [Meilicke et al., 2009, Jiménez-Ruiz and Grau, 2011], or using word embedding [Zhang et al., 2014, Kolyvakis et al., 2018], to cite but a few.

Developments in the field depend, however, on the ability to evaluate existing matching solutions. Furthermore, the large scale adoption of these solutions, in real world applications, depends on the ability to determine the quality of a system

[1] `https://lod-cloud.net/` accessed on August 2019.

in terms of its expected performance on realistic data. Hence, evaluation is an important aspect in the ontology matching field. A major and long-term goal of evaluation is to help developers to improve their solutions and to help users to evaluate the suitability of the proposed systems to their needs [Euzenat et al., 2011b]. Evaluation should also help assessing absolute results, i.e., what are the properties achieved by a system, and relative results, i.e., how do the results of one system compare to the results of other ones. Evaluation of matching systems and their alignments may be achieved in different ways [Do et al., 2002]. The most common one consists of providing matchers with two ontologies and comparing the returned alignments with a reference alignment. However, this raises the issue of the choice of ontologies and the validity of the reference alignments. Since 2004, the Ontology Alignment Evaluation Initiative (OAEI)[2] makes available a collection of test sets for evaluating ontology matching systems, covering different domains, ontology size scales and evaluation strategies. The main goal of OAEI is to compare systems and algorithms on the same basis and to allow for drawing conclusions about the best matching strategies in each scenario. OAEI's ambition is that from such evaluations, tool developers can learn and improve their systems, thus extending the state of the art in ontology matching.

The ontology matching field has reached so far some maturity, in particular with respect to approaches dealing with simple alignments involving a pair of ontologies at the same level of abstraction. However, simple correspondences are proven to be insufficient to fully cover the different types of heterogeneity between ontologies. More expressive (complex) correspondences are required instead. Moreover, in domains where several ontologies describing different but related aspects of the domain have to be linked together, matching multiple ontologies (i.e., holistic matching) simultaneously is necessary. Furthermore, while linking domain ontologies to foundational ontologies has proved to improve ontology quality and interoperability between domain ontologies, the problem of matching this kind of ontologies has been addressed to a lesser extent in the field. This is a challenging task, specially due to the different levels of abstraction of these ontologies. These matching contexts raise as well the problem of automatically evaluating the corresponding approaches, which requires dedicated evaluation strategies. Last but not least, other aspects that merit attention in the field include the multilingualism, as specific strategies have to be considered when matching ontologies with terminological layers expressed in different natural languages, and visualization and manipulation of multiples alignments, in order to help users seeking qualitative evaluation. These are the topics addressed in my contributions to the ontology matching field, as summarized below and further detailed in the next chapters.

1.2 Contributions

I have been working on the generation of complex correspondences between two ontologies (Section 1.2.1); on holistic ontology matching (Section 1.2.2); on matching ontologies whose terminological layers are expressed in different natural languages

[2] http://oaei.ontologymatching.org/

(Section 1.2.4); and matching of ontologies with different levels of abstraction, as foundational and domain ontologies (Section 1.2.3). The problem of automatic evaluation of the corresponding approaches has also been taken into account (Section 1.2.5). I have also worked on an environment to manage multiple alignments (Section 1.2.6). The contributions in topics other than ontology matching, such as relation extraction, semantic integration of Earth observations data and querying and manipulating (large) RDF statistical data, are also briefly presented (Section 1.2.7).

1.2.1 Complex ontology matching

Although the ontology matching field has fully developed in the last two decades, most ontology matching works still focus on generating simple correspondences (e.g., the correspondence $\langle Author, Writer, \equiv \rangle$ expressing that the concept `Author` in the source ontology is equivalent to the concept `Writer` in the target ontologie). These correspondences are however insufficient to fully cover the different types of heterogeneity between knowledge bases. More expressive correspondences are required instead (e.g., $\langle IRITMember, Researcher \sqcap \exists belongsToLab.\{IRIT\}, \equiv \rangle$[3], the correspondence expressing that the concept `IRITMember` in the source ontology is equivalent, in the target ontology, to a `Researcher` that belongs to `IRIT` as laboratory). We have provided a comprehensive survey on complex matching in [Thiéblin et al., 2019c].

In order to overcome this lack in the field, an approach for generating complex alignments that relies on the notion of Competency Questions for Alignment (CQAs) has been proposed. A CQA expresses the needs of a user with respect to the matching task and represents the scope of the ontologies that the alignment should cover [Thiéblin et al., 2018d]. This notion is inspired from ontology authoring, where competency questions have been introduced as *ontology's requirements in the form of questions the ontology must be able to answer* [Grüninger and Fox, 1995, Pinto and Martins, 2004, Ren et al., 2014]. Our approach takes as input a set of CQAs translated into SPARQL queries over the source ontology. The answer to each query is a set of instances retrieved from a knowledge base described by the source ontology. These instances are matched to those of a knowledge base described by the target ontology. The generation of the correspondences is performed by matching the graph-pattern from the source query to the lexically similar surroundings of the target instances.

I have also worked on exploiting complex correspondences in the task of query rewriting. First, we have proposed a mechanism for rewriting query patterns [Gillet et al., 2013] with the help of complex correspondences. Query patterns represent typical user query families (i.e., the family of queries asking for the actors playing in movies) and are expressed as knowledge subgraphs. Using complex correspondences improved the coverage of the system with respect to simple correspondences. This mechanism has been extended in order to tackle the problem of SPARQL query rewriting, whose existing solutions were basically based on simple correspondences

[3]As introduced in Chapter 2, correspondences are represented as triples.

[Thiéblin et al., 2016]. This rewriting system has been used for cross-querying agronomic classifications on the LOD [Thiéblin et al., 2019a, Thiéblin et al., 2017a].

While the matching approach able to generate complex correspondences has been exploited for pairwise matching, I have worked on the task of holistic matching, as introduced in the following section, however with the perspective of generating simple holistic alignments.

1.2.2 Holistic ontology matching

Most efforts in the ontology matching field are still dedicated to *pairwise ontology matching* (i.e., matching a pair of ontologies). However, in domains where several ontologies describing different but related aspects of the domain have to be linked together, matching multiple ontologies simultaneously, known as *holistic matching*, is required. As stated in [Pesquita et al., 2014], the increase in the matching space and the inherently higher difficulty to compute alignments pose interesting challenges to this task.

We have started working on holistic ontology matching as an extension of the work from [Berro et al., 2015], in the field of schema matching and especially designed to hierarchical structures like XML. This work has been extended to deal with ontologies rather than simpler schemes. In that context, we have worked on the LPHOM (Linear Program for Holistic Ontology Matching) approach [Megdiche et al., 2016a, Megdiche et al., 2016b]. LPHOM treats the ontology matching problem, at schema-level, as a combinatorial optimization problem. The problem is modeled through a linear program extending the maximum-weighted graph matching problem with linear constraints (matching cardinality, structural, and coherence constraints).

While this work has addressed the matching of domain ontologies, another problem in ontology matching is matching ontologies with different levels of abstraction (such as domain and foundational ontologies), as discussed in the following section (switching back to a pairwise setting, though).

1.2.3 Foundational and domain ontology matching

A *foundational* or *top-level* ontology is a high-level and domain-independent ontology whose concepts (e.g., *physical object*, *event*, *quality*, etc.) and relations (e.g., *parthood*, *participation*, etc.) are intended to be basic and universal to ensure generality and expressiveness for a wide range of domains. This kind of ontology plays an important role in the construction and integration of domain ontologies, providing a well-founded reference model that can be shared across domains. While most efforts in ontology matching have been particularly dedicated to domain ontologies, the problem of matching domain and top-level ontologies has been addressed to a lesser extent. This is a challenging task in the field, specially due to the different levels of abstraction of these ontologies. Furthermore, it is as task that requires the ability of identifying subsumption relations, as top-level ontologies refers to a higher level of abstraction. Currently, matching systems cannot deal with the task, as we could observe when running preliminary experiments on the task of matching foundational ontologies to domain ones using the OAEI matching systems [Schmidt et al., 2016].

In order to overcome this limitation, we have proposed an approach to match domain and foundational ontologies that exploits existing alignments between foundational ontologies and WordNet [Schmidt et al., 2018]. These alignments act as bridges for aligning these ontologies. This shifts the problem to finding the WordNet synset that expresses the right meaning of the domain concept. For that, we have adopted classical word disambiguation approaches and word embeddings.

Besides the problem of matching ontologies with different levels of abstraction, I have also addressed the problem of matching ontologies using different natural languages in their terminological layers, as introduced in the following.

1.2.4 Multilingual ontology matching

Matching approaches typically rely on a preliminary step of string-based comparisons of entity names and annotations (ID, labels, comments), whereby an initial estimate of the likelihood that two elements refer to the same real-world phenomenon is provided. However, this comparison restricts these matching techniques to ontologies that are labelled in the same or comparable[4] natural languages, as demonstrated by [Fu et al., 2009]. The increased awareness of the usefulness of ontologies for practical applications has lead to a situation where a number of ontologies actually used in real world applications do not use English as a base language. The existence of such ontologies pushes the ontology matching problem to a new level as the basic step used by most matching algorithms has to be adapted. I have worked on an approach exploiting translations and indirect matching approaches, relying on existing alignments, for composing cross-lingual alignments [Trojahn et al., 2010]. In [Trojahn et al., 2014], we have proposed an overview of multilingual and cross-lingual ontology matching together with a classification of different techniques and approaches. Systematic evaluations of these techniques are also discussed with an emphasis on standard and freely available data sets and systems.

Multilingual matching and the other matching scenarios discussed in the sections above require dedicated evaluation strategies, for which I have proposed solutions, as discussed in the next section. For sake of space and given that this topic is not at the core of my recent researches, the contributions to multilingual matching are out of the scope of this manuscript.

[4] An example of comparable natural languages is English and German, both belonging to the Germanic language family. Comparable natural languages can also be languages that are not from the same language family. For example, Italian belonging to the Romance language family, and German belonging to the Germanic language family can still be compared using string comparison techniques such as edit distance, as they are both alphabetic letter-based with comparable graphemes. An example of natural languages that are not comparable in this context can be Chinese and English, where the former is logogram-based and the latter is alphabetic letter-based. In this book, we consider natural languages are comparable when they contain graphemes that can be analyzed using automated string comparison techniques.

1.2.5 Ontology matching evaluation

Evaluation is a topic at the heart of any field of research. Since 2009, I have participated as organiser of the annual OAEI campaigns. I have been involved in evaluating the systems in two tracks, Benchmark [Euzenat et al., 2011a, Euzenat et al., 2010, Euzenat et al., 2009] and MultiFarm [Algergawy et al., 2018, Achichi et al., 2017, Achichi et al., 2016, Cheatham et al., 2015, Dragisic et al., 2014, Grau et al., 2013, Aguirre et al., 2012]. This evaluation task involves running the systems, comparing their results with the results of other systems, and analyzing how the proposed matching strategies fit the features of the matching cases. We have reported the lessons learned from the first six campaigns in [Euzenat et al., 2011b], both specific to ontology matching and generally relevant to the evaluation of semantic technologies in general. This has guided the execution of the OAEI evaluations, in particular in terms of automation and involvement of the user in the matching cycle.

One particular kind of evaluation is *benchmarking*. A *benchmark* is a well-defined set of tests on which the results of a system can be measured [Castro et al., 2004]. It should enable to measure the degree of achievement of proposed tasks on a well-defined scale (that can be achieved or not). The OAEI Benchmark test has been used for many years as a main reference to evaluate and compare ontology matching systems. This test, however, did not vary since 2004 and has become a relatively easy task for matchers. In order to overcome these limitations, we have been proposed a test generator based on an extensible set of alterators which may be used programmatically for generating different test sets from different seed ontologies and different alteration modalities [Euzenat et al., 2013]. This highlights the stability of results over different generations and the preservation of difficulty across seed ontologies, as well as a systematic bias towards the initial Benchmark test set and the inability of such tests to identify an overall winning matcher. This generator has been also included as part of the Alignment API [David et al., 2011].

While OAEI features a variety of different benchmark datasets covering a wide range of typical matching problems, almost all datasets included so far considered that the ontologies to be aligned use English as a common language for naming and describing concepts and relations. In order to push the development of multilingual and cross-lingual approaches and to overcome the lack of multilingual benchmarks in OAEI, we have proposed the MultiFarm dataset [Meilicke et al., 2012]. This dataset has been jointly created by the authors on the basis of the Conference dataset from the OAEI campaigns. The proposed benchmark consists of seven ontologies for which mutual reference alignments have been created manually. Each of the ontologies has been translated into eight languages other than English: Chinese, Czech, Dutch, French, German, Portuguese, Russian, and Spanish. Each combination of ontologies and languages establishes a test case for multilingual ontology matching summing up to roughly 1500 test cases. This dataset has been further extended with Italian and Arabic translations[5]. The MultiFarm dataset has also been used for constructing a multilingual comparable corpus [Granada et al., 2012], which can be further exploited in the tasks of ontology construction, population,

[5] https://www.irit.fr/recherches/MELODI/multifarm/

and distributional model training in cross-lingual and multilingual scenarios.

More recently, we have addressed the problem of evaluating complex alignments. In this context, two datasets for complex evaluation have been proposed [Thiéblin et al., 2018b]. This work was further extended by taking into account the need for a consensual generation of reference alignments. These datasets have been then used in the first complex track in OAEI task, opening new perspectives in the field [Thiéblin et al., 2018a, Algergawy et al., 2018].

We have also addressed the lack of benchmarks dedicated to holistic matching evaluation. We have proposed in [Roussille et al., 2018a] a methodology for constructing *pseudo*-holistic reference alignments from available pairwise ones. We have discussed the problem of relaxing graph cliques representing these alignments involving a different number of ontologies. We argue that fostering the development of holistic matching approaches depends on the availability of such datasets.

Last but not least, qualitative evaluation can take benefit from tools supporting the manipulation, evaluation and visualization of multiple ontology alignments, as discussed in the next section.

1.2.6 Alignment visualisation

Providing ways for visualizing alignments is required in many tasks involving users in the process (alignment evaluation, validation, comparison, consensus, repairing, etc.). As stated in [Dragisic et al., 2016], automatic generation of alignments should be viewed only as a first step towards a final alignment, with validation by one or more users being essential to ensure alignment quality. Visualizing tools play a key role in this process. We have proposed the VOAR (**V**isual **O**ntology **A**lignment Envi**R**onment)[6] [Severo et al., 2017, Severo et al., 2014], a configurable environment for manipulating multiple alignments. It provides an open Web-based environment that is not bound to any specific system and that offers a GUI for assisting users in the tasks of alignment visualization, manipulation, and evaluation. The major contribution of this version with respect to [Severo et al., 2015] is the possibility of choosing different visualization modes of multiple alignments, both at schema and instance levels, together with the possibility of storing and searching them. Users can configure their environment according to their needs and tasks, creating different profiles. For sake of space, the contributions to alignment visualization are however left out of the scope of this manuscript.

1.2.7 Other research contributions

My research activities are not limited to the ontology matching topics I have developed above. I have been working on related topics of the semantic web and knowledge representation fields, as briefly introduced in the following. These topics however are out of the scope of this manuscript.

Relation extraction from text Extracting hypernym relations from text (e.g., the relation expressing that `Toulouse` is a `City` or that an `Author` is a `Person`)

[6] `http://voar.inf.pucrs.br/` (recorded video at `https://youtu.be/wq-yPBOFN_I`)

is one of the key steps in the construction and enrichment of semantic resources. This kind of relation allows expressing the backbone structure of such resources and for assigning types to entities. Several hypernym extraction methods have been proposed in the literature (linguistic, statistical, learning based, hybrid), trying to better cover the different ways this kind of relation is expressed in written natural language. In [Kamel et al., 2017a], we have proposed a distant learning approach to extract hypernym relations from text. We have evaluated its ability to capture regularities from the corpus, without human intervention. We evaluated our approach on a subset of disambiguation pages from the French Wikipedia, with the approach outperforming lexico-syntactic patterns used as baseline. This work has been further extended in [Kamel et al., 2017b], where we deeply analyzed how complementary these approaches of different nature are when identifying hypernym relations on a structured corpus containing both well-written text and syntactically poor formulations, together with a rich formatting. The best results have been obtained when combining both approaches together. In addition, 55% of the identified relations, with respect to a reference corpus, were not expressed in the French DBpedia knowledge base and could be used to enrich this resource.

We have also contrasted different relations extraction methods on the same set of corpora and across different languages. Such comparison is important to see whether they are complementary or incremental. Also, we can see whether they present different tendencies towards recall and precision, i.e., some can be very precise but with very low recall and others can achieve better recall but low precision. Another aspect concerns to the variation of results for different languages. We have evaluated different methods (syntactic patterns, head modifier, distributional analysis, distributional inclusion, hierarchical clustering and document subsumption) on the basis of hierarchy metrics such as density and depth, and evaluation metrics such as recall and precision. The evaluation was performed over the same corpora, which consists of English and Portuguese parallel and comparable texts.

We have also addressed the problem of extracting hypernym relations from semi-structured textual elements, such as vertical enumerative structures (those using typographic and dispositional layout) [Kamel and Trojahn, 2018]. This kind of structure has been under-exploited by classical Wikipedia extractors. However, frequent in corpora, they are rich sources of specific semantic relations, such as hypernym. We have proposed a distant learning approach for extracting hypernym relations from vertical enumerative structures from the French Wikipedia, with the aim of enriching DBpedia. Our relation extraction approach achieves an overall precision of 62%, and 99% of the extracted relations can enrich DBpedia, with respect to a reference corpus.

Finally, validating extracted relations is a crucial step before integrating them into semantic resources. While manual validation is a time-consuming task requiring domain expert judges, automatic ones rely on external semantic resources (such as WordNet, BabelNet), which are usually not domain-specific, or gold standards, which may suffer from imperfections or low domain coverage. In order to overcome these limitations, we have proposed an approach for automatically validating candidate hierarchical relations extracted from parallel enumerative structures [Kamel and Trojahn, 2016]. It relies on the discursive properties of these structures and on the combination of resources of different nature: a semantic network and a distributional resource. The results show an improvement in the validation accuracy when combining both resources.

Semantic integration of Earth observation data Earth Observation (EO) provides added value to a wide variety of areas. Recently, the European Space Agency (ESA) has launched the Copernicus program[7], with two types of satellites, Sentinel-1 and Sentinel-2 providing high-quality Earth images (estimated between 8TB to 10TB of data daily), for free. The availability of these data opens up many economic opportunities through new applications in fields as diverse as agriculture, environment, urban planning, oceanography and climatology. These applications, however, have a strong need to couple these images with data on the observed areas. These data come from various measurement sensors. They are available from different sources with heterogeneous formats and distinct temporal features: they may be either static, like soil data, or dynamic, like weather observations. They can be useful for instance to indicate that an image contains a region affected by a natural phenomenon such as an earthquake or heat wave [Arenas et al., 2016b], and may be used for deciding what to do in this area or for longer-term analyses. Moreover, by exploiting the spatio-temporal characteristics of a phenomenon (its spatial imprint and its date), it becomes possible to know whether a geo-located entity within the footprint of this image (e.g., a city), has undergone the same phenomenon.

Previous works have already demonstrated the gain brought by semantic technologies to facilitate the task of EO data integration [Reitsma and Albrecht, 2005, Janowicz et al., 2012, Sukhobok et al., 2017]. In line with these works, we proposed a semantic approach to integrate data with the aim of enriching metadata of satellite imagery with data from various sources that provide EO for a particular need. It required to define a vocabulary that enables a homogeneous data representation and query, and to write mapping rules or algorithms to populate the model with data from the heterogeneous sources. In the particular case of EO data, an important aspect is that the data from diverse origins can relate through spatio-temporal topological relationships. The data integration process needs hence to properly manage the spatial and temporal properties and relationships.

As a first contribution, we have proposed a vocabulary that allows the semantic and homogeneous description of geo-spatial data as well as the metadata of satellite images as entities with spatial and temporal properties [Arenas et al., 2016a, Arenas et al., 2018c, Arenas et al., 2018a]. A subset of the geo-spatial data to be integrated to the image metadata is contextual information measured on Earth, so that it can be considered as sensor data. While in [Arenas et al., 2016a], image metadata records and meteorological observations were represented with DCAT and SSN (Semantic Sensor Network ontology) vocabularies, in [Arenas et al., 2018c, Arenas et al., 2018a] we extended this work proposing a modular ontology that specialises SOSA[8] (a light-weight but self-contained core ontology representing elementary

[7]`https://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus`
[8]`https://www.w3.org/TR/vocab-ssn/`

classes and properties of SSN), GeoSPARQL [Kolas et al., 2013] and OWL-Time[9].

As a second contribution, we defined an integration process that is based on the topology of entities [Arenas et al., 2018b]. The diversity of data sources raises different heterogeneity issues. For each data set to be integrated, we defined mapping templates and functions. Temporal properties and relations contribute to integrate dynamic data. In order to avoid duplicating static data that would tag all the images of the same area over time, the notion of tile defined by ESA is very convenient: the Earth's surface is associated with a grid where a tile represents a fixed area on the Earth's surface. As a side effect, using tiles enables to better scale up by reducing the amount of data to be handled.

We have illustrated our approach through a case study exploiting Sentinel-2 images metadata and contextual data (weather report data, Earth land cover, tiles, etc.). For this study, the semantic representation of the image metadata provided by the CNES (French National Center for Space Studies) is linked to data overlapping the image footprint and with similar capture time, in particular, meteorological data from Meteo France, Land Cover and Administrative Units. In [Dorne et al., 2018], this work has been used as a basis for a semantic representation of vegetation indices (NDVI) calculated on Sentinel-2 images. Currently, our ontology is used as a basis for modelling Earth observation change detection from deep learning image processing.

Querying and manipulating (large) RDF statistical data Multidimensional analysis is an alternative way for summarizing, aggregating and viewing RDF data on different axes (dimensions) and subjects of analysis (facts). From a RDF data collection conforming to the W3C Data Cube specification, we have formalized a multidimensional model in terms of RDF data structures following a conceptual constellation model [Saad et al., 2013]. This model clusters facts, which are studied according to several dimensions possibly shared between facts, with dimensions relating multi-hierarchies. We have shown how elementary OLAP operations can be translated into SPARQL queries using an OLAP algebra that is compliant to the constellation model. This algebra is based on a multidimensional table which displays data from one fact and two of its linked dimensions.

We have also addressed the problem of querying large volumes of RDF data cubes [Ravat et al., 2019b]. Our approach relies on combining pre-aggregation strategies and the performance of NoSQL engines to represent and manage statistical RDF data. Specifically, we have defined a conceptual modelling solution to represent original RDF data with aggregates in a multidimensional structure. We completed the conceptual modelling with a logical design process based on well-known multidimensional RDF graph and property-graph representations. We implemented our model in RDF triple stores and a property-graph NoSQL database, and we compared the querying performance, with and without aggregates. Experimental results, on real-world datasets containing 81.92 million triplets, show that pre-aggregation allows reducing query runtime in both RDF triple stores and property-graph NoSQL databases. Neo4j NoSQL database with aggregates out-

[9] `https://www.w3.org/TR/owl-time`

performs RDF Jena TDB2 and Virtuoso triple stores, speeding up to 99% query runtime. More recently, in [Ravat et al., 2019a], this work has been extended with one column-oriented NoSQL database (Cassandra) and two relational databases (MySQL and PostGreSQL). We compared the querying performance, with and without aggregates, in the six data stores.

1.3 Manuscript structure

As stated before, this manuscript presents my main contributions to the ontology matching field, with the contributions to multilingualism and alignment visualization left out of its scope. The rest of the document is organized as follows:

Chapter 2 introduces the background on ontology matching and evaluation, introducing the matching process, its parameters and dimensions, together with the evaluation process, criteria and metrics.

Chapter 3 presents the contributions on the generation of expressive (complex) alignments, in particular the CANARD approach for generating complex correspondences from user needs.

Chapter 4 presents the contributions on holistic ontology matching, extending a previous approach for holistic matching of XML structures, in order to accommodate more expressive structures as ontologies.

Chapter 5 presents the contributions on matching foundational and domain ontologies, in particular one approach that exploits existing alignments between external resources (WordNet) and foundational ontologies.

Chapter 6 discusses the creation of dedicated datasets to the task of evaluating expressive (complex) alignments and the evaluation metrics based on instance sets comparison.

Chapter 7 presents a methodology for constructing holistic alignments from existing pairwise alignments relying on graph cliques involving a different number of ontologies.

Chapter 8 discusses the effort of manually matching the OAEI Conference dataset to the SUMO foundational ontology. This is a first effort towards a reference OAEI dataset involving the task of matching foundational and domain ontologies.

Chapter 9 outlines my perspectives for future research work.

Chapter 2

Background

2.1 Introduction

Ontology matching is an essential task in the management of ontology heterogeneity in a range of applications such as ontology integration, query rewriting, and ontology evolution. Ontologies can be heterogeneous in various ways, e.g., in terms of the terminologies adopted for describing their entities; in terms of their modelling strategies, with respect to the coverage and granularity of their representations; or still in terms of human interpretations [Visser et al., 1997, Klein, 2001, Euzenat and Shvaiko, 2013]. In terms of their level of "generality" [Guarino, 1998], ontologies can also describe very general concepts (e.g., space, time, object, etc.), which are independent of a particular problem or domain (e.g., *top-level* or *foundational* ontologies) or describe the concepts and relations specific of a domain of interest (e.g, *domain* ontologies). The typically high degree of heterogeneity reflected in different ontologies makes the matching task an inherently complex task [Rahm, 2011].

In order to characterize and compare matching solutions, evaluation is required. Matching evaluation can be done in different ways [Do et al., 2002]. One classical way consists of comparing generated alignments to reference ones (gold standard). However, constructing such references is a time-consuming task that requires experts in the domain. In the absence of such resources or when dealing with large datasets, alternatives include manual labelling on sample alignments [Van Hage et al., 2007], computing the minimal set of correspondences (which can be used for computing all the other ones) for reducing the effort on manual validation [Maltese et al., 2010], or still measuring the quality of alignments in terms of coherence measurements [Meilicke and Stuckenschmidt, 2008] or conservativity principle violation [Solimando et al., 2017]. Alternatively, an alignment can be assessed regarding its suitability for a specific task or application [Isaac et al., 2008, Hollink et al., 2008, Solimando et al., 2014b]. Other approaches consider the generation of natural language questions to support end-users in the validation task [Abacha and Zweigenbaum, 2014] or validating correspondences on graph-based algorithms in a semi-automatic way [Serpeloni et al., 2011]. In this chapter, first the matching process, its parameters and dimensions are introduced (Section 2.2). Next, the evaluation process, the criteria and metrics are presented (Section 2.3).

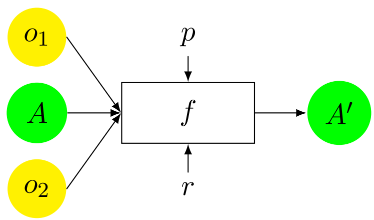

Figure 2.1: The matching process (adapted from [Euzenat and Shvaiko, 2013] (page 41)).

2.2 Matching process

The matching process consists of generating an alignment (A') from a set of ontologies Ω, usually a pair of ontologies ($\Omega = \{o_1, o_2\}$). Despite this general definition, there are various other parameters which can extend this definition. These are the use of an (partial) input alignment (A) which is to be completed by the process, the method parameters (e.g, weights) and some external resources used by the matching process (which can be general-purpose resources, e.g., lexicons). This process can be defined as follow (Figure 2.1):

Definition 1 (Matching process) *The matching process can be seen as a function f which, from a pair of ontologies o_1 and o_2 to match, an input alignment A, a set of parameters p, and a set resources r, returns an alignment A' between o_1 and o_2:*

$$A' = f(o_1, o_2, A, p, r)$$

Each of the elements featured in this definition can have specific characteristics which influence the difficulty of the alignment task. Before introducing the features of each dimension, the notions of correspondence and alignment are introduced.

2.2.1 Correspondence and alignment

An alignment (A or A') consists of a set of correspondences $\{c_1, c_2, ..., c_x\}$:

Definition 2 (Alignment) *An alignment $A_{o_1 \to o_2}$ is a set of correspondences $\{c_1, c_2, ..., c_n\}$. $A_{o_1 \to o_2}$ is directional between a source ontology o_1 and a target ontology o_2.*

A correspondence expresses a relation r between ontology entities of o_1 and o_2. Here, the ontology entities are *members* of the correspondence:

Definition 3 (Correspondence) *A correspondence c_i is a tuple (e_1, e_2, r, n). e_1 and e_2 are the members of the correspondence. They can be simple or complex expressions with entities from respectively o_1 and o_2:*

- *if the correspondence is* ***simple****, both e_1 and e_2 are simple expressions.*

- *if the correspondence is* ***complex****, at least one of* e_1 *or* e_2 *is a complex expression, involving union, intersection, disjunction, cardinality restrictions, etc.*

- r *is a relation, e.g., equivalence (*$\equiv$*), more general (*$\sqsupseteq$*), more specific (*$\sqsubseteq$*), disjointedness (*$\perp$*) holding between* e_1 *and* e_2*.*

- n *is a number in the* $[0, 1]$ *range.*

The correspondence $\langle id, e_1, e_2, n, r\rangle$ asserts that the relation r holds between the ontology entities e_1 and e_2 with confidence n. The higher the confidence value, the higher the likelihood that the relation holds.

The members of the correspondences can be a simple expression, noted s, or a complex expression, noted c. A simple correspondence is always (s:s) whereas a complex correspondence can be (s:c), (c:s) or (c:c). The (1:1), (1:n), (m:1), (m:n) notations have been used for the same purpose in the literature [Rahm and Bernstein, 2001, Zhou et al., 2018] (*1* for s and m or n for c). However, they can be misinterpreted as the alignment arity or multiplicity [Euzenat, 2003]. An alignment is said to be simple if it contains only simple correspondences. An alignment is said to be complex if it contains at least one complex correspondence.

The examples in the following (and in the rest of this manuscript) are based on the OAEI Conference ontologies [Šváb et al., 2005, Zamazal and Svátek, 2017]. This choice is motivated by the fact that this dataset is the one used across the different works presented here. Consider the fragment of the ontologies *ekaw*[1] and *cmt*[2] in Figures 2.2 and 2.3, respectively. The format used to represent the ontologies is described in [Stapleton et al., 2014]. The following correspondences (for sake of simplicity, correspondences are represented as triples in the rest of the manuscript) can been established between these two ontologies:

c_1 = (*ekaw:Paper,cmt:Paper,*$\equiv$) is a (s:s) simple correspondence.

c_2 = (*ekaw:AcceptedPaper,* $\exists$*cmt:hasDecision.cmt:Acceptance,*$\equiv$) is a (c:s) complex correspondence with constructors.

c_3 = *(ekaw:writtenBy,cmt:writePaper*$^{-}$*,* $\equiv$*)* is a (s:c) complex correspondence with the *inversion* constructor.

While the RDF alignment format[3] [David et al., 2011] is the format *de facto* used in the OAEI campaigns for representing simple alignments, correspondences can also be represented as OWL 2 subclass, equivalence, and disjointedness axioms (with confidence values represented as axiom annotations). Alternatively, the EDOAL[4] language (Expressive and Declarative Ontology Alignment Language) extends the alignment format in order to represent complex correspondences.

[1] `http://oaei.ontologymatching.org/2019/conference/data/ekaw.owl`
[2] `http://oaei.ontologymatching.org/2019/conference/data/cmt.owl`
[3] `http://alignapi.gforge.inria.fr/format.html`
[4] `http://alignapi.gforge.inria.fr/edoal.html`

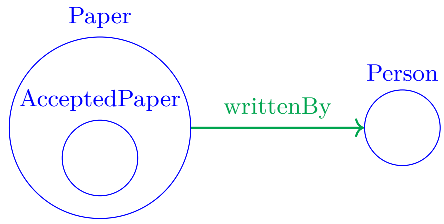

Figure 2.2: Fragment of *ekaw* ontology.

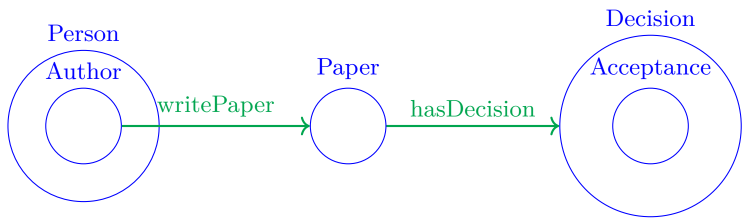

Figure 2.3: Fragment of *cmt* ontology.

2.2.2 Matching dimensions

As depicted in Figure 2.1, the matching process receives different parameters, which can have specific characteristics that influence the difficulty of the matching task, as introduced in the following [Euzenat and Shvaiko, 2007, Trojahn et al., 2009]:

Input ontologies The input ontologies (o_1, o_2) can be characterized by different dimensions: input languages and their expressiveness (e.g., OWL-Lite, OWL-DL, OWL-Full); natural language used for annotating the ontologies (labels, comments, etc); number of input ontologies (a pair or multiple ontologies); size of the input ontologies in terms of concepts, properties and instance; the coherence/consistency, the correctness (equipped or not with some incorrect axioms) and their completeness (relations modeled in detail or complete descriptions of the represented domain).

Input alignment The input alignment (A) can have the following characteristics, such as its multiplicity (how many entities of one ontology can correspond to one entity of the others, as (1:1), (1:m), (n:1) or (n:m)); its completeness (it can contain only few correspondences or nearly all correct correspondences); its coverage (cover a small fraction of the ontologies, depending on the fact that ontologies might cover different, only partially overlapping domains); or still its correctness (it can contain some erroneous correspondences). In simple scenarios, the input alignment is empty.

Parameters The parameters (p, r) of the matching process can be identified as: the kinds of resources used (including human input); and the proper parameters that are expected. A good tuning of these must be available when the method is sensitive to the variation of parameters. It can be the case that some methods are able to tune their parameters depending on the input ontologies. In such a case, the tuning process is considered part of the method. Training on some samples is very often used by methods for matching ontologies. A situation in which this makes sense is when a user provides some example of aligned instances and the system can induce the alignment from this.

In general, if human input is provided, the efficiency of systems can be expected to be better. In fact, a user may intervene in a matching process at different stages: before, during or after. Prior to the matching, the user may express the expected knowledge content of an alignment: its scope. With regards to user specification of the alignments, a definition of user knowledge need is given in [Lopez et al., 2006], where queries define the content of the expected alignment. During the process, as users can help the system, for example by detecting incorrect correspondences.

Output alignment Different constraints on the output alignment (A') can be identified: multiplicity (as for the input alignment), which can let to characterize the *mapping* as injective, surjective and total or partial on both side, defining its arity (noted with, 1 for injective and total, ? for injective, + for total and * for none and each sign concerning one mapping and its converse): ?:?, ?:1, 1:?, 1:1, ?:+, +:?, 1:+, +:1, +:+, ?:*, *:?, 1:*, *:1, +:*, *:+, *:* [Euzenat and Shvaiko, 2007]. These assertions could be provided as input (or constraint) for the alignment algorithm or be provided as a result by the same algorithm; the kind of relations covered by the alignment (only equivalence relations or other kind of relations); and the strictness of the results, which can be expressed with different confidence degrees (discrete, continuous, etc.). Justification is another aspect, as outputs are rarely annotated with justifications to the found results.

Matching process The matching process (f) itself can be constrained by: the resource constraints (in terms of amount of time or space available for computing the alignment); language restrictions (the scope limited to some kind of entities e.g., only classes); or properties that should be true, as one might want that the alignment be a consequence of the combination of the ontologies (i.e., $o_1, o_2 \models A'$) or that the initial alignment is preserved (i.e., $o_1, o_2, A' \models A$). Resource constraints can be considered either as a constraint (the amount of resource is limited) or a result (the amount consumed is measured). Constraints on the kind of language construct to be found in alignments can be designed (simple or complex expressions, for instance).

2.3 Matching evaluation

Many different techniques have been proposed for implementing the matching process. An alignment is obtained by combining these techniques towards a particular goal (obtaining an alignment with particular features, optimizing some criterion, etc). Several combination techniques are also used. This variety suggests the need to systematic evaluating these methods. Evaluation can be carried out in different ways. A usual way consists in evaluating the generated alignment in terms of its compliance with respect to a reference alignment. In this basic evaluation setting, a matcher receives two ontologies o_1 and o_2 as input and generates an alignment A'

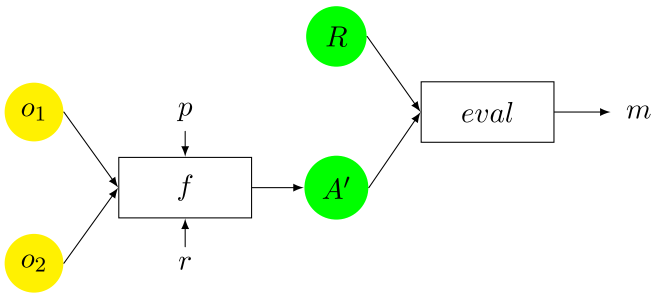

Figure 2.4: Basic evaluation process (from [Euzenat and Shvaiko, 2013] (page 289)).

using a certain set of resources and parameters. An evaluation component receives this alignment and computes a (set of) quality measure(s) m – usually precision and recall – by comparing it to the reference alignment R (Figure 2.4). This basic process is simplistic and has to be concretized in many respects. First of all, the input data in terms of the ontologies to be matched has to be defined. No single pair of ontologies can test all aspects of ontology matching. Another insight is that standard quality measures, in particular precision and recall, are not always suited for the purpose of ontology matching as they fail to completely capture the semantics of ontology alignments and different measures are needed for evaluating different aspects. Furthermore, more complex approaches are sometimes needed in certain situations, for instance, if a partial alignment exists or if no reference alignment is available.

In the following, the evaluation prerequisites in terms of datasets and the different dimensions an alignment can be evaluated are discussed.

2.3.1 Evaluation prerequisites

Good datasets are a prerequisite for a good evaluation. The nature of the datasets determines how far the evaluation design the coverage of relevant aspects and the fairness of the evaluation [Euzenat et al., 2011b]. In the case of ontology matching, a dataset typically consists of at least two ontologies and a reference alignment between these ontologies. In the following, the combination of two ontologies and, if present, a reference alignment between these ontologies is referred as a *matching task*. A dataset consists of several matching tasks.

The work in [Giunchiglia et al., 2009] proposed the following criteria for designing or selecting datasets for ontology matching evaluation:

- Complexity, i.e., that the dataset is hard for state of the art matching systems.

- Discrimination ability, i.e., that the dataset can discriminate sufficiently among various matching approaches.

- Incrementality, i.e., that the dataset allows for incrementally discovering the weaknesses of the tested systems.

- Monotonicity, i.e., that the matching quality measures calculated on the subsets of the dataset do not differ substantially from the measures calculated on the whole dataset.
- Correctness, i.e., that the dataset includes a reference alignment which allows to divide generated correspondences into correct and incorrect ones.

As in [Euzenat et al., 2011b, Euzenat and Shvaiko, 2013], there are two basic properties that determine the nature of a dataset, and thus, how well it meets the quality criteria mentioned above: the properties of the ontologies to be matched and the properties of the reference alignment, that are expected to be reproduced by the matching systems.

Ontologies Two main features of the ontologies impact the matching process are the complexity of their annotations (essentially labels and comments), that are exploited in the initial determination of candidate correspondences, and the complexity of their structures.

