Towards automated complex ontology alignment

using rule-based machine learning

by

Lu Zhou

B.S., Dalian Jiaotong University, 2011

M.S., Wright State University, 2014

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Computer Science
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Abstract

An ontology usually serves as the schema of a knowledge graph, which provides a vocabulary describing one or many domains of discourse and a specification of the meaning of terms in that vocabulary. Different parties would in general adopt different ontologies. And each ontology may have its own data vocabulary, modeling philosophy, and even language, which makes the semantic data integration process very challenging.

To facilitate interoperability between different organizations, ontology alignment has been considered as the silver bullet for many applications. Ontology alignment has been studied for over a decade, and over that time many alignment systems and methods have been developed by researchers in order to find simple 1-to-1 equivalence matches between two ontologies. However, the simple correspondences are not expressive enough to fully cover the different types of heterogeneities in real-world problems. And very few alignment systems focus on finding complex correspondences.

There are several reasons for this limitation. First, there are no widely accepted alignment benchmarks that contain such complex relationships. Second, tackling complex alignment is more challenging than finding simple alignment. It also requires experts from different domains to work together to manually generate the alignment, which is extremely time-consuming and inefficient. Third, the traditional evaluation metrics like precision, recall, and f-measure, are not fine-grained enough to evaluate the performance of complex alignment systems. Therefore, it hinders the generation and evaluation of complex ontology alignment systems.

To tackle this problem and advance the development of ontology matching and alignment, we seek to address the problem by first developing potential benchmarks that contain the complex relations from real-world ontologies. We then propose an automated complex

ontology alignment system based on association rule learning to generate not only simple correspondences but also complex ones. The algorithm can also be used in a semi-automated fashion to effectively assist users in finding potential complex alignments that they can then validate or edit. Finally, we evaluate the performance of the proposed algorithm on the benchmarks and analyze the results in detail and provide insights into the field of complex ontology alignment.

Towards automated complex ontology alignment

using rule-based machine learning

by

Lu Zhou

B.S., Dalian Jiaotong University, 2011

M.S., Wright State University, 2014

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Computer Science
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Approved by:

Major Professor
Dr. Pascal Hitzler

Copyright

Abstract

An ontology usually serves as the schema of a knowledge graph, which provides a vo-cabulary describing one or many domains of discourse and a specification of the meaning of terms in that vocabulary. Different parties would in general adopt different ontologies. And each ontology may have its own data vocabulary, modeling philosophy, and even language, which makes the semantic data integration process very challenging.

To facilitate interoperability between different organizations, ontology alignment has been considered as the silver bullet for many applications. Ontology alignment has been studied for over a decade, and over that time many alignment systems and methods have been developed by researchers in order to find simple 1-to-1 equivalence matches between two ontologies. However, the simple correspondences are not expressive enough to fully cover the different types of heterogeneities in real-world problems. And very few alignment systems focus on finding complex correspondences.

There are several reasons for this limitation. First, there are no widely accepted align-ment benchmarks that contain such complex relationships. Second, tackling complex align-ment is more challenging than finding simple alignment. It also requires experts from different domains to work together to manually generate the alignment, which is extremely time-consuming and inefficient. Third, the traditional evaluation metrics like precision, re-call, and f-measure, are not fine-grained enough to evaluate the performance of complex alignment systems. Therefore, it hinders the generation and evaluation of complex ontology alignment systems.

To tackle this problem and advance the development of ontology matching and alignment, we seek to address the problem by first developing potential benchmarks that contain the complex relations from real-world ontologies. We then propose an automated complex

ontology alignment system based on association rule learning to generate not only simple correspondences but also complex ones. The algorithm can also be used in a semi-automated fashion to effectively assist users in finding potential complex alignments that they can then validate or edit. Finally, we evaluate the performance of the proposed algorithm on the benchmarks and analyze the results in detail and provide insights into the field of complex ontology alignment.

Table of Contents

List of Figures

List of Tables

Acknowledgments

During this journey, I have had the fortune to be supported by a number of truly incredible people. First, of course, I would like to thank my advisor, Pascal Hitzler, for encouraging me at every step to question and drive forward with my own ideas, and consistently (relentlessly, even) providing me with new opportunities to grow as both an individual and researcher. Also, to my co-advisor Michelle Cheatham, for always being nice and patient to answer all my naive questions, to my committee members, William Hsu, Lior Shamir. Also to the many collaborators, Elodie Thiéblin, Cassia Trojahn, Ondřej Zamazal, Daniel Faria, and Catia Pesquita, who have pushed my boundaries in so many new and interesting directions.

To all of my friends, for supporting me even when I can't make the gathering for many times.

A special thanks to my family, for giving me the foundation to see this through to the end and quiet support inspired me to make every step count, even when I hesitated.

Chapter 1

Introduction

1.1 The Semantic Web

The Semantic Web is an extension of the World Wide Web (WWW), which is commonly known as the Web, through standards set by the World Wide Web Consortium (W3C).[1] The goal of the Semantic Web is to make Internet data machine-readable, whereas the resources on the Web are mostly limited to HyperText Markup Language (HTML) static pages, which is intended for human readability. To enable the encoding of semantics with the data, technologies such as Resource Description Framework (RDF)[2] and Web Ontology Language (OWL)[3] are used. These technologies are used to formally represent metadata and knowledge. For example, ontology can describe concepts, relationships between entities, and categories of things. These embedded semantics offer significant advantages such as reasoning over data and operating with heterogeneous data sources.

---

[1] https://www.w3.org/
[2] https://www.w3.org/TR/rdf-syntax-grammar/
[3] https://www.w3.org/TR/owl-features/

1.2 Knowledge Graph

A knowledge graph is a knowledge base that uses a graph-structured data model or topology to integrate data. Knowledge graphs are often used to store interlinked descriptions of entities, such as objects, events, situations, or abstract concepts with free-form semantics.[4] Since the development of the Semantic Web, knowledge graphs are often associated with Linked Open Data (LOD) projects, focusing on the connections between concepts and entities [1]. There are many knowledge graphs developed and published already. And we can roughly distinguish two types of knowledge graphs in practice, which are *open knowledge graphs* and *enterprise knowledge graphs* [2]. Open knowledge graphs are published online, making their content accessible for the public good. The most prominent examples are DBpedia [3], Freebase [4], Yago [5], and Wikidata [6]. These open knowledge graphs cover many domains and are either extracted from Wikipedia or built by communities of volunteers. On the other hand, enterprise knowledge graphs are typically internal to a company and applied for commercial use-cases. A variety of companies have announced the creation of their internal and proprietary knowledge graphs, such as Google [7], and Bing [8] for search engines, Wolfram Alpha,[5] Apple's Siri,[6] and Amazon Alexa [9] for knowledge-engines and question-answering services, and LinkedIn [10] and Facebook [11] for social networks.

1.3 Knowledge Graph Integration

The real power of knowledge graphs comes when the users transform their own data into RDF triples and then connect their proprietary knowledge to open global knowledge. For example, suppose two organizations would like to collaborate by sharing their knowledge graphs, or they want to enrich their knowledge graph by merging some small knowledge graphs or common resources like Wikidata. What kind of problems will we experience when

---

[4]https://www.ontotext.com/knowledgehub/fundamentals/what-is-a-knowledge-graph/
[5]https://www.wolframalpha.com/
[6]https://www.apple.com/siri/

integrating these knowledge graphs to enable the opportunities of interoperability?

Knowledge graph integration can be separated into two levels, which are schema level and instance level. First, an ontology usually serves as the schema of a knowledge graph, which provides a vocabulary describing a domain of interest and a specification of the meaning of terms in that vocabulary. Different parties would in general adopt different ontologies. Furthermore, each ontology may have its own data vocabulary, modeling philosophy, and even language, which makes the integration process very challenging.

In terms of instance level, the data integrated into the knowledge graphs are usually extracted from a variety of sources. Some refinements of the knowledge graphs, like adding new links and consolidating duplicate entities, are also very important steps to improve the quality of knowledge graphs.

1.4   Ontology Matching and Alignment

In this dissertation, we focus on the schema level matching as called ontology matching. Ontology matching has been researched for a long time. It aims to find the alignment from one ontology to another ontology and then merge the two ontologies into a larger ontology based on the alignment between them. We define ontology matching as the process of generating an alignment A between two ontologies: a source ontology O and a target ontology O', as in [12]. A is directional, denoted $A_{O \to O'}$, and is a set of correspondences $\langle e, e',$ $r, s \rangle$. Each correspondence contains a relation r (e.g., equivalence ($\equiv$), subsumption ($\leq, \geq$)) between two members e and e', and s expresses the strength or confidence (in [0;1]) of this correspondence. Each member can be a single ontology entity (class, object property, data property, individual, value) of respectively O and O' or a more complex construction that is composed of some entities using constructors or transformation functions. A constructor is a logic constructor, e.g., union, intersection, inverse or a restriction constructor, e.g., cardinality restriction, type restriction, and value restriction. A transformation function is

a function that modifies the values of a literal field. It can be an aggregation function, e.g., string concatenation, sum of integers, or a conversion function, e.g., metric conversion, etc.

Most of the current state of the art automated ontology alignment systems still focus on finding 1-to-1 simple alignments. But, Nowadays, more and more data from the real-world have been populated into the ontologies and published as Linked Open Data (LOD). The applications that utilizing these ontologies require ontology alignment and data integration. Due to the complexity of the ontologies, only identifying simple alignment is not enough to fulfill the growing demand for these applications. Therefore, there is a growing awareness that we desperately need to study complex alignment rules. Earlier works also have introduced the need for complex alignments [13, 14]. These correspondences may stand for equivalence as well as other relations, like subsumption or disjointness between ontology entities. Ontology entities usually denote the named entities of ontologies, such as classes, properties, or individuals. An alignment consists of a set of correspondences between two ontologies. There are generally two types of correspondences, which are simple correspondence and complex correspondence.

Simple Correspondence. Simple correspondence refers to a basic 1-to-1 simple mapping between two ontologies, in which the entities involved may be either classes or properties. This category not only includes 1-to-1 equivalence relations, but also 1-to-1 subsumption and 1-to-1 disjointness. For example, In Figure 1.1, `People` in Ontology A is equivalent to class `Human` in Ontology B. The property `hasName` in ontology A is a superProperty of the property `hasFirstName` in ontology B.

Complex Correspondence. Complex correspondence refers to more complex patterns, such as 1-to-n equivalence, 1-to-n subsumption, m-to-1 equivalence, m-to-1 subsumption, m-to-n equivalence, m-to-n subsumption, and m-to-n arbitrary relationship. For example, In Figure 1.1, a `Professor` with a `hasRank` property value of "assistant" in ontology A is equivalent to the class `AssistantProfessor` in ontology B. Obviously, complex correspondence can contain many entities from both ontologies.
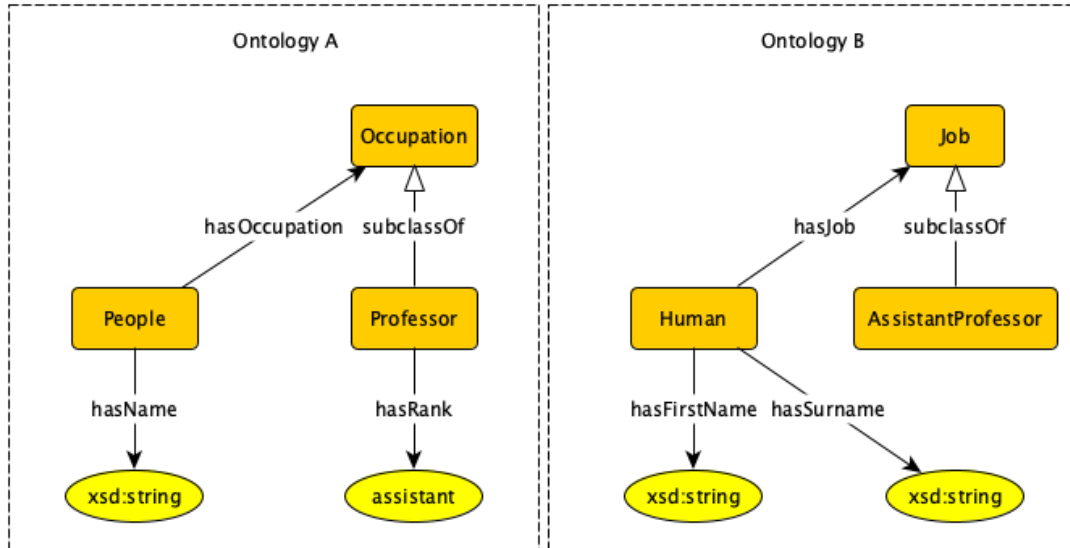
Figure 1.1: *An Example of Simple and Complex Correspondences*

1.5 Outline

This document is a cumulative dissertation that details the research towards advancing the automated complex ontology alignment methodology and evaluation. As mentioned in the above introduction, this is done to facilitate the development of ontology matching and alignment in the sense of complex alignment in the real-world scenario. This work can be divided into three research topics that incrementally build towards this ultimate goal. The dissertation is organized as follows:

Chapter 2 discusses the first research topic: detecting the complex relations from the heterogeneous ontologies and establishing the benchmarks as the gold standard. There are no widely accepted and applicable benchmarks that contain complex relations. In addition, there is also no clear guidance and a comprehensive list of the types of complex alignment in the ontology matching community. We therefore propose to search for the complex relations in the real-world datasets since these datasets are developed to solve the practical problem but not tailored for benchmarking. It is therefore of a rather unique nature and will inform complex ontology alignment research from a practical and applied perspective, rather than

artificial laboratory-like. The primary contributions reference in this section are

- A Complex Alignment Benchmark: Geolink dataset [15]

- GeoLink Dataset: A Complex Alignment Benchmark from Real-world Ontology [16]

- The First Version of the OAEI Complex Alignment Benchmark [17]

- The Enslaved Dataset: A Real-world Complex Ontology Alignment Benchmark using Wikibase [18]

Chapter 3 discusses the second research topic: creating an automated complex alignment system to effectively uncover as many complex alignments as possible. In this chapter, it first introduces the state of the art complex ontology alignment systems, in particular, their methodology, performance, and limitation. Then, it presents our contributions to automatically detecting the complex ontology alignments. The contributions are as follows:

- Alignment of surface water ontologies: a comparison of manual and automated approaches [19]

- A Journey From Simple to Complex Alignment on Real-World Ontologies [20]

- Towards Association Rule-Based Complex Ontology Alignment [21]

Chapter 4 discusses the third research topic: fairly evaluating the performance of complex alignment systems and providing more useful information than the traditional evaluation metrics. By far, classical precision and recall are the most widely used evaluation metrics to assess performance in the majority of existing work on ontology alignment. However, several complications arise in the use of these metrics when the alignments contain complex relations due to their all-or-nothing syntactic comparisons of individual mappings, which do not distinguish between correspondences that are formally incorrect but closely related to the correct correspondences and those that are completely incorrect. We propose to break the

evaluation process down into two subtasks, which are entity identification and relationship identification. This evaluation method has been used in Ontology Alignment Evaluation Initiative (OAEI) campaigns since 2019. In addition, we also discussed the strengths and weaknesses of some potential evaluation methods, but the automated evaluation of complex ontology alignment is still in the exploratory stage. The contributions are listed as follows:

- Ontology Alignment Evaluation Initiative 2018,[7] 2019,[8], 2020[9]

- GeoLink Dataset: A Complex Alignment Benchmark from Real-world Ontology [16]

- The Enslaved Dataset: A Real-world Complex Ontology Alignment Benchmark using Wikibase [18]

- Towards Association Rule-Based Complex Ontology Alignment [21]

- Towards evaluating complex ontology alignments [22]

Chapter 5 concludes the overall contributions with a discussion of potential future work in this area.

---

[7]http://oaei.ontologymatching.org/2018/results/complex/index.html
[8]http://oaei.ontologymatching.org/2019/results/complex/index.html
[9]http://oaei.ontologymatching.org/2020/results/complex/index.html

Chapter 2

Complex Ontology Alignment Benchmarks Generation

2.1 Overview

Benchmarks come in at least two varieties. On the one hand, there are artificial benchmarks that provide a kind-of laboratory setting for evaluation. On the other hand, there are benchmarks created from data as it is used in realistic use cases or even deployed scenarios. Both of these types are important, and they cover different aspects of the spectrum and may have different advantages. Artificial benchmarks can be made to be balanced or to focus on certain aspects of a problem, and sometimes they can be used to test scalability issues more easily as different versions of the same benchmark set may be easily producible. Natural benchmarks, on the other hand, may expose issues arising in practice which may easily be overlooked by designers of artificial benchmarks, in particular in a young field such as complex ontology alignment. Natural benchmarks also may come with an independently verified gold standard baseline.

For this topic, We formulated the following research questions considering these concerns are addressed in the next section.

Q1. *Can we find the complex relations in real-world datasets?*

Q2. *How do we establish benchmarks that contain complex relations?*

Q3. *How do we maintain or improve the benchmarks?*

2.2 Contributions

This section connects the individual contributions to the above research questions.

[15] Lu Zhou, Michelle Cheatham, Adila Krisnadhi, and Pascal Hitzler. A complex alignment benchmark: Geolink dataset. In Denny Vrandecic, Kalina Bontcheva, Mari Carmen Suárez-Figueroa, Valentina Presutti, Irene Celino, Marta Sabou, Lucie-Aimée Kaffee, and Elena Simperl, 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, 2018

This contribution first addresses Research Question Q1. We use the datasets collected from the project called GeoLink [23] and was funded under the U.S. National Science Foundation's EarthCube initiative. This planned decade-long endeavor is a recognition that oftentimes the most innovative and useful discoveries come at the intersection of traditional fields of research. This is particularly true in the geosciences, which often bring together disparate groups of researchers such as geologists, meteorologists, climatologists, ecologists, archaeologists, and so on. For its part, GeoLink employs semantic web technologies to support data representation, sharing, integration, and discovery [24]. In particular, seven diverse geoscience datasets have been brought together into a single data repository. In this contribution, Adila Krisnadhi created the two ontologies for the data repository, which are called GeoLink Base Ontology(GBO) and GeoLink Modular Ontology(GMO). The simple and complex reference alignment between these two ontologies were established in consultation with the domain experts and ontology engineers to guarantee high quality. Therefore, Q1 can be considered as being answered successfully. In order to address Research Question

Q2, Lu Zhou collected all the data, including the ontologies and the reference alignment, and categorized the simple and complex alignment based on their patterns found in this dataset. Lu Zhou then utilized a format called Expressive and Declarative Ontology Alignment Language (EDOAL) that is provided by the Alignment API [25] to express these types of complex relations. This format can be read and manipulated programmatically using the Alignment API and is therefore very convenient for ontology alignment researchers. Since EDOAL can be difficult for humans to parse quickly, the alignments are also expressed in using a naive rule syntax. The rule presentation is not intended for programmatic manipulation, but rather to make it easier for humans to read and understand the alignments. Michelle Cheatham and Pascal Hitzler both provided feedback through all the steps.

[16] Lu Zhou, Michelle Cheatham, Adila Krisnadhi, and Pascal Hitzler. Geolink data set: A complex alignment benchmark from real-world ontology. *Data Intell.*, 2(3):353–378, 2020

This contribution further addresses Research Question Q2 and Q3. It is the extended version of the previous contribution. In addition to the contributions introduced above, in order to extend the functionality of the benchmark for instance-based applications or alignment systems, Lu Zhou sampled 74k instance data collected from the data providers in a triple format <Subject, Predicate, Object> and populated them into the two ontologies. With respect to the maintenance of the benchmark, the proposed benchmark has been incorporated into the Ontology Alignment Evaluation Initiative (OAEI), which is a yearly event that provides system developers with various tracks that evaluate performance on different facets of the problem, such as instance matching, large ontology matching, and interactive matching, among others. Lu Zhou has been actively involved in the OAEI complex alignment track for years. We thus have an intrinsic interest in keeping the benchmark maintained and usable, which would, for example, mean that we are prepared to transfer it to a new benchmarking framework if required in the future. At the same time, based on participants' feedback, we will modify the reference alignment if necessary to perfect

the benchmark by making it more convenient to use. This may involve, e.g., making the alignment available in additional formats.

[17] Élodie Thiéblin, Michelle Cheatham, Cássia Trojahn dos Santos, Ondrej Zamazal, and Lu Zhou. The first version of the OAEI complex alignment benchmark. In Marieke van Erp, Medha Atre, Vanessa López, Kavitha Srinivas, and Carolina Fortuna, editors, *Proceedings of the ISWC 2018 Posters & Demonstrations, Industry and Blue Sky Ideas Tracks co-located with 17th International Semantic Web Conference (ISWC 2018), Monterey, USA, October 8th - to - 12th, 2018*, volume 2180 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2018

This contribution presents the first version of the OAEI complex track. It is composed of four benchmarks from different domains. Besides our proposed Geolink benchmark. Elodie Thiéblin, Cassia Trojahn, and Ondřej Zamazal generated the synthetic conference benchmark that contains complex relations. Lu Zhou and Michelle Cheatham also contributed to this benchmark generation by verifying the reference alignment. In addition, Michelle Cheatham and Lu Zhou proposed a hydrography benchmark in the surface water domain. Elodie Thiéblin and Cassia Trojahn also generated a Taxon benchmark from the plant taxonomy domain. The complex matching track starts to open a new perspective in the field of ontology matching.

[18] Lu Zhou, Cogan Shimizu, Pascal Hitzler, Alicia M. Sheill, Seila Gonzalez Estrecha, Catherine Foley, Duncan Tarr, and Dean Rehberger. The enslaved dataset: A real-world complex ontology alignment benchmark using wikibase. In *29th ACM International Conference on Information and Knowledge Management*, 2020

This contribution further addresses the Research Questions Q1 - Q3. The previous benchmarks are established based on two or more domains or proprietary ontologies. To the best of our knowledge, this contribution is the first time an OWL ontology had been mapped onto a Wikibase installation. Wikibase is the powerful knowledge base software that drives Wikidata [26]. Wikidata is an immense, crowdsourced knowledge base with persistent

data that is available for public use and consumption. It would be very difficult to have an ontology of everything, but Wikidata is probably close enough for this purpose. It contains millions of pieces of knowledge from many different domains in the world. In addition, Wikidata is crowdsourced and can act as a *common resource.*[1] People can export data to Wikidata so that it is publicly persistent in an open and transparent manner. Wikidata is an instance of Wikibase. Any organization can adapt it to their own needs, including setting up their own Wikibase repositories to host their data under different licenses, so that the other instances of Wikibase can be linked with the data on Wikidata. Therefore, it is crucial to be able to find alignments between domain or proprietary ontologies and this common resource. It is also a fact that some organizations have their own internal and proprietary knowledge graphs. They can apply their alignments to this public resource as an important tool to augment or induce new information into their own knowledge graph. And this is one of the main motivations we developed a complex alignment benchmark using Wikibase. The benchmark is based on an ongoing project entitled Enslaved: Peoples of the Historical Slave Trade[2] and funded by The Andrew W. Mellon Foundation where the focus is on tracking the movements and details of peoples in the historical slave trade. It further acts as an interchange format between several heterogeneous data formats among other projects in the digital humanities, because it has built a proof of concept for a slave data hub. At the heart of the project, the Enslaved Hub allows students, researchers, and the general public to search over numerous databases to reconstruct the lives of individuals who were part of the historical slave trade. The Enslaved project leverages Linked Open Data (LOD) techniques, including the use of Wikibase to create an innovative and compelling centralized Hub for engaging with the historical slave trade data from a variety of sources.

---

[1]https://en.wikipedia.org/wiki/Commons
[2]https://enslaved.org/

Chapter 3

Automated Complex Ontology Alignment Systems Generation

3.1 Overview

Ontology alignment has been studied for over a decade. Over that time, many alignment systems have been developed by researchers in order to find simple 1-to-1 equivalence alignments between ontologies. However, finding complex alignments, i.e., alignments that are not simple class or property equivalences, is a topic largely unexplored but with growing significance. Currently, establishing a complex alignment requires domain experts to work together to manually generate the alignment, which is extremely time-consuming and labor-intensive.