Complexity of labels. Most matching systems typically rely on string-based techniques for comparing annotations, as an initial estimate of the likelihood that two elements refer to the same real world phenomenon. Hence, the kind of labels found in an ontology influences heavily the performance of a particular matching system: simple labels vs. sentence-like labels, monolingual vs. multilingual labels. It also often makes a large difference whether labels used in an ontology can be anchored to common background knowledge sources, such as WordNet, that helps interpreting those labels. Further complexity is added if the ontologies to be matched use specific vocabularies, e.g., from the biomedical or geo-spatial applications, that are outside common language.

Complexity of structures. Almost all matching systems use the structure of the ontology entities in later stages of the matching process to propagate similarity estimations and to validate hypotheses on correct correspondences. Therefore, structures found in ontologies are also an important issue in the design of benchmark datasets. Light ontologies and taxonomies use the hierarchical structure given by subsumption, while more expressive ontologies use relations between classes that may be constrained by various kinds of axioms. On the level of instances, we can also have different levels of complexity. In particular, instances can either be described in detail using attributes and relations to other instances or can be atomic entities with no further explicit definitions or property specifications.

Reference alignments A reference alignment is another important aspect to consider: characteristics, such as the types of semantic relations used in the alignment or the coverage of the alignment, have a significant impact not only on the hardness of the task but also puts restrictions on evaluation measures that are discussed later.

Types of semantic relations As introduced in Section 2.2, an alignment consists of a set of correspondences defined by entities from the input ontologies and a relation between them. The kind of relation found in the reference alignment also determines what kind of relations the matching systems should be able to produce. The most commonly used relation is equivalence (in most cases classes and relations). The majority of available matching systems are designed to generate equivalence relations. There are exceptions, however, that should be taken into account. Other kinds of relations that were investigated are subclass [van Hage et al., 2005, Giunchiglia et al., 2007] and disjointness relations [Sabou and Gracia, 2008, Giunchiglia et al., 2007].

Formal properties of the alignment Besides the type of relations, its semantics is another relevant aspect. In particular, one have to distinguish between more and less rigorous interpretations of relations. The equivalence relation, for example, can be interpreted as logical equivalence or more informally as a high level of similarity. Using a rigorous formal interpretation of the relations has the advantage that we can enforce formal properties on the reference alignment. For example, we can claim that the merged model consisting of the two ontologies and the alignment should be coherent, i.e., it should not contain unsatisfiable classes. Enforcing such consistency conditions is not possible for less formal interpretations.

Cardinality and coverage A less obvious property with a significant influence on the evaluation results is the cardinality of the reference alignment. In principle, there is no restriction on the alignment, so the relation between elements from the different ontologies can be an (n:m) relation. In practice, however, it turns out that the alignment relation is (1:1) in most cases. Therefore, matching systems often generate (1:1) alignments. Along the same lines, the degree of overlap between the ontologies to be matched is not restricted and a dataset could consist of two ontologies with little or no overlap. Typically, however, it is assumed that the two ontologies to be matched describe the same domain. As a consequence, matching systems normally try to find a correspondence for every element in the two ontologies rather than ignoring elements.

Construction of reference alignments in general follows different strategies, including starting the alignment generation from scratch, relying on a set of initial alignments for gathering additional ones, and creating a reference from validating and selecting a set of correspondences from automatically generated correspondences from a number of matching systems. In the first category, the creation of the first reference alignment of the Conference dataset dates back to 2008, when the track organizers created a reference alignment for all possible pairs of five of the conference ontologies. The reference alignments were based on the majority opinion of three evaluators and were discussed during a consensus workshop. This dataset has evolved over the years, as described in [Zamazal and Svátek, 2017], with the feedback from the OAEI participants and has been revised in [Cheatham and Hitzler, 2014]. They re-examined the dataset with a focus on the degree of agreement between the reference alignments and the opinion of experts. With the aim of studying the way different raters evaluate correspondences, in [Tordai et al., 2011] experiments in manual evaluation have been carried out using a set of correspondences generated by different matchers between different vocabularies. Five raters evaluated alignments and talked through their decisions using the think aloud method. Their analysis showed which variables can be controlled to affect the level of agreement, including the correspondence relations, the evaluation guidelines and the background of the raters. That work refers as well to the different levels of agreements between annotators reported in the literature. While a perfect agreement between raters is reported in the Very Large Crosslingual dataset in [Euzenat et al., 2009], [Halpin et al., 2010] reported a quite different observation when establishing *owl:sameAs* relationships in the LOD. These aspects have also been discussed in [Stevens et al., 2018] for the task of integrating foundational and domain ontologies.

2.3.2 Evaluation dimensions, criteria and metrics

One can distinguish different dimensions of an evaluation (Table 2.1), i.e., different ways of viewing the problem:

Tool-oriented This dimension refers to the evaluation of the matching tool in terms of efficiency (the computational resources consumption it requires for completing the matching task), usually in terms of execution runtime and amount of required memory, and in terms of its ability to scale. Both efficiency and scalability can depend on the nature of the ontologies and more specifically on the complexity of the structures and definitions found in the ontologies. Therefore there is a strong interaction between the hardness of tests and the complexity of the input ontologies. Furthermore, the best way to measure efficiency is running all the systems under the same controlled evaluation environment.

Output-oriented This dimension refers to the evaluation of the generated alignment itself. It can be intrinsic, extrinsic or task-oriented:

- **Extrinsic** An alignment can be evaluated with regard to a manually created reference alignment. It measures the compliance of an alignment with respect to a reference one. Metrics as precision and recall are usually adopted.
- **Intrinsic** The quality of an alignment can be measured based on its intrinsic characteristics. [Meilicke and Stuckenschmidt, 2008] evaluates the logical coherence as marker of quality of an alignment while [Solimando et al., 2017] considers that a good alignment should not violate the conservativity principle.
- **Task-oriented** The quality of an alignment can also be assessed regarding to its suitability for a specific task or application, as for the evaluation of the alignments over the query rewriting scenario [Solimando et al., 2014b].

Table 2.1: Criteria and metrics for ontology matching evaluation.

Dimension		*Criteria*	*Metric*
Tool-oriented		Efficiency	execution time and required memory
		Scalability	different test sizes
Output-oriented	Extrinsic	Compliance to reference alignment	precision, recall, and generalizations
	Intrinsic	Coherence	minimal revision effort to achieve coherence
		Conservativity	alignment does not introduce new relations between concepts
	Task-oriented	Suitability to the task	task-dependent (e.g. % of well-written queries)
	User-oriented	User satisfaction	subjective satisfaction (qualitative evaluation: good, satisfactory, etc.)

User-oriented The alignment can also be qualitatively evaluated by a user (expert).

This section presents the aspects to be evaluated (criteria) and how to evaluate these aspects (metrics) according to each evaluation dimension, which are summarized in Table 2.1.

Extrinsic evaluation metrics

One natural way to measure the alignment quality is measuring precision and recall, specially because they can be interpreted easily. These measures have been used as basis in the OAEI campaigns.

Precision and recall Precision (true positive/retrieved) and recall (true positive/expected) are common measures in information retrieval. The alignment A' generated by the matching system is compared to the (manually created) reference alignment R'.

Definition 4 (Precision) *Given a reference alignment R, the precision of some alignment A is given by*

$$P(A', R) = \frac{|R \cap A'|}{|A'|}$$

Precision can also be determined without explicitly having a complete reference alignment. Only the correct alignments among the retrieved alignments have to be determined ($R \cap A'$).

Definition 5 (Recall) *Given a reference alignment R, the recall of some alignment A' is given by*

$$R(A', R) = \frac{|R \cap A'|}{|R|}$$

Traditionally, precision and recall score only exact matches, where a correspondence is scored 1 if both of its entities and its relation are syntactically equivalent to the reference correspondence and scored 0 otherwise. While precision and recall are the most widely and commonly used measures, comparing systems can be done combining these measure. For this purpose, the F-measure is used to aggregate the result of precision and recall.

Definition 6 (F-measure) *Given a reference alignment R and a number α between 0 and 1, the F-measure of some alignment A' is given by*

$$F - measure_{\alpha}(A', R) = \frac{P(A', R) \cdot R(A', R)}{(1 - \alpha) \cdot P(A', R) + \alpha \cdot R(A', R)}$$

If $\alpha = 1$, then the F-measure is equal to precision and if $\alpha = 0$, the F-measure is equal to recall. In between, the higher α, the more importance is given to precision with regard to recall. Very often, the value $\alpha = 0.5$ is used, i.e. $F - measure_{0.5}(A', R) = \frac{2 \times P(A',R) \times R(A',R)}{P(A',R)+R(A',R)}$, the harmonic mean of precision and recall.

Variations and generalizations of precision and recall Although precision and recall are standard measures for evaluating compliance of alignments, alternative measures addressing some limitations of these measures have been used. A first alternative to the traditional binary syntactic evaluation is the weighted evaluation, where the confidence scores of the alignment to evaluate and those of the reference alignment are taken into consideration. This is particularly relevant when the reference alignment is not considered ground truth and has similarity scores other than 1, such as in the approach proposed by [Cheatham and Hitzler, 2014]. This penalizes an alignment system more if it fails to identify a strong correspondence than a weak one, and rewards the alignment system if its scoring scheme approximates the confidence scores of the reference alignment. A similar methodology relying on a vector representation of the ontology alignments, has been recently proposed by [Sagi and Gal, 2018].

Furthermore, it may happen that an alignment is very close to the expected result (reference alignment) and another one is quite remote from it, although both share the same precision and recall. The reason for this is that standard metrics only compare two sets of correspondences (strict syntactic comparison) without considering if these are close or remote to each other. It may be helpful for users to know whether the found alignments are close to the expected one and easily repairable or not. It is thus necessary to measure the proximity between alignments instead of their strict equality. In order to better discriminate such systems a relaxed precision and recall measures were defined which replace the set intersection by a distance [Ehrig and Euzenat, 2005].

Relaxed precision and recall [Ehrig and Euzenat, 2005] propose to generalize precision and recall, measuring the proximity of correspondence sets rather than their strict overlap. Instead of taking the cardinal of the intersection of the two sets $|R \cap A'|$, they propose to measure their proximity (ω).

Definition 7 (Generalized Precision and Recall) *Given a reference alignment R and an overlap function ω between alignments, the precision and recall of an alignment A' are given by*

$$P_\omega(A', R) = \frac{\omega(R \cap A')}{|A'|}$$

$$R_\omega(A', R) = \frac{\omega(R \cap A')}{|R|}$$

There are different ways to design such a proximity given two sets. In [Ehrig and Euzenat, 2005] the authors propose to find correspondences matching each other and computing the sum of their proximity ω(A',R). To compute ω(A',R), it is necessary to measure the proximity between two matched correspondences (i.e., <a,r> $\in$ M(A',R)) on the basis of how close the result is from the ideal one. Each element in the tuple a = <e_{a_1},e_{a_2},r_a,n_a > will be compared with its counterpart in r = <e_{r_1},e_{r_2},r_r,n_r >. For any two correspondences (the found *a* and the reference *r*), three similarities are computed: σ_{pair}, σ_{rel}, σ_{conf}:

- σ_{pair} How is one entity pair similar to another entity pair? In ontologies, it can follow any relation which exists (e.g., subsumption, instantiation), or which can be derived in a meaningful way. The most important parameters are the relations to follow and their effect on the proximity;

- σ_{rel} The alignment relations can be other than equivalence (e.g. subsumption). Again, one has to assess the similarity between these relations. The two relations of the alignments (generated and reference) can be compared based on their distance in a conceptual neighborhood structure [Euzenat et al., 2003];

- σ_{conf} Finally, one has to decide, what to do with different levels of confidence. The similarity could simply be the difference. Unfortunately, none of the current alignment approaches have an explicit meaning attached to confidence values, which makes it rather difficult in defining an adequate proximity.

Based on these three similarities, the correspondence proximity can be defined:

Definition 8 (Correspondence Proximity) *Given two correspondences* $<e_{a_1},e_{a_2},r_a,n_a>$ *and* $<e_{r_1},e_{r_2},r_r,n_r>$*, their proximity is:*

$$\sigma(< e_{a_1}, e_{a_2}, r_a, n_a >, < e_{r_1}, e_{r_2}, r_r, n_r >) = \sigma(< e_{a_1}, e_{r_1} >, < e_{a_2}, e_{r_2} >) \times \sigma(r_a, r_r) \times \sigma(n_a, n_r)$$

Three concrete measures based on the above definitions are proposed in [Ehrig and Euzenat, 2005]: *symmetric proximity*, *correction effort*, and *oriented proximity*. The *symmetric measure* is based on computing a distance δ on the ontological entities and to weight the proximity with the help of this distance: the higher the distance between two entities in the matched correspondences the lower their proximity. Using the *correction effort* measure, the quality of alignments can be measured through the effort required for transforming the found alignment into the correct one. This measure can be implemented as an edit distance [Levenshtein, 1965], which defines a number of operations by which an object can be corrected and assigns a cost to each of these operations (the effort required to identify and repair some mistake). The cost of a sequence of operations is the sum of their cost and the distance between two objects is the cost of the less costly sequence of operations that transform one object into the other one. Such a distance is then turned into a proximity measures. Finally, *oriented-effort* measure considers two different similarities depending of their application for evaluating either precision or recall. It associates different weights to compute the proximity measure in each case.

Semantic precision and recall The measures above are based on syntactic generalizations of precision and recall. In order to design a generalization of precision and recall that is semantically grounded, [Euzenat, 2007] proposes the *semantic precision* and *recall*. In such measures, those correspondences that are consequences of the evaluated alignments have to be considered as recalled and those that are consequence of the reference alignments as correct.

The semantic extension of precision and recall consists of using the set of α-consequences (or deductive closure on the prover side) instead of |A' ∩ R|:

Definition 9 (α-consequence of aligned ontologies) *Given two ontologies o_1 and o_2 and an alignment A' between these ontologies, a correspondence σ is a α-consequence of o_1, o_2 and A' (note A' $\models \sigma$) if and only if for all models $< m_1, m_2, \gamma >$ of o_1, o_2 and A', m_1, $m_2 \models \gamma$ (the set of α-consequence is noted by Cn(A')).*

In order to deal with the problems raised by the infinite character of the set of α-consequences, it is proposed to use a deductive closure bounded by a finite set so that the result is finite. It is based on different sets of true positives as:

$$TP_P(A', R) = \{\delta \in A'; R \models \delta\} = A' \cap Cn(R)$$
$$TP_R(A,' R) = \{\delta \in R; A' \models \delta\} = Cn(A') \cap R$$

The semantic precision and recall are based on these sets:

Definition 10 (Semantic Precision and Recall) *Given a reference alignment R, the precision of some alignment A' is given by:*

$$P_{sem}(A', R) = \frac{|A' \cap Cn(R)|}{|A'|}$$

$$R_{sem}(A', R) = \frac{|Cn(A') \cap R|}{|R|}$$

Intrinsic evaluation metrics

Alignment coherence The coherence of an alignment is a quality of its own. The term alignment (in)coherence has first been introduced in a paper concerned with the task of reasoning about ontology alignments in general [Stuckenschmidt et al., 2006][5]. Measuring the degree of (in)coherence of an alignment has been proposed in [Meilicke and Stuckenschmidt, 2008]. The authors argue that the incoherence of an alignment results in different kinds of problems depending on the specific application context. Thus, coherence of an alignment is an important quality, which has to be taken into account in the evaluation context.

The approach for measuring the degree of (in)coherence measuring is based on the notion of an aligned or merged ontology. Given two ontologies o_1 and o_2 and an alignment A' between them, the merged ontology $o_1 \cup_{A'} o_2$ is the union of o_1, o_2, and A' where A' is interpreted as a set of axioms. A correspondence expressing equivalence between two concepts, for example, is thus translated into an equivalence axiom in the context of the merged ontology.

An alignment A' between two ontologies o_1 and o_2 is called incoherent, if there exists an unsatisfiable concept $i\#C_{i\in\{1,2\}}$ in $o_1 \cup_{A'} o_2$; its unsatisfiability must have (at least partially) been caused by A'.

Definition 11 (Incoherence of an alignment) *Given an alignment A' between ontologies o_1 and o_2. If there exists a concept $i\#C$ with $i \in \{1,2\}$ such that $o_1 \cup_{A'} o_2 \models \bot \sqsupseteq i\#C$ and $o_i \not\models \bot \sqsupseteq i\#C$ then A' is incoherent with respect to o_1 and o_2. Otherwise A' is coherent with respect to o_1 and o_2.*

Different metrics have been proposed in [Meilicke and Stuckenschmidt, 2008], in particular the *Unsatisfiability Measure* and the *Maximum Cardinality Measure*. The first measure is based on the idea of counting unsatisfiable concepts. It is derived from an ontology incoherence measure introduced in [Qi and Hunter, 2007]. Contrary to measuring incoherences in ontologies, it has to be distinguished between two types of concept unsatisfiability in the merged ontology: there are unsatisfiable concepts in $o_1 \cup_{A'} o_2$ which have already been unsatisfiable in o_1, respectively o_2, while there are unsatisfiable concepts which have been satisfiable in o_1, respectively o_2. These concepts have become unsatisfiable due to the impact of A'. In particular, the number of these concepts with the number of all named concepts satisfiable in o_1 or o_2 is compared. Alternatively, the *Maximum Cardinality Measure* is concerned with the effort of revising an incoherent alignment. This measure is based on the idea to remove a minimum number of correspondences to achieve the coherence of the alignment.

Measuring the degree of incoherence obviously requires full-fledged reasoning techniques. It is thus heavily linked to issues concerned with reasoning systems.

There exists a strong dependence between incoherence and inconsistency [Meilicke, 2011]. There are ontologies that are incoherent but consistent and ontologies that are coherent but inconsistent. In general, an incoherent ontology is an inconsistent ontology if one of the unsatisfiable concepts has an individual as instance

[5]More precisely, the authors referred to the corresponding notion as 'mapping inconsistency'.

[Flouris et al., 2006]. The term 'consistency' however has been used in [Jiménez-Ruiz et al., 2011] for defining the 'consistency principle' as a measure of the quality of alignments, stating that the ontology $o_1 \cup o_2 \cup A$ should be consistent and all the entities in its vocabulary should be satisfiable.

Alignment conservativity While the coherence measure proposes that correspondences should not lead to unsatisfiable concepts in the merged ontology, conservativity principle states that correspondences should not introduce new semantic relationships between concepts from the input ontologies. This principle has been introduced in [Jiménez-Ruiz et al., 2011, Solimando et al., 2014a]. In general, the conservativity principle suggests that the integrated ontology $o_{\mathcal{U}} = o_1 \cup_{A'} o_2$ should not introduce any change in the concept hierarchies of the input ontologies o_1 and o_2. That is, the deductive difference between o_1 and o_2, $diff^{\approx}_{\Sigma}(o_1, o_{\mathcal{U}})$ and $diff^{\approx}_{\Sigma}(o_2, o_{\mathcal{U}})$ must be empty for signatures $\Sigma_1 = Sig(o_1)$ and $\Sigma_2 = Sig(o_2)$, respectively.

In [Jiménez-Ruiz et al., 2011], a variant of this principle states that the alignment A' itself should not introduce new subsumption relations between the concepts from one of the input ontologies. In [Solimando et al., 2014a], it is required that the integrated ontology $o_{\mathcal{U}}$ does not introduce new subsumption relationships between concepts from one of the input ontologies, unless they were already involved in a subsumption relationship or they shared a common descendant. It is assumed the alignment to be coherent with respect to o_1 and o_2.

Definition 12 (Conservativity Violations [Solimando et al., 2014a]) *Given A,B,C atomic concepts (not including $\bot$, $\top$), let o be one of the input ontologies, let Sig(o) be the signature of o, and let $o_{\mathcal{U}}$ be the integrated ontology. The set of conservativity principle violations of $o_{\mathcal{U}}$ is a set of axioms of the form $A \sqsubseteq B$ satisfying: (i) A, B, C $\in$ Sig(o), (ii) $A \sqsubseteq B \in diff^{\approx}_{Sig(o)}(o, o_{\mathcal{U}})$, (iii) $o \not\models B \sqsubseteq A$, and (iv) there is no C s.t. $o \models C \sqsubseteq A$, and $o \models C \sqsubseteq B$.*

This variant of the conservativity principle follows the assumption of disjointness i.e., if two atomic concepts A and B from one of the input ontologies are not involved in a subsumption relationship nor share a common subconcept (excluding $\bot$) they can be considered as disjoint. Hence, the conservativity principle can be reduced to the consistency principle, if the input ontologies are extended with sufficient disjointness axioms.

In a scenario of query rewriting, as shown in [Solimando et al., 2014a], the quality of alignments in terms of the conservativity principle directly affects the quality of the query results, avoiding introducing noise in the results.

The logical-based metrics (coherence and/or conservativity) have been introduced as a measure of the alignment quality in diverse OAEI tracks, as Anatomy, Conference and LargeBio tracks.

Task-oriented evaluation

The evaluation criteria described above involve standard ways to evaluate ontology matching systems. However, the quality of an alignment can be assessed regarding its suitability for a specific task or application. In terms of measurements, it would be useful to set up experiments which do not stop at the delivery of alignments but carry on with the particular task or application. This is especially true when there is a clear measure of the success of the overall task (as thesaurus merging and data translation in [Isaac et al., 2008]). Early works have provided an analysis of the different needs for evaluation depending on specific applications [Ehrig, 2006]. [Euzenat and Shvaiko, 2013] argued that different task profiles can be established to explicitly compare matching systems for certain tasks, such as ontology evolution or query answering. Based on the analysis of such tasks, the requirements of applications can be established with regard to matching systems:

- input (for instance, applications require only a matching solution able to work without instances).
- some specific behavior of matching, such as requirements of (*i*) being *automatic*, i.e., not relying on user feedback, (*ii*) being *correct*, i.e., not delivering incorrect correspondences, (*iii*) being *complete*, i.e., delivering all the correspondences, and (*iv*) having a good *runtime* efficiency.
- the use of the matching result. In particular, how the alignment is going to be processed, e.g., by merging the data or conceptual models under consideration or by translating data instances among them.

Furthermore, the analysis in [Ehrig, 2006] can be rewritten in function of the measurements obtainable by evaluating the matchers (Table 2.2).

Application	speed	automatic	precision	recall
Ontology evolution	medium	low	high	high
Schema integration	low	low	high	high
Catalog integration	low	low	high	high
Data integration	low	low	high	high
P2P information sharing	high	low	medium	medium
Web service composition	high	high	high	low
Multi agent communication	high	high	high	medium
Context matching in ambient computing	high	high	high	medium
Semantic web browsing	high	medium	high	low
Query answering	high	medium	medium	high

Table 2.2: Application requirements interpreted as measurement weights (from [Euzenat and Shvaiko, 2013] (page 314)).

In OAEI, in particular, the approach adopted by the early library track organizers, for compensating the lack of complete reference alignments, was based on application relevance. They considered the provided alignment in the context of an *annotation translation* process supporting the re-indexing of books indexed with one vocabulary A, using concepts from the aligned vocabulary B [Isaac et al., 2008].

For each pair of vocabularies (A, B), this scenario interprets the correspondences as rules to translate existing book annotations with A into equivalent annotations with B. Based on the quality of the results for those books for which the correct annotations are known, the quality of the initial correspondences can be assessed. Later on, a new task-oriented evaluation approach was introduced in the OAEI in 2015 at the *OA4QA* track [Solimando et al., 2014b, Cheatham et al., 2015], which focused on the task of query answering. This track used a synthetically populated version of the *Conference* dataset and a set of manually constructed queries over these $\mathcal{A}boxes$. Precision and Recall will be calculated with respect to the ability of the generated alignments to answer a set of queries in a ontology-based data access scenario where several ontologies exist.

User related evaluation

So far the measures have been machine focused. In some cases algorithms or applications require some kind of user interaction [Dragisic et al., 2016]. This can range from the user utilizing the alignment results to concrete user input during the alignment process. In this case, it is even more difficult to obtain some objective evaluation.

Level of user input effort In the cases where the algorithms require user intervention, this intervention could be measured in terms of some elementary information the users provide to the system. When comparing systems which require different input or no input from the user, it will be necessary to consider a standard for elementary information to be measured. This is not an easy task. A first step towards evaluating the impact of user effort has been proposed in the OAEI anatomy track in 2008 [Caracciolo et al., 2008]. Participating systems could not only use the information encoded in the ontologies, but could also take into account a provided partial reference alignment as additional parameter. The additional information encoded in the partial reference alignment can be seen as a simulation of user input. Based on this approach it is possible to measure in how far this information can be exploited. Later on, the OAEI interactive track has been introduced in 2013 [Grau et al., 2013] with the aim at showing if user interaction can improve the matching results, which methods are most promising and how many interactions are necessary. Thus, beside the quality of the alignment, other measures like the number of interactions are used to decide which matching system is best suitable for a certain matching task.

General subjective satisfaction From a use case point of view it makes sense to directly measure the user satisfaction. As this is a subjective measure it cannot be assessed easily. Extensive preparations have to be made to ensure a valid evaluation. Almost all of the objective measures mentioned so far have a subjective counterpart. Possible measurements would be: input effort, speed, resource consumption (memory), output exactness (related to precision), output completeness (related to recall), and understandability of results (explanations). Due to its subjective nature numerical ranges as evaluation result are less appropriate than qualitative values such as very good, good, satisfactory, etc.

2.4 Conclusions

Matching ontologies consists of finding corresponding entities in different ontologies. Many different techniques have been proposed for implementing this process. An alignment (set of correspondences) is obtained by combining these techniques towards a particular goal (obtaining an alignment with particular features, optimizing some criterion, etc). The quality of the alignment can be assessed with the help of some measurement. This chapter has presented an overview of the different parameters and dimensions of the matching process together with the different criteria and metrics for ontology matching evaluation. As stated in Section 2.2, one dimension of the matching process refers to the constraints on the kind of language construct to be found in alignments from ontologies. Currently, relatively few matching systems are able to generated expressive (complex) alignments. The next chapter presents my contributions in the generation of alignments not limited to single (named) entities.

Chapter 3

Complex ontology matching

3.1 Motivation

Simple correspondences link one single entity of a source ontology to one single entity of a target ontology. This kind of correspondences however is not expressive enough to fully overcome the different kinds of heterogetities between different ontologies. The need for more expressiveness in ontology alignments has been recognized in diverse applications and domains. For mentioning a few, in the culture heritage domain, complex correspondences have been required for data translation and integration [de Boer et al., 2012, Szekely et al., 2013, Nurmikko et al., 2015]. In the agronomic domain, complex alignments help cross-querying linked open data repositories [Thiéblin et al., 2019a]. In the biomedical domain, complex alignments have also been used to build a consensual model from heterogeneous terminologies [Jouhet et al., 2017]. Complex alignments between medical ontologies have also been published [Fung and Xu, 2012, Giannangelo and Millar, 2012]. Recent work has shown that alignments between pairs of real-world ontologies contain many relations that are more complex than those targeted by current systems. As discussed in [Zhou et al., 2018], these more complex relationships often make up half or more of the relations within an alignment. Compared to approaches able to deal with simple correspondences, the generation of complex alignments is still addressed to a lesser extent in the ontology matching field.

Earlier works in field have introduced the need for complex alignments [Visser et al., 1997, Maedche et al., 2002], and different approaches for generating complex ontology alignments have been proposed in the literature afterwards. These approaches rely on diverse methods, such as correspondence patterns [Ritze et al., 2009, Ritze et al., 2010, Faria et al., 2018], knowledge-rules [Jiang et al., 2016], statistical methods [Parundekar et al., 2010, Parundekar et al., 2012, Walshe et al., 2016], genetic programming [Nunes et al., 2011] or path-finding algorithms [Qin et al., 2007]. In others fields, such as relational databases, different approaches have been proposed so far [Dhamankar et al., 2004, He et al., 2004]. Despite this variety, generating complex correspondences between ontologies (as more precise and expressive representations than taxonomies and database schemes) remains a challenge.

As introduced above, a task requiring complex correspondences is query rewriting, where most proposals address the task of rewriting SPARQL queries. A naive SPARQL rewriting approach consists in replacing the IRI of an entity of the initial query by the corresponding IRI in the correspondence, using simple correspondences. This approach is integrated in the Alignment API [David et al., 2011]. However, it does not take into account the specific kind of relation expressed in the correspondence (e.g., generalisation or specialization). The approach in [Euzenat et al., 2008] aims at writing `CONSTRUCT` SPARQL queries from complex alignments. A new knowledge base expressed with the source ontology vocabulary is populated with the instances of the target knowledge base. A rewriting approach not limited to queries of type `CONSTRUCT` and that takes advantage of complex (s:c) alignments has been proposed in [Correndo et al., 2010]. It applies a declarative formalism for expressing alignments between RDF graphs. In [Correndo and Shadbolt, 2011], a subset of EDOAL expressions are transformed into a set of rewriting rules. The expressions involving the restrictions on concepts and properties and the restrictions on property occurrences and values are not featured in the rewriting rules. In [Makris et al., 2010, Makris et al., 2012], the SPARQL-RW rewriting framework is presented. They define a set of correspondence types on which the rewriting process is based (i.e., *Class Expression*, *Object Property Expression*, *Datatype Property*, and *Individual*). Finally, in [Zheng et al., 2012] a rewriting algorithm that serves the purpose of context (i.e, units of measure) interchange for interoperability is proposed. While few works are able to deal with complex correspondences, they are limited in terms of covered constructors and transformations. Furthermore, (c:s) and (c:c) correspondences have not been fully addressed.

3.2 Contributions

We have proposed a matching matching approach able to generate complex alignments. It is based on the following assumptions: (a) the matching search space can be reduced by taking into account the specific user needs; and (b) at least one instance is shared between the knowledge bases whose ontologies has to be aligned. In fact, most matching approaches aim at fully aligning two ontologies, i.e., the output alignment aims at fully covering the common scope of the two ontologies. However, a user may not need as much coverage, as he or she may be interested by only a part of the ontology scope. Moreover, when reducing the scope of the ontologies, the matching task can be performed more efficiently and even allow for on-the-fly ontology matching [Lopez et al., 2006], in particular when dealing with large knowledge bases. The scale becomes even more problematic for complex matching where the number of possible correspondences is not $O(mn)$ as for simple matching, m and n being the number of entities from the source and target ontology, but worst than $O(2^{mn})$.

The need for complex correspondences has also been identified in the task of query rewriting in the context of an application translating natural language queries into SPARQL [Pradel et al., 2012]. This application relied on the notion of *query patterns* whose principle stated that, in real applications, the submitted queries are variations of few typical query families (i.e., the family of queries asking for the actors playing in movies). One of the main limitations of this approach is that for each data source to be queried, the corresponding query patterns have to be (manually) built. The idea was rewriting query patterns with the help of ontology alignments (i.e., rewrite a query pattern based on an ontology to query patterns based on another ontology). In a first set of experiments, as reported in [Gillet et al., 2013], run on the MusicBrainz and DBpedia collections, we used a set of simple correspondences for rewriting patterns. Despite the fact that the quality of the alignments has not bee measured, these first experiments highlighted the limitations in replacing individually the entities in query patterns. For that task, it turned out that simple correspondences were not sufficient to capture all the meaningful relations between entities of two related ontologies. This limitation has been also corroborated in the task of rewriting SPARQL queries, in a lower level of abstraction. We have hence started working on SPARQL query rewriting systems using complex correspondences and proposed an approach dealing with (s:c) complex correspondences.

The main contributions of the work on generating complex alignments and query rewriting using them are:

- a comprehensive overview of the complex matching approaches in the literature [Thiéblin et al., 2019c], as discussed in Section 3.2.1.
- an approach able to generate complex correspondences [Thiéblin et al., 2018d, Thiéblin et al., 2018c], as discussed in Section 3.2.2.
- an approach for rewriting SPARQL query patterns from (s:c) complex correspondences [Thiéblin et al., 2016], as discussed in Section 3.2.3.
- a set of manually established complex alignments between four agronomic datasets on the LOD, namely AgronomicTaxon, AGROVOC, DBpedia, and TAXREF-LD[Thiéblin et al., 2019a]. Part of this dataset has been used in the first complex track of OAEI in 2018 [Thiéblin et al., 2018a].

The work on the generation of complex correspondences has been carried out in cooperation with Élodie Thiéblin (PhD student at IRIT), co-advised with Ollivier Haemmerlé (Full Professor at University Toulouse 2 – Jean Jaurès and researcher at IRIT). The work on query rewritting is the result of a collaboration with Pascal Gillet (Master student co-advised with Ollivier Haemmerlé), Nathalie Hernandez (Assistant professor University Toulouse 2 – Jean Jaurès and researcher at IRIT), Catherine Roussey (Researcher at IRSTEA), Fabien Amarger (PhD student advised by Nathalie Hernandez, Catherine Roussey and Ollivier Haemmerlé) and Camille Pradel (PhD student advised by Nathalie Hernandez and Ollivier Haemmerlé).

In the following, the proposed classification of approaches able to generate complex correspondences is introduced (Section 3.2.1). Next, the approach for generating complex correspondences is presented (Section 3.2.2). Finally, the query rewriting approach is introduced (Section 3.2.3).

3.2.1 Classification for complex matching approaches

Diverse surveys in the literature have focused on the different aspects of schema and ontology matching [Rahm and Bernstein, 2001, Kalfoglou and Schorlemmer, 2003, Noy, 2004, Shvaiko and Euzenat, 2005, De Bruijn et al., 2006, Euzenat and Shvaiko, 2013]. However, none of them have addressed the specificities of complex matching (underlying strategy, structure of complex correspondences, etc.). We have reviewed the complex matching approaches dealing with different kinds of knowledge representation models (taxonomies, XML schemata, database schemata, formal ontologies, etc.) and a classification based on the specificities of complex alignments has been proposed (Figure 3.1). These specificities mainly rely on their output (types of correspondences) and their process (guiding structures), which are the two axes of the proposed classification.

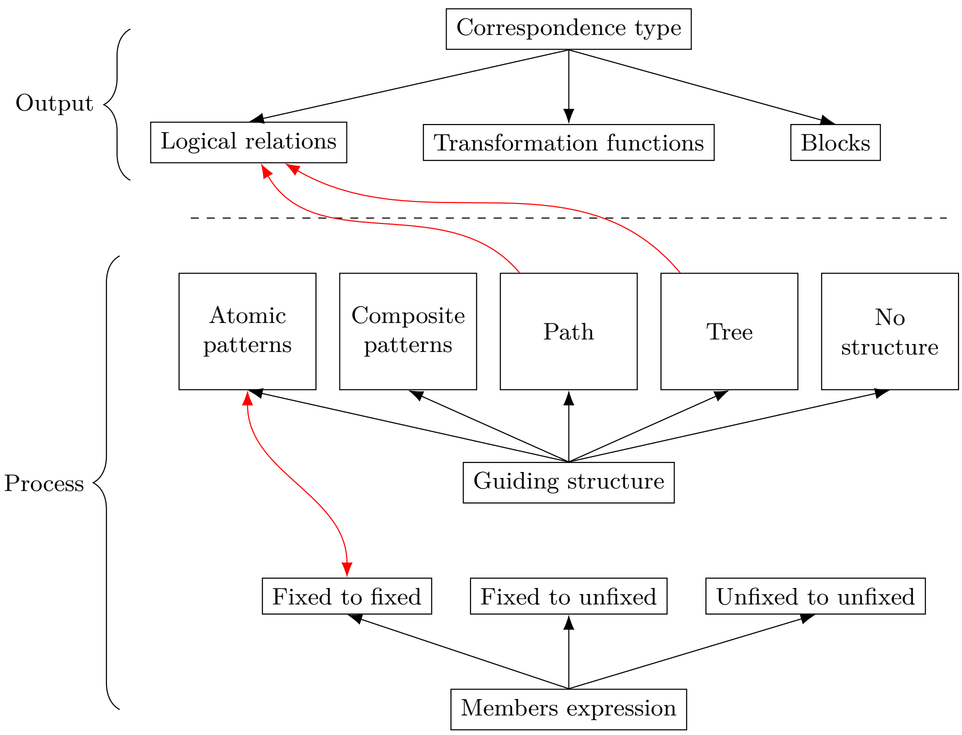

Figure 3.1: Two axes to characterize the complex matching approaches: output and process. The correlation between the categories are represented with red arrows (from [Thiéblin et al., 2019c]).

Correspondence type. The type of correspondence can be classified in *logical relations*, *transformation functions* and *blocks* categories. The *logical relations* category stands for correspondences in which the members are expressed with logical constructors. In contrast, the *transformation functions* category includes the approaches that generate correspondences with *transformation functions* in its members. The *blocks* correspondences gather entities using a grouping constructor (clusters of entities), not specifying a semantic relation between them.

Guiding structures. These categories aim at classifying the matching approaches based on their process dimension. It focuses on the structure on which the process generating the correspondences relies:

- *Atomic patterns* The approaches in this category consider the correspondence as an instantiation of an atomic pattern, such as those defined by Scharffe [Scharffe, 2009]. An atomic pattern is a template of a correspondence representing logical relation or transformation function correspondences. For example, an approach looking for correspondences following this exact pattern: (o_1:A, $\exists o_2$:$b.o_2$:C, $\equiv$) falls into this category and in the *logical relation* type of correspondence.

- *Composite patterns* The approaches in this category aim at finding repetitive compositions of an atomic pattern. For example, an approach looking for correspondences of the form (o_1:A, o_2:$B \sqcup o_2$:$C \sqcup o_2$:$D \sqcup \ldots$, $\equiv$), where o_1:A, o_2:B, o_2:C, o_2:D, *etc.* are classes and the number of unions in the target member of correspondences is not *a-priori* defined by the approach, falls into this category.

- *Path* The approaches in this category detect the correspondences using path-finding strategies. The resulting correspondence is a property path in o_1 put in relation with a path in o_2. For example, an approach looking for a path between two pairs of aligned instances described by o_1 resp. o_2 falls into this category.

- *Tree* The approaches in this category rely on tree structures inside the ontologies for correspondence detection. The ontologies are either considered as a tree or a tree-like structure is sought in an ontology graph. For example, when an XML schema is considered as a tree and the approach consists in finding the smallest equivalent tree in an ontology.

- *No structure* Contrary to the other approaches, the approaches of this category do not rely on a structure to guide the correspondence generation. Instead, they discover correspondences more freely.

As in Figure 3.1, the structures are used to guide the matching process, and therefore impact the structure of the output correspondences. For example, the correspondence (o_1:$AcceptedPaper$, $\exists o_2$:$acceptedBy.\top$, $\equiv$), could be obtained by an approach based on atomic patterns with the pattern (A, $\exists b.\top$, $\equiv$), by an approach based on composite patterns such as (A, $\exists b.\top \sqcup \exists c.\top \sqcup \ldots$, $\equiv$) or by an approach with no guiding structure.

Members expression. Finally, the *member expression* specifies whether one of the members of the correspondence is assigned a fixed structure or not before the process. Three types of pre-definition are possible: fixed to fixed, fixed to unfixed and unfixed to unfixed.

- *Fixed to fixed* category includes the matching approaches that always produce correspondences with fixed member expressions. Atomic pattern-based approaches generate fixed to fixed correspondences as both members' expressions are defined by the pattern. As shown in Figure 3.1, this category is strongly correlated to the Atomic-pattern guiding structure category.

- *Fixed to unfixed* member expression category covers the matching approaches for which one of the members of the correspondence will always follow the same expression template, while the expression of the other member may vary. For example, an approach aiming at finding for each property of an ontology a corresponding property path in the other ontology falls into this category: one of the members will always be one property while the other will be a path of a-priori undefined length.

- *Unfixed to unfixed* member expression category includes the approaches that output correspondences whose members have an undefined expression beforehand. For example, an approach aiming at finding similar paths in two ontologies falls into this category: both members have a-priori undefined length.