Most work associated with evaluating the performance of ontology alignment systems has been done in conjunction with the Ontology Alignment Evaluation Initiative (OAEI)[1]. These yearly events allow developers to test their alignment systems on various tracks that evaluate performance on different facets of the problem such as instance matching, large ontology matching, and interactive matching, among others. The first version of the complex

---

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

alignment track [17] opened new perspectives in the field of ontology matching. There are currently only two automated ontology alignment systems that focus on detecting complex alignments, which are AgreementMakerLightComplex (AMLC) [27] and Complex Alignment Need and A-box based Relation Discovery (CANARD) [28]. From the result of OAEI 2018,[2] AMLC and CANARD were able to generate correct complex correspondences on the complex Conference and Taxon benchmarks. However, the correct number of mappings found was quite limited. AMLC focuses on computing lexical or terminological similarity to decide on complex alignments, while the CANARD system relies on the instance data. In addition, the current version of CANARD is limited to finding complex correspondences that involve only classes. Therefore, we seek to explore the methods to improve the automated alignment algorithms to overcome the limitations existing in the current systems. In particular, we came up with some Research Questions listed as follows:

Q4 . *What is the performance of the traditional alignment systems on the complex alignment benchmarks?*

Q5 . *How do we utilize the traditional alignment systems in the complex alignment task?*

Q6 . *How can we effectively generate complex alignment in automated or semi-automated fashion?*

3.2   Contributions

[19] Michelle Cheatham, Dalia Varanka, Fatima Arauz, and Lu Zhou. Alignment of surface water ontologies: a comparison of manual and automated approaches. *J. Geogr. Syst.*, 22(2):267–289, 2020

In order to address the Research Question Q4, this contribution selected two traditional automatic alignment systems, which are AgreementMakerLight (AML) [29] and LogMap

---

[2]http://oaei.ontologymatching.org/2018/results/complex/index.html

[30] because of their strong performance in the OAEI over several years. We used these two automated alignment systems to perform the same alignment tasks on the Hydrography complex alignment benchmark. AML allows users to select a set of different matchers, e.g., string matcher, instance-based matcher, and background knowledge matcher, to run (or does so automatically based on a profile of the ontologies to be matched) and runs each matcher individually. The resulting alignments are combined using a greedy selection strategy and any logical inconsistencies are removed. Matchers available within AML include lexical and structural algorithms, as well as approaches that leverage background information such as from WordNet or domain-specific lexicons. On the other hand, LogMap compares two entities based on their ISUB, a flexible tool that computes a similarity score for any pair of input strings, and scope (i.e. the degree of overlap of their neighborhoods). Additionally, LogMap's approach to ontology alignment heavily involves consideration of whether or not a relation would conflict with another relation that has a higher confidence value. For example, the system either filters out or more carefully scrutinizes what it calls "dangerous" and logically inconsistent relations.

Our results show that existing alignment systems do not perform as well in this domain as they do on standard ontology alignment benchmarks. In addition, no current systems were able to find relations other than 1-to-1 equivalences. This is particularly true in the surface water domain because such ontologies frequently have less syntactic and structural (due to differing levels of abstraction) similarity than ontologies in other domains that have been a focus for alignment system developers.

[20] Lu Zhou. A journey from simple to complex alignment on real-world ontologies. In Sabrina Kirrane and Lalana Kagal, editors, *Proceedings of the Doctoral Consortium at ISWC 2018 co-located with 17th International Semantic Web Conference (ISWC 2018), Monterey, USA, October 8th - to - 12th, 2018.*, volume 2181 of *CEUR Workshop Proceedings*, pages 93–101. CEUR-WS.org, 2018

[21] Lu Zhou, Michelle Cheatham, and Pascal Hitzler. Towards association rule-based com-

plex ontology alignment. In Xin Wang, Francesca Alessandra Lisi, Guohui Xiao, and Elena Botoeva, editors, *Semantic Technology - 9th Joint International Conference, JIST 2019, Hangzhou, China, November 25-27, 2019, Proceedings*, volume 12032 of *Lecture Notes in Computer Science*, pages 287–303. Springer, 2019

These two contributions further addresses the Research Question Q5 and Q6. In cases where instance data is available for both the source and target ontologies being matched, extensional alignment approaches could be considered. Extensional alignment systems consider the overlap between instances when aligning entities at the schema level. We proposed to create an extensional matcher that leverages the FP-Growth frequent itemset mining algorithm [31] to generate rules that can be stored in lieu of the triples they are based on. While FP-Growth seeks to find any rules that can be used to find the relevance of entities, it is possible that some of these rules represent meaningful semantic relations that hold between entities. Because the FP-Growth algorithm underlying association rules mining can generate a very large number of rules, some mechanism must be put in place to choose the more semantically meaningful rules. Our planned approach for this is to choose rules that involve the entities suggested by traditional alignment systems plus the alignment patterns to finally form the simple and complex alignment. Figure 3.1 illustrates the overview of our proposed algorithm.

We first extract all triples <Subject, Predicate, Object> from the source and target ontologies and keep the triples that contain at least one entity under the source or the target ontology namespace and also the triples that contain rdf:type information. After the filtering process, the system generates the transaction database for the FP-growth algorithm based on all of the remaining triples. Then we replace the object in the triples with its rdf:type,[3] because we focus on generating schema-level (rather than instance-level) mapping rules between two ontologies, and the type information of the object is more meaningful

---

[3]Our evaluation data only has one single type. If there are multiple types of the object, it can also combine the subject and predicate as additional information to determine the correct type, or keep both types as two triples.
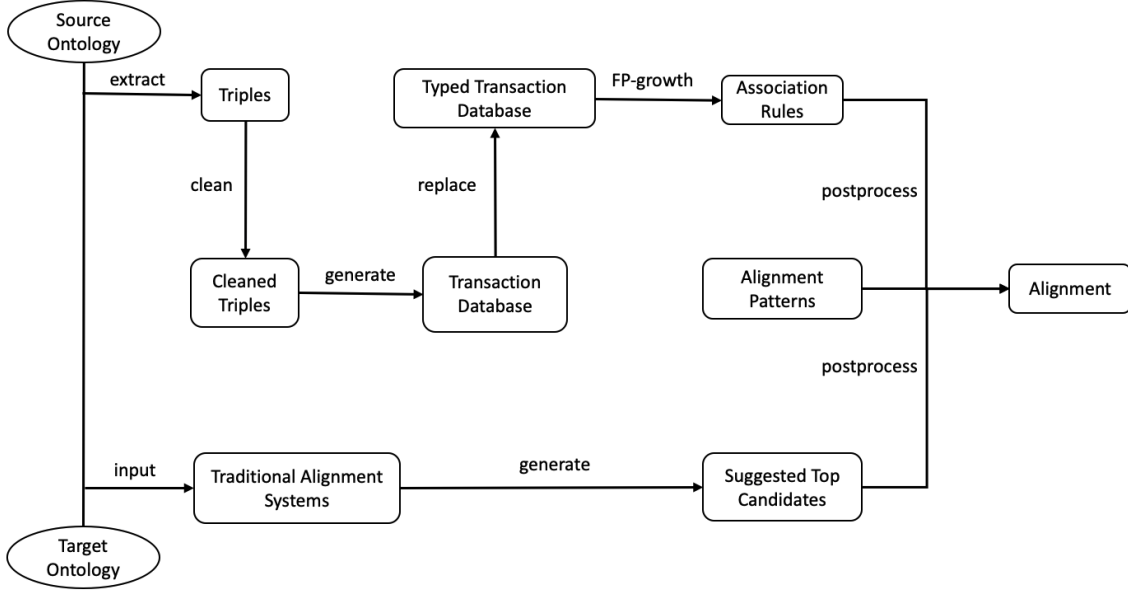
Figure 3.1: *Overview of The Proposed Alignment Algorithm*

than the original URI. We run the FP-growth algorithm on the typed transaction database
and generate a set of association rules. Since we are trying to find the mappings *between* two
ontologies, we focus on mining the rules whose antecedent only contains entities from the
source ontology and whose consequent only contains entities from the target ontology. The
association rules tell us which source entities are related to which target entities, but they
do not give us information on *how* those entities are related. In order to determine this,
we analyze the output of the association rule mining step in light of the common alignment
patterns introduced in [32, 15]. These alignment patterns also narrow down a great number
of the association rules plus the suggested top candidates from the traditional alignment
systems and finally generate more accurate alignments.

This algorithm therefore addresses Research Questions Q5 and Q6 by automatically dis-
covering potential complex correspondences that can then be presented to human experts
in order to effectively generate complex alignment between two ontologies with populated
common instance data. We evaluate the performance of our system on two benchmarks

from the complex alignment track of the OAEI 2020,[4] the GeoLink benchmark, which contains around 74k instances from the real-world dataset and the Enslaved benchmark, which contains around 32k triples from the historic slave trade data. The evaluation strategies and results are introduced in the next chapter.

---

[4]http://oaei.ontologymatching.org/2020/complex/index.html

Chapter 4

Complex Alignment Systems Evaluation

4.1 Overview

By far, classical precision and recall are the most widely used evaluation metrics to assess performance in the majority of existing work on ontology alignment. However, several complications arise in the use of these metrics when the alignments contain complex relations due to their all-or-nothing syntactic comparisons of individual mappings, which do not distinguish between correspondences that are formally incorrect but closely related to the correct correspondences and those that are completely incorrect. For example, this is a mapping in the reference alignment of the Geolink benchmark:

$$\mathsf{Award}(x) \wedge \mathsf{hasSponsor}(x, z) \leftrightarrow \mathsf{FundingAward}(x) \wedge \mathsf{providesAgentRole}(x, y)$$
$$\wedge\, \mathsf{SponsorRole}(y) \wedge \mathsf{performedBy}(y, z)$$

Moreover, below are two different mappings that were generated by two different hypothetical complex ontology alignment systems. The first mapping is:

$$\text{Award}(x) \wedge \text{hasSponsor}(x, z) \leftrightarrow \text{FundingAward}(x) \wedge \text{providesAgentRole}(x, y)$$
$$\wedge \text{AgentRole}(y) \wedge \text{performedBy}(y, z)$$

and the second one is:

$$\text{Award}(x) \wedge \text{hasSponsor}(x, z) \rightarrow \text{Program}(x) \wedge \text{providesAgentRole}(x, y)$$
$$\wedge \text{DataManagerRole}(y) \wedge \text{performedBy}(y, z)$$

The first mapping is formally incorrect compared to the reference alignment, but it is very closely related to it because `SponsorRole` is a subclass of `AgentRole`. Conversely, the second one is completely incorrect, as it contains the incorrect domain and range restrictions of `providesAgentRole` and the relationship between the two sides indicates subsumption rather than an equivalence relation. While the first system should be penalized for not producing the correct relation, but considering this result as completely incorrect lacks important nuance. For instance, this relation could be relatively easily corrected by a user in a semi-automated alignment system. Moreover, the alignment system developer would likely benefit from knowing how close the system came to generating the correct output in this case. Therefore, we seek to answer the following Research Questions:

Q7 . *How can we fairly evaluate the performance of complex alignment systems by providing the system developers with detailed clues to improve the algorithms?*

Q8 . *How can we work towards completely automated or semi-automated evaluation of the complex alignment?*

4.2   Contributions

[16] Lu Zhou, Michelle Cheatham, Adila Krisnadhi, and Pascal Hitzler. Geolink data set: A complex alignment benchmark from real-world ontology. *Data Intell.*, 2(3):353–378, 2020

[18] Lu Zhou, Cogan Shimizu, Pascal Hitzler, Alicia M. Sheill, Seila Gonzalez Estrecha, Catherine Foley, Duncan Tarr, and Dean Rehberger. The enslaved dataset: A real-world complex ontology alignment benchmark using wikibase. In *29th ACM International Conference on Information and Knowledge Management*, 2020

[21] Lu Zhou, Michelle Cheatham, and Pascal Hitzler. Towards association rule-based complex ontology alignment. In Xin Wang, Francesca Alessandra Lisi, Guohui Xiao, and Elena Botoeva, editors, *Semantic Technology - 9th Joint International Conference, JIST 2019, Hangzhou, China, November 25-27, 2019, Proceedings*, volume 12032 of *Lecture Notes in Computer Science*, pages 287–303. Springer, 2019

In order to assess the quality of the complex alignment and address Research Question Q7, the first three contributions evaluated the complex alignments by looking into two dimensions. First, we evaluate if the mapping contains the correct entities that should be involved based on the reference alignment. Another dimension is the relationship between the entities, like equivalence and subsumption. Based on this, we break the evaluation procedure down into two subtasks.

Entity Identification: For each entity in the source ontology, the alignment systems will be asked to list all of the entities that are related in some way in the target ontology. For example, in this example,

$$\mathsf{Award}(x) \wedge \mathsf{hasSponsor}(x, z) \leftrightarrow \mathsf{FundingAward}(x) \wedge \mathsf{providesAgentRole}(x, y)$$
$$\wedge \mathsf{SponsorRole}(y) \wedge \mathsf{performedBy}(y, z),$$

the expected output from an alignment system is that `hasSponsor` in the GBO is related

Table 4.1: *Similarity for Relationship Identification*

Found Relation	Correct Relation	Similarity	Comment
$=$	$=$	1	correct relation
$\subset$	$\subset$	1	correct relation
$\supset$	$\supset$	1	correct relation
$\subset$	$=$	0.8	return less information, but correct
$=$	$\supset$	0.8	return less information, but correct
$\supset$	$=$	0.6	return more information, but incorrect
$=$	$\subset$	0.6	return more information, but incorrect
$\subset$	$\supset$	0.3	incorrect relation
$\supset$	$\subset$	0.3	incorrect relation

to `FundingAward`, `providesAgentRole`, `SponsorRole` and `performedBy` in the GMO and `Award` in the GBO. Based on the two lists of entities from the reference alignment and the matcher, precision, recall, and f-measure can be calculated.

Relationship Identification: In terms of the example above, an alignment system eventually needs to determine that the relationship between two sides is an equivalence. At this stage, we would like to focus on some relationships that happen frequently in the datasets, which are equivalence and subsumption. We will further assess other complex relationships in future work. Table 4.1 shows the different similarities for different situations. We slightly penalize differently for the situations in finding less information, but all the information returned is correct, and finding more information, but part of the information is incorrect. We do not penalize the incorrect relationship by giving a ZERO value because that would completely neglect the entity identification outputs without considering whether it is a reasonable result or a completely incorrect one. In order to generate the final results, we multiply the results from the entity identification by the penalty of the relations.[1] The formulas for computing the final results are as follows:

`Relaxed_precision = Precision_entity` $\times$ `Relation_similarity`

`Relaxed_recall = Recall_entity` $\times$ `Relation_similarity`

`Relaxed_f-measure = F-measure_entity` $\times$ `Relation_similarity`

---

[1]To be accurate, it could also have been better aggregated with other aggregation functions rather than multiplication [33]. But we do not focus on this question in this dissertation.

[22] Lu Zhou, Élodie Thiéblin, Michelle Cheatham, Daniel Faria, Catia Pesquita, Cássia Trojahn dos Santos, and Ondrej Zamazal. Towards evaluating complex ontology alignments. *Knowl. Eng. Rev.*, 35:e21, 2020

This contribution further addresses Research Question Q8. It surveys existing ontology alignment evaluation metrics and analyzes their strengths and weaknesses with respect to the evaluation of complex alignments when a reference alignment is available or absent and analyzes the requirements for effective evaluation of complex ontology alignments and assesses the degree to which these requirements are met by existing approaches. It also provides a roadmap for future work on this topic taking into consideration emerging community initiatives and major challenges that need to be addressed.

Chapter 5

Conclusion and Future Work

5.1 Summary

Complex ontology alignment has been studied for a long time, but relatively little work has been done to advance the state of the art. The lack of an available complex ontology alignment benchmark is a primary reason for the slow speed of the development. In addition, most current ontology alignment benchmarks have been created by humans for the sole purpose of evaluating ontology alignment systems, and they may not always represent real-world cases.

In this dissertation, we have proposed several complex ontology alignment benchmarks based on real-world knowledge bases. In our benchmarks, the alignments not only cover 1-to-1 simple correspondences but also contain 1-to-n, m-to-1, and m-to-n complex relationships. In addition, the alignment has been evaluated by domain experts from different organizations to ensure high quality. Moreover, instance data has been published into these benchmarks, which is also important in order to support the use of the benchmarks by extensional alignment systems. Furthermore, the ontologies and alignments in both rule syntax and EDOAL format have been published in FigShare with an open-access license for reusability and can be accessed on the OAEI website as well.

In addition, we also proposed a complex alignment algorithm based on association rule mining. Our algorithm takes advantage of instance data to mine frequent patterns, which show us which entities in one ontology are related to which entities in the other. Then we apply common simple and complex patterns and the output of traditional alignment systems to arrange these related entities into the formal alignment. Finally, we proposed a novel evaluation strategy to fairly evaluate alignment systems on the complex alignment benchmarks and analyzed the results in detail to provide a better understanding of the challenges related to complex ontology alignment research.

5.2   Future Work

Some reflections on my experimental findings lead to future work. First, our system relies on instance data for mining the association rules, while the instance data is not available for most of synthetic benchmarks. However, this could possibly be resolved with automated instance data generation in the future to populate the ontologies.

Second, we incorporate some common patterns that have been widely accepted in the ontology alignment community in our proposed system. This could be another limitation since the set of mapping patterns in our system is likely not comprehensive. However, our algorithm is extensible, so more patterns can be easily added in the future as the need arises.

Third, the evaluation methods for the complex alignment are still relatively simple, because we only include limited situations in different relationships, like equivalence and subsumption. However, there are many other complicated situations that we would need to consider. Thus, we plan to work on the improvement of the evaluation methods as well.

All in all, complex ontology alignment has a positive impact on advancing the semantic data integration significantly, moving forward.

Bibliography

[1] Lisa Ehrlinger and Wolfram Wöß. Towards a definition of knowledge graphs. In Michael Martin, Martí Cuquet, and Erwin Folmer, editors, *Joint Proceedings of the Posters and Demos Track of the 12th International Conference on Semantic Systems - SEMAN-TiCS2016 and the 1st International Workshop on Semantic Change & Evolving Semantics (SuCCESS'16) co-located with the 12th International Conference on Semantic Systems (SEMANTiCS 2016), Leipzig, Germany, September 12-15, 2016*, volume 1695 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2016.

[2] Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard de Melo, Claudio Gutierrez, José Emilio Labra Gayo, Sabrina Kirrane, Sebastian Neumaier, Axel Polleres, Roberto Navigli, Axel-Cyrille Ngonga Ngomo, Sabbir M. Rashid, Anisa Rula, Lukas Schmelzeisen, Juan F. Sequeda, Steffen Staab, and Antoine Zimmermann. Knowledge graphs. *CoRR*, abs/2003.02320, 2020.

[3] Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N. Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick van Kleef, Sören Auer, and Christian Bizer. Dbpedia - A large-scale, multilingual knowledge base extracted from wikipedia. *Semantic Web*, 6(2):167–195, 2015.

[4] Kurt Bollacker, Patrick Tufts, Tomi Pierce, and Robert Cook. A platform for scalable, collaborative, structured information integration. In *Intl. Workshop on Information Integration on the Web (IIWeb'07)*, pages 22–27, 2007.

[5] Johannes Hoffart, Fabian M. Suchanek, Klaus Berberich, Edwin Lewis-Kelham, Gerard de Melo, and Gerhard Weikum. YAGO2: exploring and querying world knowledge

in time, space, context, and many languages. In Sadagopan Srinivasan, Krithi Ramamritham, Arun Kumar, M. P. Ravindra, Elisa Bertino, and Ravi Kumar, editors, *Proceedings of the 20th International Conference on World Wide Web, WWW 2011, Hyderabad, India, March 28 - April 1, 2011 (Companion Volume)*, pages 229–232. ACM, 2011.

[6] Denny Vrandecic and Markus Krötzsch. Wikidata: a free collaborative knowledgebase. *Commun. ACM*, 57(10):78–85, 2014.

[7] Amit Singhal. Introducing the knowledge graph: things, not strings. *Official google blog*, 5, 2012.

[8] Saurabh Shrivastava. Bring rich knowledge of people, places, things and local businesses to your apps. bing blogs, 2017.

[9] Xin Luna Dong. Building a broad knowledge graph for products. In *35th IEEE International Conference on Data Engineering, ICDE 2019, Macao, China, April 8-11, 2019*, page 25. IEEE, 2019.

[10] Q He, B Chen, and D Agarwal. Building the linkedin knowledge graph. *Engineering. linkedin. com*, 2016.

[11] Natasha F. Noy, Yuqing Gao, Anshu Jain, Anant Narayanan, Alan Patterson, and Jamie Taylor. Industry-scale knowledge graphs: Lessons and challenges. *ACM Queue*, 17(2):20, 2019.

[12] Jérôme Euzenat and Pavel Shvaiko. *Ontology Matching, Second Edition*. Springer, 2013.

[13] Alexander Maedche, Boris Motik, Nuno Silva, and Raphael Volz. MAFRA - A mapping framework for distributed ontologies. In Asunción Gómez-Pérez and V. Richard Benjamins, 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, 2002.

[14] Pepijn RS Visser, Dean M Jones, Trevor JM Bench-Capon, and MJR Shave. An analysis of ontology mismatches; heterogeneity versus interoperability. In AAAI 1997 Spring Symposium on Ontological Engineering, Stanford CA., USA, pages 164–72, 1997.

[15] Lu Zhou, Michelle Cheatham, Adila Krisnadhi, and Pascal Hitzler. A complex alignment benchmark: Geolink dataset. In Denny Vrandecic, Kalina Bontcheva, Mari Carmen Suárez-Figueroa, Valentina Presutti, Irene Celino, Marta Sabou, Lucie-Aimée Kaffee, and Elena Simperl, 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, 2018.

[16] Lu Zhou, Michelle Cheatham, Adila Krisnadhi, and Pascal Hitzler. Geolink data set: A complex alignment benchmark from real-world ontology. Data Intell., 2(3):353–378, 2020.

[17] Élodie Thiéblin, Michelle Cheatham, Cássia Trojahn dos Santos, Ondrej Zamazal, and Lu Zhou. The first version of the OAEI complex alignment benchmark. In Marieke van Erp, Medha Atre, Vanessa López, Kavitha Srinivas, and Carolina Fortuna, editors, Proceedings of the ISWC 2018 Posters & Demonstrations, Industry and Blue Sky Ideas Tracks co-located with 17th International Semantic Web Conference (ISWC 2018), Monterey, USA, October 8th - to - 12th, 2018, volume 2180 of CEUR Workshop Proceedings. CEUR-WS.org, 2018.

[18] Lu Zhou, Cogan Shimizu, Pascal Hitzler, Alicia M. Sheill, Seila Gonzalez Estrecha, Catherine Foley, Duncan Tarr, and Dean Rehberger. The enslaved dataset: A real-world complex ontology alignment benchmark using wikibase. In 29th ACM International Conference on Information and Knowledge Management, 2020.

[19] Michelle Cheatham, Dalia Varanka, Fatima Arauz, and Lu Zhou. Alignment of surface water ontologies: a comparison of manual and automated approaches. *J. Geogr. Syst.*, 22(2):267–289, 2020.

[20] Lu Zhou. A journey from simple to complex alignment on real-world ontologies. In Sabrina Kirrane and Lalana Kagal, editors, *Proceedings of the Doctoral Consortium at ISWC 2018 co-located with 17th International Semantic Web Conference (ISWC 2018), Monterey, USA, October 8th - to - 12th, 2018.*, volume 2181 of *CEUR Workshop Proceedings*, pages 93–101. CEUR-WS.org, 2018.

[21] Lu Zhou, Michelle Cheatham, and Pascal Hitzler. Towards association rule-based complex ontology alignment. In Xin Wang, Francesca Alessandra Lisi, Guohui Xiao, and Elena Botoeva, editors, *Semantic Technology - 9th Joint International Conference, JIST 2019, Hangzhou, China, November 25-27, 2019, Proceedings*, volume 12032 of *Lecture Notes in Computer Science*, pages 287–303. Springer, 2019.

[22] Lu Zhou, Élodie Thiéblin, Michelle Cheatham, Daniel Faria, Catia Pesquita, Cássia Trojahn dos Santos, and Ondrej Zamazal. Towards evaluating complex ontology alignments. *Knowl. Eng. Rev.*, 35:e21, 2020.

[23] Jia You. Geoscientists aim to magnify specialized web searching, 2015.