3.2.2 CANARD approach

The CANARD (Complex Alignment Need and A-box based Relation Discovery) system discovers complex correspondences between populated ontologies based on Competency Questions for Alignment (CQAs). We have introduced the notion of Competency Questions for Alignment (CQAs) as a way for representing the knowledge needs of a user and defining the scope of the alignment. This notion is inspired from the ontology authoring field, where competency questions have been introduced as *ontology's requirements in the form of questions the ontology must be able to answer* [Grüninger and Fox, 1995, Pinto and Martins, 2004, Ren et al., 2014]. They are competency questions that need to be satisfied over two or more ontologies. Our approach takes as input a set of CQAs translated into SPARQL queries over the source ontology. The answer to each query is a set of instances retrieved from a knowledge base described by the source ontology. These instances are matched with those of a knowledge base described by the target ontology. The generation of correspondences is performed by matching the graph-pattern from the source query to the lexically similar surroundings of the target instances.

Comparing the proposed approach to the approaches which involve the user (mostly for validation [Dragisic et al., 2016, Cruz et al., 2009, Noy and Musen, 2003]), or for user knowledge need expression [Lopez et al., 2006]), they do not deal with complex correspondences. On the other hand, none of the complex approaches involve the user before or during the matching process. Like the ones in [Hu et al., 2011, Parundekar et al., 2010, Parundekar et al., 2012, Walshe et al., 2016, Qin et al., 2007], the proposed solution relies on the assumption that the knowledge bases contain common instances. Furthermore, as for the matching processing in general, in particular for the complex matching approaches in [Ritze et al., 2009, Ritze et al., 2010], we rely on the assumption that the ontologies in the knowledge base have a relevant lexical layer. Differently from [Ritze et al., 2009, Ritze et al., 2010, Walshe et al., 2016, Parundekar et al., 2012, Parundekar et al., 2010], the approach does not rely on correspondence patterns. As far as we know, competency questions have neither been adapted nor used for ontology matching.

In this section, a definition of CQAs with their characteristics is introduced. Then, the proposed approach based on CQAs is presented.

CQA definition A Competency Question for Alignment (CQA) can be defined as a Competency Question (CQ) that needs to be satisfied over two or more ontologies. Therefore, an alignment is needed. CQAs can not be used for Ontology Authoring whereas CQs can be. Hence, the scope of a CQA is limited by the intersection of its source and target ontologies' scopes. Another difference is that the maximal and ideal alignment's scope is not known *a priori* (as it is the purpose of the alignment). The characteristics defined by [Ren et al., 2014] for ontology authoring (question type, element visibility, question polarity, predicate arity, modifier, domain independent element) are applicable CQAs except the *predicate arity* which depends on the associated SPARQL query. Indeed, a CQA has not only one but as many associated SPARQL queries as ontologies that it should cover.

For example, the CQA "What are the accepted papers ?" can be represented by `SELECT ?x WHERE {?x a o1:AcceptedPaper.}` in which there is only a unary predicate (`o1:AcceptedPaper`) with only explicit elements or by `SELECT ?x WHERE {?x a o2:Paper. ?x o2:hasDecision o2:accept.}` in which `o2:hasDecision` is a binary predicate and an implicit element of the query. We chose to adapt only the definition of predicate arity for the CQA into **question arity**. The **question arity** represents the arity of the expected answers to a CQA.

- A *unary* question expects a set of instances or values, e.g., "What are the accepted papers?" *(paper1), (paper2).*

- A *binary* question expects a set of instances or value pairs, e.g., "Who is the reviewer of a paper?" *(reviewer1, paper1), (reviewer1, paper2).*

- A *n-ary* question expects a tuple of size 3 or more, e.g., "What is the decision of a paper given by a reviewer?" *(paper1, reviewer1, accept), (paper3, reviewer2, reject).*

Matching approach The proposed approach takes as input a set of CQAs in the form of SPARQL queries over the source ontology. It requires that the source and target ontologies have an $\mathcal{A}box$ with at least a common instance. The answer to each input query is a set of instances, which are matched with those of a knowledge base described by the target ontology. The matching is performed by finding the lexically similar surroundings of the target instances. CQAs for the approach are limited to questions of *select* type, positive polarity and no modifier. The choice of the *select* question type, comes from the fact that *binary* and *counting* questions have a corresponding *select* question. With regards to the question polarity, a negative question implies that a "positive" information is being negated, therefore, the questions can be limited to positive polarity only. We make the assumption that the user knows the source ontology and is able to write each CQA into a SPARQL on the source ontology. The approach is articulated in 11 steps, as depicted in Figure 3.2:

1. Extract source DL formula e_s from SPARQL CQA (e.g., o_1*:AcceptedPaper*)

2. Extract lexical information from the CQA, L_s set labels of atoms from the DL formula (e.g., "accepted paper")

3. Extract source instances $inst_s$ (e.g., o_1*:paper1*)

4. Find equivalent or similar (same label) target instances $inst_t$ to the source instances $inst_s$ (e.g. o_1*:paper1* $\sim$ o_2*:paper3*)

5. Retrieve the description of target instances: the set of triples in which the target instances appear as well as the object/subject type of the triple (e.g. in DL, the description of o_2*:paper3* would be:
$\langle$*(*o_2*:paper3,*o_2*:accept):*o_2*:hasDecision;* o_2*:accept:*o_2*:Decision*$\rangle$;
$\langle$*(*o_2*:reviewer1,*o_2*:paper3):*o_2*:reviewerOf;* o_2*:reviewer1:*o_2*:Reviewer*$\rangle$

6. For each triple, retrieve L_t labels of entities (e.g., o_2*:hasDecision* $\rightarrow$ "decision", o_2*:accept* $\rightarrow$ "accept", o_2*:Decision* $\rightarrow$ "decision")

7. Compare L_s and L_t using a string comparison metric (e.g., Levenshtein distance with a threshold)

8. Keep the triples with the summed similarity of their labels above a threshold τ. Keep the object/subject type if its similarity is better than the one of the object/subject (e.g. sim(o_2*:accept*, L_s) > sim(o_2*:Decision*,L_s) so we only keep o_2*:accept* in the triple)

9. Express the triple into a DL formula (e.g., $\exists$ o_2*:hasDecision.*{o_2*:accept*})

10. Aggregate the formulas into an explicit or implicit form. If two DL formulas have a common atom in their right member (target member)

11. Put the DL formulae e_s and e_t together in a correspondence (e.g., o_1*:AcceptedPaper* $\equiv$ $\exists$ o_2*:hasDecision.*{o_2*:accept*}) and express this correspondence in EDOAL

The instance matching phase of the step (4) is based on existing *owl:sameAs*, *skos:closeMatch*, *skos:exactMatch*. In the absence of such links, an exact label matching is applied. The similarity between the sets of labels L_s and L_t of step (7) is the cartesian product of the string similarities between the labels of L_s and L_t (Equation 3.1).

$$sim(L_s, L_t) = \sum_{l_s \in L_s} \sum_{l_t \in L_t} strSim(l_s, l_t) \tag{3.1}$$

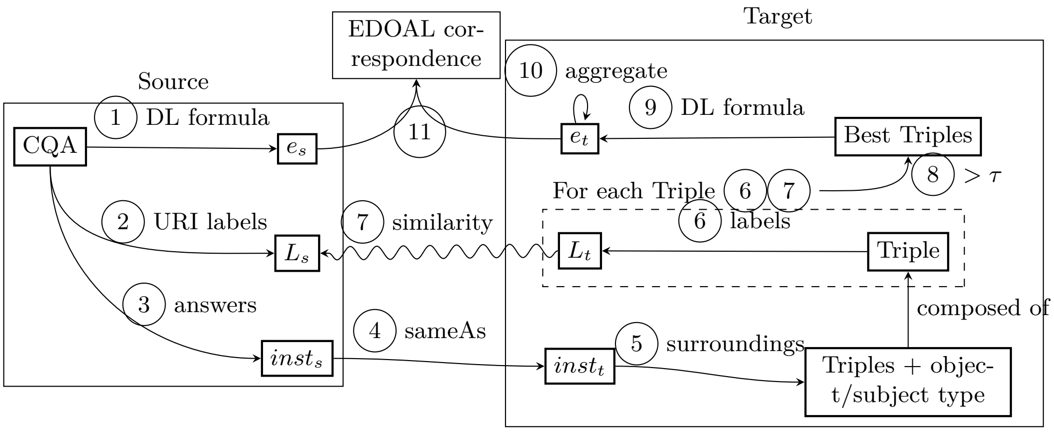

Figure 3.2: Schema of the general approach (from [Thiéblin et al., 2018c]).

$strSim$ is the string similarity between two labels l_s and l_t (equation 3.2). τ is the threshold for the similarity measure.

$$strSim(l_s, l_t) = \begin{cases} \sigma \text{ if } \sigma > \tau, \text{ where } \sigma = 1 - \dfrac{levenshteinDist(l_s, l_t)}{\max(|l_s|, |l_t|)} \\ 0 \text{ otherwise} \end{cases} \tag{3.2}$$

The confidence value given to the final correspondence, step (11), is the similarity of the triple it comes from or average similarity if it comes from more than one triple. The confidence value is reduced to 1 if it is initially calculated over 1. Currently, the approach has been extended using the notion of counter-examples, which are used for reassessing the similarity of each DL formula. The counter-examples are common instances of the two knowledge bases which are described by the target DL formula but not by the original CQA. The percentage of counter-examples is taken into account for calculating the final correspondences confidence.

Evaluation A first version of the approach, limited to unary queries, has been firstly evaluated in the OAEI complex track 2018 [Thiéblin et al., 2018c]. The track consisted of four datasets: conference organization, hydrography, geoscience, and plant taxonomies. Each dataset was evaluated in a different way and only plant taxonomy dataset has instanciated ontologies. For this dataset, the evaluation was two-fold. First, the precision of the output alignment with respect to exact match against the reference was manually assessed. Then, a set of source queries was rewritten using the output alignment. Each rewritten target query was then manually classified as correct or incorrect. A source query was considered successfully rewritten if at least one of the target queries was semantically equivalent to it. As the datasets does not cover CQAs, we developed a CQA generator that was integrated to the version of the system used in the evaluation.

CANARD could only output correspondences for the Taxon dataset of the Complex track. CANARD was able to generate different kinds of correspondences: (s:s), (s:c) and (c:c). The best precision was obtained for the pair agronomicTaxon-agrovoc with a precision of 0.57. CANARD did not output any correspondence for 4 oriented pairs (agrovoc-agronomicTaxon, dbpedia-agronomicTaxon, taxref-agronomicTaxon and taxref-dbpedia). These empty results can be due to the fail of the instance matching phase of our approach. We could observe that with TaxRef as the source knowledge base, no correspondence could be generated. The exception is the pair *taxref-agrovoc* where 8 correspondences were found but only involving *skos:exactMatch* or *skos:closeMatch* properties in the constructions. Looking for the query rewriting task in Taxon, CANARD's alignment was used to rewrite the most queries (best *qwr*). As CANARD does not deal with binary CQAs, none of the 3 binary queries × 12 pairs of ontologies = 36 binary query cases could be dealt with. Out of the 2 unary queries × 12 pairs = 24 unary query cases, CANARD could deal with 6 unary cases needing a complex correspondence and 2 needing simple correspondences for a total of (8/24) 33% of unary query cases. Overall, for the query cases needing complex correspondences, 14% were covered by CANARD. For all the query cases, the CANARD system could provide an answer to 13% of all cases. Currently, the approach is being deeply evaluated in the context of the PhD thesis of Élodie Thiéblin.

3.2.3 Query rewriting with complex correspondences

We have proposed a set of rules for automatically rewriting SPARQL queries based on (s:c) complex alignments in [Thiéblin et al., 2016]. Differently from [Euzenat et al., 2008], our approach rewrites SPARQL queries instead of writing them from a complex alignment. Unlike [Correndo and Shadbolt, 2011], the proposed mechanism can handle restrictions on concepts and relations. In comparison with the approach in [Makris et al., 2010], EDOAL is an alternative to represent alignments in a more expressive (thus complete) way. For instance, we propose occurrence and property datatype restrictions translation rules. However, our approach is limited to (s:c) complex alignments and does not handle initial SPARQL queries containing filters, unions, or other SPARQL options. The approach is based on the assumption that the queries to be transformed aim at retrieving new instances to meet a certain need. This is why only $\mathcal{T}box$ elements are taken into account. [Zheng et al., 2012] focuses on context correspondences while our approach intends to translate all $\mathcal{T}box$ elements of a query. More recently, an instance-based rewriting system has been proposed as part of the approach for evaluation complex alignments, as further discussed in Chapter 6. This system can deal with (c:c) correspondences but cannot combine correspondences in the rewriting process, *i.e.*, if more than one correspondence is needed to rewrite the query, the system can not deal with it. The approaches for query rewriting based on (s:c) and (c:c) complex correspondences are briefly introduced in the following.

SPARQL query rewriting approach using (s:c) correspondences The proposal consists in a set of rules for automatically rewriting (`SELECT`) SPARQL queries based on (s:c) complex alignments. It focuses on a subset of SELECT SPARQL queries of the type: $Q_{o_1} = SELECT\ ?(\,Var+\,|\,'*'\,)\ WHERE\ \{\ T_{Q_{o_1}}\ \}$ where Var corresponds to the set of variables used as projection attributes and $T_{Q_{o_1}}$ stands for the query pattern made of triples expressed using the source ontology o_1. A triple t of $T_{Q_{o_1}}$ is composed of a subject s, a predicate p and an object o. We only consider triples where s is a variable. The aim is to produce the set $T_{Q_{o_2}}$ that contains the triples expressed according to entities of the ontology o_2, by using the complex alignment $A_{o_1 \rightarrow o_2}$. The approach is limited to complex correspondences establishing an equivalence relation between entities of same nature (classes, relations or properties). We also make the assumption that the alignment is complete and covers all the correspondences required to transform the entities of $T_{Q_{o_1}}$. We define rules that take the set of triples T_{Q_o} as input and generate a SPARQL query. Three types of triples in T_{Q_o} are considered: *Class Object Triples*, *Predicate Triples* and *Other Triples*.

Algorithm 1 depicts the SPARQL query rewriting process. The **rewriteClassObject** and **rewritePredicate** functions apply a set of rules specific to each kind of element in the triple. These functions are recursive and can call each other. If a triple is not a *Class Object Triples* or a *Predicate Triples*, it means that its subject s is a variable, its predicate is an object property or a data property for which no correspondence is needed and its object is either a literal or a variable. This kind of triple does not need any transformation and is directly added to the final query. An example of such triple is `?s rdfs:label "a literal"`.

Algorithm 1 Rewriting algorithm.

```
new_query ← " "
for all triple t = ⟨s, p, o⟩ in query do
  if t is a Class Object Triple then
    new_query ← new_query + rewriteClassObject(s,p,o_o2)
  else if t is a Predicate Triple then
    new_query ← new_query + rewritePredicate(s,p_o2,o)
  else
    new_query ← new_query + t
  end if
end for
return new_query
```

Table 3.1 presents an example of a query transformation based on the complex correspondences in Section 2.2.1 between the ontologies *ekaw* (Figure 2.2) and *cmt* (Figure 2.3). This correspondence involves a *class object triple* with a *ClassRestriction*, in particular a *DomainRestriction*, which limits the range of a relation expression to a class expression. The **rewritePredicate** function is called with the relation $relation(o_{o_2})$ between the subject s and an intermediate variable v. The **rewriteClassObject** function is called to assert that v is an instance of the $range(o_{o_2})$ class expression: **rewritePredicate**$(s, relation(o_{o_2}), v)$ + **rewriteClassObject**$(v,$ `rdf:type`$, range(o_{o_2}))$.

Validation The approach has been validated on two datasets. The first one was built to meet the needs of agriculture experts willing to find cross knowledge about

Table 3.1: Transformation of a class triple based on the correspondences from the ontologies *ekaw* (Figure 2.2) and *cmt* (Figure 2.3) between a *class object triple* and a *class expression*.

Source query
`SELECT ?z WHERE {` `?z rdf:type ekaw:Accepted_Paper.}`
Rewritten query

`SELECT ?z WHERE {`
`?z rdf:type cmt:Paper.`
`?z cmt:hasDecision ?var_temp.`
`?var_temp rdf:type cmt:Acceptance.}`

agronomic taxon between DBpedia and AgronomicTaxon[1], a dedicated knowledge base. The second dataset was inspired from a subset of queries from the OAEI oa4qa[2] task data set. Our validation was based on the manual comparison of the set of results returned from the automatically rewritten query with respect to the results of the reference query. 10 complex correspondences (and 1 simple) have been manually produced between Agronomic Taxon and DBpedia. 8 simple and 6 complex correspondences have been manually produced between the three ontologies of the Conference data set. The alignments are available online[3]. Even though the reference query and the rewritten one differ in terms of syntax, they retrieve the same set of instances. It is the case for all the considered queries. The whole set of rewritten queries is available online[4]. More recently, in [Thiéblin et al., 2019a], this work has been extended with two datasets (AGROVOC[5] and TAXREF-LD [Michel et al., 2017]) and the complex alignments have been manually created by domain experts. The findings in this experiment highlight the high heterogeneity for representing taxon classifications in the LOD. Queries expecting a part of ontology as a result or using the structure of the query itself do not give good results with the complex alignments used here. The granularity heterogeneity between ontologies also affected the semantic equivalence between two queries. Because of the scope heterogeneity between ontologies, some queries could not been rewritten. These findings could be further exploited in an iterative process of alignment construction.

SPARQL query rewriting approach using (c:c) correspondences The second system is based on instances and has been used in the task of evaluating complex alignments, as further developed in Chapter 6. For each correspondence c_i of A', the instances represented by its source member e_s are retrieved over the source knowledge base *skb*. If $I_{e_s}^{skb} \equiv I_{query_s}^{skb}$, then, the target member of c_i is transformed into a query and added to Q_T. For example the source query from Table 3.1 retrieves

[1] `http://ontology.irstea.fr/AgronomicTaxon`
[2] `http://oaei.ontologymatching.org/2015/oa4qa/index.html`
[3] `https://www.irit.fr/recherches/MELODI/telechargements/alignements.zip`
[4] `https://www.irit.fr/recherches/MELODI/telechargements/requetesgenerees.zip`
[5] `http://aims.fao.org/aos/agrovoc/`

a set of accepted paper instances in the o_1 ontology. This set of instances is then compared to the set of instances described by the source member of each correspondence. In this case, o_1*:AcceptedPaper* describes exactly the same set of instances as the source member of c_1. The target member of c_1 can therefore be transformed into the rewritten query in Table 3.1. This rewriting system can deal with (c:c) correspondences but cannot combine correspondences in the rewriting process, *i.e.*, if more than one correspondence is needed to rewrite the query, the system can not deal with it. This system requires however knowledge bases regularly populated and that sense it is suitable to the task of evaluating complex correspondences.

3.3 Conclusions

This chapter has presented my contributions in the generation of complex alignments and the use of this kind of alignment in the task of SPARQL query rewriting. While most solutions in the field have been dedicated to the generation of simple alignments, we proposed an approach that takes into account the user needs, in the form of competency questions for alignment, in order to guide the discovery of more expressive correspondences. This 'discovery' process is based on common instances between the ontologies to be matched, adopting a naive form of generalization (i.e., based on the lexical comparison of annotations of surrounding entities). In parallel, query rewriting approaches are mostly limited to simple alignment, while the ones dealing with complex correspondences are mostly limited to specific kinds of them (s:c). In order to fully cover the expressivity of complex correspondences in this task, we have proposed query rewriting systems dealing with (s:c) and (c:c) correspondences, but still failing in combining multiple (c:c) correspondences together.

The matching approach presented here can be extended in several directions: one could consider exploring more sophisticated instance-based matching approaches and, alternatively, conditional keys [Symeonidou et al., 2017] or link keys [Atencia et al., 2014] (systems generating keys could also benefit from complex correspondences to improve their results); designing a purely T-Box strategy based on both linguistic and semantic properties of the ontologies and CQAs; or still dividing the problem in sub-tasks through ontology partitioning (given the inherent high search space in this task). Last but not least, incoherence resolution systems for complex alignments are scarce (the proposal by [Meilicke, 2011] is dedicated to simple and pairwise alignments).

In the next chapter, the task of holistic matching and my contributions in this topic are presented. However, they are limited to the generation of simple holistic alignments.

Chapter 4

Holistic ontology matching

4.1 Motivation

Efforts in ontology matching have been mostly dedicated to the *pairwise ontology matching* task (i.e., matching a single pair of ontologies). However, with the increasing amount of knowledge bases being published on the Linked Open Data cloud, covering different aspects of overlapping domains, the ability of simultaneously matching different ontologies, a task so-called *holistic ontology matching* [Rahm, 2011, Rahm, 2016, Megdiche et al., 2016a], is more than ever required. It is typically the case in complex domains, such as bio-medicine, where several ontologies describing different but related phenomena have to be linked together [Oliveira and Pesquita, 2015]. As stated in [Pesquita et al., 2014], the increase in the matching space and the inherently higher difficulty to compute alignments pose interesting challenges to this task.

We can see the pairwise matching as a special case of holistic ontology matching. The *holistic ontology matching* takes a set $\Omega = \{o_1, ..., o_N\}$ of ontologies with N ≥ 2. It consists in determining a set of correspondences as $A'_{1...N} = \{c_1, c_2, ..., c_M\}$. Each correspondence c_i is defined as $< \{e_1, ..., e_N\}, r, n >$ such as $\forall j \in [1..N], e_j \in o_j$. For instance, if $\Omega = \{o_1, o_2, o_3\}$, then the alignment is defined as $A' = A'_{12} \cup A'_{13} \cup A'_{23}$ where:

- $A'_{12} = \{< e_1, e_2, r_{12}, c_{12} > | e_1 \in o_1 \wedge e_2 \in o_2\}$,
- $A'_{13} = \{< e_1, e_3, r_{13}, c_{13} > | e_1 \in o_1 \wedge e_3 \in o_3\}$,
- $A'_{23} = \{< e_2, e_3, r_{23}, c_{23} > | e_2 \in o_2 \wedge e_3 \in o_3\}$.

Triple correspondences between entities of o_1, o_2, and o_3 can be deduced from A' by detecting *cliques*; *e.g.*, each subset of adjacent correspondences $< e_1, e_2, r, c_{12} >$, $< e_1, e_3, r, c_{13} >$ and $< e_2, e_3, r, c_{23} >$. The main limitation of the pairwise approaches regard to the holistic approaches is that in the former, A' is considered as a local solution depending of the order with which the ontology matching is carried out; *e.g.* $A_{12} \cup A_{(12)3} \neq A_{13} \cup A_{(13)2} \neq A_{23} \cup A_{(23)1}$. Thus the set of correspondences in A' differs according to the order the ontology matching pairwise approach is applied. This opens interesting challenges in the field, such as dealing with the logical coherence, optimization of the process when dealing with large ontologies, etc.

Early works on the database field have addressed the problem of holistic schema matching, in particular the works on attribute matching [He and Chang, 2003, He et al., 2004, Su et al., 2006, Saleem et al., 2008]. In [He and Chang, 2003], a probabilistic framework determines an underlying model capturing the correspondences between attributes in different schemes. For dealing with complex attribute correspondences, the approach in [He et al., 2004] exploits co-occurrence information across schemes and a correlation mining method. This approach has been extended in [Su et al., 2006] improving accuracy and efficiency, by reducing the number of synonymous candidates. In [Saleem et al., 2008], the approach aims at incrementally merging 2-way schemes by clustering the nodes based on linguistic similarity and a tree mining technique. These approaches however handle relatively simple structures compared to the more structured schemes of ontologies.

Emerging works have addressed the problem of holistic matching of more expressive structures. In [Gruetze et al., 2012], the proposal relies on a cross-domain holistic matching approach for aligning large ontologies by grouping concepts in topics that are aligned locally. More recently, a cluster-based distributed holistic approach for data linking has been proposed in [Nentwig et al., 2017], which is based on a clustering of entities representing the same real-world object. In [Pesquita et al., 2014], the holistic AML-Compound system extends the pairwise AML system adapting WordNet similarities and Jaccard indexes.

4.2 Contributions

We have proposed a holistic ontology matching approach that extends a previous contribution in the field of schema matching, especially designed to hierarchical schema structures like XML. Differently from [He and Chang, 2003, He et al., 2004, Su et al., 2006, Gruetze et al., 2012], the approach is not restricted to attributes, while we do not perform cross-domain holistic matching as [Gruetze et al., 2012]. Compared to [He et al., 2004], the proposed approach can also return simple and multiple correspondences and it is extensible to new constraints, differently from [Pesquita et al., 2014]. As some pairwise matchers [Jean-Mary et al., 2009, Jiménez-Ruiz and Grau, 2011], we adopt constraints that reduce the possibility of generating incoherent alignments. With respect to the matching strategies we apply, while the selection strategy in [Xiang et al., 2015a] is based on paths in the graph, we reduce the selection to the maximum-weighted bipartite graph matching (MWGM) problem like OLA [Euzenat and Valtchev, 2004] and we adopt a different structural similarity strategy from [Hu et al., 2005]. Compared to OLA we do not compute structural similarities but encode structural properties as linear constraints. Unlike CODI whose pairwise approach is reduced to a NP-Hard problem, our solution extends a polynomial problem in both pairwise and holistic versions [Megdiche et al., 2016a]. In a holistic and monolingual setting, we apply a combinatorial optimisation problem using linear programming, as done in [Prytkova et al., 2015] in pairwise. The constraints proposed by [Prytkova et al., 2015] for multiple correspondences, can be simply added to our model to enhance the matching of multiple correspondences in the relaxed version of our model.

Summing up, the main contributions of the work on holistic matching are as follows:

- an approach to determine holistic correspondences between multiple ontologies [Megdiche et al., 2016a, Megdiche et al., 2016b]. We model the approach within a linear program by reducing the ontology matching problem to the maximum-weighted graph matching problem, which is solvable in polynomial time.
- the approach is extensible with different structural similarity strategies and several linear constraints [Roussille et al., 2018b], insuring mostly coherent alignments.

This work has been carried out in cooperation with Philippe Roussille (post-doc at IRIT), co-advised with Imen Megdiche (Assistant professor at Institute Jean-Francois Champollion and researcher at IRIT) and Olivier Teste (Full professor at University Toulouse 2 – Jean Jaurès and researcher at IRIT).

In the next section the proposed approach is briefly introduced.

4.2.1 LPHOM: holistic approach

The LPHOM approach (Linear Program for Holistic Ontology Matching) is based on a well-known combinatorial optimisation problem, the maximum-weighted graph matching (MWGM) problem [Schrijver, 2003]. The idea consists in generalizing the pairwise matching on a set of N input ontologies through generic decision variables and generic linear constraints modelled in a linear program. The MWGM problem aims at finding a set of disjoint edges having the maximum weights in a weighted graph G. Indeed, we consider that G expresses the potential candidate correspondences between the input ontologies and has (i) three types of nodes representing classes, object and data properties and (ii) edges representing virtual connections between the same types of nodes (i.e classes related to classes, object properties to object properties and data properties to data properties). These edges have weights that represent similarities between the nodes and can be established using different strategies. In our approach, the similarities are calculated in a pre-processing step. In this setting, searching simple correspondences (1:1) with a maximum weight on similarities is equivalent to find a set of disjoint edges with a maximum weight in the MWGM problem.

Approach The proposed approach involves a pre-processing step and a processing step. In the pre-processing step, we apply element-level matchers and then aggregate the results in order to produce similarities between the entities of the ontologies. In the processing step, we instantiate the different elements of the linear program (decision variables and linear constraints) and then resolve the model by using the CPLEX solver. Overall, LPHOM follows the execution workflow composed of four main steps:

1. The first step consists in ontology loading and flattening. After loading the N different ontologies, we flatten every ontology entity (classes, object properties and data properties) in a same structure, named *Node*. Classes, object properties and data properties inherit from *Node*. The idea behind flattening the ontologies is to simplify the access to all information about each entity, which can be seen near to the structure of document-oriented NoSql databases. But actually, as duplication and treatment are done in memory, pre-processing is not very performant.

2. The second step consists of similarity matrices construction. For a set of N ontologies, we compute $N(N-1)/2$ similarity matrices representing the average results of different element-level matchers. These matrices are computed between each pair of ontologies and for each type of entity (classes, object properties and data properties). Similarity matrices have been constructed with *character-based metrics* [Sun et al., 2015] (ISUB and 3-gram to compute similarity between tokens then generalized Mongue-Elkan method on these metrics to get the similarity between entities) and *token-based category* (Jaccard).

3. The third step consists of constructing the linear program, which is detailed in [Megdiche et al., 2016a]. The algorithm was developed in Java by the mean of the methods proposed by the Java API of the CPLEX Solver[1]. For constructing the linear program, we consider only the pairs of correspondences (our decision variables), which similarity measure is higher than 0.65 (empirically chosen). We highlight also that the used threshold is the same for each type of entity (classes, object properties and data properties).

4. The fourth step consists of resolving the linear program using the CPLEX solver. The solution represents the set of final correspondences.

Evaluation The approach has been firstly evaluated in both pairwise and holistic matching settings, on the OAEI Conference dataset (results reported in [Megdiche et al., 2016a]). In the pairwise setting, we compared the results of our approach with the results of the 14 matchers participating in the 2015 OAEI campaign. Overall, our approach reaches intermediate results for its first comparison with regard to the pairwise ontology matching problem. Our model is more efficient when we use all the proposed constraints. The interaction between constraints leads to semantically significant results closer to gold references which are illustrated by a good recall on semantic distances. The constraints proposed for reducing incoherence are experimentally efficient. We applied the approach from [Meilicke, 2011] to evaluate if there is incoherence in our results and we get the following average results (for the 21 combinations): for ra1-m1 (only classes) we have 0,95 removed correspondences; for ra1-m2 (only properties) we have 0 removed correspondences; and for ra1-m3 (classes and properties) we have 0,85 removed correspondences.

[1] http://www.ibm.com/support/knowledgecenter/SSSA5P_12.6.2/ilog.odms.cplex.help/refjavacplex/html/index.html

For the holistic matching evaluation, given the lack of benchmarks dedicated to the evaluation of holistic ontology matching, we analyze: (i) the differences between cliques manually deduced from reference alignments and the cliques generated by our holistic approach; (ii) the differences between the results of iterative pairwise and holistic matching settings. In the following, we denote a clique as $Cl_i =< e_1, ..., e_N >$, such as each e_j belongs to ontology o_j. For the seven available ontologies in the Conference track, which are classified into types (Tool, Insider and Web), we selected three ontologies from the 'Tool' type (*cmt*, *conf-of*, *edas*). Indeed, in order to maximize the chance to have cliques in the reference alignments, we have tried to find $N \geq 2$ ontologies of the same type. The only combination of ontologies verifying that was *cmt*, *conf-of*, and *edas*, which the reference alignments are available. From the reference alignments, we have manually identified the following four cliques (Figure 4.1).

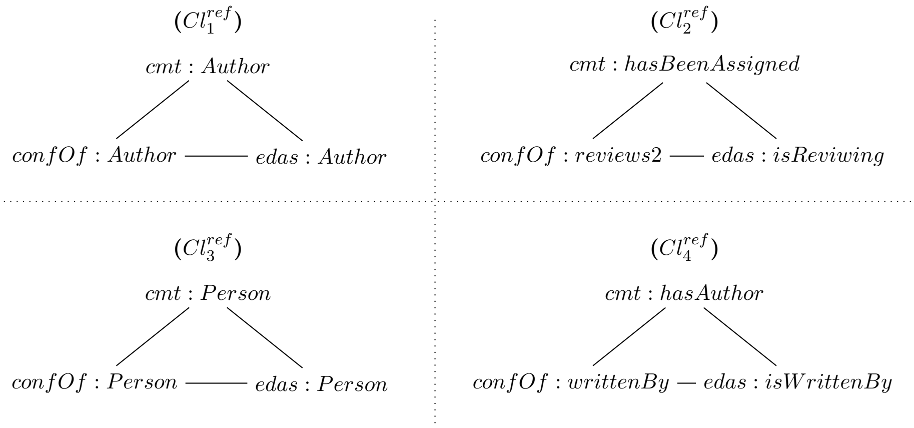

Figure 4.1: Reference cliques (Cl_1^{ref}), (Cl_2^{ref}), (Cl_3^{ref}), and (Cl_4^{ref}).

Our approach was able to identify 6 cliques, with 3 of them in the reference cliques set (Cl_1^{ref}, Cl_3^{ref} and Cl_4^{ref}). Two of the remaining 3 cliques (Figure 4.2), however, are somehow close to a possible clique. Cl_2 (Figure 4.2) is composed of the same concept *Paper* occurring in all ontologies. In the reference alignments, however, the correspondences in which *Paper* occur does not form a clique. The Cl_5 clique is particularly interesting since that the properties of Cl_5 are the inverse of the properties of Cl_4^{ref}. Finally, Cl_6 is composed of similar data properties which is also relevant but not provided in the reference alignments.

We have also analyzed the differences between the results of pairwise and holistic matching settings, applied to the ontologies *cmt*, *sigkdd* and *confOf*. Holistic approach discovers simultaneously alignments for N ontologies, from all combinations of pairs of input ontologies. The resulting alignments are collected from a simultaneous resolution of $A_{cmt-sigkdd}$, $A_{cmt-confOf}$ and $A_{sigkdd-confOf}$, as shown in Table 4.1.

When performing the holistic matching for *cmt*, *sigkdd* and $confOf$, we get the following alignments (Table 4.2). From these alignments, two cliques are deduced (Table 4.3). These results presented show the subtleties between a local and

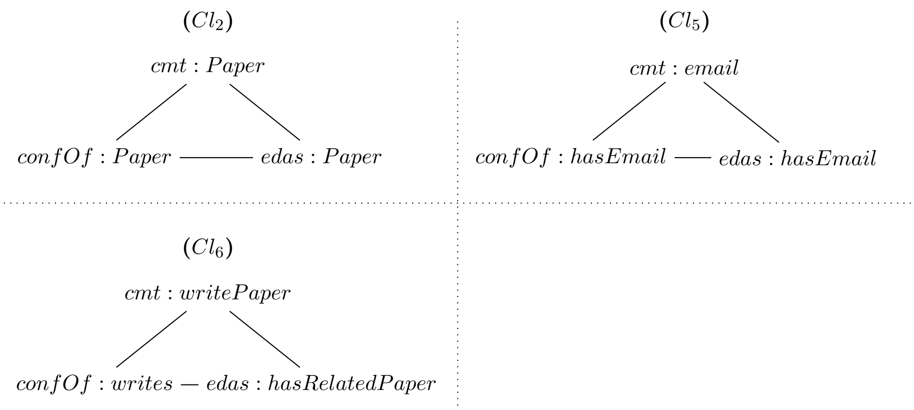

Figure 4.2: Cliques found by the proposed holistic approach.

Table 4.1: Resolution of $A_{cmt-sigkdd}$, $A_{cmt-confOf}$ and $A_{sigkdd-confOf}$.

$A_{cmt-sigkdd}$	$\{< Conference, ConferenceHall, \equiv, 0.63 >$ $< ConferenceMember, Conference, \equiv, 0.66 >$ $< Paper, Paper, \equiv, 1 >\}$
$A_{(cmt-sigkdd)-confOf}$	$\{< Conference, Conference, \equiv, 1 >$ $< Paper, Paper, \equiv, 1 >\}$
$A_{cmt-confOf}$	$\{< Paper, ShortPaper, \equiv, 0.63 >,$ $< PaperFullVersion, Paper, \equiv, 0.66 >,$ $< Conference, Conference, \equiv, 1 >\}$
$A_{(cmt-confOf)-sigkdd}$	$\{< Conference, Conference, \equiv, 1 >,$ $< ShortPaper, AuthorOfPaper, \equiv, 0.5 >,$ $< Paper, Paper, \equiv, 1 >\}$

global investigations on $N \geq 2$ ontologies, which confirm the usefulness of holistic approaches for ontology matching.

Table 4.2: Holistic matching for *cmt*, *sigkdd* and *confOf*.

$A_{cmt-sigkdd}$	$\{< Conference, Conference, \equiv, 1 >,$ $< Paper, Paper, \equiv, 1 > \}$
$A_{sigkdd-confOf}$	$\{< Conference, Conference, \equiv, 1 >,$ $< Paper, Paper, \equiv, 1 >\}$
$A_{cmt-confOf}$	$\{< Conference, Conference, \equiv, 1 >,$ $< Paper, Paper, \equiv, 1 >\}$

Table 4.3: Deduced cliques from the alignment in Table 4.2.

Cl_1	$\{< Paper_{cmt}, Paper_{sigkdd}, Paper_{confOf} >\}$
Cl_2	$\{< Conference_{cmt}, Conference_{sigkdd}, Conference_{confOf} >\}$

The approach has been also evaluated in other OAEI datasets in the context of the OAEI campaigns [Megdiche et al., 2016b]. In particular, we can observe that our tool is quite slow to perform the Anatomy track, and takes about 26 min (the faster system took 20 seconds). The non-scalability of our tool is closely dependant on the non-optimized pre-processing steps. Furthermore, we observed that some incoherent results have been obtained for this track. In fact, the constraints we have proposed are mainly limited to non-disjoint entities. We should may add some new constraints in our model in order to tackle the incoherence generated in this track. For the Conference track, an interesting aspect on our results concerns conservativity and consistency violation, with no conservativity principle violation nor consistency violation being reported.

Finally, in [Roussille et al., 2018b], Hontology (the successor of LPHOM) has been participated in the OAEI campaign, with improvements in the pre-processing step, but limited to a string-based matching approach. With this new configuration, for the Anatomy track we observe that globally the quality of results decreases (F-measure looses 0.3 points). However, we observe that Holontology is 8 times faster that LPHOM. For the Conference track, we observed an improvement in the results with respect to LPHOM. The tool however needs additional efforts to handle data and object properties. For the Knowledge graph track, Holontology proceeded faster than the other systems (including the baseline). However, it has not be able to deal with properties.

4.3 Conclusions

This chapter has presented an approach for holistic matching that extends a proposal designed to match XML schemes. The proposed approach is modeled within a linear program by reducing the ontology matching problem to the maximum-weighted graph matching problem, which is solvable in polynomial time. The approach is extensible with different linear constraints handling classes and properties of ontologies. These constraints are used to reduce the logical incoherence in generated alignments, what is not systematically taken into account by matching systems. The approach presented here could be extended in several directions (besides a deeply study of the impact of each constraint and their combination), as considering additional hypothesis concerning incoherence, and dealing with instances (holistic instance matching and exploitation of the instance matching results in holistic ontology matching and vice-verse).

In the next chapter, back to a pairwise setting, the problem of matching ontologies with different levels of abstraction, such as domain and foundational ontologies is discussed.

Chapter 5

Foundational and domain ontology matching

5.1 Motivation

Ontologies can be classified according to their "level of generality" [Guarino, 1998] in *foundational ontologies* and *domain ontologies.* Foundational ontologies describe general concepts (e.g., physical object, event) and relations (e.g., parthood, participation), which are independent of a particular domain. These ontologies, also named *upper* or *top-level*, are usually equipped with a rich axiomatic layer. They play an important role in the construction and integration of domain ontologies, providing a well-founded reference model that can be shared across domains. While the clarity in semantics and a rich formalization of foundational ontologies are important requirements for ontology development [Mika et al., 2004, Keet, 2011] improving ontology quality, they may also act as semantic bridges supporting interoperability between domain ontologies [Mascardi et al., 2010, Keet, 2011, Nardi et al., 2013]. Recently, the lack of ontological distinctions and the sparse axiomatisation in Linked Open Data knowledge bases have been addressed in [Asprino et al., 2018]. As stated by the authors, *distinctions such as whether an entity is inherently a class or an individual, or whether it is a physical object or not, are hardly expressed in the data, although they have been largely studied and formalised by foundational ontologies.* Such distinctions are however key aspects in many applications in Artificial Intelligence.

There are two approaches for the use of foundational ontologies [Semy et al., 2004]. With a *top-down approach*, the foundational ontology is used as a reference for deriving domain concepts, taking advantage of the knowledge and experience already encoded in it. In a *bottom-up approach*, one usually matches an existing domain ontology to the foundational ontology. The latter is more challenging since inconsistencies may exist between domain and foundational ontologies and one has to deal with different levels of abstraction in the matching process. As reported in [Khan and Keet, 2012, Keet, 2011], methodologies for constructing ontologies should not neglect the use of foundational ontologies and should better address it in a *top-down* approach. In the absence of systematic adoption of foundational ontologies within the domain ontology development process, *bottom-up* approaches have to be applied instead. In this task, matching foundational and domain ontologies plays a key role.

Matching ontologies from different levels of abstraction, as domain and foundational ontologies, is still an early tackled challenge in the ontology matching field. This is a complex task, even manually, that requires the deep identification of the semantic context of concepts and, in particular, the identification of subsumption relations. The latter is largely neglected by most state-of-the-art matchers. For instance, the concept `Document` from the *ekaw* ontology (Figure 2.2), a super-concept of the concept `Paper`, can be seen as a sub-concept of `FactualText` from SUMO foundational ontology (*Suggested Upper Merged Ontology*) [Niles and Pease, 2001, Pease, 2011]. The main problem of matching foundational and domain ontologies using these matching systems is that, despite the variety of approaches, most of them typically rely on string-based techniques as an initial estimate of the likelihood that two elements refer to the same real world phenomenon, hence the found correspondences represent equivalences with concepts that are equally or similarly written. However, in many cases, this correspondence is wrong [Schmidt et al., 2016]. In fact, when having different levels of abstraction it might be the case that the matching process is rather capable of identify subsumption correspondences than equivalence, since the foundational ontology has concepts at a higher level. Relatively few matching systems are able to discover other relations than equivalence. The examples are AML, BLOOM, S-Match, TaxoMap and Aroma, many depending on background knowledge as WordNet), with few other propositions in the literature [Vennesland, 2017, Zong et al., 2015].

Approaches dealing with the task of matching foundational and domain ontologies are mostly based on manual matching [Brodaric and Probst, 2008, Mika et al., 2004]. In [Brodaric and Probst, 2008], Geoscience ontologies have been manually aligned to DOLCE (*Descriptive Ontology for Linguistic and Cognitive Engineering*) [Gangemi et al., 2002][1] and incompatibilities issues have been discussed. In [Mika et al., 2004], DOLCE has also been manually aligned to a domain ontology describing services, in order to address its conceptual ambiguity, poor axiomatization, loose design and narrow scope. In [Damova et al., 2010], several schemata of FactForge (which enables SPARQL querying over the Linked Open Data cloud) have been manually aligned to PROTON (PROTo ONtology) [Terziev et al., 2005][2] in order to provide a unified way to access the data. Manually alignments have also been established between biomedical ontologies and BFO (*Basic Formal Ontology*)[3] [Grenon et al., 2004, Arp et al., 2015] in [Silva et al., 2011]. More recently, in [Jezek, 2019], the alignment between the T-PAS resource (Typed Predicate-Argument Structures [Jezek et al., 2014]) and DOLCE categories has been manually established, highlighting the distinctions and similarities between the two resources from a cognitive and application-based perspective. One of the few automatic approaches is BLOOMS+ [Jain et al., 2011], which has been used to automatically align PROTON to LOD datasets using as gold standard the alignments provided

[1] `http://www.loa.istc.cnr.it/old/DOLCE.html`
[2] `http://ontotext.com/proton`
[3] `https://github.com/bfo-ontology/BFO/wiki`

in [Damova et al., 2010]. BLOOMS+ first uses Wikipedia to construct a set of category hierarchy trees for each class in the source and target ontologies. It then determines which classes to align using 1) similarity between classes based on their category hierarchy trees; and 2) contextual similarity between these classes to support (or reject) an alignment. More recently, in [Asprino et al., 2018], automatic classification of foundational distinctions (class vs. instance or physical vs. non-physical objects) of LOD entities is done with two strategies: an (unsupervised) alignment approach and a (supervised) machine learning approach. The alignment approach, in particular, relies on the linking structure of alignments between DBpedia, DOLCE, and lexical linked data, using resources such as BabelNet [Navigli and Ponzetto, 2012], YAGO [Rebele et al., 2016] and OntoWordNet [Gangemi et al., 2003b]. For instance, they use the paths of alignments and taxonomical relations in these resources and automated inferences to classifying whether a DBpedia entity is a physical object or not.

Complementary, while the purpose of a foundational ontology is to solve interoperability issues among ontologies, the development of different foundational ontologies re-introduces the ontology interoperability problem, as stated in [Khan and Keet, 2013a]. Early works addressed this problem [Grenon, 2003, Seyed, 2009, Temal et al., 2010] on different perspectives. While fundamental issues and primitive relations between BFO and DOLCE have been studied in [Grenon, 2003] and [Seyed, 2009], respectively, [Temal et al., 2010] established an alignment between these ontologies in order to conciliate their respective realistic and cognitive points of view. In [Muñoz and Grüninger, 2016], the core characterization of mereotopology of SUMO and DOLCE has been studied, relating their axiomatizations via ontology alignments, while in [Oberle et al., 2007] alignments between DOLCE and SUMO have been established for supporting domain ontology integration. In [Khan and Keet, 2013a, Khan and Keet, 2013b], alignments between BFO, DOLCE and GFO were built both with automatic matching tools (H-Match, PROMPT, LogMap, YAM++, HotMatch, Hertuda and Optima) and manually, with substantially fewer alignments found by the matching tools. During the process, it was found that differences in foundational ontologies, such as their hierarchical structure, conflicting axioms due to complement and disjointness, and incompatible domain and range restriction, cause logical inconsistencies in foundational ontology alignments, thereby greatly reducing the number of correspondences. While the accuracy and percentage of alignments that were found vary greatly among the tools, exploiting the aligned entities whilst keeping a consistent ontology reduces the feasible set of alignments. The resulting alignments have been made available on the ROMULUS platform [Khan and Keet, 2013b][4]. Aligning foundational ontologies reveals also the problem of matching their different versions (as for domain ontologies). In [Seppälä et al., 2014], a method for tracking, explaining and measuring changes between successive versions of BFO 1.0, BFO 1.1, and BFO 2.0 was applied. Automation in these tasks have been however very limited.

[4] http://www.thezfiles.co.za/ROMULUS/

5.2 Contributions

We have proposed an automatic approach for matching domain and foundational ontologies that exploits existing alignments between WordNet [Miller, 1995] and foundational ontologies, as an intermediate layer [Schmidt et al., 2017]. It reduces the problem to the matching of domain concepts and WordNet synsets. For that, as for classical approaches on word sense disambiguation, the notion of context is adopted. Contexts are constructed from all information about an ontology entity (e.g., entity naming, annotations and information on the neighbors of entities) and are used for disambiguating the senses that better express the meaning of ontology entities in WordNet. After selecting an appropriated synset for a given domain ontology, a relation between that synset and a foundational concept is identified, via existing alignments between WordNet and the foundational ontology. While in [Schmidt et al., 2017], an adaptation of the Lesk measure [Lesk, 1986] has been used for word sense disambiguation, this work has been further extended in [Schmidt et al., 2018] using word embeddings [Mikolov et al., 2013]. Most strategies we apply here, in particular indirect matching [Kachroudi et al., 2015, Jung et al., 2009, Zhang and Bodenreider, 2005], WordNet-based matching [Lin and Sandkuhl, 2008, Yatskevich and Giunchiglia, 2004], the classical notion of context [Wang, 2011, Schadd and Roos, 2012, David, 2011] and word-sense disambiguation [Navigli, 2009], have been already exploited in different ways in the field. However, we argue that their combination remains under-explored in the specific task of matching top-level and domain ontologies. The use of word embedding for the matching task is, however, has been recently studied [Zhang et al., 2014, Vieira and Revoredo, 2017, Kolyvakis et al., 2018]. With respect to the few automatic approaches dealing with the task, as [Asprino et al., 2018] we exploit existing alignments between lexical resources and foundation ontologies (even if limited to WordNet), however, we are not limited to specific kinds of foundational distinctions. As in [Jain et al., 2011] we naturally adopt the notion of context, but still do not exploit Wikipedia hierarchies.

With respect to the matching of foundational ontologies, in [Schmidt et al., 2019b] we have analyzed the behaviour of matching systems in this task. Our work extended the one from [Khan and Keet, 2013a, Khan and Keet, 2013b] in two ways: it considers more recent matching systems, those participating in the Ontology Alignment Evaluation Initiative (OAEI) 2018, and it considers a new pair of aligned foundational ontologies SUMO and DOLCE [Oberle et al., 2007], which consists essentially of subsumption relations. The alignments in [Khan and Keet, 2013a] and [Oberle et al., 2007] served as a reference alignment in order to evaluate the matchers. The findings in [Schmidt et al., 2019b] are in line to what has been reported when evaluating the behaviour of matchers in the task of matching domain and foundational ontologies. Current tools fail not only dealing with the different levels of abstraction between foundational and domain ontologies but fail in dealing with the generality level of foundational ontologies. From these findings, we have proposed an approach for matching foundational ontologies involving subsumption relations [Kamel et al., 2019]. We argue here that the knowledge encoded in the ontologies has to be further exploited. In that way, we propose to borrow approaches from relation extraction from text in NLP in order to establish subsumption relations between the ontologies to be matched. While the approach is not completely new, as NLP techniques are often used to extract knowledge from text, their exploitation in ontology matching brings some novelty. Relation extraction in ontology matching has been considered in few works (in particular hypernym relation extraction between terms, which in most cases can be transferred to subsumption between concepts). In [Spiliopoulos et al., 2010], a supervised method learns patterns of subsumption evidences, while in [Beisswanger, 2010] the approach relies on free-text parts of Wikipedia in order to help detecting hypernym, even without clear evidence in the input ontologies themselves. Hearst patterns [Hearst, 1992] have been adopted in [van Hage et al., 2005] and [Vazquez and Swoboda, 2007], with the former using them to eliminate noise in matching results. In this first version of our approach, we exploit lexico-syntactic patterns from Hearst and evidences of hypernym relation carried out in definitions layout.

The main contributions of the work on matching foundational and domain ontologies are:

- we evaluated how a set of available matching tools, applying different matching strategies, performs in this task. Even though they were not exactly developed for that purpose, their output might help us to investigate the problem [Schmidt et al., 2016].

- an approach for automatically matching foundational and domain ontologies using existing WordNet to foundational ontology alignments (DOLCE and SUMO) as intermediary layer [Schmidt et al., 2017, Schmidt et al., 2018].

- we quantitatively analyzed the alignments provided by current (OAEI 2018) matching systems in the task of matching foundational ontologies [Schmidt et al., 2019b].

- an approach to match foundational ontologies (and that could be applied to domain matching) relying on lexico-syntactic patterns and evidences of hypernym relation carried out in definitions layout [Kamel et al., 2019].

The work on matching domain and foundational ontologies has been carried out in collaboration with Daniela Schmidt (PhD student at Pontificia Catolica Universidade do Rio Grande do Sul, Brazil), co-advised with Renata Vieira (Full Professor at Faculty of Informatics at Pontificia Catolica Universidade do Rio Grande do Sul, Brazil). Mouna Kamel (Assistant professor at University of Perpignan and researcher at IRIT) has also collaborated with us in the task of matching foundational ontologies.

The proposed matching approaches are briefly introduced in the following.

5.2.1 Domain and foundational ontology matching

Our approach for matching foundational and domain ontologies has two main steps. The first step constructs the context of each domain concept and selects the most appropriated WordNet synset (disambiguation). The second steps matches the domain concept to the foundational concept via existing correspondences between

WordNet and the foundational ontologies, as detailed below. Our approach focuses on DOLCE and SUMO foundational ontologies.

Synset disambiguation In order to select the synset that better expresses the meaning behind the domain ontology concept, a *context* is constructed from all information available about an ontology entity, including entity naming (ID), annotation properties (usually labels and comments) and information on the neighbours (super and sub-concepts). Given $Sup(e)$ and $Sub(e)$, the sets of terms denoting the super-concepts and sub-concepts of the entity e, and $Ann(e)$ the set of terms from its annotations, a naive strategy considers these sets as a *bag of words*:

$$context(e) = \{e, w|w \in Sup(e) \cup w \in Sub(e) \cup w \in Ann(e)\}$$

This context is used to find the closer synset using two strategies, as above.

Lesk measure The Lesk measure for word sense disambiguation [Lesk, 1986] relies on the calculation of the word overlap between the sense definitions of two or more target words. Given a word w, it identifies the sense of w whose textual definition has the highest overlap with the words in the context of w. Here, we overlap the $context(e)$ with the context of each WordNet synset:

$$context(synset) = \{w|w \in Terms(synset) \cap w \in Gloss(synset)\}$$

where $Terms(synset)$ the set of terms in a *synset* and $Gloss(synset)$ the corresponding set of terms from the gloss (i.e, textual description containing definitions and examples) associated to the synset. We hence retrieve the highest overlap between $context(e)$ and $context(synset)$

$$score'_{Lesk}(e) = |context(e) \cap context(synset)|$$

Word embeddings The second similarity measure compares contexts of entities $context(e)$ and of WordNet synsets $context(synset)$ (represented as vectors of words). The comparison is based on the distance of contexts in vector spaces. This method adopts the cosine distance between two words generated by the word embedding model to identify the similarity between them. We retrieve the similarity between $context(e)$ and $context(synset)$, then we calculate the average similarity. After calculating this average to all elements of the context, we calculate the average of the context, considering the context length. The synset with the higher average is selected.

Identification of correspondences to foundational ontologies In this step we perform the identification of the foundational concept. This step relies on the representation of the given existing alignments.

DOLCE correspondence identification This step uses existing alignments between DOLCE-Lite-Plus and WordNet 1.6 from OntoWordNet [Gangemi et al., 2003b].

The authors assume that the hyponymy relation could be aligned to the subsumption relation and the synset notion could be aligned to the notion of concept. Figure 5.1 presents a fragment of WordNet synsets (as concepts) linked to DOLCE-Lite-Plus concepts. The first-level concepts (in lower case) correspond to a DOLCE-Lite-Plus concept. The upper case concepts represent WordNet synsets. Each concept in OntoWordNet is associated to an annotation containing the corresponding gloss of the synset in WordNet.

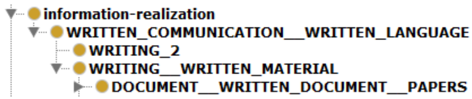

Figure 5.1: Example of WordNet synsets linked to DOLCE.

For each concept of the domain ontology, we use the selected synset (step 1) to identify the corresponding concept in OntoWordNet. To select the concept in OntoWordNet we compare the WordNet synset to each concept c in OntoWordNet. After finding the OntoWordNet concept c corresponding to the synset, the higher level concept h^c of c is retrieved, h^c corresponds to the DOLCE concept.

SUMO correspondence identification Similarly to the correspondence identification in DOLCE, this step uses existing alignments between SUMO and WordNet 1.6 (a more recent release considers WordNet 3.0), in order to identify the domain and foundational concepts correspondences. As SUMO-WordNet alignment is a file containing the synset ID, terms, gloss, and the alignment to foundational concept (as example below), we search for the domain selected synset in this file and, if the synset is found, we match the domain concept with the foundational concept related to the synset. An example of the structure of a correspondence representing a synonymy relation can be seen below, for one of the synsets associated to the term "document" in WordNet. In the example, "06470073 ... 0 papers" corresponds to the synset, followed by the gloss "writing .. nature)", with "FactualText" corresponding to the SUMO concept and the signal "+" is the suffix indicating the hyponymy relation between the synset and the SUMO concept.

```
06470073 10 n 03 document 0 written_document 0 papers -- writing that
provides information (especially information of an official nature)
&%FactualText+
```

Evaluation The matching approach has been evaluated to match domain ontologies to DOLCE and SUMO. This choice is motivated by the fact that they are the most used top-level ontologies and serve as a reference model for the modelling and integration of ontologies. We align them to ontologies from three domains (SSN[5], CORA [Prestes et al., 2013], and OAEI Conference ontologies[6]). These ontologies sum up 501 concepts, however, we consider in our experiments the first-level concept

[5] https://www.w3.org/TR/vocab-ssn/
[6] http://oaei.ontologymatching.org/2017/conference/index.html

of the hierarchies, what corresponds to 70 concepts (assuming that the other concepts will inherit their alignment with top ontologies from their roots). All generated correspondences are available in `https://github.com/danielasch/top-match`.

Word embedding models We used pre-trained models, GloVe [Pennington et al., 2014] and GoogleNews[7]. GloVe is an *unsupervised learning algorithm to obtain vector representations for words*[8]. The training phase uses the Wikipedia 2014 and Gigaword5 corpora. It has 6 billions tokens, 400 thousand vocabulary size and neural network dimension of 200. The GoogleNews model is trained on part of Google News dataset (about 100 billion words). The model contains 300-dimensional vectors for 3 million words and phrases.

Reference alignments For the pairs involving SSN and CORA, given that these ontologies are already aligned to the DOLCE and SUMO, respectively, we adopt these existing alignments as reference. We note that SSN is originally aligned with a different version of DOLCE. We hence consider the results in an interpreted way which consists at looking each generated correspondence and identify if they are the exact correspondence or related to the previous alignment via a subsumption relation. In the same way, we observe that some found correspondences from CORA and SUMO, were not exact the same of the adopted reference, however, they are hierarchically related, hence, we also adopted the interpreted evaluation. For the Conference dataset, which is not equipped with reference alignments to DOLCE and SUMO, the generated correspondences were manually evaluated by three specialists. Firstly, one evaluator analysed each correspondence, after, the results were discussed with all evaluators, maintaining or changing the initial analysis. For this dataset, we made the hypothesis that, for each high-level domain concept, a corresponding WordNet synset exists. Hence, we were able to compute both precision and recall. This alignment has been further extended with new annotators, as described in Chapter 8.

Baseline Our baseline corresponds to the results of a set of matching tools participating in OAEI 2017, with exception only of those specialised in instance matching (Legato, I-match and njuLink) and one specialised in the bio domain (Yam-bio). The matchers that were tested in our experiment are: ALIN, AML, CroLOM, KEPLER, LogMap, LogMap-Lite, ONTMAT, POMap, SANOM, WikiV3, WikiMatch and XMap. Even though they were not exactly developed for that purpose, their results were the only available for comparison, and we set that as a baseline.

Results We run our system with the Lesk similarity (*lesk*) and word embedding models (*WE-GloVe* and *WE-GloogleNews*) and the OAEI tools for 16 matching tasks (SSN and DOLCE, CORA and SUMO, and 7 Conference ontologies with DOLCE and SUMO). They have been evaluated in terms of precision and recall. The best results were obtained for the conference domain, with .80 of F-measure

[7] https://code.google.com/archive/p/word2vec/
[8] https://nlp.stanford.edu/projects/glove/

with WE-GlloogleNews. We observe that overall WE-GlloogleNews performs better than Lesk and WE-GloVe. However, looking at the SSN and CORA domain ontologies, the obtained results are lower than for Conference. Our hypothesis is that concepts from the conference ontology are more general (common sense) than these other domains. Note that the selected word embedding models were trained with general domain texts. The better performance obtained with the WE-GoogleNews model over the WE-GloVe model could be explained by the larger coverage of the first with respect to the training set.

Regarding the number of correspondences, our approach was able to find 69 out of 70 correspondences from the Conference ontologies (we were not able to find the correspondences for 1 concept, for which there is no entry in WordNet). Considering Lesk and WE-GloVe, 51 correct correspondences were found when aligned to SUMO and 49 correct with DOLCE. This number increased up when using WE-GlloogleNews (57 and 56, respectively). For SSN-DOLCE, we have 5 correct correspondences out of 8 considering Lesk, and 3 correct with WE-GloVe and WE-GlloogleNews. For CORA-SUMO, 12 correct in a total of 29 correspondences considering Lesk, 11 correct with WE-GloVe and 6 correct with WE-GlloogleNews. Although our approach was able to found a high number of correspondences for the three domains, in some cases, the generated correspondences were wrong. First, as we adopt the context of concepts, this seems not to be enough to disambiguate the sense of the domain concept (Conference domain ontologies are not equipped of comments and labels). This can be improved by enriching the terminological layer. Second, we can observe that word embedding based on Google News model contributes to the disambiguation step, mainly with the Conference ontologies. However, for SSN and CORA it is still not able to retrieve the right synsets. In order to overcome this weaknesses, one direction is to use domain-specific embedding models. Third, word sense disambiguation here is still based on the overlapping of words, and more sophisticated word sense disambiguation techniques have be used instead.

With to respect to the matching systems, only 4 tools (AML, LogMap, LogMapLite, and POMap) were able to find correspondences for 6 pairs of ontologies. Considering the correspondences found by these tools, 13 domain concepts from conference (out of 70) were aligned. Regarding the number of correspondences, AML was able to find 12 correspondences, and 7 of them were correct. POMap found 7 correspondences, and 6 were correct. LogMap and LogMapLite found 6 correspondences respectively, and 5 of them were correct. Related to CORA, 1 correspondence was correctly found by POMap. As shown, our approach outperforms all system in terms of Recall and F-measure. Looking at WE-GlloogleNews, the results are quite similar in terms of precision and better than all in terms of recall and F-measure. As somehow expected, while the tools perform well in terms of precision, they retrieve a limited number of correspondences.

5.2.2 Foundational ontology matching

Our approach for matching foundational ontologies relies on two main steps: (i) hypernym extraction from ontology annotations and (ii) subsumption generation between ontology concepts:

Hypernym extraction The hypernym relation extraction takes as input the ontology annotations as concept definitions (what are common in foundational ontologies). A *definition* attaches a meaning to a term denoting the concept. The term that is to be defined is called the *definiendum*, and the term or action that defines it is called the *definiens*. In the example below, the *definiendum* = “Product” and the *definiens*=“An Artifact that is produced by Manufacture and that is intended to be sold”. Many linguistic studies show that definitions mostly express one of the main lexical relations e.g., hypernymy, meronymy or synonymy, between *definiens* and *definiendum* [Malaisé et al., 2004, Navigli et al., 2010].

```
<owl:Class rdf:ID= "Product">
  <rdfs:comment> An Artifact that is produced by Manufacture and
  that is intended to be sold.</rdfs:comment>
</owl:Class>
```

Different strategies are exploited for extracting the hypernym relations:

Hypernym relations expressed using definitions layout We focus on cases where the *definiens* starts by expressing an entity (denoted by a term and different from the *definiendum*) which have some properties. In the above example, the entity in the *definiens* is “Artifact” and the property is “that is produced by Manufacture and that is intended to be sold”. Thus the *definiendum* (Product) is an *hyponym* of the definiens (Artifact). When no property is expressed, it is usually a synonym relation, as below:

```
<owl:Class rdf:about="#Quale">
  <rdfs:comment> An atomic region. </rdfs:comment>
</owl:Class>
```

Hypernym relations lexically expressed in text annotations OWL class definitions may also be more fine grained exploited, as comment paragraphs may contain well-written text. We then exploit this text using a set of lexico-syntactic patterns from Hearst:

[`NP such as {NP ,}* {or|and} NP`], [`NP like {NP ,}* {or|and} NP`], [`NP which is an example of NP`], [`NP including {NP ,}* {or|and} NP`], [`NP is called NP if`], [`NP is an NP that`].

For instance, the pattern [`NP like {NP ,}* {or|and} NP`] means that a noun phrase (`NP`) must be followed by the word “like”, which must be followed by an NP or by a list of NPs separated by comma, having before the last NP “or” or “and”. When applied on the definition below, the hypernym relations (Self Connected Object, planet), (Self Connected Object, star) and (Self Connected Object, asteroid) can be identified.

```
<owl:Class rdf:about="#AstronomicalBody">
```

```
  <rdfs:comment> The Class of all astronomical objects of
     significant size. It includes Self Connected Objects
     like planets, stars, and asteroids ...
 </rdfs:comment>
</owl:Class>
```

Hypernym relations carried out by the concept identifier Hypernym relations may also be identified from modifiers of a head of a compound noun denoting the identifier of the OWL class. In the example above, the hypernym relation (astronomical body, body) can be identified thanks to this strategy.

Subsumption generation Having extracted all the hypernym relations from both ontologies to be matched, we verify if the terms appearing as hyponyms and hypernyms denote concepts in the ontologies. In the example above, as the alignment is directional, "Product" denotes a concept in the source ontology and "Artifact" in the target ontology, hence this hypernym pair is kept.

Evaluation We used the foundational ontologies DOLCE and SUMO. The reference alignment involving 41 subsumption correspondences comes from [Oberle et al., 2007]. The approach has been implemented with GATE: to extract concepts and their associated comments from the ontology OWL file and restructuring them according to an XML format; to identify terms using first the TermoStat term extractor, and then expanding the recognition of terms using JAPE rules (for instance, the sequence made of a TermoStat term preceded or followed by adjectives, constitutes a new term); to annotate the XML corpus with different NLP tools (ANNIE Tokenizer, Stanford POS, Stanford parser, Gazeteer of identified terms); and to identify hypernym relations.

We evaluate each approach individually (layout, patterns, head modifier) and as somehow expected, their combination brings better results compared to the results of each individual approach. Patterns are very precise while head modifier provides good results in terms of recall with respect to the other strategies. Comparing the approach to the OAEI 2018 matchers, besides the fact that we do not distinguish subsumption and equivalence relations when computing precision and recall, no matcher were able to find the correspondences. From the 41 reference correspondences, only one correspondence refers to similar terms (`dolce:geographical-object` and `sumo:GeographicArea`) and 5 of them could be found via a head modifier method (e.g., `dolce:organization` and `sumo:PoliticalOrganization`). In order to see how close the generated alignments were to the reference, we have calculated the relaxed precision and recall (Section 6.1), that measure the closeness of the results to the reference. While the results of our approach are not that close to the reference, in terms of recall we obtain results similar than the relaxed recall for all matchers. This first approach has to be combined with current matching strategies to better deal with the task.

5.3 Conclusions

Linking domain and foundational ontologies allows for improving the clarity in semantics of domain ontologies. These links can also be further exploited as bridges in the task of matching domain ontologies. If ontologies are not constructed on the basis of foundational ontologies, they have to be aligned to them afterwards, what is unfortunately mostly the case in ontology construction. The problem of matching domain and foundational ontologies is a challenging task specially due to the different levels of abstraction of these ontologies. This chapter has presented an approach to match domain and foundational ontologies exploiting alignments between WordNet and foundational ontologies. We have also presented an approach exploiting symbolic hypernym relation extraction approaches for generating subsumption correspondences between foundational ontologies. This approach can as well be applied to domain ontologies.

Currently, we are working on an extension of the presented approach that exploits the hierarchy of external resources, as BabelNet and YAGO (as they have also alignments to SUMO), in order to avoiding using existing WordNet alignments as bridges, as WordNet may suffer of low domain coverage for some domains (as corroborated in our experiments). Besides that, the generation of expressive correspondences could be also applied to the task of matching domain and foundational ontologies (Chapter 5), as it has demonstrated to be a scenario where the relationships between entities from different ontologies require rather full fledged axioms, as pointed out in [Reed and Lenat, 2002, Damova et al., 2010]. As the example from [Damova et al., 2010], *professions* are modeled as instances of the class *Profession* in PROTON, and the single entity of DBPedia is matched to an expression in PROTON which restricts the property *hasProfession* to the value of the profession of interest: *dbp:Architect,∃pupp:hasProfession.p-ext:Architect,*$\sqsupseteq$). Complementary, as Nicola Guarino pointed out in our presentation in ONTOBRAS 2019, this task has to be rather done via axiom matching [Fürst and Trichet, 2005], which can be further revised to the generation of complex correspondences.

Matching domain and foundational ontologies is one of the scenarios where dedicated benchmarks (not yet fully developed) or dedicated evaluation strategies are still to be developed. In the next chapter, the evaluation of complex alignments is discussed, before discussing the construction of domain and foundational ontologies reference alignments in the subsequent chapter.

Chapter 6

Complex matching evaluation

6.1 Motivation

Automatic support for evaluating approaches able to generate complex correspondences has still not been addressed in the literature. This requires having benchmarks that contain complex correspondences and appropriate metrics with which to evaluate the quality of the alignments on such benchmarks.

With respect to benchmarks, although a large spectrum of matching cases has been proposed in OAEI, e.g., involving synthetically generated or real world datasets with large and domain-specific ontologies, these datasets are limited to alignments with simple correspondences. Recently, we have contributed to the first OAEI complex track [Thiéblin et al., 2018a], which opens new perspectives for the evaluation in the field. This track contained four datasets about different domains: Conference, Hydrography, GeoLink and Taxon. In particular, the complex Conference dataset results from a consensus between three raters manually generating the complex correspondences, with a special focus on the task of query rewriting. This consensual dataset extends the dataset we have presented in [Thiéblin et al., 2018b], where two (non-consensual) alignment sets for two task purposes (ontology merging and query rewriting) were proposed. The closer approach to ours is from [Jiang et al., 2016] who also extended the conference dataset with complex alignments to evaluate their knowledge-rule based approach. However, the methodology used for the construction of the dataset is not specified and the dataset is not publicly available.

With respect to the evaluation metrics, evaluation of most existing approaches has been done by manually calculating the precision of the alignments generated by the systems [Ritze et al., 2009, Ritze et al., 2010, Parundekar et al., 2012, Walshe et al., 2016]. In order to be able to measure recall, specific tailored datasets have been constructed. The approach of [Parundekar et al., 2012] estimated their recall based on the recurring pattern between DBpedia and Geonames, while in [Qin et al., 2007] a set of reference correspondences between two ontologies was manually created, involving nine reference correspondences from which only two cannot be expressed with simple correspondences. In [Walshe et al., 2016] the authors proposed an algorithm to create an evaluation data set that is composed of a synthetic ontology containing the specific *Class-by-attribute-value* correspondence patterns.

As stated in Section , ontology alignment evaluation is often performed by comparing a generated alignment to a reference one and computing precision and recall based on this. Most of the OAEI tracks use this kind of evaluation. While these types of evaluation are developed and automatised for simple alignments, there exist difficulties inherent to complex alignment evaluation. The main difficulty in complex evaluation resides in the comparison of the alignments with respect to a reference (usually, a reference alignment). For example, in the case of the aforementioned conference ontologies, these three correspondences can be considered as true positive: (o_1:AcceptedPaper,$\exists o_2$:hasDecision.o_2:Acceptance,$\equiv$), ($\exists o_2$:accepted.{true},$\exists o_3$:hasDecision.o_3:Acceptance, $\equiv$), or (o_1:AcceptedPaper,$\exists o_2$: acceptedBy.$\top$,$\equiv$). However, comparing them to a reference alignment (what highly depends on the choices of the reference) requires more sophisticated ways than for simple alignment evaluation.

As introduced in Section for simple alignments, from a syntactic perspective, the implementation of more sophisticated and promising approaches to deal with complex ones is a challenge. Rule-based and particularly edit-distance metrics have, for instance, to cope with a potentially high search space of possible combinations of constructions and transformations. This means that such evaluation approaches would have to adopt non-naive techniques to reduce the search space, and contemplate only the more plausible combinations of constructions in order to ensure efficiency. From a semantic perspective, transformations cannot be expressed in OWL at all. This means that semantic approaches relying on existing OWL reasoners would only be able to evaluate correspondences with constructions supported by those reasoners, which would limit their applicability. In contrast, instance-based approaches are unaffected by the complexity of the correspondences, and could be the most realistic way to address the complex alignment evaluation problem, by shifting from the comparison of correspondences into the comparison of sets of instances.

6.2 Contributions

We have addressed the lack of benchmarks for evaluating complex correspondences, by extending the OAEI Conference dataset with complex correspondences, together with the creation of complex alignments involving different taxon classifications from LOD datasets. This resulted in the the first OAEI complex track. With respect to the evaluation metrics, in order to overcome the limitations of syntactic and semantic strategies for comparing members of complex correspondences, we proposed to shift the problem to the comparison of set of instances in a task of query rewriting. Our approach relies on two main assumptions: a) complex alignments are highly relevant to the task of query rewriting and b) an alignment should be able to cover a knowledge need in terms of competence questions, reducing the search space. The alignment to be evaluated is used to rewrite a set of reference source queries whose results (set of instances) are compared to the ones returned by the corresponding target reference queries. While those metrics show the overall coverage of the alignment with respect to the knowledge needs and the best rewritten query, a balancing strategy consists in calculating the alignment precision based on common instances.

Summing up, the work on complex ontology evaluation has the following contributions:

- we propose two (non-consensual) datasets of complex alignments between 10 pairs of ontologies from the OAEI Conference simple alignment dataset [Thiéblin et al., 2018b]. The methodology for creating the alignment sets is described and takes into account the use of the alignments for two tasks: ontology merging and query rewriting. We extended the work presented in [Thiéblin et al., 2017c] and in [Thiéblin et al., 2017b] by enriching the alignment sets with new pairs of ontologies and by considering the task for which the alignment is needed. Building benchmark suites is highly valuable not just for the group of people that participates in the contests, but for all the research community.
- we propose in [Thiéblin et al., 2019b] a consensual complex dataset for the Conference dataset that results from the adoption of an adapted methodology from [Thiéblin et al., 2018b], by three domain experts, with the same level of expertise on the domain of conference organization.
- we provide an evaluation of state-of-the-art matching systems on the consensual dataset, extending the evaluation that has been reported in the first OAEI complex track [Thiéblin et al., 2018a, Algergawy et al., 2018] by including additional matchers able to generate complex correspondences.
- we propose a populated version of the OAEI Conference dataset [Thiéblin and Trojahn, 2019, Thiéblin et al., 2019d]. A subset of Conference ontologies have been populated, both synthetically and with real data. It has been based on the notion of *competence questions for alignment* (CQA). The use of CQAs ensures that the populating is homogeneous across ontologies. Thanks to this dataset, we could apply the proposed metrics for evaluating complex alignments.
- we survey and analyze the requirements for effective evaluation of complex ontology alignments and evaluate the degree to which these requirements are met by existing approaches [Zhou et al., 2019]. We also provide a roadmap for future work on this topic taking into consideration emerging community initiatives and major challenges that need to be addressed.
- we propose an automatic approach for evaluating complex alignments, shifting the problem to the comparison of instances in a task of query rewriting targeting user needs [Thiéblin et al., 2019d]. The proposed evaluation strategy consists of two measures, both relying on the comparison of instances. While the *CQA coverage* measure relies on pairs of equivalent SPARQL queries and measures how well an alignment covers these queries, the *intrinsic precision* compares the instances of the correspondences members. Intrinsic precision

balances the CQA coverage by like precision balances recall in classical matching evaluation.

The work on the generation of evaluation benchmarks has been carried out in collaboration with Élodie Thiéblin (PhD student at IRIT), co-advised with Ollivier Haemmerlé), Michelle Cheatham (Assistant professor, Wright State University, USA), Nathalie Hernandez (Assistant professor, UT2J/IRIT), and Ondrej Ŝváb Zamazal (Assistant professor, University of Economics, Czech Republic). For the evaluation metrics, I have also collaborated with Cátia Pesquita (Assistant professor, Faculdade de Ciências, Universidade de Lisboa, Portugal), Daniel Faria (Researcher, Instituto Gulbenkian de Ciência, Portugal) and Lu Zhou (PhD student, Kansas State University, USA).

In the following, the datasets with complex alignments are discussed (Section 6.2.1). Next, the proposed evaluation metrics are detailed (Section 6.2.2).

6.2.1 Complex evaluation datasets

This section presents the different datasets we have proposed for supporting the evaluation of complex alignments: the two task-oriented datasets (Section 6.2.1), the consensual Conference dataset (Section 6.2.1) and the populated version of the Conference dataset Section 6.2.1).

Task-oriented dataset

There is a clear lack of complex alignment datasets for systematic evaluation of complex alignments. We have addressed this lack by proposing two datasets taking into account the use of the alignments for two tasks: ontology merging and query rewriting. We argue that different tasks may have different correspondence expressiveness needs. While in ontology merging, for decidability reasons, the expressiveness of the correspondences has to be $\mathcal{SROIQ}$ (the decidable fragment of OWL [Horrocks et al., 2006]), for query rewriting, there is no expressiveness constraint. For example, the correspondence stating that o_2*:name* is equivalent to the concatenation of o_1*:firstname* and o_1*:lastname* is not applicable for ontology merging but is adequate for query rewriting.

The overall methodology followed for creating the datasets was:

1. Find simple equivalence correspondences between o_1 and o_2.

2. Create the complex correspondences based on simple correspondences so that the complex correspondences fit the purpose of the alignment.

3. Express the correspondences in a reusable format (e.g., EDOAL).

This overall methodology has been adapted to take into account the specificity of each task. For ontology merging, in particular, we added a step of coherence verification of the merge ontology using a reasoner.

The proposed datasets extend the OAEI Conference dataset. This choice is motivated by the fact that the ontologies are real ontologies (as opposed to synthetic

	Simple	Complex	TOTAL	Nb patterns
Ontology merging	259	54	313	9
Query rewriting	240	191	431	17

Figure 6.1: Number of correspondences per alignment set.

ones), they are expressive and largely used for evaluation in the field. This dataset has also been extended with different proposals [Cheatham and Hitzler, 2014, Meilicke et al., 2012]. Moreover, the reference alignments of simple correspondences between these ontologies are available. We chose five ontologies among the ones in the reference simple alignment for their different number of classes: *cmt*, *conference* (Sofsem), *confOf* (confTool), *edas* and *ekaw*. The reference simple alignment set was modified during the first step of the methodology.

The **ontology merging dataset** is composed of 313 correspondences with 54 complex correspondences from 9 different patterns (some patterns are composite). The **query rewriting dataset** is composed of 431 correspondences with 191 complex correspondences from 17 different patterns (some patterns are composite). The patterns are used *a posteriori* for analyzing the alignments, not as a basis for the correspondence creation. An extensive list of the patterns can be found in [Scharffe, 2009].

Figure 6.1 details the number of correspondences per alignment set. The ontology merging alignment set has no correspondences implementing domain or range restrictions, transformation functions, inverse properties, union of object or data properties, or negation. Indeed, these correspondences are either not in $\mathcal{SROIQ}$ (domain restriction, range restriction, union of properties) or were already entailed by previous correspondences (inverse property, negation). The number of subsumptions also differs in both alignment sets because of the adopted methodology (*top-down* subsumption in ontology merging and *bottom-up* subsumption in query rewriting). Nevertheless, the subsumption correspondences are frequent in both sets. As above, we argue that complex correspondences come as a complement to simple correspondences. Their need may be different depending on the task purpose of the alignment. For query rewriting for instance, complex correspondences represent 44% of all correspondences whereas they only represent 17% of all correspondences for ontology merging. This dataset is available online[1].