[24] Adila Alfa Krisnadhi, Yingjie Hu, Krzysztof Janowicz, Pascal Hitzler, Robert A. Arko, Suzanne Carbotte, Cynthia Chandler, Michelle Cheatham, Douglas Fils, Tim Finin, Peng Ji, Matthew B. Jones, Nazifa Karima, Kerstin A. Lehnert, Audrey Mickle, Tom Narock, Margaret O'Brien, Lisa Raymond, Adam Shepherd, Mark Schildhauer, and Peter Wiebe. The geolink framework for pattern-based linked data integration. In Serena Villata, Jeff Z. Pan, and Mauro Dragoni, editors, *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.*, volume 1486 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2015.

[25] Jérôme David, Jérôme Euzenat, François Scharffe, and Cássia Trojahn dos Santos. The alignment API 4.0. *Semantic Web*, 2(1):3–10, 2011.

[26] Denny Vrandecic and Markus Krötzsch. Wikidata: a free collaborative knowledgebase. *Commun. ACM*, 57(10):78–85, 2014.

[27] Daniel Faria, Catia Pesquita, Booma Sowkarthiga Balasubramani, Teemu Tervo, David Carriço, Rodrigo Garrilha, Francisco M. Couto, and Isabel F. Cruz. Results of AML participation in OAEI 2018. In Pavel Shvaiko, Jérôme Euzenat, Ernesto Jiménez-Ruiz, Michelle Cheatham, and Oktie Hassanzadeh, editors, *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.*, volume 2288 of *CEUR Workshop Proceedings*, pages 125–131. CEUR-WS.org, 2018.

[28] Élodie Thiéblin, Ollivier Haemmerlé, and Cássia Trojahn. CANARD complex matching system: Results of the 2019 OAEI evaluation campaign. In Pavel Shvaiko, Jérôme Euzenat, Ernesto Jiménez-Ruiz, Oktie Hassanzadeh, and Cássia Trojahn, editors, *Proceedings of the 14th International Workshop on Ontology Matching co-located with the 18th International Semantic Web Conference (ISWC 2019), Auckland, New Zealand, October 26, 2019*, volume 2536 of *CEUR Workshop Proceedings*, pages 114–122. CEUR-WS.org, 2019.

[29] Daniel Faria, Catia Pesquita, Emanuel Santos, Matteo Palmonari, Isabel F. Cruz, and Francisco M. Couto. The agreementmakerlight ontology matching system. In Robert Meersman, Hervé Panetto, Tharam S. Dillon, Johann Eder, Zohra Bellahsene, Norbert Ritter, Pieter De Leenheer, and Dejing Dou, editors, *On the Move to Meaningful Internet Systems: OTM 2013 Conferences - Confederated International Conferences: CoopIS, DOA-Trusted Cloud, and ODBASE 2013, Graz, Austria, September 9-13, 2013. Proceedings*, volume 8185 of *Lecture Notes in Computer Science*, pages 527–541. Springer, 2013.

[30] Ernesto Jiménez-Ruiz and Bernardo Cuenca Grau. Logmap: Logic-based and scalable ontology matching. In Lora Aroyo, Chris Welty, Harith Alani, Jamie Taylor, Abraham Bernstein, Lalana Kagal, Natasha Fridman Noy, and Eva Blomqvist, editors, *The Semantic Web - ISWC 2011 - 10th International Semantic Web Conference, Bonn, Germany, October 23-27, 2011, Proceedings, Part I*, volume 7031 of *Lecture Notes in Computer Science*, pages 273–288. Springer, 2011.

[31] J. Han et al. Mining frequent patterns without candidate generation: A frequent-pattern tree approach. *Data Min. Knowl. Discov.*, 8(1):53–87, 2004.

[32] Dominique Ritze, Christian Meilicke, Ondrej Sváb-Zamazal, and Heiner Stuckenschmidt. A pattern-based ontology matching approach for detecting complex correspondences. In Pavel Shvaiko, Jérôme Euzenat, Fausto Giunchiglia, Heiner Stuckenschmidt, Natalya Fridman Noy, and Arnon Rosenthal, 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, 2009.

[33] Marc Ehrig and Jérôme Euzenat. Relaxed precision and recall for ontology matching. In Benjamin Ashpole, Marc Ehrig, Jérôme Euzenat, and Heiner Stuckenschmidt, editors, *Integrating Ontologies '05, Proceedings of the K-CAP 2005 Workshop on Integrating Ontologies, Banff, Canada, October 2, 2005*, volume 156 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2005.

Appendix A

Contributions

In the following pages, all contributions made for this dissertation are listed in order of the appearance in this document.

A Complex Alignment Benchmark: GeoLink Dataset

Lu Zhou[1], Michelle Cheatham[1], Adila Krisnadhi[2], and Pascal Hitzler[1]

[1] DaSe Lab, Wright State University, Dayton OH 45435, USA,
{zhou.34, michelle.cheatham, pascal.hitzler}@wright.edu
[2] Faculty of Computer Science, Universitas Indonesia,
Depok, Jawa Barat 16424, Indonesia
adila@cs.ui.ac.id

Abstract. Ontology alignment has been studied for over a decade, and over that time many alignment systems and methods have been developed by researchers in order to find simple 1-to-1 equivalence matches between two ontologies. However, very few alignment systems focus on finding complex correspondences. One reason for this limitation may be that there are no widely accepted alignment benchmarks that contain such complex relationships. In this paper, we propose a real-world dataset from the GeoLink project as a potential complex alignment benchmark. The dataset consists of two ontologies, the GeoLink Base Ontology (GBO) and the GeoLink Modular Ontology (GMO), as well as a manually created reference alignment, that were developed in consultation with domain experts from different institutions. The alignment includes 1:1, 1:n, and m:n equivalence and subsumption correspondences, and is available in both EDOAL and rules syntax.

1 Introduction

Ontology alignment is an important step in enabling computers to query and reason across the many linked datasets on the semantic web. This is a difficult challenge because the ontologies underlying different linked datasets can vary in terms of subject area coverage, level of abstraction, ontology modeling philosophy, and even language. Due to the importance and difficulty of the ontology alignment problem, it has been an active area of research for over a decade [12].

Ideally, alignment systems should be able to uncover any entity relationships across two ontologies that can exist within a single ontology. Such relationships have a wide range of complexity, from basic 1-to-1 equivalence, such as a Person in one ontology being equivalent to a Human in another ontology, to arbitrary m-to-n relationships, such as a Professor with a hasRank property value of "Assistant" in one ontology being a subclass of the union of the Faculty and TenureTrack classes in another. Unfortunately, though, the majority of the research activities in the field of ontology alignment remains focused on the simplest end of this scale – finding 1-to-1 equivalence relations between ontologies.

33

Part of the reason for this may be that there are no widely used and accepted ontology alignment benchmarks that involve complex relations.

This paper seeks to take a step in that direction by proposing a complex alignment benchmark based on two ontologies which were developed by domain experts jointly with the reference alignment, and which in fact were developed for deployment on major ocean science data repository platforms, i.e., without the actual intention to develop an alignment benchmark. For this reason, the benchmark, including the reference alignment, can be considered to be (a) objective, in that it was created for deployment and not for benchmarking, (b) realistic, in that it captures an application use case developed for deployment, and (c) a valid ground truth alignment, in that the two ontologies and the reference alignment were developed together, by domain experts. We argue that it is therefore of rather unique nature and will inform complex ontology alignment research from a practical and applied, rather than artificial laboratory-like, perspective. The benchmark, coincidently, as this was the requirement of the use case, has a particular focus on relationships involving properties, which is particularly interesting because those have been shown to be rather difficult to handle for current alignment approaches [1].

The main contributions of this paper are therefore the following:

- Presentation of two ontologies to support data representation, sharing, integration, and discovery for the geoscience research domain.
- Creation of an alignment of these two ontologies that includes 1:1, 1:n, and m:n correspondences, and given the creation history and usage of the alignment, it is fair to say that the alignment constitutes a gold-standard reference.
- Publication of the benchmark alignment in both rule and EDOAL[3] syntax at a persistent URL[4] under a CC-BY license.

In addition, we have analyzed and categorized the mapping rules constituting the alignment. We found several which had not been classified or discussed previously, and we will present and discuss our analysis.

This paper is organized as follows. Section 2 discusses the few existing ontology alignment benchmarks that involve relationships other than 1-to-1 equivalence. Section 3 gives further background on the GeoLink modeling process, including why two different but related ontologies were developed. Section 4 discusses the alignment between the two GeoLink ontologies, along with some descriptive statistics and an analysis of the types of mapping rules constituting the alignment. Section 5 concludes with a discussion of potential future work in this area.

---

[3] `http://alignapi.gforge.inria.fr/edoal.html`
[4] `http://doi.org/10.6084/m9.figshare.5907172`

2   Related Work

Most work associated with evaluating the performance of ontology alignment systems has been done in conjunction with the Ontology Alignment Evaluation Initiative (OAEI)[5]. These yearly events allow developers to test their alignment systems on various tracks that evaluate performance on different facets of the problem such as instance matching, large ontology matching, and interactive matching, among others. Currently, most of these tracks involve the identification of 1-to-1 equivalence relationships, such as a Participant being equivalent to an Attendee. In 2009, the OAEI ran an "oriented" matching track that challenged systems to find subsumption relationships such as a Book is a subclass of a Publication. However, this track was abandoned after one year. Some system developers complained that the quality of the reference alignment was low [2]. This frustrated system developers and hindered participation. A discussion at the last two Ontology Matching workshops[6] made it clear that the community is interested in complex alignment, but that lack of applicable benchmarks is hindering progress. Our proposed benchmark seeks to address this concern by providing a reference alignment as a benchmark, and by addressing the quality issue of the previous benchmark by the fact that the process leading to the reference alignment guarantees its high quality.

In addition to using the OAEI benchmark, alignment systems that attempt to identify subsumption relations have sometimes used their own manually developed (and sometimes unpublished) reference alignments [5]. Other subsumption systems have evaluated the precision of their approach by manually validating relations produced by their system, while foregoing an assessment of recall [13]. Other related work has centered on developing a benchmark for compound alignments, which the authors define as mappings between class or property expressions involving more than two ontologies [10]. Their first step in this direction was to create a set of reference alignments containing relations of the form $< X, Y, Z, R, M >$, where X, Y and Z are classes from three different ontologies and R is a relation between Y and Z that results in a class expression that is related to X by the relation M. For example, a DisabledVeteran (X) is equivalent to (M) the intersection (R) of Veteran (Y) and Disabled (Z). This benchmark is based on cross-products among the OBO Foundry biomedical ontologies, which have been manually validated by at least two experts.

The work presented herein differs from these approaches by considering a wider range of relationship types (beyond subsumption and the type of ternary relation described in [10]), as they naturally arose out of the application from which the reference alignment was taken.

More related work is currently being undertaken by Thieblin and her colleagues, who are creating a complex alignment benchmark using the Conference track ontologies within the OAEI [14]. This work is partially completed, and at the time of this writing it covers three of the seven ontologies. In addition, we are

---

[5] http://oaei.ontologymatching.org
[6] http://www.ontologymatching.org/

collaborating with them (under their direction) to complete the dataset and prepare a new task in OAEI to evaluate complex alignment systems. The reference alignment we describe herein differs from the effort by Thieblin et al. in that the GeoLink ontologies and alignment constitute real-world datasets designed for practice and applied by geoscientists, rather than being an artificial artifact designed solely for alignment benchmarking. Furthermore, data from seven geoscience repositories have been published according to the GeoLink schema and they are available online[7]. This instance data can in the future be used by alignment systems that employ extensional matching techniques. In contrast to this, significant instance data is not readily available for most of the OAEI Conference Track ontologies.

3   The GeoLink Modeling Process

Benchmarks come in at least two varieties. On the one hand there are artificial benchmarks which provide a kind-of laboratory setting for evaluation. On the other hand there are benchmarks created from data as it is used in realistic use cases or even deployed scenarios. Both of these types are important, and they cover different aspects of the spectrum, and may have different advantages. Artificial benchmarks can be made to be balanced, or to focus on certain aspects of a problem, and sometimes they can be used to test scalability issues more easily as different versions of the same benchmark set may be easily producible. Natural benchmarks, on the other hand, may expose issues arising in practice which may easily be overlooked by designers of artificial benchmarks, in particular in a young field such as complex ontology alignment. Natural benchmarks also may come with an independently verified gold standard baseline, as in our case.

The project this benchmark arose from is called GeoLink [15] and was funded under the U.S. National Science Foundation's EarthCube initiative. This planned decade-long endeavor is a recognition that oftentimes the most innovative and useful discoveries come at the intersection of traditional fields of research. This is particularly true in the geosciences, which often bring together disparate groups of researchers such as geologists, meteorologists, climatologists, ecologists, archaeologists, and so on. For its part, GeoLink employs semantic web technologies to support data representation, sharing, integration, and discovery [9]. In particular, seven diverse geoscience datasets have been brought together into a single data repository.

At the beginning of the project, some providers' data resided in relational databases while others' had been published as RDF triples and exposed via a SPARQL endpoint. Because each provider had their own schema, the first step in the GeoLink project was to develop a unified schema according to which all data providers could publish their data [9]. Creating a unified schema for independently developed datasets is sometimes difficult, and the final product often ends up requiring providers to shoehorn their data into a schema that does not quite
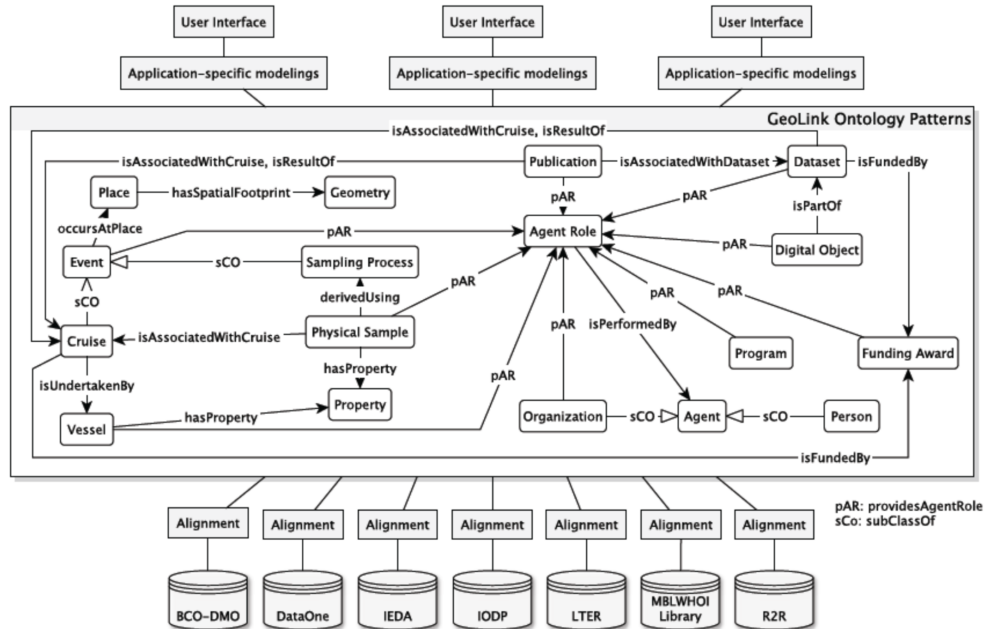
---

[7] `http://data.geolink.org`

Fig. 1: Intended usage of the GMO

fit. GeoLink uses an approach that relies on ontology design patterns (ODP) in an attempt to avoid this issue [4]. An ODP represents a reusable solution to a recurring modeling problem. An ODP generally encodes a specific abstract notion, such as a process, event, agent, etc. These are frequently the small areas of semantic overlap that exist between datasets from different subfields of the same high-level domain. ODPs provide a structured and application-neutral representation of the key concepts within a domain. Throughout the first year of the project, geoscientists, data providers and ontologists worked together to identify and model the important concepts within the geosciences that recurred across two or more datasets. The result of this were what we call ontology modules, based on ODPs, and eventually they were stitched together to form the GeoLink Modular Ontology (GMO) [7].

As shown in Figure 1, the GMO allows data providers to publish only those aspects of their data modeled by the GMO according to that schema. Any data the provider has that is not covered by that schema can be published using the provider's own schema, since no other providers have similar content. For example, in the figure, the provider R2R has data related mostly to the cruise and vessel modules in the lower left of the figure, and so it publishes its related data using that terminology. R2R also has data not modeled by the GMO and so it uses its own terminology when publishing that information. This freedom is intended to make the publishing process easier; however, some problems still remained.
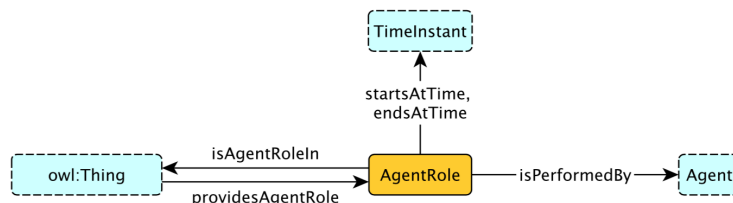
Fig. 2: The Agent Role pattern

Some of the patterns contain a rather complicated structure, mostly due to reification, which was employed to accommodate different perspectives (e.g., based on granularity) on the data. For example, many of the data providers have information about the sponsor of a project, and R2R has a native relation in their schema called hasSponsor with domain Award and range Organization. However following best practices, it leads to a more versatile model if being a sponsor is recognized (and thus modeled) as a role which an agent (in this case an organization) can assume. Creating a distinct relation for each type of role on a project (sponsor, chief scientist, research assistant, etc.) is brittle, in the sense that if new roles will be added later, potentially due to the inclusion of a new dataset, then the schema will need to be edited by adding new vocabulary for new roles together with (possibly complex) role relationships. Another issue with using a relation such as hasSponsor is that a more fine-grained data repository may have additional temporal information related to the sponsor role, and then it is not clear how to add this temporal information to the hasSponsor model without punning. Essentially, hasSponsor should better be expressed as a ternary relation between award, organization, and the type of relation (in this case, being a sponsor) expressed using an individual which can be reused in all sponsor relationships. In terms of ODPs, this is realized by reusing the Agent Role pattern, shown in abstract form in Figure 2. This approach both allows new roles to be added easily (by subclassing AgentRole) and supports temporal queries if desired.

Unfortunately, while the data providers recognized the utility of this modeling approach, they found it cumbersome to map their data to it. Looking at their own schemas, they found nothing equivalent to AgentRole, and looking at the GMO, they found no obvious way to model the Sponsor field in their database. Additionally, reification led to the generation of blank nodes and the need to create and maintain many URIs. A simpler interface for the data providers was therefore requested.

To accommodate this, a second ontology, together with a manual alignment between this ontology and the GMO, was created to bridge the gap via an intermediate schema that is "flatter" than the patterns and closer to the data providers' own schemas, but still easy to align to the GMO modules because it has been developed directly out of the GMO. This ontology is referred to as the GeoLink Base Ontology (GBO). The providers publish their data according to

the GBO and then SPARQL construct queries which encode the alignment can be used to map data to the GMO. From the very beginning, it was intended that the data integration process would be based on manual, and thus high-quality, mappings between different schemas. As a consequence, ontology alignment systems were not employed to make these mappings, not even to inform human decisions. All mappings were established as a collaborative effort between the data repository providers, the domain experts, and the ontology engineers involved in the modeling and deployment process. Because the GBO was manually engineered directly from the GMO in order to serve this particular purpose, the alignment is guaranteed to be precisely the one intended by the developers. I.e. the alignment is guaranteed to contain all of the relations necessary to solve this real-world alignment problem and no superfluous relations have been included. We argue that this characteristic makes the GeoLink ontologies a good example of a complex ontology alignment problem that can be used as a benchmark for systems that attempt to automate such alignment processes: While it is not a synthetic benchmark, it reflects complex alignment issues encountered in practice.

The example below illustrates the use of the GBO and its alignment to the GMO. In the GBO, there is a relation called hasSponsor with a domain that includes Award and range Organization. This mirrors many of the providers' existing schemas. Providers publish triples either according to the GMO schema (e.g., if they have temporal information), or according to the GBO schema.

```
x:award1 a                view:Award ;
         view:hasSponsor x:org1 .
x:org1   a                view:Organization .
```

Then, the GBO-oriented triples are converted into the GMO schema using this SPARQL construct:

```
PREFIX view: <http://schema.geolink.org/dev/view#>
CONSTRUCT {
  ?x    a                   :FundingAward ;
        :providesAgentRole _:bn1 .
  _:bn1 :isPerformedBy      ?y ;
        a                   :SponsorRole .
  ?y    a                   :Organization .
} WHERE {
  ?x    a                   view:Award ;
        view:hasSponsor     ?y .
  ?y    a                   view:Organization
}
```

Let us look at this by means of a schema diagram. In Figure 3, the three nodes and the two solid arrows indicate the graph pattern used to express the sponsoring organization role in the GMO. The dashed arrow is that is sometimes called a *shortcut* [8]. This shortcut (which is not part of the GMO) "flattens"
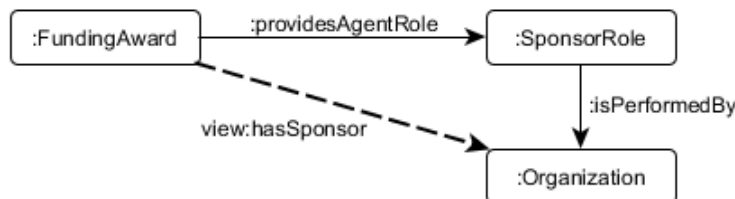
Fig. 3: A schema diagram to explain an example alignment

this part of the GMO, and in the GBO, the `:SponsorRole` node is removed, but the shortcut is added (and `:FundingAward` and `:Organization` have been replaced by the local `view:Award` and `view:Organization`, respectively).

Note that there is no doubt here about the intended alignment between the corresponding parts of the GBO and the GMO: `view:Award` and `:FundingAward` should be mapped to each other (as equivalent), as should `view:Organization` and `:Organization`. It is also clear that that the relation `view:hasSponsor` between an `view:Award` and an `view:Organization` should be aligned (as equivalence) to the concatenation of `:providesAgentRole` and `:isPerformedBy`, provided the entity shared by the two relation expressions is of type `:SponsorRole`, and the chain starts at a `:FundingAward` and ends at a `:Organization`. I.e. a complex alignment is required to express this very natural relationship between these two ontology snippets. Below we will give more examples of complex alignments arising from our setting, when we discuss the different alignment patterns we have identified. The example above is a "Typed Property Chain Equivalence" in our classification, and below we discuss this example further.

More information about the GMO and the project is available from [6] and from the project website[8].

4   The GeoLink Complex Alignment Benchmark

4.1   Dataset

In order to prepare the GeoLink ontologies for use as a complex alignment benchmark, some changes to the namespaces were required. As we introduced in the previous section, several ODPs and modules were created to represent the frequently recurring concepts in the GeoLink datasets, and these were stitched together to form the GeoLink Modular Ontology (GMO). During this process, the namespace of some entities was changed from one that reflected its originating pattern to the namespace of the GMO, which is `http://gmo#`. For example, the class FundingAward was originally in the fundingaward pattern, with the namespace `http://schema.geolink.org/1.0/pattern/fundingaward#`. After merging these modules, the namespace of the class FundingAward became `http:`

---

[8] `http://www.geolink.org/`

Ontology	Classes	Object Properties	Data Properties
GeoLink Base Ontology	40	149	49
GeoLink Modular Ontology	156	124	46

Table 1: The number of classes, object properties, and data properties in both GeoLink ontologies

`//gmo#`. This has been applied to all entities except those that are imported from other ontologies, which retain their original namespace. For example, the namespace of the class Instant, which is imported from `http://www.w3.org/2006/time#`, remains unchanged. Additionally, the namespace of entities in the GeoLink Base Ontology (GBO) has been changed from `http://schema.geolink.org/1.0/base/main#` to `http://gbo#`.

Table 1 shows the number of classes and properties in both ontologies. Both ontologies are comparable in size to ontologies currently used by the OAEI, meaning that they are within the capabilities of most current ontology alignment systems to handle.

4.2   Simple and Complex Correspondences

In order to understand the correspondences in the benchmark, we give the formal definition of simple and complex correspondences.

Simple Correspondence. Simple correspondence refers to basic 1-to-1 simple alignment between two ontologies, including class and property. It not only includes 1-to-1 equivalence, but also contains 1-to-1 subsumption, and 1-to-1 disjointness.