Consensual complex dataset

The process of manual construction of reference alignments is rarely documented. However, this is a hard and time-consuming task that ideally should require multiple raters and the ability to reconcile the differences in the interpretation of ontology entities and their relations, between (usually) ill-defined natural language definitions. As stated in [Tordai et al., 2011], the manual creation of alignments is by no means an easy task and the ontology alignment community should be careful in the construction and use of reference alignments. The complexity of the problem becomes

[1] `https://figshare.com/articles/Complex_alignment_dataset_on_conference_organisation/4986368/7`

worse when dealing with complex correspondences. While the datasets in [Thiéblin et al., 2018b] have been constructed by one experts (with a post-evaluation by two other experts), it represents one single interpretation of the problem. We extended the methodology from [Thiéblin et al., 2018b] for constructing complex alignments, with a focus on the query rewriting task and presented a consensual dataset that results from the adoption of the proposed methodology by three domain experts. While gathering annotators in the field is difficult, we argue that three annotators are reasonable for this task.

The overall methodology is articulated in the following steps:

1. Agree on the simple equivalence correspondences between o_1 and o_2 to rely on.

2. Individually create the complex correspondences based on the simple correspondences so that the complex correspondences fit the purpose of the alignment; and express the correspondences in First Order Logic (FOL)[2].

3. Collaboratively validate the set of found complex correspondences.

During the creation of the complex correspondences, some annotators did not exactly follow the methodology. The correspondences that they created were all annotated by the others annotators even if not compliant with the methodology. This was, in particular, related to the lack of direction in our current methodology regarding creation of (m:n) correspondences. This resulted in three alignment sets:

All Alignment containing all of the correspondences created by the annotators.

Methodo Alignment containing all of the correspondences created by the annotators and compliant with the methodology.

Logic methodo Alignment containing the correspondences with logic expressions as members created by the annotators and compliant with the methodology (all correspondences from Methodo except the value transformation function correspondences).

The observed agreements for the three datasets are shown in Table 6.1. Note that this agreement has been calculated over the consensus dataset. Overall, we observe a higher agreement, with a slight lower agreement for the *Methodo* and *methodo* involving the pairs cmt-ekaw.

Table 6.2 shows the differences between the methodology-compliant consensual alignment and the query-rewriting one from [Thiéblin et al., 2018b] (following the same methodology). One can notice that for some ontology pairs such as *cmt-ekaw*, few changes were made, whereas for others, such as *ekaw-conference*, we observe a higher number of changes. By comparing the alignments, for some cases, a change in the simple correspondences implies changes for the complex correspondences. This was the case for the *conference-ekaw* and *conference-cmt* correspondences, in which a simple equivalence correspondence (e.g., *cmt:Paper*

[2]This choice is motivated by the fact it is a common representation language which has a good balance of expressiveness and readability.

Table 6.1: Observed agreement between raters, for each pair of ontologies, for each version of the alignment (All, Methodo, Logic methodo).

	All (1)	Methodo (2)	Logic methodo (3)
cmt-conference	93%	87%	86%
conference-cmt	89%	91%	90%
cmt-ekaw	73%	67%	67%
ekaw-cmt	78%	100%	100%
conference-ekaw	79%	77%	77%
ekaw-conference	90%	91%	91%
Average	83%	85%	85%

Table 6.2: Differences between the methodology-compliant consensus alignment and the query-rewriting alignment from [Thiéblin et al., 2018b]. It shows the number of correspondences which are identical (=), have been added (+), deleted (-) or whose relation (r) was changed from the query-rewriting alignment to obtain the consensus alignment.

	complex				simple			
	=	+	-	*r* changed	=	+	-	*r* changed
cmt-conference	11		4	2	13	1	1	1
conference-cmt	6			4	18	2	3	
cmt-ekaw	8	1			11	2		1
ekaw-cmt	6	1	3		11	3	1	
conference-ekaw	10	2	6	5	20		5	2
ekaw-conference	12	10	5	1	21		3	

$\equiv$ *conference:Written_contribution*) was found in the consensus alignment to be a subsumption ($\sqsubseteq$), leading to complex correspondences with different relations from the query-rewriting alignment from [Thiéblin et al., 2018b]. Overall, the simple correspondences are more easily consensual than the complex correspondences. 79% of the simple correspondences from the consensus and the query-rewriting one [Thiéblin et al., 2018b] are identical whereas only 55% of the complex ones are.

From this work, we highlight that:

- ontology interpretations and their propagation has a strong impact in the generated correspondences;
- reaching the consensus needs a strong collaboration; keeping track of the usage of the alignments and of the evaluation metric to optimise helps in their construction;
- there is a strong relation between the kind of task and the expressiveness of the correspondences;
- existing alignment representation languages does not cover all possible constructions and transformations;

- last but not least, there is a lack of tools supporting the whole process; the process of manual alignment creation could also benefit of tools supporting the traceability (in that sense, the M-Gov framework described in [Singh et al., 2017] could help describing the metadata related to the users involved and theirs discussions during the generation of alignments).

Populated Conference dataset

As stated before, the Conference dataset has became one of the most used ones in matching evaluation [Zamazal and Svátek, 2017] and has been extended in different proposals [Cheatham and Hitzler, 2014, Meilicke et al., 2012]. This data set however is not equipped with instances, limiting the evaluation of matching approaches relying on them. While in [Solimando et al., 2014b], a partially populated version of the dataset has been used to evaluate alignments on the query rewriting task, the resulting dataset is limited to the scope of the queries used in the evaluation (only ontology concepts corresponding to 18 queries).

We have proposed a populated version of the data set [Thiéblin and Trojahn, 2019, Thiéblin et al., 2019d]. The ontologies have been populated both synthetically and with real data. It has been based on the notion of *competence questions for alignment* (CQA). The use of CQAs ensures that the populating is homogeneous across ontologies. Thanks to this data set, it will be possible to automatise the evaluation process of complex matchers using an evaluation strategy based on the comparison of instances in a query rewriting setting, as detailed in Section 6.2.2.

The methodology followed for populating the dataset has the main steps below:

1. Create a set of CQAs based on an application scenario in order to guide the ontology interpretation by the experts. Examples of CQAs include: "What are the accepted papers?" (unary CQA) or "Which are the authors of accepted papers?" (binary CQA).

2. Create a pivot format (e.g., JSON schema) for covering the CQAs from step 1 (e.g. covering attributes describing specific types of objects, such as papers or people):

```
{ "id": "10",
  "title": "User-Centric Ontology Population",
  "authors": ["K. Clarkson", ...],
  "type": "Research track",
  "decision":"accept" }
```

3. For each ontology of the data set, create SPARQL INSERT queries from the pivot format (here, an ontology may not cover the whole pivot format).

```
INSERT DATA {
      {{pap}} a :Camera_ready_contribution.
      {{pap}} rdfs:label {{paptitle}}.
      {{pap}} :is_submitted_at {{conf}}.
      {{pap}} :has_authors {{auth}}.
      ...  }
```

4. Instantiate the pivot format with real-life or synthetic data.

5. Populate the ontologies with the instantiated pivot format using the SPARQL INSERT queries.

6. Run a reasoner to verify the consistency of the populated ontologies. If an exception occurs, try to change the interpretation of the ontology and iterate over steps 3 to 5.

The methodology above has been followed to populate 5 ontologies from the Conference data: *cmt*, *conference* (Sofsem), *confOf* (confTool), *edas* and *ekaw* (Table 6.3). This choice is motivated by the fact that these ontologies have been also the ones used in the complex version of this dataset. A total of 152 CQAs have been created by an expert using as basis the ESWC 2018 conference scenario (whose data were fully open) and expanded by ontology exploration. The pivot format was first instantiated with data from the ESWC 2018 website and an automatic instantiating script of the pivot format was developed taking into account some statistics (e.g, proportion of members of the program committee author of articles, etc.). The dataset and instantiations of the pivot format have been made available[3].

In addition to the ESWC 2018 dataset, 6 other datasets (with 25 artificial conference) have been generated in order to cover the cases where ontologies share common instances. In these artificial datasets, each ontology has been populated with 5 pivot instantiation data. In the "dataset 0%" all ontologies were populated with 5 different pivot format instantiations; in the "dataset 20%", the ontologies were populated with 1 identical and 4 different instantiations; the other datasets (40%, 60%, 80%, and 100%) followed the same strategy. Since the size of each instantiation may differ, the percentage of common instances between two ontologies varies. For example, in the dataset 20%, the instances *Papers* common to the ontologies represent between 7% instances of *Papers* of *ekaw* and 11% of instances of *Papers* of *cmt*.

Table 6.3: Populated entities/total entities by ontology. Number of CQAs covered by each ontology.

	cmt	conference	confOf	edas	ekaw
Classes	26 / 30	51 / 60	29 / 39	42 / 104	57 / 74
Obj. prop.	43 / 49	37 / 46	10 / 13	17 / 30	26 / 33
Data prop.	7 / 10	13 / 18	10 / 23	11 / 20	0 / 0
CQAs	46	90	67	60	84

6.2.2 Metrics for complex alignment evaluation

As stated above, a realistic way to address the complex alignment evaluation problem is to shifting from the comparison of correspondences into the comparison of sets of instances. We propose two evaluation measures. While the *CQA coverage* measure relies on pairs of equivalent SPARQL queries (source and target queries) and measures how well an evaluated alignment covers these queries, the *intrinsic*

[3] `https://framagit.org/IRIT_UT2J/conference-dataset-population`

precision compares the instances of the correspondences members. Intrinsic precision balances the CQA coverage by like precision balances recall in information retrieval. CQA coverage, in particular, requires however a way for rewriting the source query into the target query, in terms of the evaluated alignment.

Before detailing the two proposed measures, the overall evaluation workflow is introduced.

Evaluation workflow

Overall, the steps followed in the evaluation process (for both proposed measures) are (Figure 6.2): (i) anchor selection; (ii) comparison; (iii) scoring; and (iv) aggregation. While this workflow also applies to the evaluation of simple alignments, these steps applied to the evaluation of complex alignments are introduced in the following.

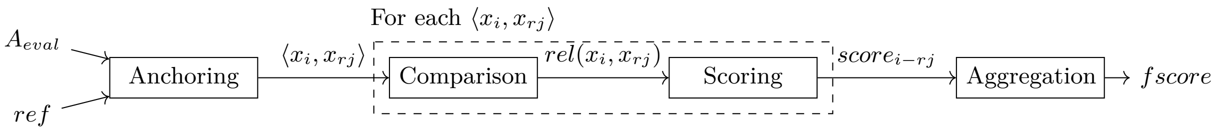

Figure 6.2: Evaluation process of the alignment A_{eval} with a reference ref.

Anchoring The anchor selection step consists in outputting a pair of comparable objects $\langle x_i, x_{rj} \rangle$. x_i is an object related to the evaluated alignment A_{eval} and x_{rj} is an object related to the reference ref. The objects depend on the type of reference. Having queries as references, the anchoring phase can consist in translating a source query based on the evaluated alignment into the target query (as for the CQA measure introduced below).

Comparison The purpose of the comparison step is to output a relation $rel(x_i, x_{rj})$ for each pair previously obtained $\langle x_i, x_{rj} \rangle$. Here, we reduce this comparison to the comparison of instances (instead of a syntactic or semantic comparison, given the drawbacks discussed above). The relation can be an equivalence, a subsumption, an overlap, a disjoint, *etc.* Given a reference set of instances I_{ref} and an evaluated set of instances I_{eval}, the possible relations between I_{ref} and I_{eval} are represented in Equation 6.1.

$$rel(x_i, x_{rj}) = rel(I_{ref}, I_{eval}) = \begin{cases} \equiv \text{ if } I_{eval} \equiv I_{ref} \\ \sqsubseteq \text{ if } I_{eval} \subseteq I_{ref} \\ \sqsupseteq \text{ if } I_{eval} \supset I_{ref} \\ \between \text{ if } I_{eval} \cap I_{ref} \neq \emptyset \\ \overset{\emptyset}{=} \text{ if } I_{eval} = I_{ref} = \emptyset \\ \perp \text{ if } I_{eval} \cap I_{ref} = \emptyset \text{ and } I_{eval} \cup I_{ref} \neq \emptyset \end{cases} \tag{6.1}$$

In particular, for the CQA coverage, this comparison can also allow for measuring not only the relation between the instance sets but also indicating the *query precision* (QP) an *query recall* (QR). These measures are introduced here because they are also useful for the anchoring phase in the CQA coverage computation, as introduced below.

$$QP = \frac{|I_{eval} \cap I_{ref}|}{|I_{eval}|} \qquad QR = \frac{|I_{eval} \cap I_{ref}|}{|I_{ref}|} \tag{6.2}$$

Scoring The scoring step associates a score to each relation found in the comparison step. Thus, the scoring functions are directly impacted by the relation $rel(x_i, x_{rj})$ found between the objects. The scoring function gives a score between 0 (for incorrect) and 1 (for correct). Different scoring metrics such as relaxed, recall or precision-oriented metrics (as introduced in Section 2.3) can be adopted. This allows for measuring the quality of the alignment under different but complementary perspectives. Different measures suit different evaluation goals. If one want to improve the system, it is better to have as many indicators as possible. For instance, if the evaluation measures how well an alignment allows for retrieving all results for a given query, regardless of the precision, a *recall-oriented* is preferred. If the purpose of the evaluation is to measure the exactitude of an alignment, then a classical function (1 if correct, 0 if incorrect) can be applied. Based on the relation between the instance sets, we propose a set of scoring functions. The **classical** (Equation 6.3), **recall-oriented** (Equation 6.4) and **precision-oriented** (Equation 6.5) scoring functions are used in state-of-the-art works to emphasize whether the alignment favours precision or recall (as introduced in Section 2.3). We introduce the **overlap** metric to represent whether two set have at least one common instance (Equation 6.6). The **not disjoint** metric gives a 1 score to all the overlapping sets or the sets where I_{ev} and I_{ref} are empty sets.

$$classical(I_{ref}, I_{ev}) = \begin{cases} 1 & \text{if } rel(I_{ref}, I_{eval}) = \ \equiv \\ 0 & \text{otherwise} \end{cases} \tag{6.3}$$

$$recall\ oriented(I_{ref}, I_{ev}) = \begin{cases} 1 & \text{if } rel(I_{ref}, I_{eval}) = \ \sqsupseteq \\ 0.5 & \text{if } rel(I_{ref}, I_{eval}) = \ \sqsubseteq \\ 0 & \text{otherwise} \end{cases} \tag{6.4}$$

$$precision\ oriented(I_{ref}, I_{ev}) = \begin{cases} 1 & \text{if } rel(I_{ref}, I_{eval}) = \ \sqsubseteq \\ 0.5 & \text{if } rel(I_{ref}, I_{eval}) = \ \sqsupseteq \\ 0 & \text{otherwise} \end{cases} \tag{6.5}$$

$$overlap(I_{ref}, I_{ev}) = \begin{cases} 1 & \text{if } rel(I_{ref}, I_{eval}) = \ \between \\ 0 & \text{otherwise} \end{cases} \tag{6.6}$$

$$not\ disjoint(I_{ref}, I_{ev}) = \begin{cases} 1 & \text{if } rel(I_{ref}, I_{eval}) = \between \ \ or\ rel(I_{ref}, I_{eval}) = \ \overset{\emptyset}{=} \\ 0 & \text{otherwise} \end{cases} \tag{6.7}$$

For the CQA coverage, note that **queryFmeasure** is used as scoring function in order to represent how close I_{ev} is to I_{ref}, based on Equation 6.2.

$$queryFmeasure(I_{ref}, I_{ev}) = 2 \times \frac{QR \times QP}{QR + QP} \tag{6.8}$$

Aggregation The scores are locally and globally aggregated to give the $fscore$ (Figure 6.2). The local aggregation aggregates all scores for a given object. There can be different local aggregations. For example, there can be an aggregation over the evaluated object and one over the reference object. The global aggregation aggregates all the locally-aggregated scores. For example, if the local aggregation was performed over the reference object, all the reference objects were given a score. The reference object scores can be aggregated into a final score. A final score locally aggregated over the evaluated objects is often referred to as the *precision* score. A final score locally aggregated over the reference objects is often referred to as the *recall* score.

We have identified in [Thiéblin et al., 2019d, Zhou et al., 2019] that **anchor selection** and **comparison** are the most difficult steps to automatise for complex alignment. The instance-based comparison seems adequate if run on a dedicated dataset. Using equivalent SPARQL CQAs as reference would ensure that the two compared objects are equivalent because they model the same piece of knowledge. In the following, an instance set will be represented as such I_e^{KB}, e being the entity which represent the instances (formula or query) and KB the knowledge base in which this instance set has been retrieved (e.g., source KB_s or target KB_t knowledge bases).

CQA coverage

This metrics evaluates how an alignment covers a set of knowledge needs represented by CQAs. The reference for the CQA Coverage is a set of equivalent CQAs in the form of SPARQL queries. Each source CQA cqa_s has an equivalent target CQA cqa_t. An evaluated alignment A_{eval} is used to rewrite each source CQA cqa_s. The rewritten queries are then compared to the reference target CQA cqa_t.

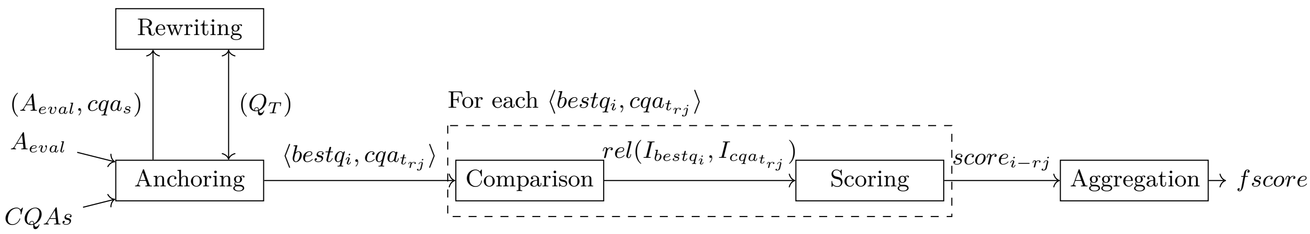

Figure 6.3: CQA-coverage evaluation process.

Source CQA anchoring In the anchoring step, each source cqa_s is rewritten using the generated alignment A_{eval}. The rewriting phase outputs all the possible rewritten target queries from the rewriting systems as the set $Q_T = rewrite(cqa_s, A_{eval}, KB_s)$. For each rewritten query q_t in Q_T, a pair (q_t, cqa_t) is formed. We considered two rewriting systems in this work (both introduced in Section 3.2.3). None of these systems take account of the correspondence relation or confidence value. The first system was proposed in [Thiéblin et al., 2016]. Each triple of cqa_s is rewritten using A_{eval}. When the predicate or object of the triple appears as the source member of a correspondence in A_{eval}, the target member of this correspondence is transformed into a SPARQL subgraph and put in the triple's place in the query. This system only deals with (s:c) correspondences. If a triple can be rewritten with different correspondences, all the possible combinations are added into Q_T. The second system is based on instances. The instances $I_{cqa_s}^{KB_s}$ of cqa_s are retrieved from KB_s. For each correspondence c_i of A_{eval}, the instances represented by its source member e_s are retrieved over KB_s. If $I_{e_s}^{KB_s} \equiv I_{cqa_s}^{KB_s}$, then, the target member of c_i is transformed into a query and added to Q_T. For instance, the query `SELECT ?s WHERE ?s a o1:AcceptedPaper.` retrieves a set of accepted paper instances in the o_1 ontology. This set of instances is then compared to the set of instances described by the source member of each correspondence. In this case, o_1*:AcceptedPaper* describes exactly the same set of instances as the source member of (o_1*:AcceptedPaper,* $\exists o_2$*:hasDecision.*o_2*:Acceptance,*$\equiv$). This system can deal with (c:c) correspondences but cannot combine correspondences in the rewriting process, *i.e.*, if more than one correspondence is needed to rewrite the query, the system can not deal with it.

The rewriting phase outputs all possible queries regardless of the correspondence relation. A lot of noise can therefore be introduced. Moreover, the same query can be output by both rewriting systems. Consequently, the best *queryFmeasure* score is applied to over the rewritten queries to select the best one (Equation 6.8), noted $bestq_t$ (Equation 6.9, where *rewrite* is a query rewriting function rewriting cqa_s into cqa_t using the A_{eval}).

$$bestq_t = \underset{q_t \in rewrite(cqa_s, A_{eval}, SKB)}{\text{argmax}} queryFmeasure(I_{cqa_t}^{KB_t}, I_{q_t}^{KB_t}) \quad (6.9)$$

Comparison and scoring The instance-based comparison and scoring are performed as presented above, with $I_{bestq_t}^{KB_t} = I_{eval}$ and $I_{cqa_t}^{KB_t} = I_{ref}$. A chosen scoring function $f \in \{$*classical, recall oriented, precision oriented, overlap, not disjoint, queryFmeasure*$\}$ is applied to the instance sets of $bestq_t$ and cqa_t: $f(I_{cqa_t}^{KB_t}, I_{bestq_t}^{KB_t})$. If a source CQA could not be rewritten by the alignment, its scores are all 0. An average function is then performed to aggregated the scores per pair of CQAs into a final score.

Aggregation A chosen scoring function is applied to each $bestq_t$ and these scores are averaged to give the CQA Coverage score. The final equation of the CQA Coverage is shown in Equation 6.10. cqa_{pairs} the set of pairs of CQAs as equivalent

SPARQL queries and f the chosen scoring function (chosen between Equations 6.3 – 6.8).

$$coverage(A_{eval}, cqa_{pairs}, KB_s, KB_t, f) = \underset{\langle cqa_s, cqa_t \rangle \in cqa_{pairs}}{\text{average}} f(I_{cqa_t}^{KB_s}, I_{bestq_t}^{KB_t}) \quad (6.10)$$

Intrinsic instance-based precision

In order to counterbalance the CQA Coverage score, we propose to measure the *intrinsic instance-based precision* of an alignment. In the OAEI 2018 Taxon evaluation [Algergawy et al., 2018], each correspondence of the alignment has been manually evaluated and classified as true positive or false positive. The proposal here is to automatize this process based on the comparison of instances (Figure 6.4). For each correspondence $c_i = (e_s, e_t)$ in the A_{eval}, the instances $I_{e_t}^{KB_t}$ represented by the target member e_t are compared (using the scores presented above) to the instance $I_{e_s}^{KB_s}$ represented by the source member e_s. We arbitrarily chose that the reference instance set (I_{ref}) is $I_{e_s}^{KB_t}$ and the evaluated one (I_{eval}) is $I_{e_t}^{KB_t}$. This decision affects the *recall-oriented* and *precision-oriented* scores which are directional.

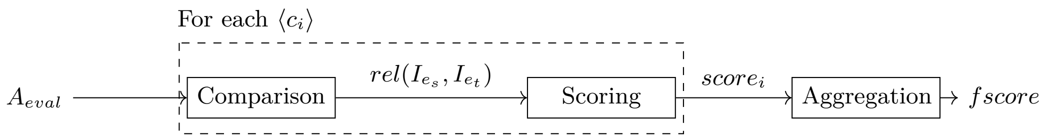

Figure 6.4: Evaluation process for intrinsic instance-based precision

The scores of the correspondences are then averaged to give the *Intrinsic Precision* score of the evaluated alignment A_{eval}. Equation 6.11 shows the calculation of the Intrinsic Precision for an evaluated alignment A_{eval}:

$$precision(A_{eval}, KB_s, KB_t, f) = \underset{\langle e_s, e_t \rangle \in A_{eval}}{\text{average}} f(I_{e_s}^{KB_s}, I_{e_t}^{KB_t}) \quad (6.11)$$

The limitations of the intrinsic instance-based precision are however manifold. First, the relation of the correspondence is not taken into account in the comparison. Then, the population of the ontologies clearly impacts the score. For example, if an ontology class o_1*:Document* is only populated with *Paper* instances, and another o_2*:Document* is only populated with *Review* instances, the correspondence (o_1:Document,o_2:Document,$\equiv$) will have a 0 score for all the metrics we proposed. On a data set where two common classes are either populated with the same instances, not populated or share at least a subclass with the same instances, this metric may give a lower and upper bound for the precision of the alignment. The lower bound is given by the *classical* score in which only equivalent members are considered correct. The upper bound is given by the *not disjoint* score in which all correspondences with overlapping or empty members are considered correct.

6.3 Conclusions

This chapter has presented evaluation benchmarks on which the approaches generating complex correspondences can be evaluated, together with metrics for evaluating complex alignments. In fact, alignment evaluation is often performed by comparing a generated alignment to a reference one. While this comparison is straightforward for simple alignments, this step becomes harder when dealing with complex correspondences. While syntactic-oriented evaluation metrics (measuring the effort to transform a correspondence into another) would fail in covering the high space of possible combinations between constructors in complex correspondences, semantic-oriented approaches would restrict the expressiveness of correspondences to those supported by current reasoners, leaving aside for instance, transformation functions. Hence, comparison of instance sets seems to be reasonable. Our proposal shifts the problem to the comparison of instances in a task of query rewriting targeting user needs. The alignment to be evaluated is used to rewrite a set of reference source queries whose results (set of instances) are compared to the ones returned by the corresponding target reference queries. This coverage metric balances with the alignment precision based on common instances.

Evaluating complex ontology alignments, however, is a too broad challenge to be tackled with a single approach, as there are multiple aspects to take into account. A complementary approach to the instance-based one proposed in this chapter could be an edit-distance approach that would reflect the effort involved in human validation. The approach should be also scalable, and avoid the need to do all-vs-all correspondence comparisons. This could also be achieved by considering the possibility of computing minimal complex correspondences (or key complex correspondences, which can be used for computing all the other ones), in line with the work of [Maltese et al., 2010]. In order to cover ontologies of various sizes and domains, developing a query generation system able to automatically generate queries adequate in coverage and scope to the evaluation of complex alignments could also help in the evaluation task.

In the next chapter, the evaluation of holistic but simple alignments is addressed, in particular on the perspective of creating holistic benchmarks.

Chapter 7

Holistic matching evaluation

7.1 Motivation

As introduced in Chapter 4, novel approaches dedicated to holistic ontology matching have emerged in the literature, still relatively few if compared to pairwise approaches. While systematic evaluation of matching approaches has been dedicated to pairwise systems, there is a lack of reference alignments on which holistic approaches can be systematically evaluated. We argue that fostering the development of holistic approaches depends on the availability of such data sets. According to [Oliveira and Pesquita, 2015], producing such kind of alignments could be potentially useful to support a next generation of semantic technologies.

Current holistic approaches have been mostly manually evaluated on data sets used in the context of the tool development. In [Pesquita et al., 2014], the authors propose to exploit OBO cross-products to create ternary compound alignments between ontologies, in order to create a benchmark. They have created a set of seven cross-products collections each with at least 100 definitions corresponding to ternary compound correspondences. More recently, in [Nentwig et al., 2017], a reference alignment for multi-source clustering of large data sets from the geographic and music domains has been proposed. They evaluate the efficiency and scalability of the distributed holistic clustering for large data sets with millions of entities from the two domains on the task of link discovery.

7.2 Contributions

We have been working on an approach for constructing *pseudo*-holistic reference alignments from available pairwise ones, as a first step towards the creating of reference holistic benchmarks. We discussed the problem of relaxing graph cliques representing these alignments involving a different number of ontologies. In fact, we can see the pairwise matching as a special case of holistic ontology matching. The main contributions of this work are [Roussille et al., 2018a]:

- we propose an approach for constructing holistic (reference) datasets from existing (and depending on) pairwise alignments, relying on different levels of relaxation of graph cliques.

- we applied our approach on the OAEI Conference data set, aiming to produce a baseline for our work in order to produce a similar matching task, as there is no current track providing holistic alignment challenges.
- we evaluate existing matching tools participating in OAEI tracks on this new task and discuss the pertinence of having such a holistic dataset.

This work has been carried out in collaboration with Philippe Roussille (post-doc at IRIT), co-advised with Imen Megdiche (Assistant professor at Institute Jean-Francois Champollion and researcher at IRIT) and Olivier Teste (Full professor at University Toulouse 2 – Jean Jaurès and researcher at IRIT).

In the following, the methodology for constructing the data set is introduced, followed by the evaluation of matching systems on this first version of the dataset.

7.2.1 Building holistic alignments

The methodology we followed to automatically constructing holistic alignments from available pairwise alignments is composed of two main steps:

1. building a graph of all combinations of correspondences from existing pairwise alignments;
2. building the holistic alignments according to different levels of relaxation with respect to complete graphs (cliques): clique-strict method (level 1) and clique-relaxed subgraph method (level 2). In case of level 2, we proposed two sub-methods: the first method is a systematic relaxation of cliques, and the second one handles the intra-ontology choice of entities based on ontology relations.

Step 1: building the graph of N pairwise alignments This step aims at building a holistic graph $G_H = (V_H, E_H)$ where nodes are entities from the ontologies to be aligned, and edges are correspondences from pairwise alignments, such as:

- $V_H = \left\{e_{i_k} | e_{i_k} \in \cup_{i=1}^{N} o_i\right\}$,
- $E_H = \{(e_i, e_j) | \exists < \{e_i, e_j\}, r, n > \in A_{1..N}\}$, with $A_{1..N} = \cup_{k=1,l=k+1}^{N-1} A_{kl}$.

Considering, for instance, $N = 4$, this leads to the group of $A_{12} \cup A_{13} \cup A_{14} \cup A_{23} \cup A_{24} \cup A_{34}$.

Step 2: building holistic alignments For building the holistic alignments $A_{1...N} = \{c_1, c_2, .., c_i, ..| i \in N\}$, each correspondence should cover the N input ontologies and should be (1:1) holistic alignment to conserve the (1:1) requirements of the pairwise alignments. In the following, the two levels of relaxation we adopt to generate the holistic alignments are introduced.

Clique-strict method (level 1) The first method concerns the generation of holistic alignments composed of cliques. The cliques are complete graphs extracted from the holistic graph G_H. The algorithm consists in searching complete subgraphs composed of N nodes belonging to the N input ontologies. A clique is considered as the most strongest holistic correspondence, hence it has the confidence value 1. The structure of the graph G_H built upon (1:1) pairwise alignments guarantees that each ontology is present only once in the cliques.

Clique-relaxed subgraph method (level 2) The clique-strict method is too strict because we are faced most commonly to incomplete graphs that should be part of the solution of holistic alignments. This is what is depicted in Figure 7.1. Figure 7.1(b) is part of the solution of the complete graph of Figure 7.1(a). Hence, we can infer from the subgraph of Figure 7.1(b) a holistic alignment with a lower level of confidence corresponding to its incompleteness with respect to the clique.

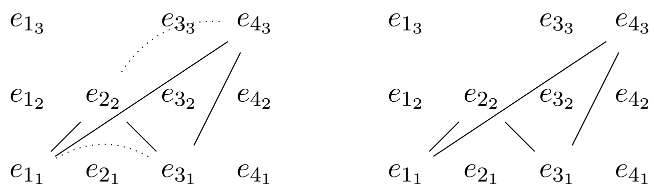

Figure 7.1: (a) Example of a clique subgraph; (b) The clique-relaxed subgraph.

In order to compute the confidence of the clique-relaxed subgraph, we define the notion of *clique-likeness*, which is the geometric distance of a subgraph compared to a clique; for instance, the level of confidence of the graph of figure 7.1(a) is $\frac{2}{3}$. The formula is as the following for a subgraph denoted $G_i = (V_i, E_i)$:

$$clique\_likeness(G_i) = \frac{2 * |E_i|}{|V_i| * (|V_i| - 1)}$$

We search all the subgraphs of G_H with respect to two conditions, namely that all ontologies should be represented by at least one node, and that each subgraph G_i is maximal. Based on the content of these subgraphs, we provide two methods to generate the holistic alignments.

Method 1 (level 2): Clique-relaxed holistic alignment algorithm This method is a systematic relaxation of cliques, which means that the subgraphs are incomplete cliques composed exactly of one node from the N input ontologies.

Method 2 (level 2): clique-relaxed subgraphs based on intra-ontology relations This method handles the case when the subgraphs are composed of one or several nodes from ontologies o_i, for some or all $i \in [1, N]$. The proposed method will then select only one tuple of nodes based on the intra-ontology relations and the best confidence value of *clique_likeness*.

By taking the example of Figure 7.2, we notice that the subgraph have two nodes from o_1, noted e_{1_1} and e_{1_2}, so we have to choose either the solution 1, composed of the clique-relaxed $= \{e_{1_1}, e_{2_1}, e_{3_2}, e_{4_3}\}$ or solution 2, composed of the clique-relaxed $= \{e_{1_2}, e_{2_1}, e_{3_2}, e_{4_3}\}$.

- For solution 1 (a), the $clique\_likeness(G_i) = \frac{1}{3}$.
- For solution 2 (b), we propose that we can use the relationship between e_{1_1} and e_{1_2} to infer new correspondences for e_{1_2}. As the $e_{1_2} \subseteq e_{1_1}$ (subclassof relation) and $< \{e_{1_1}, e_{2_1}\}, \equiv, 1 >$ thus we can infer the pairwise correspondences $< \{e_{1_2}, e_{2_1}\}, \equiv, 1 >$. Therefore, the $clique\_likeness(G_i) = \frac{1}{2}$.

Based on the *clique_likeness* score, we choose the solution 2 because of its higher confidence value.

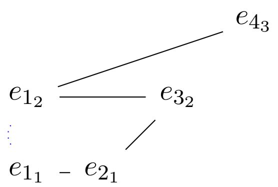

Figure 7.2: Example of intra-ontology multiple choice. The circled elements belongs to the same ontologies, the black vertices shows the extra-ontological links while the blue dotted vertices shows intra-ontological links.

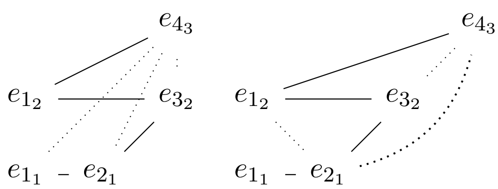

Figure 7.3: (a) Solution 1 and (b) solution 2 from Figure 7.2.

In the example of Figure 7.4, we illustrate the case of $N = 4$ ontologies from the OAEI Conference Track (*cmt*, *conference*, *iasted* and *edas*). We notice two possible solutions that can be proposed for the subgraph composed of the entities "Submission" (iasted), "Submitted_contribution" and "Paper" (conference), "Paper" (edas), and "Paper" (cmt). In order to find the alignment, we compute the score of the two potential clique-relaxed subgraphs which contains either the entity "Paper" or "Submitted_contribution" (conference). The retained holistic alignment is solution 1, which has the highest score; its confidence value is $\frac{3}{6} = 50\%$.

Evaluation Our holistic dataset has been constructed from the OAEI Conference data set. We have applied the 3 methods described above for generating the holistic reference alignments on the basis of available pairwise reference alignment (ra1). All the generated alignments and the code for generating them are available online[1].

[1] `https://github.com/PhilippeRoussilleIRIT/EKAW-2018-holistic`

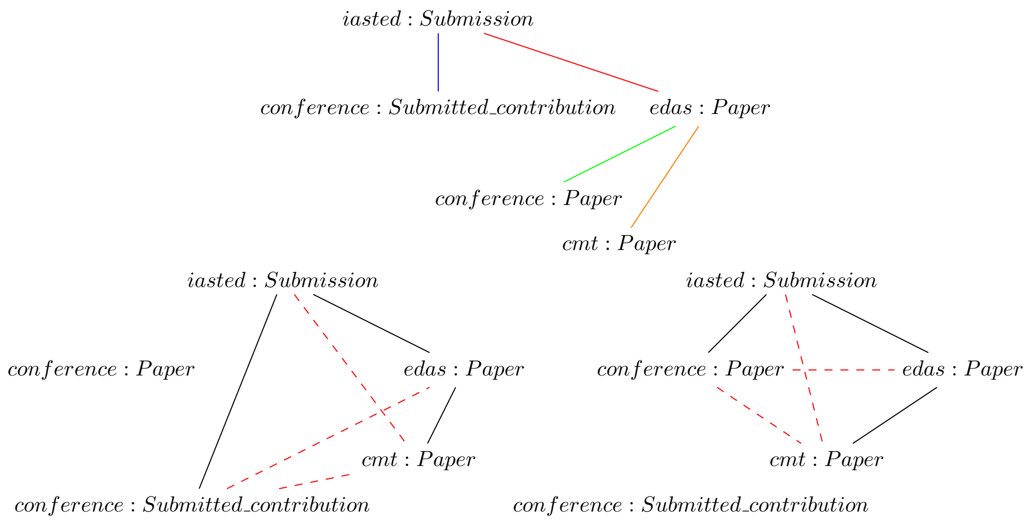

Figure 7.4: (a) Original extracted subgraph (top), (b) method 1 (left), (c) method 2 (right).

We have also applied our methodology to generate holistic alignments from the available results of OAEI 2017 participating tools and compared their results with the LPHOM system (Chapter 4). The available results for the following tools were considered: ALIN, AML, KEPLER, LogMap, LogMapLt, ONTMAT, POMap, SANOM, WikiV3 and XMap. Even though these tools were not developed for that purpose, their results were the only available for a baseline comparison. To the best of our knowledge very few holistic systems are available. We have run the AML-Compound tool[2], but it was not able to generate any alignment for this data set.

We could clearly distinguish two types of behaviours:

- for $N = 3$ and $N = 4$, with few exceptions, we can observe that the clique-strict method results are closer to both relaxed methods. It shows that with few number of ontologies, regardless the kind of method, the correspondences generated by the methods are close. The structural difference given a clique compared to a relaxed clique is smaller the fewer nodes in the sub-graph.

- for $N = 5$ and $N = 6$, we can observe that the method clique-strict is better than both relaxed methods. These cliques allows for identifying the common entities shared across the ontologies. In the case of the Conference dataset, by manually examining the outputs, the clique-strict alignments are composed of exact matches. It corroborates the intuition that increasing the number of nodes in a subgraph, increases the differences between their structures (cliques and relaxed cliques structures).

With respect to the OAEI matching systems, although this evaluation setting

[2] https://github.com/AgreementMakerLight/AML-Compound

may introduce a bias in the evaluation, in the lack of available fully holistic tools, it is the material we have for comparison. Looking at first to the holistic tool, we can observe that, although the LPHOM holistic approach does not perform very well for a small number of ontologies, it is in the top-3 (F-measure) for $N = 6$ ontologies (for all methods). As expected, the tools specifically designed for the pairwise task better perform for $N = 2$. Their performance however mostly decreases with the increasing of N (some are not able to generated alignments for $N = 6$), while LPHOM relatively maintains its performance.

Overall, in terms of precision, LPHOM (.56) is of the top-4 systems (AML and ALIN .59, XMap .58 and PopMap .57). The holistic approach privileges precision in detriment of recall (.35), with coherent generated alignments. In terms of F-measure, the given results are intermediate, about .10 points (.42) compared to the best system, which is AML (.52). However, we have to keep in mind that our approach here is *pseudo*-holistic, and thus heavily influenced by the number of ontologies. As the number of ontologies increases, reaching up to 6, the F-measure decreases, showing that there are room for improvements. This can be explained due to the structures of the tasks and the way the tools work: as the matching structures differ from a strict clique approach (which, in a pairwise context, is kept all the time as pairwise alignments are cliques), the limits between matches become blurrier. Most tools will easily find a similarity between two entities, and two groups of entities, but the transient aspect of the pseudo-holistic relaxation cannot be easily translated in terms of strictness. As such, when trying to assess all ontologies at once, only the main and nearly exact matches remain; while when computed pairwise, this information cannot be extrapolated as the similarity matrix does not incorporate the new similarities. Finally, we are ware that the performance of the different matchers compared to LPHOM are not as significant as if our experiments were ran using specifically holistic matchers. However, they are significant enough to show that the holistic matching task has inherent properties.

7.3 Conclusions

This chapter has presented a methodology for constructing holistic alignments from existing pairwise alignments. The approach relies on graph cliques involving a different number of ontologies. This is a first step towards the construction of reference holistic alignments in order to overcome the lack of such data sets in the field. While the focus on this chapter has been the creation of (reference) datasets, there is room for further improvements in terms of holistic evaluation metrics. Semantic measures for evaluating this kind of alignment, in particular with respect to the logical incoherence could be further investigated. This could take inspiration from what has been done in reasoning with network of alignments [Zimmermann and Le Duc, 2008, Klai et al., 2016]. Furthermore, current evaluation infrastructures (as the Alignment API and its Alignment format, evaluation bridges and clients, and semantic evaluation metrics) have to be further extended to accommodated the possibility of having multiple ontologies as input in the matching process. The evaluation of holistic alignments could also benefit from benchmark generators.

In line in terms of creating reference alignments, the next chapter discusses the manual creation of alignments between foundational and domain ontologies.

Chapter 8

Foundational and domain ontology evaluation

8.1 Motivation

As stated in Chapter 5, matching foundational and domain ontologies is far from being a trivial task and most approaches still rely on manually or semi-automatic strategies. This has been corroborated in [Stevens et al., 2018], where manually classifying domain entities under foundational ontology concepts is reported to be difficult to do correctly. The findings in [Stevens et al., 2018] also point out the need for improving the methodological process of manual integration of domain and foundational ontologies, in accord with what has been stated in [Keet, 2011]. We argue that fostering the task can be done by boosting automatic systems to take into account this kind of ontologies. However, despite the variety of datasets in OAEI, it still lacks matching tasks involving foundational ontologies.

While different ontology alignments have been constructed from manual analysis, involving a different number of experts and resulting in different levels of agreement, the focus has mostly been on describing the resulting alignment rather than on the details of the manual process. Guidelines for constructing alignments are in fact scarce in the field, though there are more general discussions on the qualities of a good benchmark in other research fields [Sim et al., 2003, Dekhtyar and Hayes, 2006]. It may be obvious that the fundamental problem of aligning ontologies is determining what is the meaning of the terms that are candidates for alignment. If the meaning is implicit, and one must resort to the domain knowledge of human matchers, then only an automatic suggestion is feasible. This is even more required when dealing with foundational ontologies.

With respect to the creation of evaluation benchmarks, as knowledge on foundational ontologies is highly specialized, it is important that such alignments consider the participation of different experts in the area.

8.2 Contributions

We have been working on creating a reference alignment between the Conference ontologies and SUMO (*Suggested Upper Merged Ontology*) [Niles and Pease, 2001, Pease, 2011]. The choice for aligning this OAEI dataset to a foundational ontology is motivated by the fact that it has became one of the most used in matching evaluations [Zamazal and Svátek, 2017]. As a complete manual alignment between SUMO and WordNet has been previously provided [Niles and Pease, 2003] and continually updated since the original effort, we argue here that using these alignments as bridges to matching domain ontologies to SUMO can facilitate the matching task. We have chosen SUMO for several reasons. It is the only formal ontology that has a complete set of manually-performed correspondences to all 117,000 word senses in WordNet. It is also one of the few ontologies that has a detailed formalization in an expressive logical language. Most ontologies are still simple taxonomies and frame systems, and so assessing the meaning of their terms requires human intuition based on term names and relationships. SUMO includes a computational toolset [Pease and Benzmüller, 2013] that allows users to test the logical consistency of its definitions, which provides a guarantee of quality and correctness than just testing type constraints. Lastly, SUMO is large and comprehensive at roughly 20,000 terms and 80,000 hand-written logical axioms, exceeding the size of other open source foundational ontologies by several orders of magnitude.

The main contribution of the work on matching foundational and domain ontologies is a reference alignment between SUMO and the domain ontologies from the OAEI Conference track [Schmidt et al., 2019a]. This work is a first step toward the construction of a dataset involving foundational ontologies that can serve to evaluation systems in the context of OAEI campaigns. These alignments can also be explored as semantic bridges in domain ontology matching. Currently, this work is being done with DOLCE. Assuming that the right synsets have been selected and that these synsets have been also manually aligned to DOLCE, we reproduce the alignment reported with SUMO.

This work has been carried out in collaboration with Daniela Schmidt (PhD student at Pontificia Catolica Universidade do Rio Grande do Sul, Brazil), co-advised with Renata Vieira (Full Professor at Faculty of Informatics at Pontificia Catolica Universidade do Rio Grande do Sul, Brazil), and Adam Pease (Infosys, Foothill Research Center, USA), the creator of the SUMO foundational ontology.

8.2.1 Foundational and domain alignment dataset

Before describing the dataset, we introduce the overall methodology we have followed to create the consensual alignment between SUMO and the domain ontologies. We have reduced the problem to the first-level concepts of the hierarchies from the domain ontologies. This has resulted in 70 first-level domain concepts (Table 8.1). For each first level concept of the domain ontology, a foundational specific concept is associated. The cost of doing manual alignment with first level concepts is smaller, as it is reduced to the number of concepts at the first level. Four evaluators have been involved in the task of aligning the 70 top-level domain concepts to

Ontology	Type	#Concepts	#Top
Cmt	Tool	36	8
ConfTool	Tool	38	7
Edas	Tool	104	16
Ekaw	Insider	74	6
Iasted	Web	140	10
Sigkdd	Web	49	9
Sofsem	Insider	60	14

Table 8.1: Number of concepts, top-concept and relations in the reference alignment.

the WordNet synsets. The evaluators are researchers, therefore all have common-sense knowledge about conferences (the domain ontology), they have background in Computer Science and are well acquainted with ontology matching. One of the evaluators is the creator of the SUMO ontology.

The overall methodology is articulated in the following two steps:

- Individually generating the alignments between domain concepts and WordNet synsets;
- Collaboratively validating the set of found correspondences.

Individual generation of correspondences In this first step, each evaluator aligned each of 70 domain concepts to WordNet synsets. To that extent, each domain concept and the corresponding WordNet synsets, resulting from searching in WordNet for the term associated to the domain concept, were listed in a spreadsheet to the evaluators. In the absence of entries in WordNet for the terms, a head modifier strategy has been applied (i.e., 'WrittenPaper' is a 'Paper'). Only one concept had not corresponding entry in WordNet (*sigkdd#Sponzor*[1]). In order to help the evaluator to understand the context of the domain concept, their sub-concepts were also presented. As the domain ontologies are equipped with very few comments or labels, we have completed the description of the concept using the definitions from the Cambridge Dictionary[2]. However, we are aware that the found definitions may not reflect the exact semantic of the concept. Each evaluator then was asked to select the right WordNet synset for each domain concept. The evaluators were instructed to select one option for each domain concept, however, in some cases more than one sense was selected. This happens because the domain concept was not clear enough, or the senses available in WordNet were very general. Evaluators were also invited to comment their decisions. Table 8.2 shows a fragment of the spreadsheet for the domain concept *ekaw:Document*.

Validating the correspondences After the individual annotation of each domain concept with the WordNet synset, the annotators were able to see the annotations of each other and identify the conflicts. Based on the views on the other

[1] http://oaei.ontologymatching.org/2018/conference/data/sigkdd.owl
[2] http://dictionary.cambridge.org/us/

Table 8.2: Example of spreadsheet adopted by the evaluators.

Domain concept	WordNet synsets	SUMO concept
ekaw#Document	**S1**: writing that provides information (especially information of an official nature) **S2**: anything serving as a representation of a person's thinking by means of symbolic marks **S3**: a written account of ownership or obligation **S4**: (computer science) a computer file that contains text (and possibly formatting instructions) using seven-bit ASCII characters	**1 – FactualText+** 2 – Text= 3 – Certificate+ 4 – ComputerFile+

Table 8.3: Examples of conflicts solved after discussion between annotators. Boldface concepts represent the conflicts and final results are indicated with a underline.

Domain concept	SUMO concept
cmt#Decision	1: **Learning+** 2: **Deciding+** 3: TraitAttribute+ 4: ConstantQuantity+ 5: ConstantQuantity+
cmt#Preference	1: **IntentionalRelation+** 2: SubjectiveAssessmentAttribute+ 3: PsychologicalAttribute+ 4: **PsychologicalAttribute+**
edas#PersonalHistory	1: **PastFn=** 2: History= 3: **HistoricalAccount+** 4: PastFn= 5: Proposition+
sigkdd#Award	1: UnilateralGiving+ 2: **ContentBearingObject+** 3:**UnilateralGiving+**
Conference#Review_preference	1: **IntentionalRelation+** 2: SubjectiveAssessmentAttribute+ 3: PsychologicalAttribute+ 4: **PsychologicalAttribute+**

annotators (and their comments), each one was able to change their initial annotations. For those conflicts where the comments were not enough for understanding the annotation, an online discussion took place. From the 70 annotated domain concepts, 8 of them have been annotated with different WordNet synsets. Table 8.3 lists some examples. All generated alignments are available in the Alignment API format[3].

Discussion During the process of alignment construction, several difficulties arose for interpreting the real meaning that the concept represents in the domain ontology. For instance, the concepts `Bid` and `Preference` (Table 8.3) in `cmt` ontology had no description clarifying its use, and no sub or super concepts which could be used to clarify their meaning. In these cases, the evaluators discussed and considered the proper meaning according to their own interpretation of the domain, however, such cases may interfere with the quality of the resulting reference alignment because there is no objective standard for what the meaning, and therefore the correct correspondence must be. We have only the consensus guess about intended meaning

[3]https://github.com/danielasch/ReferenceAlignment

among human evaluators. In addition, some concepts represented in the ontology present other kind of problems such as doubts regarding ontology elements' adequacy, for example, the concept `ReviewRating` in `edas` ontology, which according to the discussion raised by the evaluators, a rating could be a relationship between a thing, an agent and a rating value. In the same way, the concept `Deadline` in `sigkdd` ontology could be a relationship between the conference and a date. They are however defined in those ontologies as concepts, rather than relationships.

In other cases, sub-concepts are different from first-level concepts and therefore they represent different information, as the concept `Event` in `ConfOf` ontology. Some of their sub-concepts `Social_event/Banquet`, `Working_event/Conference`, `Working_event/Workshop` are in line with the main concept, however others such as `Administrative_event/Camera_Ready_event` seems out of the context. In fact, it should not be a Process at all but a deadline for doing something (submitting a version of a paper, for instance).

In contrast, one can examine a SUMO definition of a term such as `FormalMeeting`[4] and see that it is necessarily a `Meeting` that is not a `SocialParty`, that it must be temporally preceded by a `Planning` that has the `result` of creating the meeting, as well as constraints that other events like a `Resolution` to be considered such, may only occur at a `FormalMeeting`. Something like a modern dictionary, but with the definitions expressed in logic, rather than human language, so that a machine can perform computation (and consistency checking) with those definitions. The cases described above consist of ontological representation problems commonly present in lightweight ontologies, and hinder the reuse and reliability of the represented knowledge. In addition, they highlight the importance of advancing in research that uses top-level ontologies to give more formalization to domain ontologies.

The challenges in aligning the OAEI ontologies should highlight two elements that are lacking in the majority of most of current ontology practice. The first element is the degree of reuse. Ontologies that are created from scratch suffer from the fact that their terms have only a small number of relationships to other terms. The point of having an ontology is to have a shared meaning among its users. When domain ontologies are created in isolation, rather than as extensions to widely used comprehensive ontologies they miss an opportunity for sharing common meaning. Modern software development, for example in Java or Python, means reusing vast amounts of existing code, such as extensive language libraries and other packages like web servers, databases, device drivers etc. Ontology development needs to follow the same practice to achieve the same efficiency of process as procedural software development.

A second element is the expressivity of definitions. If, as with several of the OAEI ontologies, one must guess at the intended meaning of a term only by its name, then there isn't much chance for shared meaning amongst its users. Each user will just be making a guess. If each term has only a set of binary relationships to other terms then it should still be clear that issues like mutual constraints on values and boundary cases are left unformalized and also at risk of being in conflict among its

[4] `http://sigma.ontologyportal.org:8080/sigma/Browse.jsp?lang=EnglishLanguage&flang=SUO-KIF&kb=SUMO&term=FormalMeeting`

users. Additionally, without a computational formalization of such constraints, the computer will not be able to test or enforce them. Comments in natural language, no matter how extensive or precise, will not overcome the need for computational definitions, and our experience in this matching effort has been that comments are often not even present, and rarely extensive or precise.

We were not engaged in correcting the ontologies, since they are part of public datasets. However we consider that a discussion about the problems identified is necessary. Perhaps more robust alignment processes would inherently require modifications in target domain ontologies but also certainly a more detailed formalization. Given the paucity of definitions, we are limited primarily to linguistically-based matches and use of WordNet is a suitable choice for assisting with this sort of match.

8.3 Conclusions

Systematic evaluation of matching systems still lacks benchmarks involving foundational ontologies. This chapter has discussed the effort of manually matching the well-known OAEI Conference dataset to the SUMO foundational ontology. We argue that boosting the development of matching systems able to better deal with this task depends on the availability of dedicated datasets on which these approaches can be systematically evaluated (over the time). The alignment presented here is a first step towards this objective.

Chapter 9

Perspectives

I have presented in this manuscript my main contributions in the last seven years to the domain of ontology matching. This domain has received a growing interest in the last decades and has so far reached some maturity, in particular with respect to approaches dealing with simple alignments involving a pair of ontologies at the same level of abstraction. In order to overcome some limitations in the field, I have worked on:

- generating complex alignments, as a way of generating more expressive correspondences between two ontologies than the correspondences involving single entities;
- holistic ontology matching, as a way of considering aligning multiple ontologies simultaneously; and
- the task of matching ontologies with different levels of abstraction, such as foundational and domain ontologies.

These matching contexts raised as well the problem of automatically evaluating the corresponding approaches, both in terms of benchmarks and evaluation measures, which I have also addressed in my work.

While the works presented here constitute significant contributions towards the maturity of the field, they can be further improved in different directions as it has been discussed in the conclusions of each chapter. Beyond the individual lines of research, several transverse aspects in complex, holistic and foundational matching generation and evaluation could also be further addressed:

- dealing with multilingualism;
- involving the user in the loop (mostly addressing simple correspondence in current proposals);
- better managing alignments repositories (which raises the problem of maintainability) [Arnold and Rahm, 2015] and the provenance of alignments (which could further extend the work by [Singh et al., 2017] to accommodate metadata related to the generation of alignments);

- developing tools for supporting visualization and collaborative creation and edition of alignments;
- generating of correspondence patterns (both for simple and complex correspondences, extending those from [Scharffe, 2009, Scharffe et al., 2014]) from ontology design patterns;
- dealing with uncertainty on alignments (which involves computing the correspondence confidence reflecting the uncertainty, explicitly representing it in the target language representing the alignments [Calì et al., 2008], and interpreting [Atencia et al., 2012] and reasoning on them (as [Al-Bakri et al., 2016] for inferring sameAs facts);
- or still managing the evolution of alignments [Hartung et al., 2013, Euzenat, 2016, Kozierkiewicz and Pietranik, 2019] (as ontologies and their instances potentially evolve over time, raising interesting challenges in those specific scenarios).

Last but not least, while I have worked on topics other than ontology matching (Section 1.2.7), I notice that there is a clear interface between them and ontology matching. While relation extraction from text can be exploited as a support to establish correspondences between ontologies (as I will further develop in the next section), there are challenges in integrating geographic knowledge bases that can be investigated under the perspective of ontology matching and entity alignment [Sun et al., 2019]. Furthermore, recent work has revised findings in OBDA with the help of alignments, where the notion of repair is transferred to the OBDA context [Bienvenu, 2018]. This could be useful when dealing with hybrid querying of multidimensional RDF data.

In the following, I discuss some potential work directions, such as interfacing NLP, machine learning, and ontology matching (Section 9.1), foundational distinctions of LOD datasets (Section 9.2) and alignment of multiple structured and unstructured sources (Section 9.3).

9.1 Interfacing NLP, machine learning and ontology matching

Despite the variety of matching approaches, most of them still fail on delivering other kinds of correspondence relations than equivalence. However, many tasks such as ontology merging, ontology evolution or data transformation require exploring relations such as subsumption, part-of, and disjointedness. For instance, it could be useful to identify that a given piece is part of a given airplane structure, when integrating airplane constructor's knowledge bases. In parallel, these relations (specially hypernym and meronym relations between terms, which correspond in many cases to the subsumption and part-of relations between concepts, respectively) have been largely studied in Natural Language Processing (NLP), in particular for the task of relation extraction from texts. Relation extraction is an active area and a variety of methods (linguistic, statistical, learning based, hybrid) have been proposed so far (reviews of them can be found in [Cui et al., 2017, Wang et al., 2017, Smirnova and Cudré-Mauroux, 2018]). It is a key step of ontology learning. In this field, the pioneering linguistic method is that of Hearst which defined a set of lexico-syntactic patterns specific to the hypernym relation for English, with extensions to this work focusing on other relations, such as meronym [Berland and Charniak, 1999, Gemechu et al., 2016]. With respect to statistical approaches, supervised learning [Bunescu and Mooney, 2005], distant learning [Mintz et al., 2009], or unsupervised learning [Fader et al., 2011] have been exploited, including distributional analyses [Lenci and Benotto, 2012, Fabre et al., 2014]. Complementary approaches also take advantage of the source layout, such as the documents structure [Kamel and Aussenac-Gilles, 2009, O'Connor and Das, 2011], or specific structures such as tables, categories [Chernov et al., 2006, Suchanek et al., 2007] or infoboxes [Auer et al., 2007] from Wikipedia pages. A tendency in the field is the adoption of high-dimension vector spaces and deep learning algorithms [Lin et al., 2016, Sorokin and Gurevych, 2017, Subasic et al., 2019].

Relatively few approaches have exploited relation extraction in ontology matching. In [Arnold and Rahm, 2014b], a set of patterns of hypernym and meronym relations are applied to Wikipedia definitions in order to transfer those relations to concepts. Hearst patterns have been adopted in [van Hage et al., 2005] and [Vazquez and Swoboda, 2007], with the former using them to eliminate noise in matching results. In [Spiliopoulos et al., 2010], a supervised method learns patterns of subsumption evidences, while in [Beisswanger, 2010] the approach relies on free-text parts of Wikipedia and a dependency feature-based relation classifier in order to help detecting subsumption relations. In [Arnold and Rahm, 2014a], an enrichment strategy refines initial correspondences by combining a set of strategies (head modifier, background knowledge, itemization, explicit structure of ontologies, and multiple linkage of concepts) to establish is-a (subsumption), inverse is-a, part-of and inverse part-of relations. Complementary, in [Zhang et al., 2012], a learning-based relation extraction method requiring minimal supervision is based on a matching approach generating (complex) correspondences between the target relation and the background knowledge base. This is one of the few works exploring complex correspondences, defined using database join, union, project and select operators.

While early works have exploited machine learning in ontology matching, such as supervised machine learning [Mao et al., 2008] or multi-strategy learning [Doan et al., 2004], following the same trend as in relation extraction, emerging matching approaches exploit vector spaces and deep learning algorithms [Zhang et al., 2014, Kolyvakis et al., 2018]. While [Zhang et al., 2014] train word2vec vectors on Wikipedia, incorporating word embeddings into the computation of semantic similarities, [Nkisi-Orji et al., 2019] use a random forest classifier approach relying on word embeddings for determining the semantic similarities between concepts. [Kolyvakis et al., 2018] refine pre-trained word vectors aiming at deriving ontological entity descriptions (as done in [Xiang et al., 2015b]), that are further exploited to compute the entities' semantic distances using a variant of a document similarity metric. In [Chen et al., 2018], cross-lingual entity alignment is performed via co-training of two embedding models (a multilingual model and a multilingual literal description embedding model) on Wikipedia.

First of all, the studies on symbolic relation extraction focusing on specific relations, as hypernym and meronym, could be further exploited in ontology matching in particular in the matching scenarios described in this manuscript, combined to the 'classical' matching approaches. In fact, complex, holistic and foundational ontology matching are unexploited in the matching approaches described above. While specific kinds of relations are of special interest, open relation extraction could be also further exploited [Niklaus et al., 2018]. Another aspect to be taken into account is multilingualism. In that sense, our experiences in looking for hypernym relations in multilingual background knowledge resources (as BabelNet) show that these resources suffer from unbalanced knowledge representation in different languages. This can be the case, for instance, of exploiting the English lexical network to compensate the lack of expressed relations in Portuguese or French. As well as quality problems derived from the semi-automatic procedures used to integrate heterogeneous linguistic resources in several languages. Another direction could also be better combining symbolic NLP approaches to learning approaches in the matching task. While the trend seems to go in the learning direction, in line with the findings in [Roller et al., 2018], symbolic approaches should not be throw away.

The perspectives described in this section could be developed in a new research line in our MELODI team (Ontology Matching and Learning), which counts with the expertise of colleagues in both machine learning (Phillipe Muller, Tim Van de Cruys, Stergos Afantenos) and relation extraction (Mouna Kamel, Nathalie Aussenac-Gilles, Farah Benamara). Confronting symbolic and learning approaches, distant and unsupervised learning (together with the exploitation of multilingual resources) in the task of ontology matching seems like a promising research direction. This can be the beginning of a long journey.

9.2 Foundational distinctions in LOD datasets

Ontologies are in the core of a plethora of (complex) systems in a range of application domains, including the Semantic Web. While systematic studies of ontological representations are at the center of the formal and applied ontology field focusing on a large spectrum of foundational issues (types of entities, formal relations, space, time, etc.), most Linked Open Data still lack such ontological distinction, as recently stated in [Asprino et al., 2018]. This has been further corroborated in [Bennett and Baclawski, 2017], where it is stated that in the Semantic Web, there is an increasingly need for serious engagement with ontology, understood as a general theory of the types of entities and relations making up their respective domains of inquiry. However, there is still little interaction between the communities (as also discussed at the IAOA General Meeting at JOWO 2019), despite the fact that they share common ambitions in terms of knowledge understanding.

In fact, most (core) Linked Open datasets are constructed from existing (encyclopedic) databases such as DBpedia, the nucleus of the LOD cloud and one of the most exploited resources in the semantic web field. Some exceptions include YAGO and BabelNet which provide (incomplete) alignments to SUMO. In particular, the Linked Open Data in general could take better advantage of the knowledge from foundational ontologies. As discussed in Chapter 5, foundational ontologies can also act as bridges improving interoperability between domain ontologies. One of the few works analysing the foundational coverage of DBPedia is the one by [Paulheim and Gangemi, 2015], where correspondences between DBpedia ontology and DOLCE-Zero [Gangemi et al., 2003a], a module of DOLCE, are used to identify inconsistent statements in DBpedia. The authors focus on finding systematic errors or anti-patterns in DBpedia. They argued that by aligning these ontologies and by combining reasoning and clustering of the reasoning results, errors affecting statements can be identified with minimal human workload. More recently, as discussed in Chapter 5, the lack of ontological distinctions (class or instance, physical object or non physical object) has been addressed in [Asprino et al., 2018], using a matching approach. In line with this indirect matching approach, which relies on existing alignments, networks of alignments (and instance links) could be exploited in this task.

Moreover, several other foundational issues have to be taken into account such as formal relations (parthood, dependence, constitution, causality, instantiation), and space, time, and change. While the approaches above follow a bottom-up approach, the lack of foundational distinctions in the LOD could be taken from a methodological perspective, including studying the role of reference ontologies ([Menzel, 2003, Ruy et al., 2017]), formal (axiom-based) comparison of ontologies, and the relation between language, semantics and context, in the process of construction, alignment and publication of (six stars) LOD data.

The perspectives above could be further developed in the context of a collaborative effort between the semantic web and applied and foundational ontology communities. First, this topic could be developed in the context of European projects involving experts in both communities. Second, it could involve the organization of dedicated workshops at FOIS and ESWC/ISWC (Foundational and Applied Ontology meets Linked Open Data). At the local scale, developments in the topic could start by an ANR project (and a PhD thesis) involving the experts on foundational and applied ontologies in our team (Laure Vieu and Adrien Barton).

9.3 Alignment of multiple structured and unstructured sources

Cognitive systems must be able to understand, identify, and extract contextual elements such as meaning, syntax, time, and location and have to be designed to weigh information from multiple sources [Kelly III, 2015, Freund, 2017]. In this perspective, the ability to establish a relationship between different forms of expression of knowledge (from structured and unstructured sources) and its meaning or intent is crucial [Matuschek and Gurevych, 2013]. This scenario reflects a unifying framework of a wide range of solutions from a variety of domains, including NLP and semantic web. Several works in the literature are interested in establishing such ‘relationships’, from different angles. In fact, different variants of the notion of 'alignment' have been adopted in a range of areas, covering their specific needs and relying on a variety of techniques in their own area of research. It opens the perspective for reusing approaches from one context to another.

On the one hand, proposals focus on the alignment of structures of the same type, such as paraphrase study in texts [Bannard and Callison-Burch, 2005] and relational paraphrases [Grycner and Weikum, 2014], multilingual text alignment [Tufiş et al., 2004], database alignment [Cole et al., 2009] or ontology alignment. Ontology alignment is a special case of alignment of structures where the semantic aspect is strongest and the knowledge more formally expressed. Other approaches focus on the alignment of heterogeneous structures, involving annotation of text with ontologies [Erdmann et al., 2000], alignment of ontologies and thesaurus [Kless et al., 2012], alignment of dictionaries and ontologies [Dalvi et al., 2015], alignment between lexicons and dictionaries [Caselli et al., 2014], or still alignments between relational databases and ontologies [Hu and Qu, 2007, De Uña et al., 2018] (in line with what is done in OBDA [Xiao et al., 2018]).

These alignment approaches, however, take little account of the alignment of multiple structures. Regarding the first aspect, holistic approaches are rare. This type of approach is becoming increasingly necessary to manage the growing volume of unstructured information sources available on the Web (encyclopedias such as Wikipedia, social media data, etc.) and LOD knowledge bases. In addition, the approaches are mostly developed for the English language. These needs have to be addressed through a global vision of alignment that takes into account a multiplicity of structures in which knowledge can be expressed. Aligning together both ontologies and other structures such as text, relational databases, lexicons and formal ontologies, together with mechanisms to manage the evolution of these alignments could be an interesting direction. In this context, different phenomena have to be addressed, to cite a few (i) the resolution of linguistic phenomena such as polysemy [Jezek and Vieu, 2014]; (ii) taking into account the different levels of abstraction, expressivity, and granularity of the sources to be aligned in a holistic vision; (iii) consideration of logical coherence and source matching for integration; and (iv) the evolution of alignments given the potential evolution of the knowledge sources, in a holistic context [Hartung et al., 2013, Euzenat, 2016].

This task finds its relevance in different domains, such as offering support to the tasks of research and reading, including question answering systems. Reading text or searching for specialized knowledge can benefit from the inter-relationship between knowledge expressed in natural language and the knowledge expressed in dictionaries and ontologies. Thus, an answer may contain the use of a concept in the text, the formal expression in an ontology and its definition in a dictionary. In an interactive system, these options can for example be presented to the user, who can refine his request, to obtain a better solution. This semantic integration requires not only to align keywords or terms of the query, but also to solve the problems of ambiguity and context.

This is however a more long-term perspective.

Bibliography

[Abacha and Zweigenbaum, 2014] Abacha, A. B. and Zweigenbaum, P. (2014). Means: une approche sémantique pour la recherche de réponses aux questions médicales. *TAL*, 55(1):71–104. (Cited in page 20.)

[Achichi et al., 2017] Achichi, M., Cheatham, M., Dragisic, Z., Euzenat, J., Faria, D., Ferrara, A., Flouris, G., Fundulaki, I., Harrow, I., Ivanova, V., Jiménez-Ruiz, E., Kolthoff, K., Kuss, E., Lambrix, P., Leopold, H., Li, H., Meilicke, C., Mohammadi, M., Montanelli, S., Pesquita, C., Saveta, T., Shvaiko, P., Splendiani, A., Stuckenschmidt, H., Thiéblin, É., Todorov, K., Trojahn, C., and Zamazal, O. (2017). Results of the ontology alignment evaluation initiative 2017. In *OM@ISWC*, CEUR Workshop Proceedings, pages 61–113. (Cited in page 14.)

[Achichi et al., 2016] Achichi, M., Cheatham, M., Dragisic, Z., Euzenat, J., Faria, D., Ferrara, A., Flouris, G., Fundulaki, I., Harrow, I., Ivanova, V., Jiménez-Ruiz, E., Kuss, E., Lambrix, P., Leopold, H., Li, H., Meilicke, C., Montanelli, S., Pesquita, C., Saveta, T., Shvaiko, P., Splendiani, A., Stuckenschmidt, H., Todorov, K., Trojahn, C., and Zamazal, O. (2016). Results of the ontology alignment evaluation initiative 2016. In *OM@ISWC*, CEUR Workshop Proceedings, pages 73–129. (Cited in page 14.)

[Aguirre et al., 2012] Aguirre, J., Eckert, K., Euzenat, J., Ferrara, A., van Hage, W. R., Hollink, L., Meilicke, C., Nikolov, A., Ritze, D., Scharffe, F., Pavel Shvaiko, O. S.-Z., Trojahn, C., Jiménez-Ruiz, E., Grau, B. C., and Zapilko, B. (2012). Results of the ontology alignment evaluation initiative 2012. In *Proceedings of the 7th International Workshop on Ontology Matching, collocated with the 11th International Semantic Web Conference, Boston, MA USA*, CEUR Workshop Proceedings. (Cited in page 14.)

[Al-Bakri et al., 2016] Al-Bakri, M., Atencia, M., David, J., Lalande, S., and Rousset, M.-C. (2016). Uncertainty-sensitive reasoning for inferring sameAs facts in linked data. In *22nd european conference on artificial intelligence (ECAI)*, pages 698–706, Der Haague, Netherlands. IOS press. albakri2016a. (Cited in page 99.)

[Algergawy et al., 2018] Algergawy, A., Cheatham, M., Faria, D., Ferrara, A., Fundulaki, I., Harrow, I., Hertling, S., Jiménez-Ruiz, E., Karam, N., Khiat, A., Lambrix, P., Li, H., Montanelli, S., Paulheim, H., Pesquita, C., Saveta, T., Schmidt, D., Shvaiko, P., Splendiani, A., Thiéblin, É., Trojahn, C., Vatascinová, J., Zamazal, O., and Zhou, L. (2018). Results of the ontology alignment evaluation initiative 2018. In *OM@ISWC*, CEUR Workshop Proceedings, pages 76–116. (Cited in pages 14, 15, 72 and 83.)

[Arenas et al., 2016a] Arenas, H., Aussenac-Gilles, N., Comparot, C., and Trojahn, C. (2016a). Semantic integration of geospatial data from earth observations. In *EKAW (Satellite Events)*, volume 10180 of *Lecture Notes in Computer Science*, pages 97–100. Springer. (Cited in page 17.)

[Arenas et al., 2016b] Arenas, H., Aussenac-Gilles, N., Comparot, C., and Trojahn, C. (2016b). Visualisation d'événements enrichis par des images satellites et des données ouvertes : application aux données sismiques. In *Spatial Analysis and Geomatics 2016 (SAGEO 2016), Nice, France, December 6-9, 2016 (demo session)*. (Cited in page 17.)

[Arenas et al., 2018a] Arenas, H., Aussenac-Gilles, N., Comparot, C., and Trojahn, C. (2018a). Ontologie pour l'intégration de données d'observation de la terre et contextuelles basée sur les relations topologiques. In *IC*, pages 5–20. (Cited in page 17.)

[Arenas et al., 2018b] Arenas, H., Aussenac-Gilles, N., Comparot, C., and Trojahn, C. (2018b). Relations topologiques pour l'intégration sémantique de données et images d'observation de la Terre (regular paper). In *Congrès Informatique des Organisations et Systèmes d'Information et de Décision (INFORSID 2018)*, pages 63–78. (Cited in page 18.)

[Arenas et al., 2018c] Arenas, H., Trojahn, C., Comparot, C., and Aussenac-Gilles, N. (2018c). Un modèle pour l'intégration spatiale et temporelle de données géolocalisées. *Revue Internationale de Géomatique*, 28(2):243. (Cited in page 17.)

[Arnold and Rahm, 2014a] Arnold, P. and Rahm, E. (2014a). Enriching ontology mappings with semantic relations. *Data Knowl. Eng.*, 93:1–18. (Cited in page 100.)

[Arnold and Rahm, 2014b] Arnold, P. and Rahm, E. (2014b). Extracting semantic concept relations from wikipedia. In *Proceedings of the 4th International Conference on Web Intelligence, Mining and Semantics (WIMS14)*, WIMS '14, pages 26:1–26:11, New York, NY, USA. ACM. (Cited in page 100.)

[Arnold and Rahm, 2015] Arnold, P. and Rahm, E. (2015). Semrep: A repository for semantic mapping. In *Datenbanksysteme für Business, Technologie und Web (BTW), 16. Fachtagung des GI-Fachbereichs "Datenbanken und Informationssysteme" (DBIS), 4.-6.3.2015 in Hamburg, Germany. Proceedings*, pages 177–194. (Cited in page 98.)

[Arp et al., 2015] Arp, R., Smith, B., and Spear, A. (2015). *Building Ontologies with Basic Formal Ontology.* MIT Press. (Cited in page 59.)

[Asprino et al., 2018] Asprino, L., Basile, V., Ciancarini, P., and Presutti, V. (2018). Empirical analysis of foundational distinctions in linked open data. In

Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden., pages 3962–3969. (Cited in pages 58, 60, 61, 101 and 102.)

[Atencia et al., 2012] Atencia, M., Borgida, A., Euzenat, J., Ghidini, C., and Serafini, L. (2012). A formal semantics for weighted ontology mappings. In *The Semantic Web – ISWC 2012*, pages 17–33, Berlin, Heidelberg. Springer Berlin Heidelberg. (Cited in page 99.)

[Atencia et al., 2014] Atencia, M., David, J., and Euzenat, J. (2014). Data interlinking through robust linkkey extraction. In *ECAI*, pages 15–20. (Cited in page 50.)

[Auer et al., 2007] Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., and Ives, Z. (2007). Dbpedia: A nucleus for a web of open data. In *The semantic web*, pages 722–735. Springer. (Cited in page 100.)

[Bannard and Callison-Burch, 2005] Bannard, C. and Callison-Burch, C. (2005). Paraphrasing with bilingual parallel corpora. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pages 597–604. Association for Computational Linguistics. (Cited in page 103.)

[Beisswanger, 2010] Beisswanger, E. (2010). Exploiting relation extraction for ontology alignment. In *Proceedings of the International Semantic Web Conference*, pages 289–296. (Cited in pages 62 and 100.)

[Bennett and Baclawski, 2017] Bennett, M. and Baclawski, K. (2017). The role of ontologies in linked data, big data and semantic web applications. *Applied Ontology*, 12(3-4):189–194. (Cited in page 101.)

[Berland and Charniak, 1999] Berland, M. and Charniak, E. (1999). Finding parts in very large corpora. In *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics*, pages 57–64. Association for Computational Linguistics. (Cited in page 100.)

[Berners-Lee et al., 2001] Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The semantic web. *Scientific American*, 284(5):34–43. (Cited in page 8.)

[Berro et al., 2015] Berro, A., Megdiche, I., and Teste, O. (2015). A linear program for holistic matching: Assessment on schema matching benchmark. In *Proceedings of the International Conference on Database and Expert Systems Applications*, pages 383–398. (Cited in page 12.)

[Bienvenu, 2018] Bienvenu, M. (2018). Inconsistency-tolerant ontology-based data access revisited: Taking mappings into account. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*, pages 1721–1729. International Joint Conferences on Artificial Intelligence Organization. (Cited in page 99.)

[Brodaric and Probst, 2008] Brodaric, B. and Probst, F. (2008). DOLCE ROCKS: Integrating Geoscience Ontologies with DOLCE. In *Semantic Scientific Knowledge Integration*, pages 3–8. (Cited in page 59.)

[Bunescu and Mooney, 2005] Bunescu, R. C. and Mooney, R. J. (2005). A shortest path dependency kernel for relation extraction. In *Proceedings of the conference on human language technology and empirical methods in natural language processing*, pages 724–731. Association for Computational Linguistics. (Cited in page 100.)

[Calì et al., 2008] Calì, A., Lukasiewicz, T., Predoiu, L., and Stuckenschmidt, H. (2008). Tightly integrated probabilistic description logic programs for representing ontology mappings. In Hartmann, S. and Kern-Isberner, G., editors, *Foundations of Information and Knowledge Systems*, pages 178–198, Berlin, Heidelberg. Springer Berlin Heidelberg. (Cited in page 99.)

[Caracciolo et al., 2008] Caracciolo, C., Euzenat, J., Hollink, L., Ichise, R., Isaac, A., Malaisé, V., Meilicke, C., Pane, J., Shvaiko, P., Stuckenschmidt, H., Sváb-Zamazal, O., and Svátek, V. (2008). Results of the ontology alignment evaluation initiative 2008. In *OM*. (Cited in page 36.)

[Caselli et al., 2014] Caselli, T., Strapparava, C., Vieu, L., and Vetere, G. (2014). Aligning an italian wordnet with a lexicographic dictionary: Coping with limited data. In *Global WordNet Conference - GWC*, pages 290–298, Tartu, Estonia. (Cited in page 103.)

[Castro et al., 2004] Castro, R. G., Maynard, D., Foxvog, D., Wache, H., and González-Cabero, R. (2004). Specification of a methodology, general criteria, and benchmark suites for benchmarking ontology tools. Deliverable D2.1.4, Knowledge web NoE. (Cited in page 14.)

[Cheatham et al., 2015] Cheatham, M., Dragisic, Z., Euzenat, J., Faria, D., Ferrara, A., Flouris, G., Fundulaki, I., Granada, R., Ivanova, V., Jiménez-Ruiz, E., Lambrix, P., Montanelli, S., Pesquita, C., Saveta, T., Shvaiko, P., Solimando, A., Trojahn, C., and Zamazal, O. (2015). Results of the ontology alignment evaluation initiative 2015. In *Proceedings of the 10th International Workshop on Ontology Matching collocated with the 14th International Semantic Web Conference, Bethlehem, PA, USA*, pages 60–115. (Cited in pages 14 and 36.)

[Cheatham and Hitzler, 2014] Cheatham, M. and Hitzler, P. (2014). Conference v2. 0: An uncertain version of the oaei conference benchmark. In *International Semantic Web Conference*, pages 33–48. Springer. (Cited in pages 27, 30, 74 and 77.)

[Chen et al., 2018] Chen, M., Tian, Y., Chang, K.-W., Skiena, S., and Zaniolo, C. (2018). Co-training embeddings of knowledge graphs and entity descriptions for cross-lingual entity alignment. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, IJCAI'18, pages 3998–4004. AAAI Press. (Cited in page 100.)

[Chernov et al., 2006] Chernov, S., Iofciu, T., Nejdl, W., and Zhou, X. (2006). Extracting semantics relationships between wikipedia categories. *SemWiki*, 206. (Cited in page 100.)

[Cole et al., 2009] Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., Kulam-Syed-Mohideen, A., McGarrell, D. M., Marsh, T., Garrity, G. M., et al. (2009). The ribosomal database project: improved alignments and new tools for rrna analysis. *Nucleic acids research*, 37(suppl 1):D141–D145. (Cited in page 103.)

[Correndo et al., 2010] Correndo, G., Salvadores, M., Millard, I., Glaser, H., and Shadbolt, N. (2010). Sparql query rewriting for implementing data integration over linked data. In *Proceedings of the 2010 EDBT/ICDT Workshops*, pages 4:1–4:11, Lausanne, Switzerland. ACM. (Cited in page 39.)