Complex Correspondence. Complex correspondence refers to more complex patterns, such as 1-to-n equivalence, 1-to-n subsumption, m-to-n equivalence, m-to-n subsumption, and m-to-n arbitrary relationship.

We have identified 12 different kinds of simple and complex correspondence patterns in the GeoLink complex alignment benchmark. Table 2 presents these different patterns and the corresponding number and category in the whole dataset. As the table shows, the alignment consists predominantly of complex relationships. In the following, we explain these alignment types, from simple 1-to-1 correspondence to complex m-to-n correspondence, with a formal pattern and example each.

Class Equivalence. The first pattern is just simple 1-to-1 class equivalence. Classes C_1 and C_2 are from ontology O_1 and ontology O_2, respectively.

Formal Pattern: $C_1(x) \leftrightarrow C_2(x)$
Example: $\mathsf{Award}(x) \leftrightarrow \mathsf{FundingAward}(x)$

Pattern	Occurrences	Category
Class Equivalence	10	1:1
Class Subsumption	2	1:1
Property Equivalence	7	1:1
Property Equivalence Inverse	5	1:1
Class typecasting Equivalence	4	1:n
Class typecasting Subsumption	1	1:n
Property typecasting Subsumption	5	1:n
Property typecasting Subsumption Inverse	5	1:n
Typed property chain Equivalence	26	m:n
Typed property chain Equivalence Inverse	17	m:n
Typed property chain Subsumption	17	m:n
Typed property chain Subsumption Inverse	12	m:n

Table 2: The alignment pattern types found in the GeoLink complex alignment benchmark, along with the number of times each occurs and the type of relation.

Class Subsumption. This pattern is very similar to the first pattern. But, instead of class equivalence, this pattern describes simple 1-to-1 class subsumption.

Formal Pattern: $C_1(x) \rightarrow C_2(x)$
Example: $\mathsf{GeoFeature}(x) \rightarrow \mathsf{Place}(x)$

Property Equivalence. Property alignment is also an important part of ontology alignment research [8]. This pattern captures simple 1-to-1 property equivalence. Property p_1 and property p_2 are from ontology O_1 and ontology O_2, respectively. The property can be either a data property or an object property.

Formal Pattern: $p_1(x,y) \leftrightarrow p_2(x,y)$
Example: $\mathsf{hasAward}(x,y) \leftrightarrow \mathsf{fundedBy}(x,y)$

Property Equivalence Inverse. This pattern is similar to the previous one, just that the domain and range values of a property are switched when it aligns to a property in another ontology.

Formal Pattern: $p_1(x,y) \leftrightarrow p_2(y,x)$
Example: $\mathsf{isAwardOf}(x,y) \leftrightarrow \mathsf{fundedBy}(y,x)$

Class Typecasting Equivalence. This pattern is more specific than the previous ones. The idea of typecasting, and why it is important in ontology mod-

eling, is formally introduced and discussed in [8]. The pattern indicates that individuals of type C_1 in one ontology are cast into a subclass of C_2 in the other ontology. Note that punning is employed here – x is treated as an individual on the left hand side of the rule and as a class on the right hand side. For example, an instance of PlaceType in the GBO might be 'ocean'. This is cast into a subclass of Place in the GMO. The reverse is also true: if the GMO has a subclass of Place called Island, then 'island' is an instance of the class PlaceType in the GBO.

Formal Pattern: $C_1(x) \leftrightarrow$ rdfs:subclassOf(x, C_2)
Example: PlaceType$(x) \leftrightarrow$ rdfs:subclassOf$(x, $Place$)$

Class Typecasting Subsumption. This pattern is almost identical to the one above, except that the rule only holds in one direction. In the example, a GeoFeatureType (which comes from the General Bathymetric Chart of the Oceans[9] vocabulary) is always a type of Place, but there are types of Places that are not GeoFeatureType.

Formal Pattern: $C_1(x) \rightarrow$ rdfs:subclassOf(x, C_2)
Example: GeoFeatureType$(x) \rightarrow$ rdfs:subclassOf$(x, $Place$)$

Property Typecasting Subsumption. This pattern is similar in spirit to the Class Typecasting patterns mentioned above. However in this case, a property is cast into a class assignment statement. In a sense, this alignment drops information, as y does not occur on the right hand side.

Formal Pattern: $p_1(x, y) \rightarrow$ rdf:type(x, C_2)
Example: hasPlaceType$(x, y) \rightarrow$ rdf:type$(x, $Place$)$

We note here that some rules that fall under this category are not exact translations of the underlying SPARQL queries, due to expressibility constraints in EDOAL (see section 4.3 below). For instance, instead the example above, which states that the hasPlaceType object property is subsumed by an rdf:type statement with the range value of Place, we would actually like to state the following, which reflects the SPARQL query:

Formal Pattern: $p_1(x, y) \leftrightarrow$ rdf:type$(x, y) \wedge$ rdfs:subclassOf(y, C_2)
Example: hasPlaceType$(x, y) \leftrightarrow$ rdf:type$(x, y) \wedge$ rdfs:subclassOf$(y, $Place$)$

For instance, we would like a rule that implies that the GBO statement has-PlaceType(Honolulu,Island) is equivalent to stating that Honolulu is a type of Island *and* that Island is a subclass of Place in the GMO. In other words, one of the individuals occurring as a property filler on the GBO side is cast into a class on the GMO side. At the same time, the other property filler on the GBO

---

[9] https://www.gebco.net

side is asserted to be an instance of this class. However, this is not possible because the statement requires a variable (y), and that is not supported by the core EDOAL language. The EDOAL specification does mention a pattern language that might enable this type of statement, but it does not appear to be fully supported at this time.

Property Typecasting Subsumption Inverse. This pattern is the same as the one above, except that the property fillers are flipped.

Formal Pattern: $p_1(x, y) \rightarrow \mathsf{rdf:type}(y, C_2)$
Example: $\mathsf{isPlaceTypeOf}(x, y) \rightarrow \mathsf{rdf:type}(y, \mathsf{Place})$

Again, in some cases we would actually like to state the following, which cannot be fully expressed in EDOAL, to the best of our knowledge:

Formal Pattern: $p_1(x, y) \rightarrow \mathsf{rdf:type}(y, x) \wedge \mathsf{rdfs:subclassOf}(x, C_2)$
Example: $\mathsf{isGeoFeatureTypeOf}(x, y) \rightarrow \mathsf{rdf:type}(y, x) \wedge \mathsf{rdfs:subclassOf}(x, \mathsf{Place})$

Typed Property Chain Equivalence. A property chain is a classical complex pattern that was introduced in [11]. This pattern captures the situation related to the hasSponsor property discussed in detail in Section 3. The pattern applies when a property, together with a type restriction on one or both of its fillers, in one ontology have been used to "flatten" the structure of the other ontology by short-cutting a property chain in that ontology. The pattern also ensures that the types of the property fillers involved in the property chain are typed appropriately in the other ontology. The formal pattern and example are shown below. The classes D_i and property r are from ontology O_1, and classes C_i and properties p_i are from ontology O_2.

Formal Pattern:
$$D_1(x_1) \wedge r(x_1, x_{n+1}) \wedge D_2(x_{n+1}) \leftrightarrow C_1(x_1) \wedge p_1(x_1, x_2) \wedge C_2(x_2)$$
$$\wedge \cdots \wedge p_n(x_n, x_{n+1}) \wedge C_{n+1}(x_{n+1})$$

Example[10]:
$$\mathsf{Award}(x) \wedge \mathsf{hasSponsor}(x, z) \leftrightarrow \mathsf{FundingAward}(x) \wedge \mathsf{providesAgentRole}(x, y)$$
$$\wedge \mathsf{SponsorRole}(y) \wedge \mathsf{performedBy}(y, z)$$

Note that in this and all following patterns, any of the D_i or C_i may be omitted (in which case they are essentially $\top$). Also, for the left-to-right direction, we assume that $x_2, \ldots x_n$ are existentially quantified variables.

---

[10] In contrast to the example discussed in Figure 3, we leave out `:Organization` and `view:Organization`, because it is possible, in principle, that a non-organization agent (e.g., an individual) may sponsor.

Typed Property Chain Equivalence Inverse. This pattern is the same as the one above, except that the property fillers are flipped.

Formal Pattern:
$$D_1(x_1) \wedge r(x_1, x_{n+1}) \wedge D_2(x_{n+1}) \leftrightarrow C_1(x_{n+1}) \wedge p_1(x_{n+1}, x_n) \wedge C_2(x_n)$$
$$\wedge \cdots \wedge p_n(x_2, x_1) \wedge C_{n+1}(x_1)$$

Example:
$$\mathsf{Award}(z) \wedge \mathsf{isSponsorOf}(x, z) \leftrightarrow \mathsf{FundingAward}(z) \wedge \mathsf{provideAgentRole}(z, y)$$
$$\wedge \mathsf{SponsorRole}(y) \wedge \mathsf{performedBy}(y, x)$$

Typed Property Chain Subsumption. This is identical to the Typed Property Chain Equivalence pattern except that the relationship only holds in one direction.

Formal Pattern:
$$D_1(x_1) \wedge r(x_1, x_{n+1}) \wedge D_2(x_{n+1}) \rightarrow C_1(x_1) \wedge p_1(x_1, x_2) \wedge C_2(x_2)$$
$$\wedge \cdots \wedge p_n(x_n, x_{n+1}) \wedge C_{n+1}(x_{n+1})$$

Example:
$$\mathsf{Cruise}(x) \wedge \mathsf{hasChiefScientist}(x, z) \rightarrow \mathsf{Cruise}(x) \wedge \mathsf{providesAgentRole}(x, y)$$
$$\wedge \mathsf{AgentRole}(y) \wedge \mathsf{performedBy}(y, z)$$

Typed Property Chain Subsumption Inverse. This pattern is the same as the one above, except that the property fillers are flipped.

Formal Pattern:
$$D_1(x_1) \wedge r(x_1, x_{n+1}) \wedge D_2(x_{n+1}) \rightarrow C_1(x_{n+1}) \wedge p_1(x_{n+1}, x_n) \wedge C_2(x_n)$$
$$\wedge \cdots \wedge p_n(x_2, x_1) \wedge C_{n+1}(x_1)$$

Example:
$$\mathsf{Cruise}(z) \wedge \mathsf{isChiefScientistOf}(x, z) \rightarrow \mathsf{Cruise}(z) \wedge \mathsf{providesAgentRole}(z, y)$$
$$\wedge \mathsf{AgentRole}(y) \wedge \mathsf{performedBy}(y, x)$$

In [11], four alignment types were identified, some of which are subsumed by ours. We do not at all claim that our classification above is exhaustive, but we consider it a refinement of the ones listed in [11]. We conjecture that there are many more important ones of relevance to other use cases. Mapping out the space of complex alignment types is, in our understanding, helpful for further research into complex alignment algorithms.

4.3   Format in EDOAL and Rule syntax

As mentioned previously, SPARQL construct queries are used to convert data published by the data providers according to the GBO into the schema described

in the GMO, because the GMO employs modeling practices that enhance extensibility and facilitate reasoning. However, most ontology alignment benchmarks are not formatted in SPARQL but rather according to the format provided by the Alignment API [5]. The standard alignment format is not expressive enough to capture complex relations. However, the Alignment API also provides a format called Expressive and Declarative Ontology Alignment Language (EDOAL) that can be used to express these types of relations. This format can be read and manipulated programmatically using the Alignment API, and is therefore very convenient for ontology alignment researchers. In addition, EDOAL is already accepted by the ontology alignment community. It has been used by others when proposing new alignment benchmarks [10] and [14], and we continue that approach here. Because EDOAL can be difficult for humans to parse quickly, we have also expressed the alignments in using a naive rule syntax. The rule presentation is not intended for programmatic manipulation, but rather to make it easier for humans to read and understand the alignments. Both versions of the alignment, along with the GBO and GMO ontologies, can be downloaded from `http://doi.org/10.6084/m9.figshare.5907172` under a CC-BY License. We have merged the two ontologies according to this reference alignment and used HermiT [3] to verify that there are no inconsistencies. The GeoLink website[11] contains detailed documentation of the dataset and provides users with more insights about the resource, such as all entities, patterns, and relationships between them in both ontologies.