[Correndo and Shadbolt, 2011] Correndo, G. and Shadbolt, N. (2011). Translating expressive ontology mappings into rewriting rules to implement query rewriting. In Shvaiko, P., Euzenat, J., Heath, T., Quix, C., Mao, M., and Cruz, I., editors, *Proceedings of the 6th International Workshop on Ontology Matching (OM-2011) in conjunction with the International Semantic Web Conference (ISWC2011)*, volume 814, Bonn, Germany. ceur-ws.org. (Cited in pages 39 and 47.)

[Cruz et al., 2009] Cruz, I., Antonelli, F., and Stroe, C. (2009). AgreementMaker: efficient matching for large real-world schemas and ontologies. In *Proceedings of the International Conference on Very Large Databases*, pages 1586–1589. (Cited in page 43.)

[Cui et al., 2017] Cui, M., Li, L., Wang, Z., and You, M. (2017). A survey on relation extraction. In Li, J., Zhou, M., Qi, G., Lao, N., Ruan, T., and Du, J., editors, *Knowledge Graph and Semantic Computing. Language, Knowledge, and Intelligence*, pages 50–58, Singapore. Springer Singapore. (Cited in page 100.)

[Dalvi et al., 2015] Dalvi, B., Minkov, E., Talukdar, P. P., and Cohen, W. W. (2015). Automatic gloss finding for a knowledge base using ontological constraints. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, pages 369–378. ACM. (Cited in page 103.)

[Damova et al., 2010] Damova, M., Kiryakov, A., Simov, K. I., and Petrov, S. (2010). Mapping the central LOD ontologies to PROTON upper-level ontology. In *Workshop on Ontology Matching*. (Cited in pages 59, 60 and 69.)

[David, 2011] David, J. (2011). AROMA results for OAEI 2011. In *Proc. of the 6th Intern. Workshop on Ontology Matching*, pages 122–125. (Cited in page 61.)

[David et al., 2011] David, J., Euzenat, J., Scharffe, F., and Trojahn, C. (2011). The Alignment API 4.0. *Semantic Web Journal: Semantic Web Tools and Systems*, 2(1):3–10. (Cited in pages 14, 22 and 39.)

[de Boer et al., 2012] de Boer, V., Wielemaker, J., van Gent, J., Hildebrand, M., Isaac, A., van Ossenbruggen, J., and Schreiber, G. (2012). Supporting linked data production for cultural heritage institutes: The amsterdam museum case study. In *The Semantic Web: Research and Applications - 9th Extended Semantic Web Conference, ESWC 2012, Heraklion, Crete, Greece, May 27-31, 2012. Proceedings*, volume 7295 of *Lecture Notes in Computer Science*, pages 733–747. Springer. (Cited in page 38.)

[De Bruijn et al., 2006] De Bruijn, J., Ehrig, M., Feier, C., Martín-Recuerda, F., Scharffe, F., and Weiten, M. (2006). Ontology mediation, merging and aligning. *Semantic web technologies*, pages 95–113. (Cited in pages 9 and 41.)

[De Uña et al., 2018] De Uña, D., Rümmele, N., Gange, G., Schachte, P., and Stuckey, P. J. (2018). Machine learning and constraint programming for relational-to-ontology schema mapping. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, IJCAI'18, pages 1277–1283. AAAI Press. (Cited in page 103.)

[Dekhtyar and Hayes, 2006] Dekhtyar, A. and Hayes, J. H. (2006). Good benchmarks are hard to find: Toward the benchmark for information retrieval applications in software engineering. In *Proceedings of the 22th International Conference on Software Maintenance.* (Cited in page 92.)

[Dhamankar et al., 2004] Dhamankar, R., Lee, Y., Doan, A., Halevy, A. Y., and Domingos, P. M. (2004). imap: Discovering complex mappings between database schemas. In Weikum, G., König, A. C., and Deßloch, S., editors, *Proceedings of the ACM SIGMOD International Conference on Management of Data, Paris, France, June 13-18, 2004*, pages 383–394. ACM. (Cited in page 38.)

[Do et al., 2002] Do, H.-H., Melnik, S., and Rahm, E. (2002). Comparison of schema matching evaluations. In *Net. ObjectDays: International Conference on Object-Oriented and Internet-Based Technologies, Concepts, and Applications for a Networked World*, pages 221–237. Springer. (Cited in pages 10 and 20.)

[Doan and Halevy, 2005] Doan, A. and Halevy, A. Y. (2005). Semantic integration research in the database community: A brief survey. *AI Magazine*, 26(1):83–94. (Cited in page 9.)

[Doan et al., 2004] Doan, A., Madhavan, J., Domingos, P., and Halevy, A. (2004). *Ontology Matching: A Machine Learning Approach*, pages 385–403. Springer Berlin Heidelberg, Berlin, Heidelberg. (Cited in page 100.)

[Dorne et al., 2018] Dorne, J., Aussenac-Gilles, N., Comparot, C., Hugues, R., Planès, J.-G., and Trojahn, C. (2018). Une approche sémantique pour représenter l'indice de végétation d'images Sentinel-2 et son évolution (short paper). In Teisseire, M. and Roche, M., editors, *Spatial Analytics and GEOmatics (SAGEO 2018), Montpellier, 06/11/2018-09/11/2018*, pages 49–54, http://www.cnrs.fr. CNRS. (Cited in page 18.)

[Dragisic et al., 2014] Dragisic, Z., Eckert, K., Euzenat, J., Faria, D., Ferrara, A., Granada, R., Ivanova, V., Jiménez-Ruiz, E., Kempf, A. O., Lambrix, P., Montanelli, S., Paulheim, H., Ritze, D., Shvaiko, P., Solimando, A., Trojahn, C., Zamazal, O., and Grau, B. C. (2014). Results of the ontology alignment evaluation initiative 2014. In *Proceedings of the 9th International Workshop on Ontology Matching collocated with the 13th International Semantic Web Conference, Riva del Garda, Trentino, Italy*, pages 61–104. (Cited in page 14.)

[Dragisic et al., 2016] Dragisic, Z., Ivanova, V., Lambrix, P., Faria, D., Jiménez-Ruiz, E., and Pesquita, C. (2016). User validation in ontology alignment. In *The Semantic Web - ISWC 2016 - 15th International Semantic Web Conference, Kobe, Japan, October 17-21, 2016, Proceedings, Part I*, pages 200–217. (Cited in pages 9, 15, 36 and 43.)

[Ehrig, 2006] Ehrig, M. (2006). *Ontology alignment: bridging the semantic gap.* PhD thesis, Universität Fridericiana zu Karlsruhe, Karlsruhe (DE). (Cited in page 35.)

[Ehrig and Euzenat, 2005] Ehrig, M. and Euzenat, J. (2005). Relaxed precision and recall for ontology matching. In Ashpole, B., Ehrig, M., Euzenat, J., and Stuckenschmidt, H., editors, *Proceedings of the Workshop on Integrating Ontologies*, volume 156, page 8. CEUR-WS.org. (Cited in pages 30 and 31.)

[Erdmann et al., 2000] Erdmann, M., Maedche, A., Schnurr, H.-P., and Staab, S. (2000). From manual to semi-automatic semantic annotation: About ontology-based text annotation tools. In *Proceedings of the COLING-2000 Workshop on Semantic Annotation and Intelligent Content*, pages 79–85. Association for Computational Linguistics. (Cited in page 103.)

[Euzenat, 2003] Euzenat, J. (2003). Towards composing and benchmarking ontology alignments. In *Proceedings of the Semantic Integration Workshop Collocated with the Second International Semantic Web Conference (ISWC-03), Sanibel Island, Florida, USA, 2003*, volume 82, pages 165–166. CEUR-WS.org. (Cited in page 22.)

[Euzenat, 2007] Euzenat, J. (2007). Semantic precision and recall for ontology alignment evaluation. In *Proc. 20th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 248–253, Hyderabad (IN). (Cited in page 32.)

[Euzenat, 2016] Euzenat, J. (2016). Knowledge change, failure, adaptation, and evolution. In *Proceedings of the Joint Ontology Workshops 2016 Episode 2: The French Summer of Ontology co-located with the 9th International Conference on Formal Ontology in Information Systems (FOIS 2016), Annecy, France, July 6-9, 2016.* (Cited in pages 99 and 103.)

[Euzenat et al., 2009] Euzenat, J., Ferrara, A., Hollink, L., Isaac, A., Joslyn, C., Malaisé, V., Meilicke, C., Nikolov, A., Pane, J., Sabou, M., Scharffe, F., Shvaiko, P., Spiliopoulos, V., Stuckenschmidt, H., Sváb-Zamazal, O., Svátek, V., Trojahn, C., Vouros, G. A., and Wang, S. (2009). Results of the ontology alignment evaluation initiative 2009. In *Proceedings of the 4th International Workshop on Ontology Matching (OM-2009) collocated with the 8th International Semantic*

Web Conference (ISWC-2009) Chantilly, USA, October 25, 2009, volume 551. (Cited in pages 14 and 28.)

[Euzenat et al., 2010] Euzenat, J., Ferrara, A., Meilicke, C., Pane, J., Scharffe, F., Shvaiko, P., Stuckenschmidt, H., Sváb-Zamazal, O., Svátek, V., and Trojahn, C. (2010). Results of the ontology alignment evaluation initiative 2010. In *Proceedings of the 5th International Workshop on Ontology Matching collocated with the 9th International Semantic Web Conference, Shanghai, China*. (Cited in page 14.)

[Euzenat et al., 2011a] Euzenat, J., Ferrara, A., van Hage, W. R., Hollink, L., Meilicke, C., Nikolov, A., Scharffe, F., Shvaiko, P., Stuckenschmidt, H., and Trojahn, O. S.-Z. C. (2011a). Results of the ontology alignment evaluation initiative 2011. In *Proceedings of the Sixth International Workshop on Ontology Matching (OM-2011) collocated with the 10th International Semantic Web Conference (ISWC-2011) Bonn, Germany*, CEUR Workshop Proceedings. CEUR-WS.org. (Cited in page 14.)

[Euzenat et al., 2003] Euzenat, J., Layaïda, N., and Dias, V. (2003). A semantic framework for multimedia document adaptation. In *Proc. 18th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 31–36, Acapulco (MX). (Cited in page 31.)

[Euzenat et al., 2011b] Euzenat, J., Meilicke, C., Stuckenschmidt, H., Shvaiko, P., and Trojahn, C. (2011b). Ontology Alignment Evaluation Initiative: six years of experience. *Journal on Data Semantic (JoDS)*, XV:158–192. (Cited in pages 10, 14, 25 and 26.)

[Euzenat et al., 2008] Euzenat, J., Polleres, A., and Scharffe, F. (2008). Processing ontology alignments with SPARQL. In *International Conference on Complex, Intelligent and Software Intensive Systems*, pages 913–917. (Cited in pages 39 and 47.)

[Euzenat et al., 2013] Euzenat, J., Rosoiu, M., and Trojahn, C. (2013). Ontology matching benchmarks: generation, stability, and discriminability. *Web Semantics: Science, Services and Agents on the World Wide Web*, 20:30–48. (Cited in page 14.)

[Euzenat and Shvaiko, 2007] Euzenat, J. and Shvaiko, P. (2007). *Ontology matching*. Springer, Heidelberg (DE). (Cited in pages 23 and 24.)

[Euzenat and Shvaiko, 2013] Euzenat, J. and Shvaiko, P. (2013). *Ontology Matching, Second Edition*. Springer, 2nd edition. (Cited in pages 9, 20, 21, 25, 26, 35 and 41.)

[Euzenat and Valtchev, 2004] Euzenat, J. and Valtchev, P. (2004). Similarity-based ontology alignment in OWL-Lite. In *Proceedings of the 16th European Conference on Artificial Intelligence*, pages 333–337. (Cited in page 52.)

[Fabre et al., 2014] Fabre, C., Hathout, N., Ho-Dac, L.-M., Morlane-Hondère, F., Muller, P., Sajous, F., Tanguy, L., and Van De Cruys, T. (2014). Présentation de l'atelier SemDis 2014 : sémantique distributionnelle pour la substitution lexicale et l'exploration de corpus spécialisés. In *21e Conférence sur le Traitement Automatique des Langues Naturelles (TALN 2014)*, pages 196–205, Marseille, France. (Cited in page 100.)

[Fader et al., 2011] Fader, A., Soderland, S., and Etzioni, O. (2011). Identifying relations for open information extraction. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1535–1545. Association for Computational Linguistics. (Cited in page 100.)

[Faria et al., 2018] Faria, D., Pesquita, C., Balasubramani, B. S., Tervo, T., Carriço, D., Garrilha, R., Couto, F. M., and Cruz, I. F. (2018). Results of aml participation in oaei 2018. In *Proceedings of the 13th International Workshop on Ontology Matching co-located with the 17th International Semantic Web Conference, OM@ISWC 2018, Monterey, CA, USA, October 8, 2018*, pages 125–131. (Cited in page 38.)

[Flouris et al., 2006] Flouris, G., Huang, Z., Pan, J., Plexousakis, D., and Wache, H. (2006). Inconsistencies, negations and changes in ontologies. In *Proceedings of AAAI Conference on Artificial Intelligence 2006 (18).* (Cited in page 34.)

[Freund, 2017] Freund, R. J. (2017). Cognitive computing and managing complexity in open innovation model. In *Managing Complexity.* Bellemare J., Carrier S., Nielsen K., Piller F(eds).Springer. (Cited in page 102.)

[Fu et al., 2009] Fu, B., Brennan, R., and O'Sullivan, D. (2009). Cross-lingual Ontology Mapping - An Investigation of the Impact of Machine Translation. In *Proceedings of the 4th Annual Asian Semantic Web Conference (ASWC 2009)*, pages 1–15. (Cited in page 13.)

[Fung and Xu, 2012] Fung, K. W. and Xu, J. (2012). Synergism between the mapping projects from snomed ct to icd-10 and icd-10-cm. In *AMIA 2012, American Medical Informatics Association Annual Symposium, Chicago, Illinois, USA, November 3-7, 2012.* American Medical Informatics Association. (Cited in page 38.)

[Fürst and Trichet, 2005] Fürst, F. and Trichet, F. (2005). Axiom-based ontology matching. In *Proceedings of the 3rd International Conference on Knowledge Capture*, K-CAP '05, pages 195–196, New York, NY, USA. ACM. (Cited in page 69.)

[Gangemi et al., 2003a] Gangemi, A., Guarino, N., Masolo, C., and Oltramari, A. (2003a). Sweetening WORDNET with DOLCE. *AI Magazine*, 24(3):13–24. (Cited in page 102.)

[Gangemi et al., 2002] Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., and Schneider, L. (2002). Sweetening Ontologies with DOLCE. In *13th Conference on Knowledge Engineering and Knowledge Management*, pages 166–181. (Cited in page 59.)

[Gangemi et al., 2003b] Gangemi, A., Navigli, R., and Velardi, P. (2003b). *The OntoWordNet Project: Extension and Axiomatization of Conceptual Relations in WordNet*, pages 820–838. (Cited in pages 60 and 63.)

[Gemechu et al., 2016] Gemechu, D. T., Zock, M., and Abate, S. T. (2016). Combining syntactic patterns and wikipedia's hierarchy of hyperlinks to extract meronym relations. In *SRW@HLT-NAACL*. (Cited in page 100.)

[Giannangelo and Millar, 2012] Giannangelo, K. and Millar, J. (2012). Mapping snomed ct to icd-10. In *Quality of Life through Quality of Information - Proceedings of MIE2012, The XXIVth International Congress of the European Federation for Medical Informatics, Pisa, Italy, August 26-29, 2012*, volume 180 of *Studies in Health Technology and Informatics*, pages 83–87. IOS Press. (Cited in page 38.)

[Gillet et al., 2013] Gillet, P., Trojahn, C., Haemmerlé, O., and Pradel, C. (2013). Complex correspondences for query patterns rewriting. In *Proceedings of the 8th International Workshop on Ontology Matching (OM-2013) collocated with the 12th International Semantic Web Conference (ISWC-2013) Sydney, Australia, October 21st*. (Cited in pages 11 and 40.)

[Giunchiglia et al., 2009] Giunchiglia, F., Paolo, M. Y., Avesani, and Shvaiko, P. (2009). A large scale dataset for the evaluation of ontology matching systems. *The Knowledge Engineering Review Journal (KER)*, 24(2):137–157. (Cited in page 25.)

[Giunchiglia et al., 2007] Giunchiglia, F., Yatskevich, M., and Shvaiko, P. (2007). Semantic matching: Algorithms and implementation. *Journal on Data Semantics*, IX:1–38. (Cited in page 27.)

[Granada et al., 2012] Granada, R., Lopes, L., Ramisch, C., Trojahn, C., Vieira, R., and Villavicencio, A. (2012). A comparable corpus based on aligned multilingual ontologies. In *Proceedings of the First Workshop on Multilingual Modeling*, MM '12, pages 25–31, Stroudsburg, PA, USA. Association for Computational Linguistics. (Cited in page 14.)

[Grau et al., 2013] Grau, B. C., Dragisic, Z., Eckert, K., Euzenat, J., Ferrara, A., Granada, R., Ivanova, V., Jimenez-Ruiz, E., Kempf, A. O., Lambrix, P., Meilicke, C., Nikolov, A., Paulheim, H., Ritze, D., Scharffe, F., Shvaiko, P., Trojahn, C., and Zamazal, O. (2013). Results of the ontology alignment evaluation initiative 2013. In *Proceedings of the 8th International Workshop on Ontology Matching (OM-2013) collocated with the 12th International Semantic Web Conference (ISWC-2013) Sydney, Australia, October 21st.* (Cited in pages 14 and 36.)

[Grenon, 2003] Grenon, P. (2003). *BFO in a Nutshell: A Bi-categorial Axiomatization of BFO and Comparison with DOLCE.* Leipzig. (Cited in page 60.)

[Grenon et al., 2004] Grenon, P., Smith, B., and Goldberg, L. (2004). Biodynamic Ontology: Applying BFO in the Biomedical Domain. In *Stud. Health Technol. Inform*, pages 20–38. IOS Press. (Cited in page 59.)

[Gruetze et al., 2012] Gruetze, T., Böhm, C., and Naumann, F. (2012). Holistic and scalable ontology alignment for linked open data. In *Proceedings of the 5th Linked Data on the Web Workshop at the 21th WWW.* (Cited in page 52.)

[Grüninger and Fox, 1995] Grüninger, M. and Fox, M. S. (1995). Methodology for the design and evaluation of ontologies. international joint conference on artificial inteligence. In *Workshop on Basic Ontological Issues in Knowledge Sharing*, volume 15, page 34. (Cited in pages 11 and 43.)

[Grycner and Weikum, 2014] Grycner, A. and Weikum, G. (2014). HARPY: Hypernyms and alignment of relational paraphrases. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 2195–2204, Dublin, Ireland. Dublin City University and Association for Computational Linguistics. (Cited in page 103.)

[Guarino, 1998] Guarino, N. (1998). *Formal Ontology in Information Systems: Proceedings of the 1st Intern. Conference.* IOS Press, 1st edition. (Cited in pages 20 and 58.)

[Halpin et al., 2010] Halpin, H., Hayes, P. J., McCusker, J. P., McGuinness, D. L., and Thompson, H. S. (2010). When owl:sameas isn't the same: An analysis of identity in linked data. In *Proceedings of the 9th International Semantic Web Conference*, pages 305–320. (Cited in page 28.)

[Hartung et al., 2013] Hartung, M., Groí, A., and Rahm, E. (2013). Conto-diff: Generation of complex evolution mappings for life science ontologies. *J. of Biomedical Informatics*, 46(1):15–32. (Cited in pages 99 and 103.)

[He et al., 2004] He, B., Chang, K. C., and Han, J. (2004). Discovering complex matchings across web query interfaces: a correlation mining approach. In Kim, W., Kohavi, R., Gehrke, J., and DuMouchel, W., editors, *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, Washington, USA, August 22-25, 2004*, pages 148–157. ACM. (Cited in pages 38 and 52.)

[He and Chang, 2003] He, B. and Chang, K. C.-C. (2003). Statistical schema matching across web query interfaces. In *Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data*, SIGMOD '03, pages 217–228. ACM. (Cited in page 52.)

[Hearst, 1992] Hearst, M. A. (1992). Automatic acquisition of hyponyms from large text corpora. In *Proceedings of the 14th Conference on Computational Linguistics*, pages 539–545. (Cited in page 62.)

[Hecht et al., 2015] Hecht, T., Buche, P., Dibie, J., Ibanescu, L., and dos Santos, C. T. (2015). Ontology alignment using web linked ontologies as background knowledge. In *Advances in Knowledge Discovery and Management - Volume 6 [Best of EGC 2014, Rennes, France / Best of EGC 2015, Luxembourg].*, pages 207–227. (Cited in page 9.)

[Hollink et al., 2008] Hollink, L., Van Assem, M., Wang, S., Isaac, A., and Schreiber, G. (2008). Two Variations on Ontology Alignment Evaluation: Methodological Issues. In *5th European Semantic Web Conference*, pages 388–401. (Cited in page 20.)

[Horrocks et al., 2006] Horrocks, I., Kutz, O., and Sattler, U. (2006). The even more irresistible sroiq. *Kr*, 6:57–67. (Cited in page 73.)

[Hu et al., 2011] Hu, W., Chen, J., Zhang, H., and Qu, Y. (2011). Learning complex mappings between ontologies. In *Joint International Semantic Technology Conference*. Springer. (Cited in page 43.)

[Hu et al., 2005] Hu, W., Jian, N., Qu, Y., and Wang, Y. (2005). Gmo: A graph matching for ontologies. In *K-Cap 2005 Workshop on Integrating Ontologies 2005*, pages 43–50. (Cited in page 52.)

[Hu and Qu, 2007] Hu, W. and Qu, Y. (2007). Discovering simple mappings between relational database schemas and ontologies. In *The Semantic Web*, pages 225–238, Berlin, Heidelberg. Springer Berlin Heidelberg. (Cited in page 103.)

[Isaac et al., 2008] Isaac, A., Matthezing, H., van der Meij, L., Schlobach, S., Wang, S., and Zinn, C. (2008). Putting Ontology Alignment in Context: Usage Scenarios, Deployment and Evaluation in a Library Case. In *5th European Semantic Web Conference*, pages 402–417. (Cited in pages 20 and 35.)

[Jain et al., 2011] Jain, P., Yeh, P., Verma, K., Vasquez, R., Damova, M., Hitzler, P., and Sheth, A. (2011). Contextual Ontology Alignment of LOD with an Upper Ontology: A Case Study with PROTON. In *Proceedings of the 8th Extended Semantic Web Conference*, pages 80–92. (Cited in pages 59 and 61.)

[Janowicz et al., 2012] Janowicz, K., Scheider, S., Pehle, T., and Hart, G. (2012). Geospatial Semantics and Linked Spatiotemporal Data – Past , Present , and Future. *Semantic Web*, (3):321–332. (Cited in page 17.)

[Jean-Mary et al., 2009] Jean-Mary, Y., Shironoshita, E., and Kabuka, M. (2009). Ontology matching with semantic verification. *Web Semantics: Science, Services and Agents on the World Wide Web*, 7(3). (Cited in page 52.)

[Jezek, 2019] Jezek, E. (2019). Sweetening ontologies cont'd: Aligning bottom-up with top-down ontologies. In *Proceedings of the Contextual Representation of Events and Objects in Language Workshop (CREOL), co-located with the Joint Ontology Workshops, JOWO-2019, Graz, Austria.* (Cited in page 59.)

[Jezek et al., 2014] Jezek, E., Magnini, B., Feltracco, A., Bianchini, A., and Popescu, O. (2014). T-pas; A resource of typed predicate argument structures for linguistic analysis and semantic processing. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation, LREC 2014, Reykjavik, Iceland, May 26-31, 2014.*, pages 890–895. (Cited in page 59.)

[Jezek and Vieu, 2014] Jezek, E. and Vieu, L. (2014). Distributional analysis of copredication: towards distinguishing systematic polysemy from coercion. In *First Italian Conference on Computational Linguistics (CLiC-it)*, volume 1, pages pp–219. (Cited in page 103.)

[Jiang et al., 2016] Jiang, S., Lowd, D., Kafle, S., and Dou, D. (2016). Ontology matching with knowledge rules. In *Transactions on Large-Scale Data-and Knowledge-Centered Systems XXVIII*, volume 28, pages 75–95. Springer. (Cited in pages 38 and 70.)

[Jiménez-Ruiz and Grau, 2011] Jiménez-Ruiz, E. and Grau, B. C. (2011). Logmap: Logic-based and scalable ontology matching. In *Proceedings of the 10th International Semantic Web Conference*, pages 273–288. (Cited in pages 9 and 52.)

[Jiménez-Ruiz et al., 2011] Jiménez-Ruiz, E., Grau, B. C., Horrocks, I., and Llavori, R. B. (2011). Logic-based assessment of the compatibility of UMLS ontology sources. *J. Biomedical Semantics*, 2(S-1):S2. (Cited in page 34.)

[Jiménez-Ruiz et al., 2012] Jiménez-Ruiz, E., Grau, B. C., Zhou, Y., and Horrocks, I. (2012). Large-scale interactive ontology matching: Algorithms and implementation. In *Proceedings of the 20th European Conference on Artificial Intelligence*, ECAI'12, pages 444–449, Amsterdam, The Netherlands, The Netherlands. IOS Press. (Cited in page 9.)

[Jouhet et al., 2017] Jouhet, V., Mougin, F., Bréchat, B., and Thiessard, F. (2017). Building a model for disease classification integration in oncology, an approach based on the national cancer institute thesaurus. *Journal of Biomedical Semantics*, 8(1):6:1–6:12. (Cited in page 38.)

[Jung et al., 2009] Jung, J. J., Håkansson, A., and Hartung, R. (2009). Indirect Alignment between Multilingual Ontologies: A Case Study of Korean and Swedish Ontologies. In *Agent and Multi-Agent Systems: Technologies and Applications: 3rd KES Intern. Symposium*, pages 233–241. (Cited in pages 9 and 61.)

[Kachroudi et al., 2015] Kachroudi, M., Chelbi, A., Souid, H., and Ben Yahia, S. (2015). Fiona: A framework for indirect ontology alignment. In *Foundations of Intelligent Systems*, pages 224–229. (Cited in page 61.)

[Kalfoglou and Schorlemmer, 2003] Kalfoglou, Y. and Schorlemmer, M. (2003). Ontology mapping: the state of the art. *The Knowledge Engineering Review*, 18(1):1–31. (Cited in pages 9 and 41.)

[Kamel and Aussenac-Gilles, 2009] Kamel, M. and Aussenac-Gilles, N. (2009). How can document structure improve ontology learning? In *Workshop on Semantic Annotation and Knowledge Markup collocated with K-CAP*. (Cited in page 100.)

[Kamel et al., 2019] Kamel, M., Schmidt, D., Trojahn, C., and Vieira, R. (2019). Hypernym relation extraction for establishing subsumptions between foundational ontologies. In *Proceedings of the 14th International Workshop on Ontology Matching*. (Cited in pages 61 and 62.)

[Kamel and Trojahn, 2016] Kamel, M. and Trojahn, C. (2016). Exploiter la structure discursive du texte pour valider les relations candidates d'hyperonymie issues de structures énumératives parallèles. In *IC 2016 : 27es Journées francophones d'Ingénierie des Connaissances (Proceedings of the 27th French Knowledge Engineering Conference), Montpellier, France, June 6-10, 2016.*, pages 111–122. (Cited in page 16.)

[Kamel and Trojahn, 2018] Kamel, M. and Trojahn, C. (2018). Towards enriching dbpedia from vertical enumerative structures using a distant learning approach. In *EKAW*, volume 11313 of *Lecture Notes in Computer Science*, pages 179–194. Springer. (Cited in page 16.)

[Kamel et al., 2017a] Kamel, M., Trojahn, C., Ghamnia, A., Aussenac-Gilles, N., and Fabre, C. (2017a). A distant learning approach for extracting hypernym relations from wikipedia disambiguation pages. In *KES*, volume 112 of *Procedia Computer Science*, pages 1764–1773. Elsevier. (Cited in page 16.)

[Kamel et al., 2017b] Kamel, M., Trojahn, C., Ghamnia, A., Aussenac-Gilles, N., and Fabre, C. (2017b). Extracting hypernym relations from wikipedia disambiguation pages : comparing symbolic and machine learning approaches. In *IWCS(1)*. The Association for Computer Linguistics. (Cited in page 16.)

[Keet, 2011] Keet, C. (2011). The use of foundational ontologies in ontology development: An empirical assessment. In *Proceedings of the 8th Extended Semantic Web Conference*, pages 321–335. (Cited in pages 58 and 92.)

[Kelly III, 2015] Kelly III, D. J. (2015). Computing, cognition and the future of knowing: How humans and machines are forging a new age of understanding. Technical report, IBM. (Cited in page 102.)

[Khan and Keet, 2012] Khan, Z. and Keet, C. (2012). ONSET: Automated Foundational Ontology Selection and Explanation. In *Proceedings of the 18th International Conference on Knowledge Engineering and Knowledge Management*, pages 237–251. (Cited in page 58.)

[Khan and Keet, 2013a] Khan, Z. and Keet, C. (2013a). Addressing issues in foundational ontology mediation. In *Proceedings of the Inter. Conference on Knowledge Engineering and Ontology Development*, pages 5–16. (Cited in pages 60 and 61.)

[Khan and Keet, 2013b] Khan, Z. and Keet, C. (2013b). The Foundational Ontology Library ROMULUS. In *Proceedings of the 3rd International Conference on Model and Data Engineering*, pages 200–211. (Cited in pages 60 and 61.)

[Klai et al., 2016] Klai, S., Zimmermann, A., and Khadir, M. T. (2016). Multilevel networked knowledge: representation and dl-reasoning. *IJMSO*, 11(1):1–15. (Cited in page 90.)

[Klein, 2001] Klein, M. (2001). Combining and relating ontologies: an analysis of problems and solutions. In Gómez-Pérez, A., Gruninger, M., Stuckenschmidt, H., and Uschold, M., editors, *Proceedings of the IJCAI-01 Workshop on Ontologies and Information Sharing Seattle, USA, August 4-5, 2001.*, volume 47 of *CEUR Workshop Proceedings*. CEUR-WS.org. (Cited in page 20.)

[Kless et al., 2012] Kless, D., Jansen, L., Lindenthal, J., and Wiebensohn, J. (2012). A method for re-engineering a thesaurus into an ontology. In *FOIS*, pages 133–146. (Cited in page 103.)

[Kolas et al., 2013] Kolas, D., Perry, M., and Herring, J. (2013). Getting started with GeoSPARQL. Technical report, OGC. (Cited in page 18.)

[Kolyvakis et al., 2018] Kolyvakis, P., Kalousis, A., and Kiritsis, D. (2018). DeepAlignment: Unsupervised ontology matching with refined word vectors. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 787–798, New Orleans, Louisiana. Association for Computational Linguistics. (Cited in pages 9, 61 and 100.)

[Kozierkiewicz and Pietranik, 2019] Kozierkiewicz, A. and Pietranik, M. (2019). Updating ontology alignment on the concept level based on ontology evolution. In Welzer, T., Eder, J., Podgorelec, V., and Kamišalić Latifić, A., editors, *Advances in Databases and Information Systems*, pages 201–214, Cham. Springer International Publishing. (Cited in page 99.)

[Lee et al., 2007] Lee, Y., Sayyadian, M., Doan, A., and Rosenthal, A. S. (2007). etuner: tuning schema matching software using synthetic scenarios. *The VLDB Journal*, 16(1):97–122. (Cited in page 9.)

[Lenci and Benotto, 2012] Lenci, A. and Benotto, G. (2012). Identifying hypernyms in distributional semantic spaces. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation*, pages 75–79. Association for Computational Linguistics. (Cited in page 100.)

[Lesk, 1986] Lesk, M. (1986). Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone. In *Proc. of the 5th Intern. Conf. on Systems Documentation*, pages 24–26. (Cited in pages 61 and 63.)

[Levenshtein, 1965] Levenshtein, V. (1965). Binary codes capable of correcting deletions, insertions, and reversals. *Doklady akademii nauk SSSR*, 163(4):845–848. In Russian. English Translation in Soviet Physics Doklady, 10(8) p. 707–710, 1966. (Cited in page 32.)

[Lin and Sandkuhl, 2008] Lin, F. and Sandkuhl, K. (2008). A Survey of Exploiting WordNet in Ontology Matching. In *Proc. of the 20th Intern. Federation for Information Processing*, pages 341–350. (Cited in pages 9 and 61.)

[Lin et al., 2016] Lin, Y., Shen, S., Liu, Z., Luan, H., and Sun, M. (2016). Neural relation extraction with selective attention over instances. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2124–2133, Berlin, Germany. Association for Computational Linguistics. (Cited in page 100.)

[Lopez et al., 2006] Lopez, V., Sabou, M., and Motta, E. (2006). Powermap: mapping the real semantic web on the fly. In *International Semantic Web Conference*, pages 414–427. Springer. (Cited in pages 24, 39 and 43.)

[Maedche et al., 2002] Maedche, A., Motik, B., Silva, N., and Volz, R. (2002). MAFRA - A mapping framework for distributed ontologies. In Gómez-Pérez, A. and Benjamins, V. R., editors, *Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web, 13th International Conference, EKAW 2002, Siguenza, Spain, October 1-4, 2002, Proceedings*, volume 2473 of *Lecture Notes in Computer Science*, pages 235–250. Springer. (Cited in page 38.)

[Makris et al., 2012] Makris, K., Bikakis, N., Gioldasis, N., and Christodoulakis, S. (2012). SPARQL-RW: transparent query access over mapped RDF data sources. In *Proceedings of the 15th International Conference on Extending Database Technology - EDBT'12*, pages 610–613, Berlin, Germany. ACM Press. (Cited in page 39.)

[Makris et al., 2010] Makris, K., Gioldasis, N., Bikakis, N., and Christodoulakis, S. (2010). Ontology Mapping and SPARQL Rewriting for Querying Federated RDF Data Sources. In Meersman, R., Dillon, T., and Herrero, P., editors, *On the Move to Meaningful Internet Systems, OTM 2010*, volume 6427 of *LNCS*, pages 1108–1117, Hersonssisos, Crete, Greece. Springer. (Cited in pages 39 and 47.)

[Malaisé et al., 2004] Malaisé, V., Zweigenbaum, P., and Bachimont, B. (2004). Detecting semantic relations between terms in definitions. *3rd International Workshop on Computational Terminology.* (Cited in page 67.)

[Maltese et al., 2010] Maltese, V., Giunchiglia, F., and Autayeu, A. (2010). Save up to 99% of your time in mapping validation. In *On the Move to Meaningful Internet Systems, OTM 2010 - Confederated International Conferences: CoopIS, IS, DOA and ODBASE, Hersonissos, Crete, Greece, October 25-29*, pages 1044–1060. (Cited in pages 20 and 84.)

[Mao et al., 2008] Mao, M., Peng, Y., and Spring, M. (2008). Ontology mapping: As a binary classification problem. In *2008 Fourth International Conference on Semantics, Knowledge and Grid*, pages 20–25. (Cited in page 100.)

[Mascardi et al., 2010] Mascardi, V., Locoro, A., and Rosso, P. (2010). Automatic Ontology Matching via Upper Ontologies: A Systematic Evaluation. *IEEE Trans. on Knowl. and Data Eng.*, 22(5):609–623. (Cited in page 58.)

[Matuschek and Gurevych, 2013] Matuschek, M. and Gurevych, I. (2013). Dijkstra-wsa: A graph-based approach to word sense alignment. *TACL*, 1:151–164. (Cited in page 102.)

[Megdiche et al., 2016a] Megdiche, I., Teste, O., and Trojahn, C. (2016a). An extensible linear approach for holistic ontology matching. In *The Semantic Web - ISWC 2016 - 15th International Semantic Web Conference, Kobe, Japan, October 17-21, 2016, Proceedings, Part I*, pages 393–410. (Cited in pages 12, 51, 52, 53 and 54.)

[Megdiche et al., 2016b] Megdiche, I., Teste, O., and Trojahn, C. (2016b). LPHOM results for OAEI 2016. In *Proceedings of the 11th International Workshop on Ontology Matching co-located with the 15th International Semantic Web Conference (ISWC 2016), Kobe, Japan, October 18, 2016*, pages 190–195. (Cited in pages 12, 53 and 56.)

[Meilicke, 2011] Meilicke, C. (2011). Alignment incoherence in ontology matching. (Cited in pages 33, 50 and 54.)

[Meilicke et al., 2012] Meilicke, C., Garcia-Castro, R., Freitas, F., van Hage, W. R., Montiel-Ponsoda, E., de Azevedo, R. R., Stuckenschmidt, H., Sváb-Zamazal, O., Svaátek, V., Tamilin, A., Trojahn, C., and Wang, S. (2012). Multifarm: A benchmark for multilingual ontology matching. *Journal Web Semantics*, 15:62–68. (Cited in pages 14, 74 and 77.)

[Meilicke and Stuckenschmidt, 2008] Meilicke, C. and Stuckenschmidt, H. (2008). Incoherence as a basis for measuring the quality of ontology mappings. In *3rd International Conference on Ontology Matching-Volume 431*, pages 1–12. (Cited in pages 20, 28 and 33.)

[Meilicke et al., 2009] Meilicke, C., Stuckenschmidt, H., and Tamilin, A. (2009). Reasoning support for mapping revision. *J. Log. Comput.*, 19(5):807–829. (Cited in page 9.)

[Menzel, 2003] Menzel, C. (2003). Reference ontologies - application ontologies: Either/or or both/and? In *Proceedings of the KI2003 Workshop on Reference Ontologies and Application Ontologies, Hamburg, Germany, September 16, 2003.* (Cited in page 102.)

[Michel et al., 2017] Michel, F., Gargominy, O., Tercerie, S., and Faron-Zucker, C. (2017). A Model to Represent Nomenclatural and Taxonomic Information as Linked Data.Application to the French Taxonomic Register, TAXREF. In Algergawy, A., Karam, N., Klan, F., and Jonquet, C., editors, *Proceedings of the 2nd International Workshop on Semantics for Biodiversity (S4BioDiv 2017) co-located with 16th International Semantic Web Conference (ISWC 2017)*, volume 1933, Vienna, Austria. CEUR-WS.org. (Cited in page 49.)

[Mika et al., 2004] Mika, P., Oberle, D., Gangemi, A., and Sabou, M. (2004). Foundations for Service Ontologies: Aligning OWL-S to DOLCE. In *Proceedings of the 13th Conference on World Wide Web*, pages 563–572. (Cited in pages 58 and 59.)

[Mikolov et al., 2013] Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). Efficient estimation of word representations in vector space. *CoRR*. (Cited in page 61.)

[Miller, 1995] Miller, G. A. (1995). WordNet: A Lexical Database for English. *Commun. ACM*, 38(11):39–41. (Cited in page 61.)

[Mintz et al., 2009] Mintz, M., Bills, S., Snow, R., and Jurafsky, D. (2009). Distant supervision for relation extraction without labeled data. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 1003–1011. (Cited in page 100.)

[Muñoz and Grüninger, 2016] Muñoz, L. and Grüninger, M. (2016). Verifying and mapping the mereotopology of upper-level ontologies. In *Proceedings of the Inter. Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, pages 31–42. (Cited in page 60.)

[Nardi et al., 2013] Nardi, J. C., de Almeida Falbo, R., and Almeida, J. P. A. (2013). Foundational ontologies for semantic integration in EAI: A systematic literature review. In *Proceedings of the 12th IFIP WG Conference on e-Business, e-Services, and e-Society, I3E*, pages 238–249. (Cited in page 58.)

[Navigli, 2009] Navigli, R. (2009). Word Sense Disambiguation: A Survey. *ACM Comput. Surv.*, 41(2):10:1–10:69. (Cited in page 61.)

[Navigli and Ponzetto, 2012] Navigli, R. and Ponzetto, S. P. (2012). BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *AI*, 193:217–250. (Cited in page 60.)

[Navigli et al., 2010] Navigli, R., Velardi, P., and Ruiz-Martínez, J. M. (2010). An annotated dataset for extracting definitions and hypernyms from the web. In *Proceedings of the International Conference on Language Resources and Evaluation*. (Cited in page 67.)

[Nentwig et al., 2017] Nentwig, M., Groß, A., Möller, M., and Rahm, E. (2017). Distributed holistic clustering on linked data. In *OTM 2017 Conferences: Confederated International Conferences: CoopIS, C&TC, and ODBASE 2017, Rhodes, Greece, October 23-27*, pages 371–382. (Cited in pages 52 and 85.)

[Niklaus et al., 2018] Niklaus, C., Cetto, M., Freitas, A., and Handschuh, S. (2018). A survey on open information extraction. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 3866–3878, Santa Fe, New Mexico, USA. Association for Computational Linguistics. (Cited in page 101.)

[Niles and Pease, 2001] Niles, I. and Pease, A. (2001). Towards a Standard Upper Ontology. In *Conference on Formal Ontology in Information Systems*, pages 2–9. (Cited in pages 59 and 93.)

[Niles and Pease, 2003] Niles, I. and Pease, A. (2003). Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. In *Proceedings of the Inter. Conference on Knowledge Engineering*, pages 412–416. (Cited in page 93.)

[Nkisi-Orji et al., 2019] Nkisi-Orji, I., Wiratunga, N., Massie, S., Hui, K.-Y., and Heaven, R. (2019). Ontology alignment based on word embedding and random forest classification. In Berlingerio, M., Bonchi, F., Gärtner, T., Hurley, N., and Ifrim, G., editors, *Machine Learning and Knowledge Discovery in Databases*, pages 557–572, Cham. Springer International Publishing. (Cited in page 100.)

[Noy, 2004] Noy, N. F. (2004). Semantic integration: a survey of ontology-based approaches. *ACM Sigmod Record*, 33(4):65–70. (Cited in pages 9 and 41.)

[Noy and Musen, 2003] Noy, N. F. and Musen, M. A. (2003). The prompt suite: interactive tools for ontology merging and mapping. *International Journal of Human-Computer Studies*, 59(6). (Cited in page 43.)

[Nunes et al., 2011] Nunes, B. P., Caraballo, A. A. M., Casanova, M. A., Breitman, K. K., and Leme, L. A. P. P. (2011). Complex matching of RDF datatype properties. In Shvaiko, P., Euzenat, J., Heath, T., Quix, C., Mao, M., and Cruz, I. F., editors, *Proceedings of the 6th International Workshop on Ontology Matching, Bonn, Germany, October 24, 2011*, volume 814 of *CEUR Workshop Proceedings*. CEUR-WS.org. (Cited in page 38.)

[Nurmikko et al., 2015] Nurmikko, T., Page, K. R., Willcox, P., Jett, J., Maden, C., Cole, T. W., Fallaw, C., Senseney, M., and Downie, J. S. (2015). Building complex research collections in digital libraries: A survey of ontology implications. In II, P. L. B., Allard, S., Mercer, H., Beck, M., Cunningham, S. J., Goh, D. H., and Henry, G., editors, *Proceedings of the 15th ACM/IEEE-CE Joint Conference on Digital Libraries, Knoxville, TN, USA, June 21-25, 2015*, pages 169–172. ACM. (Cited in page 38.)

[Oberle et al., 2007] Oberle, D., Ankolekar, A., Hitzler, P., Cimiano, P., Sintek, M., Kiesel, M., Mougouie, B., Baumann, S., Vembu, S., Romanelli, M., Buitelaar, P., Engel, R., Sonntag, D., Reithinger, N., Loos, B., Zorn, H., Micelli, V., Porzel, R., Schmidt, C., Weiten, M., Burkhardt, F., and Zhou, J. (2007). DOLCE Ergo SUMO: On Foundational and Domain Models in the SmartWeb Integrated Ontology (SWIntO). *Web Semantics*, 5(3):156–174. (Cited in pages 60, 61 and 68.)

[O'Connor and Das, 2011] O'Connor, M. J. and Das, A. (2011). Acquiring owl ontologies from xml documents. In *Proceedings of the Sixth International Conference on Knowledge Capture*, pages 17–24, New York, NY, USA. ACM. (Cited in page 100.)

[Oliveira and Pesquita, 2015] Oliveira, D. and Pesquita, C. (2015). Compound matching of biomedical ontologies. In *Proceedings of the International Conference on Biomedical Ontology, ICBO 2015, Lisbon, Portugal, July 27-30*. (Cited in pages 51 and 85.)

[Parundekar et al., 2010] Parundekar, R., Knoblock, C. A., and Ambite, J. L. (2010). Linking and building ontologies of linked data. In Patel-Schneider, P. F., Pan, Y., Hitzler, P., Mika, P., Zhang, L., Pan, J. Z., Horrocks, I., and Glimm, B., editors, *The Semantic Web - ISWC 2010 - 9th International Semantic Web Conference, ISWC 2010, Shanghai, China, November 7-11, 2010, Revised Selected Papers, Part I*, volume 6496 of *Lecture Notes in Computer Science*, pages 598–614. Springer. (Cited in pages 38, 43 and 44.)

[Parundekar et al., 2012] Parundekar, R., Knoblock, C. A., and Ambite, J. L. (2012). Discovering concept coverings in ontologies of linked data sources. In Cudré-Mauroux, P., Heflin, J., Sirin, E., Tudorache, T., Euzenat, J., Hauswirth, M., Parreira, J. X., Hendler, J., Schreiber, G., Bernstein, A., and Blomqvist, E., editors, *The Semantic Web - ISWC 2012 - 11th International Semantic Web Conference, Boston, MA, USA, November 11-15, 2012, Proceedings, Part I*, volume 7649 of *Lecture Notes in Computer Science*, pages 427–443. Springer. (Cited in pages 38, 43, 44 and 70.)

[Paulheim and Gangemi, 2015] Paulheim, H. and Gangemi, A. (2015). Serving dbpedia with dolce – more than just adding a cherry on top. In *The Semantic Web - ISWC 2015*, pages 180–196, Cham. Springer International Publishing. (Cited in page 102.)

[Pease, 2011] Pease, A. (2011). *Ontology: A Practical Guide*. Articulate Software Press, Angwin, CA. (Cited in pages 59 and 93.)

[Pease and Benzmüller, 2013] Pease, A. and Benzmüller, C. (2013). Sigma: An Integrated Development Environment for Logical Theories. *AI Communications*, 26:9–97. (Cited in page 93.)

[Pennington et al., 2014] Pennington, J., Socher, R., and Manning, C. D. (2014). Glove: Global vectors for word representation. In *Empirical Methods in Natural Language Processing*, pages 1532–1543. (Cited in page 65.)

[Pesquita et al., 2014] Pesquita, C., Cheatham, M., Faria, D., Barros, J., Santos, E., and Couto, F. M. (2014). Building reference alignments for compound matching of multiple ontologies using obo cross-products. In *Proceedings of the 9th International Workshop on Ontology Matching*, pages 172–173. (Cited in pages 12, 51, 52 and 85.)

[Pinto and Martins, 2004] Pinto, H. S. and Martins, J. P. (2004). Ontologies: How can They be Built? *Knowledge and Information Systems*, 6(4):441–464. (Cited in pages 11 and 43.)

[Pradel et al., 2012] Pradel, C., Haemmerlé, O., and Hernandez, N. (2012). A Semantic Web Interface Using Patterns: The SWIP System. In *Graph Structures for Knowledge Representation and Reasoning*, LNCS, pages 172–187. Springer Berlin Heidelberg. (Cited in page 39.)

[Prestes et al., 2013] Prestes, E., Carbonera, J. L., Fiorini, S. R., et al. (2013). Towards a core ontology for robotics and automation. *Robotics and Autonomous Systems*, 61(11):1193 – 1204. (Cited in page 64.)

[Prytkova et al., 2015] Prytkova, N., Weikum, G., and Spaniol, M. (2015). Aligning multi-cultural knowledge taxonomies by combinatorial optimization. In *Proceedings of the 24th International Conference on World Wide Web*, pages 93–94. ACM. (Cited in page 52.)

[Qi and Hunter, 2007] Qi, G. and Hunter, A. (2007). Measuring incoherence in description logic-based ontologies. In *Proceedings of the 6th International Semantic Web Conference*, Busan, Korea. (Cited in page 33.)

[Qin et al., 2007] Qin, H., Dou, D., and LePendu, P. (2007). Discovering executable semantic mappings between ontologies. In Meersman, R. and Tari, Z., editors, *On the Move to Meaningful Internet Systems 2007: CoopIS, DOA, ODBASE, GADA, and IS, OTM Confederated International Conferences CoopIS, DOA, ODBASE, GADA, and IS 2007, Vilamoura, Portugal, November 25-30, 2007, Proceedings, Part I*, volume 4803 of *Lecture Notes in Computer Science*, pages 832–849. Springer. (Cited in pages 38, 43 and 70.)

[Rahm, 2011] Rahm, E. (2011). *Schema Matching and Mapping*, chapter Towards Large-Scale Schema and Ontology Matching, pages 3–27. Springer Berlin Heidelberg. (Cited in pages 20 and 51.)

[Rahm, 2016] Rahm, E. (2016). The case for holistic data integration. In *Advances in Databases and Information Systems - 20th East European Conference, ADBIS 2016, Prague, Czech Republic, August 28-31, 2016, Proceedings*, pages 11–27. (Cited in page 51.)

[Rahm and Bernstein, 2001] Rahm, E. and Bernstein, P. A. (2001). A survey of approaches to automatic schema matching. *The VLDB Journal*, 10(4):334–350. (Cited in pages 9, 22 and 41.)

[Ravat et al., 2019a] Ravat, F., Song, J., Teste, O., and Trojahn, C. (2019a). Efficient querying of multidimensional rdf data with aggregates: comparing nosql, rdf and relational data stores. *International Journal of Information Management (submitted)*. (Cited in page 19.)

[Ravat et al., 2019b] Ravat, F., Song, J., Teste, O., and Trojahn, C. (2019b). Improving the performance of querying multidimensional RDF data using aggregates. In *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, SAC 2019, Limassol, Cyprus, April 8-12, 2019*, pages 2275–2284. (Cited in page 18.)

[Rebele et al., 2016] Rebele, T., Suchanek, F. M., Hoffart, J., Biega, J., Kuzey, E., and Weikum, G. (2016). YAGO: A multilingual knowledge base from wikipedia, wordnet, and geonames. In *The Semantic Web - ISWC 2016 - 15th International Semantic Web Conference, Kobe, Japan, October 17-21, 2016, Proceedings, Part II*, pages 177–185. (Cited in page 60.)

[Reed and Lenat, 2002] Reed, S. and Lenat, D. (2002). Mapping Ontologies into Cyc. In *Proceedings of the Workshop on Ontologies For The Semantic Web*, pages 1–6. (Cited in page 69.)

[Reitsma and Albrecht, 2005] Reitsma, F. and Albrecht, J. (2005). Modeling with the semantic web in the geosciences. *IEEE Intelligent Systems*, 20(2):86–88. (Cited in page 17.)

[Ren et al., 2014] Ren, Y., Parvizi, A., Mellish, C., Pan, J. Z., van Deemter, K., and Stevens, R. (2014). Towards Competency Question-Driven Ontology Authoring. In *The Semantic Web: Trends and Challenges*, volume 8465, pages 752–767. (Cited in pages 11, 43 and 44.)

[Ritze et al., 2009] Ritze, D., Meilicke, C., Sváb-Zamazal, O., and Stuckenschmidt, H. (2009). A pattern-based ontology matching approach for detecting complex correspondences. In Shvaiko, P., Euzenat, J., Giunchiglia, F., Stuckenschmidt, H., Noy, N. F., and Rosenthal, A., editors, *Proceedings of the 4th International Workshop on Ontology Matching (OM-2009) collocated with the 8th International Semantic Web Conference (ISWC-2009) Chantilly, USA, October 25, 2009*, volume 551 of *CEUR Workshop Proceedings*. CEUR-WS.org. (Cited in pages 38, 43, 44 and 70.)

[Ritze and Paulheim, 2011] Ritze, D. and Paulheim, H. (2011). Towards an automatic parameterization of ontology matching tools based on example mappings. In *Proceedings of the 6th International Workshop on Ontology Matching, Bonn, Germany, October 24, 2011.* (Cited in page 9.)

[Ritze et al., 2010] Ritze, D., Völker, J., Meilicke, C., and Sváb-Zamazal, O. (2010). Linguistic analysis for complex ontology matching. In Shvaiko, P., Euzenat, J., Giunchiglia, F., Stuckenschmidt, H., Mao, M., and Cruz, I. F., editors, *Proceedings of the 5th International Workshop on Ontology Matching (OM-2010), Shanghai, China, November 7, 2010*, volume 689 of *CEUR Workshop Proceedings*. CEUR-WS.org. (Cited in pages 38, 43, 44 and 70.)

[Roller et al., 2018] Roller, S., Kiela, D., and Nickel, M. (2018). Hearst patterns revisited: Automatic hypernym detection from large text corpora. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 358–363, Melbourne, Australia. Association for Computational Linguistics. (Cited in page 101.)

[Roussille et al., 2018a] Roussille, P., Megdiche, I., Teste, O., and Trojahn, C. (2018a). Boosting holistic ontology matching: Generating graph clique-based relaxed reference alignments for holistic evaluation. In *EKAW*, volume 11313 of *Lecture Notes in Computer Science*, pages 355–369. Springer. (Cited in pages 15 and 85.)

[Roussille et al., 2018b] Roussille, P., Megdiche, I., Teste, O., and Trojahn, C. (2018b). Holontology: results of the 2018 OAEI evaluation campaign. In

OM@ISWC, volume 2288 of *CEUR Workshop Proceedings*, pages 167–172. CEUR-WS.org. (Cited in pages 53 and 57.)

[Ruy et al., 2017] Ruy, F. B., Guizzardi, G., Falbo, R. A., Reginato, C. C., and Santos, V. A. (2017). From reference ontologies to ontology patterns and back. *Data & Knowledge Engineering*, 109:41 – 69. Special issue on conceptual modeling — 34th International Conference on Conceptual Modeling. (Cited in page 102.)

[Saad et al., 2013] Saad, R., Trojahn, C., and Teste, O. (2013). Olap manipulations on rdf data following a constellation model. In *Proceedings of the 1st International Workshop on Semantic Statistics (SemStats 2013) collocated with the 12th International Semantic Web Conference (ISWC-2013) Sydney, Australia, October 22th, 2013*. (Cited in page 18.)

[Sabou and Gracia, 2008] Sabou, M. and Gracia, J. (2008). Spider: bringing non-equivalence mappings to OAEI. In *Proceedings of the ISWC 2008 Workshop on Ontology Matching*, pages 199–205, Karlsruhe (DE). (Cited in page 27.)

[Sagi and Gal, 2018] Sagi, T. and Gal, A. (2018). Non-binary evaluation measures for big data integration. *VLDB J.*, 27(1):105–126. (Cited in page 30.)

[Saleem et al., 2008] Saleem, K., Bellahsene, Z., and Hunt, E. (2008). PORSCHE: Performance ORiented SCHEma mediation. *Information Systems*, 33(7-8):637–657. (Cited in page 52.)

[Schadd and Roos, 2012] Schadd, F. C. and Roos, N. (2012). Coupling of Wordnet Entries for Ontology Mapping Using Virtual Documents. In *Proc. of the 7th Conf. on Ontology Matching*, pages 25–36. (Cited in page 61.)

[Scharffe, 2009] Scharffe, F. (2009). *Correspondence Patterns Representation*. PhD thesis, Faculty of Mathematics, Computer Science and University of Innsbruck. (Cited in pages 42, 74 and 99.)

[Scharffe et al., 2014] Scharffe, F., Zamazal, O., and Fensel, D. (2014). Ontology alignment design patterns. *Knowl. Inf. Syst.*, 40(1):1–28. (Cited in page 99.)

[Schmidt et al., 2017] Schmidt, D., Basso, R., Trojahn, C., and Vieira, R. (2017). Matching domain and top-level ontologies via ontowordnet. In *OM@ISWC*, volume 2032 of *CEUR Workshop Proceedings*, pages 225–226. CEUR-WS.org. (Cited in pages 61 and 62.)

[Schmidt et al., 2018] Schmidt, D., Basso, R., Trojahn, C., and Vieira, R. (2018). Matching domain and top-level ontologies exploring word sense disambiguation and word embedding. In *ISWC (Best Workshop Papers)*, volume 36 of *Studies on the Semantic Web*, pages 27–38. IOS Press. (Cited in pages 13, 61 and 62.)

[Schmidt et al., 2019a] Schmidt, D., Pease, A., Trojahn, C., and Vieira, R. (2019a). Aligning Conference Ontologies with SUMO: A Report on Manual Alignment via WordNet. In *Proceedings of the 3rd Workshop on Foundational Ontology (FOUST III), co-located with the Joint Ontology Workshops, JOWO-2019, Graz, Austria.* (Cited in page 93.)

[Schmidt et al., 2016] Schmidt, D., Trojahn, C., and Vieira, R. (2016). Analysing top-level and domain ontology alignments from matching systems. In *Proceedings of the 11th International Workshop on Ontology Matching co-located with the 15th International Semantic Web Conference (ISWC 2016), Kobe, Japan, October 18, 2016*, pages 13–24. (Cited in pages 13, 59 and 62.)

[Schmidt et al., 2019b] Schmidt, D., Trojahn, C., and Vieira, R. (2019b). Matching bfo, dolce, gfo and sumo: an evaluation of oaei 2018 matching systems. In *Proceedings of the 12th Seminar on Ontology Research in Brazil (ONTOBRAS)*. (Cited in pages 61 and 62.)

[Schrijver, 2003] Schrijver, A. (2003). *Combinatorial Optimization - Polyhedra and Efficiency*. Springer. (Cited in page 53.)

[Semy et al., 2004] Semy, S., Pulvermacher, M., and Obrst, L. (2004). Toward the use of an upper ontology for U.S. government and U.S. military domains: An evaluation. Technical report, MTR 04B0000063, The MITRE Corporation. (Cited in page 58.)

[Seppälä et al., 2014] Seppälä, S., Smith, B., and Ceusters, W. (2014). Applying the realism-based ontology-versioning method for tracking changes in the basic formal ontology. In *Proceedings of FOIS*, pages 227–240. (Cited in page 60.)

[Serpeloni et al., 2011] Serpeloni, F., Moraes, R., and Bonacin, R. (2011). Ontology mapping validation. *International Journal of Web Portals*, 3(3):1–11. (Cited in page 20.)

[Severo et al., 2014] Severo, B., Trojahn, C., and Vieira, R. (2014). Voar: A visual and integrated ontology alignment environment. In *Proceedings of the 9th Language Resources and Evaluation Conference (LREC)*. (Cited in page 15.)

[Severo et al., 2015] Severo, B., Trojahn, C., and Vieira, R. (2015). A GUI for visualising and manipulating multiple ontology alignments. In *Proceedings of the ISWC 2015 Posters & Demonstrations Track co-located with the 14th International Semantic Web Conference (ISWC-2015), Bethlehem, PA, USA, October 11, 2015*. (Cited in page 15.)

[Severo et al., 2017] Severo, B., Trojahn, C., and Vieira, R. (2017). VOAR 3.0 : a configurable environment for manipulating multiple ontology alignments. In *International Semantic Web Conference (Posters, Demos & Industry Tracks)*, volume 1963 of *CEUR Workshop Proceedings*. CEUR-WS.org. (Cited in page 15.)

[Seyed, 2009] Seyed, A. (2009). BFO/DOLCE Primitive Relation Comparison. In *Nature proceedings*. (Cited in page 60.)

[Shvaiko and Euzenat, 2005] Shvaiko, P. and Euzenat, J. (2005). A Survey of Schema-Based Matching Approaches. In Spaccapietra, S., editor, *Journal on Data Semantics IV*, number 3730 in Lecture Notes in Computer Science, pages 146–171. Springer Berlin Heidelberg. (Cited in pages 9 and 41.)

[Silva et al., 2011] Silva, V., Campos, M., Silva, J., and Cavalcanti, M. (2011). An Approach for the Alignment of Biomedical Ontologies based on Foundational Ontologies. *Information and Data Management*, 2(3):557–572. (Cited in page 59.)

[Sim et al., 2003] Sim, S. E., Easterbrook, S., and Holt, R. C. (2003). Using benchmarking to advance research: A challenge to software engineering. In *Proceedings of the 25th International Conference on Software Engineering*, pages 74–83. (Cited in page 92.)

[Singh et al., 2017] Singh, A., Debruyne, C., Brennan, R., Meehan, A., and O'Sullivan, D. (2017). Extension of the m-gov ontology mapping framework for increased traceability. In *Proceedings of the 12th International Workshop on Ontology Matching co-located with the 16th International Semantic Web Conference (ISWC 2017), Vienna, Austria, October 21*, pages 49–60. (Cited in pages 77 and 98.)

[Smirnova and Cudré-Mauroux, 2018] Smirnova, A. and Cudré-Mauroux, P. (2018). Relation extraction using distant supervision: A survey. *ACM Comput. Surv.*, 51(5):106:1–106:35. (Cited in page 100.)

[Solimando et al., 2014a] Solimando, A., Jiménez-Ruiz, E., and Guerrini, G. (2014a). Detecting and correcting conservativity principle violations in ontology-to-ontology mappings. In Mika, P., Tudorache, T., Bernstein, A., Welty, C., Knoblock, C., Vrandečić, D., Groth, P., Noy, N., Janowicz, K., and Goble, C., editors, *The Semantic Web – ISWC 2014*, pages 1–16, Cham. Springer International Publishing. (Cited in page 34.)

[Solimando et al., 2017] Solimando, A., Jimenez-Ruiz, E., and Guerrini, G. (2017). Minimizing conservativity violations in ontology alignments: Algorithms and evaluation. *Knowledge and Information Systems*, 51(3):775–819. (Cited in pages 20 and 28.)

[Solimando et al., 2014b] Solimando, A., Jiménez-Ruiz, E., and Pinkel, C. (2014b). Evaluating ontology alignment systems in query answering tasks. In *Proceedings of the 2014 International Conference on Posters & Demonstrations Track-Volume 1272*, pages 301–304. CEUR-WS. org. (Cited in pages 20, 28, 36 and 77.)

[Sorokin and Gurevych, 2017] Sorokin, D. and Gurevych, I. (2017). Context-aware representations for knowledge base relation extraction. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1784–1789, Copenhagen, Denmark. Association for Computational Linguistics. (Cited in page 100.)

[Spiliopoulos et al., 2010] Spiliopoulos, V., Vouros, G. A., and Karkaletsis, V. (2010). On the discovery of subsumption relations for the alignment of ontologies. *Journal of Web Semantics*, 8(1):69 – 88. (Cited in pages 62 and 100.)

[Stapleton et al., 2014] Stapleton, G., Howse, J., Bonnington, A., and Burton, J. (2014). A vision for diagrammatic ontology engineering. In Ivanova, V., Kauppinen, T., Lohmann, S., Mazumdar, S., Pesquita, C., and Xu, K., editors, *Proceedings of the International Workshop on Visualizations and User Interfaces for Knowledge Engineering and Linked Data Analytics co-located with 19th International Conference on Knowledge Engineering and Knowledge Management, VISUAL@EKAW 2014, Linköping, Sweden, November 24, 2014*, volume 1299 of *CEUR Workshop Proceedings*, pages 1–13. CEUR-WS.org. (Cited in page 22.)

[Stevens et al., 2018] Stevens, R., Lord, P., Malone, J., and Matentzoglu, N. (2018). Measuring expert performance at manually classifying domain entities under upper ontology classes. *Journal of Web Semantics*. (Cited in pages 28 and 92.)

[Stuckenschmidt et al., 2006] Stuckenschmidt, H., Serafini, L., and Wache, H. (2006). Reasoning about ontology mappings. In *Proceedings of the ECAI workshop on contextual representation and reasoning*. (Cited in page 33.)

[Su et al., 2006] Su, W., Wang, J., and Lochovsky, F. (2006). Holistic schema matching for web query interfaces. In *Proceedings of the 10th International Conference on Advances in Database Technology*, EDBT'06, pages 77–94. Springer-Verlag. (Cited in page 52.)

[Subasic et al., 2019] Subasic, P., Yin, H., and Lin, X. (2019). Building knowledge base through deep learning relation extraction and wikidata. In *AAAI Spring Symposium: Combining Machine Learning with Knowledge Engineering*. (Cited in page 100.)

[Suchanek et al., 2007] Suchanek, F. M., Kasneci, G., and Weikum, G. (2007). Yago: a core of semantic knowledge. In *Proceedings of the 16th international conference on World Wide Web*, pages 697–706. ACM. (Cited in page 100.)

[Sukhobok et al., 2017] Sukhobok, D., Sánchez, H., Estrada, J., and Roman, D. (2017). Linked data for common agriculture policy: Enabling semantic querying over sentinel-2 and lidar data. In *Proceedings of the ISWC 2017 Posters & Demonstrations and Industry Tracks co-located with 16th Int. Semantic Web Conference, Vienna, Austria, Oct. 23-25, 2017*. (Cited in page 17.)

[Sun et al., 2019] Sun, K., Zhu, Y., and Song, J. (2019). Progress and challenges on entity alignment of geographic knowledge bases. *ISPRS International Journal of Geo-Information*, 8:77. (Cited in page 99.)

[Sun et al., 2015] Sun, Y., Ma, L., and Shuang, W. (2015). A comparative evaluation of string similarity metrics for ontology alignement. *Journal of Information & Computational Science*, 12(3):957 – 964. (Cited in page 54.)

[Šváb et al., 2005] Šváb, O., Svátek, V., Berka, P., Rak, D., and Tomášek, P. (2005). Ontofarm: Towards an experimental collection of parallel ontologies. *Poster Track of ISWC*, 2005. (Cited in page 22.)

[Symeonidou et al., 2017] Symeonidou, D., Galárraga, L., Pernelle, N., Saïs, F., and Suchanek, F. M. (2017). VICKEY: mining conditional keys on knowledge bases. In *ISWC*, pages 661–677. (Cited in page 50.)

[Szekely et al., 2013] Szekely, P., Knoblock, C. A., Yang, F., Zhu, X., Fink, E. E., Allen, R., and Goodlander, G. (2013). Connecting the Smithsonian American Art Museum to the Linked Data Cloud. In Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Mitchell, J. C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M. Y., Weikum, G., Cimiano, P., Corcho, O., Presutti, V., Hollink, L., and Rudolph, S., editors, *The Semantic Web: Semantics and Big Data*, volume 7882, pages 593–607. Springer Berlin Heidelberg. (Cited in page 38.)

[Temal et al., 2010] Temal, L., Rosier, A., Dameron, O., and Burgun, A. (2010). Mapping BFO and DOLCE. In *Proceedings of the World Congress on Medical Informatics*, pages 1065–1069. (Cited in page 60.)

[Terziev et al., 2005] Terziev, I., Kiryakov, A., and Manov, D. (2005). Base Upper-level Ontology (BULO) Guidance. Deliverable 1.8.1, sekt project. (Cited in page 59.)

[Thiéblin et al., 2016] Thiéblin, É., Amarger, F., Haemmerlé, O., Hernandez, N., and Trojahn, C. (2016). Rewriting SELECT SPARQL queries from 1: n complex correspondences. In *Proceedings of the 11th International Workshop on Ontology Matching co-located with the 15th International Semantic Web Conference (ISWC 2016), Kobe, Japan, October 18, 2016*, pages 49–60. (Cited in pages 12, 40, 47 and 82.)

[Thiéblin et al., 2017a] Thiéblin, É., Amarger, F., Hernandez, N., Roussey, C., and Trojahn, C. (2017a). Cross-querying LOD datasets using complex alignments: An application to agronomic taxa. In *MTSR*, volume 755 of *Communications in Computer and Information Science*, pages 25–37. Springer. (Cited in page 12.)

[Thiéblin et al., 2019a] Thiéblin, É., Amarger, F., Hernandez, N., Roussey, C., and Trojahn, C. (2019a). Cross-querying LOD data sets using complex alignments: an experiment using AgronomicTaxon, Agrovoc, DBpedia and TAXREF-LD. *International Journal of Metadata, Semantics and Ontologies.* (Cited in pages 12, 38, 40 and 49.)

[Thiéblin et al., 2019b] Thiéblin, É., Cheatham, M., Trojahn, C., and Zamazal, O. (2019b). Complex Conference Benchmark: Consensus Dataset. *KER (submitted).* (Cited in page 72.)

[Thiéblin et al., 2018a] Thiéblin, É., Cheatham, M., Trojahn, C., Zamazal, O., and Zhou, L. (2018a). The first version of the OAEI complex alignment benchmark. In *International Semantic Web Conference (P&D/Industry/BlueSky)*, volume 2180 of *CEUR Workshop Proceedings*. CEUR-WS.org. (Cited in pages 15, 40, 70 and 72.)

[Thiéblin et al., 2017b] Thiéblin, É., Haemmerlé, O., Hernandez, N., and dos Santos, C. T. (2017b). Towards a complex alignment evaluation dataset. In *OM@ISWC*, volume 2032 of *CEUR Workshop Proceedings*, pages 217–218. CEUR-WS.org. (Cited in page 72.)

[Thiéblin et al., 2017c] Thiéblin, É., Haemmerlé, O., Hernandez, N., and Trojahn, C. (2017c). Un jeu de données d'évaluation de correspondances complexes entre ontologies. In *IC*, pages 68–79. (Cited in page 72.)

[Thiéblin et al., 2018b] Thiéblin, É., Haemmerlé, O., Hernandez, N., and Trojahn, C. (2018b). Task-oriented complex ontology alignment: Two alignment evaluation sets. In *ESWC*, volume 10843 of *Lecture Notes in Computer Science*, pages 655–670. Springer. (Cited in pages 15, 70, 72, 75 and 76.)

[Thiéblin et al., 2019c] Thiéblin, É., Haemmerle, O., Hernandez, N., and Trojahn, C. (2019c). Survey on complex ontology matching. *Semantic Web Journal (to appear)*. (Cited in pages 11, 40 and 41.)

[Thiéblin et al., 2018c] Thiéblin, É., Haemmerlé, O., and Trojahn, C. (2018c). CANARD complex matching system: results of the 2018 OAEI evaluation campaign. In *OM@ISWC*, volume 2288 of *CEUR Workshop Proceedings*, pages 138–143. CEUR-WS.org. (Cited in pages 40 and 46.)

[Thiéblin et al., 2018d] Thiéblin, É., Haemmerlé, O., and Trojahn, C. (2018d). Complex matching based on competency questions for alignment: a first sketch. In *OM@ISWC*, volume 2288 of *CEUR Workshop Proceedings*, pages 66–70. CEUR-WS.org. (Cited in pages 11 and 40.)

[Thiéblin et al., 2019d] Thiéblin, É., Haemmerlé, O., and Trojahn, C. (2019d). Automatic evaluation of complex alignment: a query-oriented approach. *Semantic Web Journal (submitted)*. (Cited in pages 72, 77 and 81.)

[Thiéblin and Trojahn, 2019] Thiéblin, É. and Trojahn, C. (2019). Conference v3.0 : A populated version of the conference dataset. In *Proceedings of the ISWC 2019 Posters & Demonstrations Track co-located with the 18th International Semantic Web Conference (ISWC-2019)*. (Cited in pages 72 and 77.)

[Tordai et al., 2011] Tordai, A., van Ossenbruggen, J., Schreiber, G., and Wielinga, B. (2011). Let's agree to disagree: On the evaluation of vocabulary alignment. In *Proceedings of the 6th International Conference on Knowledge Capture*, pages 65–72. (Cited in pages 28 and 74.)

[Trojahn et al., 2009] Trojahn, C., Euzenat, J., Meilicke, C., and Stuckenschmidt, H. (2009). Evaluation design and collection of test data for matching tools. Technical Report D12.1, SEALS Project. (Cited in page 23.)

[Trojahn et al., 2014] Trojahn, C., Fu, B., Zamazal, O., and Ritze, D. (2014). State-of-the-art in multilingual and cross-lingual ontology matching. In *Towards the Multilingual Semantic Web, Principles, Methods and Applications*, pages 119–135. (Cited in page 13.)

[Trojahn et al., 2010] Trojahn, C., Quaresma, P., and Vieira, R. (2010). An API for Multilingual Ontology Macthing. In Chair), N. C. C., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., and Tapias, D., editors, *Proceedings of the 7th Edition of the Language Resources and Evaluation Conference*, pages 3830–3835, Malta. European Language Resources Association (ELRA). (Cited in page 13.)

[Tufiş et al., 2004] Tufiş, D., Barbu, A. M., and Ion, R. (2004). Extracting multilingual lexicons from parallel corpora. *Computers and the Humanities*, 38(2):163–189. (Cited in page 103.)

[Van Hage et al., 2007] Van Hage, W. R., Isaac, A., and Aleksovski, Z. (2007). Sample evaluation of ontology-matching systems. In *EON*, volume 2007, pages 41–50. (Cited in page 20.)

[van Hage et al., 2005] van Hage, W. R., Katrenko, S., and Schreiber, G. (2005). A method to combine linguistic ontology-mapping techniques. In *International Semantic Web Conference*, pages 732–744. (Cited in pages 27, 62 and 100.)

[Vazquez and Swoboda, 2007] Vazquez, R. and Swoboda, N. (2007). Combining the semantic web with the web as background knowledge for ontology mapping. In *Meaningful Internet Systems*, pages 814–831. (Cited in pages 62 and 100.)

[Vennesland, 2017] Vennesland, A. (2017). Matcher composition for identification of subsumption relations in ontology matching. In *Proceedings of the Conference on Web Intelligence*, pages 154–161. (Cited in page 59.)

[Vieira and Revoredo, 2017] Vieira, R. and Revoredo, K. (2017). Using word semantics on entity names for correspondence set generation. In *Workshop on Ontology Matching*, pages 223–224. (Cited in page 61.)

[Visser et al., 1997] Visser, P. R., Jones, D. M., Bench-Capon, T. J., and Shave, M. (1997). An analysis of ontology mismatches: heterogeneity versus interoperability. In *AAAI 1997 Spring Symposium on Ontological Engineering, Stanford CA., USA*, pages 164–72. (Cited in pages 20 and 38.)

[Walshe et al., 2016] Walshe, B., Brennan, R., and O'Sullivan, D. (2016). Bayes-recce: A bayesian model for detecting restriction class correspondences in linked open data knowledge bases. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 12(2):25–52. (Cited in pages 38, 43, 44 and 70.)

[Wang et al., 2017] Wang, C., He, X., and Zhou, A. (2017). A short survey on taxonomy learning from text corpora: Issues, resources and recent advances. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1190–1203, Copenhagen, Denmark. Association for Computational Linguistics. (Cited in page 100.)

[Wang, 2011] Wang, P. (2011). Lily Results on SEALS Platform for OAEI 2011. In *Proc. of the 6th Intern. Workshop on Ontology Matching*, pages 156–162. (Cited in page 61.)

[Xiang et al., 2015a] Xiang, C., Chang, B., and Sui, Z. (2015a). An ontology matching approach based on affinity-preserving random walks. In *Proceedings of the*

24th International Conference on Artificial Intelligence, pages 1471–1477. (Cited in page 52.)

[Xiang et al., 2015b] Xiang, C., Jiang, T., Chang, B., and Sui, Z. (2015b). ERSOM: A structural ontology matching approach using automatically learned entity representation. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 2419–2429, Lisbon, Portugal. Association for Computational Linguistics. (Cited in page 100.)

[Xiao et al., 2018] Xiao, G., Calvanese, D., Kontchakov, R., Lembo, D., Poggi, A., Rosati, R., and Zakharyaschev, M. (2018). Ontology-based data access: A survey. In Lang, J., editor, *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden*, pages 5511–5519. ijcai.org. (Cited in page 103.)

[Yatskevich and Giunchiglia, 2004] Yatskevich, M. and Giunchiglia, F. (2004). Element level semantic matching using WordNet. In *Meaning Coordination and Negotiation Workshop*, pages 37–48. (Cited in pages 9 and 61.)

[Zamazal and Svátek, 2017] Zamazal, O. and Svátek, V. (2017). The Ten-Year OntoFarm and its Fertilization within the Onto-Sphere. *Web Semantics: Science, Services and Agents on the World Wide Web*, 43:46–53. (Cited in pages 22, 27, 77 and 93.)

[Zhang et al., 2012] Zhang, C., Hoffmann, R., and Weld, D. S. (2012). Ontological smoothing for relation extraction with minimal supervision. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, AAAI'12, pages 157–163. AAAI Press. (Cited in page 100.)

[Zhang and Bodenreider, 2005] Zhang, S. and Bodenreider, O. (2005). Alignment of multiple ontologies of anatomy: Deriving indirect mappings from direct mappings to a reference. In *American Medical Informatics Association Annual Symposium*. (Cited in pages 9 and 61.)

[Zhang et al., 2014] Zhang, Y., Wang, X., Lai, S., He, S., Liu, K., Zhao, J., and Lv, X. (2014). Ontology matching with word embeddings. In Sun, M., Liu, Y., and Zhao, J., editors, *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data*, pages 34–45. (Cited in pages 9, 61 and 100.)

[Zheng et al., 2012] Zheng, X., Madnick, S. E., and Li, X. (2012). SPARQL Query Mediation over RDF Data Sources with Disparate Contexts. In Bizer, C., Heath, T., Berners-Lee, T., and Hausenblas, M., editors, *WWW Workshop on Linked Data on the Web (LDOW2012)*, volume 937, Lyon, France. CEUR-WS.org. (Cited in pages 39 and 47.)

[Zhou et al., 2018] Zhou, L., Cheatham, M., Krisnadhi, A., and Hitzler, P. (2018). A complex alignment benchmark: Geolink dataset. In Vrandecic, D., Bontcheva, K., Suárez-Figueroa, M. C., Presutti, V., Celino, I., Sabou, M., Kaffee, L., and

Simperl, E., editors, *The Semantic Web - ISWC 2018 - 17th International Semantic Web Conference, Monterey, CA, USA, October 8-12, 2018, Proceedings, Part II*, volume 11137 of *Lecture Notes in Computer Science*, pages 273–288. Springer. (Cited in pages 22 and 38.)

[Zhou et al., 2019] Zhou, L., Thiéblin, É., Cheatham, M., Faria, D., Pesquita, C., Trojahn, C., and Zamazal, O. (2019). Towards Evaluating Complex Ontology Alignments. *KER (submitted)*. (Cited in pages 72 and 81.)

[Zimmermann and Le Duc, 2008] Zimmermann, A. and Le Duc, C. (2008). Reasoning with a network of aligned ontologies. In Calvanese, D. and Lausen, G., editors, *Web Reasoning and Rule Systems*, pages 43–57, Berlin, Heidelberg. Springer Berlin Heidelberg. (Cited in page 90.)

[Zong et al., 2015] Zong, N., Nam, S., Eom, J.-H., Ahn, J., Joe, H., and Kim, H.-G. (2015). Aligning ontologies with subsumption and equivalence relations in linked data. *Knowledge Based Systems*, 76(1):30–41. (Cited in page 59.)