5   Conclusion

Complex alignment has been discussed for a long time, but relatively little work has been done to advance the state of the art of complex ontology alignment. The lack of an available complex alignment benchmark may be a primary reason for the slow speed of development. In addition, most current alignment benchmarks have been created by humans for the sole purpose of evaluating alignment systems, and they may not always represent real-world cases. In this paper, we have proposed a complex alignment benchmark based on the real-world GeoLink dataset. The two ontologies and the reference alignment were designed and created by ontologists and geoscience domain experts to support data representation, sharing, integration and discovery. We take advantage of these ontologies to create a complex alignment benchmark. In our dataset, the alignments not only cover 1:1 simple correspondences, but also contain 1:n and m:n complex relations. All correspondences required to convert between the two ontologies (a key goal of ontology alignment) are guaranteed to be present, because one ontology was consciously created from the other, with SPARQL queries to mitigate each change. In addition, the alignment has been evaluated by domain experts from different organizations to ensure high quality. Moreover, the ontologies and alignments in both rule and EDOAL syntax have been published in FigShare with an open access license for reusability.

---

[11] `http://schema.geolink.org/`

As future work in this area, we plan to put forth this alignment problem as a potential new track within the OAEI. After that, based on participants' feedback, we will modify the reference alignment if necessary to perfect the benchmark by making it more convenient to use. This may involve, for example, making the alignment available in additional formats. Furthermore, we also plan to create an automated alignment system to tackle the alignment problem set forth by this benchmark.

Acknowledgments We would like to thank the geosciences data providers for sharing the data, and the domain experts for helping understand the concepts to create the ontologies and evaluate the reference alignment. In addition, we would also like to show our gratitude to Jerome Euzenat for providing advice regarding the conversion of rules to EDOAL.

References

1. Cheatham, M., Hitzler, P.: The properties of property alignment. In: Proceedings of the 9th International Workshop on Ontology Matching collocated with the 13th International Semantic Web Conference (ISWC 2014), Riva del Garda, Trentino, Italy, October 20, 2014. pp. 13–24 (2014)
2. David, J.: AROMA results for OAEI 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 (2009)
3. Glimm, B., Horrocks, I., Motik, B., Stoilos, G., Wang, Z.: Hermit: An OWL 2 reasoner. J. Autom. Reasoning **53**(3), 245–269 (2014)
4. Hitzler, P., Gangemi, A., Janowicz, K., Krisnadhi, A., Presutti, V. (eds.): Ontology Engineering with Ontology Design Patterns - Foundations and Applications, Studies on the Semantic Web, vol. 25. IOS Press (2016)
5. Jain, P., Hitzler, P., Sheth, A.P., Verma, K., Yeh, P.Z.: Ontology alignment for linked open data. In: The Semantic Web - ISWC 2010 - 9th International Semantic Web Conference, ISWC 2010, Shanghai, China, November 7-11, 2010, Revised Selected Papers, Part I. pp. 402–417 (2010)
6. Krisnadhi, A.: Ontology Pattern-Based Data Integration. Ph.D. thesis, Wright State University (2015)
7. Krisnadhi, A., Hu, Y., Janowicz, K., Hitzler, P., Arko, R.A., Carbotte, S., Chandler, C., Cheatham, M., Fils, D., Finin, T.W., Ji, P., Jones, M.B., Karima, N., Lehnert, K.A., Mickle, A., Narock, T.W., O'Brien, M., Raymond, L., Shepherd, A., Schildhauer, M., Wiebe, P.: The geolink modular oceanography ontology. In: The Semantic Web - ISWC 2015 - 14th International Semantic Web Conference, Bethlehem, PA, USA, October 11-15, 2015, Proceedings, Part II. pp. 301–309 (2015)
8. Krisnadhi, A.A., Hitzler, P., Janowicz, K.: On the capabilities and limitations of OWL regarding typecasting and ontology design pattern views. In: Ontology Engineering - 12th International Experiences and Directions Workshop on OWL, OWLED 2015, co-located with ISWC 2015, Bethlehem, PA, USA, October 9-10, 2015, Revised Selected Papers. pp. 105–116 (2015)
9. Krisnadhi, A.A., Hu, Y., Janowicz, K., Hitzler, P., Arko, R.A., Carbotte, S., Chandler, C., Cheatham, M., Fils, D., Finin, T., Ji, P., Jones, M.B., Karima, N., Lehnert, K.A., Mickle, A., Narock, T., O'Brien, M., Raymond, L., Shepherd, A., Schildhauer, M., Wiebe, P.: The geolink framework for pattern-based linked data integration. 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. (2015)

10. Pesquita, C., Cheatham, M., Faria, D., Barros, J., Santos, E., Couto, F.M.: Building reference alignments for compound matching of multiple ontologies using OBO cross-products. In: Proceedings of the 9th International Workshop on Ontology Matching collocated with the 13th International Semantic Web Conference (ISWC 2014), Riva del Garda, Trentino, Italy, October 20, 2014. pp. 172–173 (2014)

11. Ritze, D., Meilicke, C., Sváb-Zamazal, O., Stuckenschmidt, H.: A pattern-based ontology matching approach for detecting complex correspondences. 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 (2009)

12. Shvaiko, P., Euzenat, J.: Ontology matching: State of the art and future challenges. IEEE Trans. Knowl. Data Eng. **25**(1), 158–176 (2013)

13. Suchanek, F.M., Abiteboul, S., Senellart, P.: PARIS: probabilistic alignment of relations, instances, and schema. PVLDB **5**(3), 157–168 (2011)

14. Thiéblin, É., Haemmerlé, O., Hernandez, N., dos Santos, C.T.: Towards a complex alignment evaluation dataset. 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, 2017. pp. 217–218 (2017)

15. You, J.: Geoscientists aim to magnify specialized web searching (2015)

GeoLink Dataset: A Complex Alignment Benchmark from Real-world Ontology

Lu Zhou[a], Michelle Cheatham[b], Adila Krisnadhi[c], Pascal Hitzler[a]

[a]*DaSe Lab, Kansas State University, Manhattan, KS 66506, USA*
[b]*Wright State University, Dayton, OH 45435, USA*
[c]*Faculty of Computer Science, Universitas Indonesia, Depok, Jawa Barat 16424, Indonesia*

Abstract

Ontology alignment has been studied for over a decade, and over that time many alignment systems and methods have been developed by researchers in order to find simple 1-to-1 equivalence matches between two ontologies. However, very few alignment systems focus on finding complex correspondences. One reason for this limitation may be that there are no widely accepted alignment benchmarks that contain such complex relationships. In this paper, we propose a real-world dataset from the GeoLink project as a potential complex ontology alignment benchmark. The dataset consists of two ontologies, the GeoLink Base Ontology (GBO) and the GeoLink Modular Ontology (GMO), as well as a manually created reference alignment that was developed in consultation with domain experts from different institutions. The alignment includes 1:1, 1:n, and m:n equivalence and subsumption correspondences, and is available in both Expressive and Declarative Ontology Alignment Language (EDOAL) and rule syntax. The benchmark has been expanded from its original version to contain real-world instance data from seven geoscience data providers that has been published according to both ontologies. This allows it to be used by extensional alignment systems or those that require training data. This benchmark has been incorporated into the Ontology Alignment Evaluation Initiative (OAEI) complex track to help researchers test their automated alignment systems and algorithms. This paper also analyzes the challenges inherent in effectively generating, detecting, and evaluating complex ontology alignments and provides a road map for future work on this topic.

Keywords: Complex Ontology Alignment, Real-world Ontology, Ontology Population, Benchmark

1. Introduction

Ontology alignment is an important step in enabling computers to query and reason across the many linked datasets on the semantic web. This is a difficult challenge because the ontologies underlying different linked datasets can vary in terms of subject area coverage, level of abstraction, ontology modeling philosophy, and even language. Due to the importance and difficulty of the ontology alignment problem, it has been an active area of research for over a decade [1].

Preprint submitted to Data Intelligence   *December 6, 2019*

Ideally, alignment systems should be able to uncover any entity relationship across two ontologies that can exist within a single ontology. Such relationships have a wide range of complexity, from basic 1-to-1 equivalence, such as a Person in one ontology being equivalent to a Human in another ontology, to arbitrary m-to-n relationships, such as a Professor with a hasRank property value of "Assistant" in one ontology being a subclass of the union of the Faculty and TenureTrack classes in another. Unfortunately, the majority of the research activities in the field of ontology alignment remains focus on the simplest end of this scale – finding 1-to-1 equivalence relations between ontologies. Part of the reason for this may be that there are no widely used and accepted ontology alignment benchmarks that involve complex relations.

This paper seeks to take a step in that direction by proposing a complex alignment benchmark based on two ontologies which were developed by domain experts jointly with the reference alignment, and which in fact were developed for deployment on major ocean science data repository platforms, i.e., without the actual intention to develop an alignment benchmark. For this reason, the benchmark, including the reference alignment, can be considered to be (a) objective, in that it was created for deployment and not for benchmarking, (b) realistic, in that it captures an application use case developed for deployment, and (c) a valid ground truth alignment, in that the two ontologies and the reference alignment were developed together, by domain experts. We argue that it is therefore of rather unique nature and will inform complex ontology alignment research from a practical and applied perspective, rather than artificial laboratory-like. The benchmark, coincidentally, as this was the requirement of the use case, has a particular focus on relationships involving properties, which is particularly interesting because those have been shown to be rather difficult to handle for current alignment approaches [2]. In addition, we have analyzed and categorized the mapping rules constituting the alignment. We found several which had not been classified or discussed previously, which we will present and discuss in our analysis.

The main contributions of this paper are therefore the following:

- Presentation of two ontologies to support data representation, sharing, integration, and discovery for the geoscience research domain.

- Creation of an alignment between these two ontologies that includes 1:1, 1:n, and m:n correspondences, and given the creation history and usage of the alignment, it is fair to say that the alignment constitutes a gold-standard reference.

- Publication of the benchmark alignment in both rule syntax and EDOAL format[1] at a persistent URL[2] under a CC-BY license.

- Population of the Abox information supported by data providers to extend the functionality of the benchmark in instance-based applications.

- Incorporation of the benchmark into the OAEI complex track[3] in order to help researchers to test and improve their complex ontology alignment systems and algorithms.

- Discussion of the challenges related to the generation, detection, and evaluation of the complex ontology alignment and the potential methods for future work in this area.

---

[1] http://alignapi.gforge.inria.fr/edoal.html
[2] http://doi.org/10.6084/m9.figshare.5907172
[3] http://oaei.ontologymatching.org/2018/complex/index.html

This paper is an extended version of the one presented at the International Semantic Web Conference 2018 [3]. The final three bullet points above represent new material.

The paper is organized as follows. Section 2 discusses the few existing ontology alignment benchmarks that involve relationships other than 1-to-1 equivalence and methods to detect them. Section 3 gives further background on the GeoLink modeling process, including why two different but related ontologies were developed. Section 4 discusses the alignment between the two GeoLink ontologies, along with some descriptive statistics and an analysis of the types of mapping rules constituting the alignment, and the instance data population process. Section 5 introduces the simplified version of the benchmark used in the OAEI complex track and presents the evaluation results. Section 6 discusses the challenges that we faced in the research and provides potential approaches to solve them. Section 7 concludes with a discussion of potential future work in this area.

2. Related Work

Most work associated with evaluating the performance of ontology alignment systems has been done in conjunction with the Ontology Alignment Evaluation Initiative (OAEI)[4]. These yearly events allow developers to test their alignment systems on various tracks that evaluate performance on different facets of the problem such as instance matching, large ontology matching, and interactive matching, among others. Currently, most of these tracks involve the identification of 1-to-1 equivalence relationships, such as a Participant being equivalent to an Attendee. In 2009, the OAEI ran an "oriented" matching track that challenged systems to find subsumption relationships such as a Book is a subclass of a Publication. However, this track was abandoned after one year. Some system developers complained that the quality of the reference alignment was low [4]. This frustrated system developers and limited participation. Discussions at the last two Ontology Matching workshops[5] made it clear that the community is interested in complex alignment, but that lack of applicable benchmarks is hindering progress. Our proposed benchmark seeks to address this concern by providing a reference alignment as a benchmark, and by addressing the quality issue of the previous benchmark by the fact that the process leading to the reference alignment guarantees its high quality.

Related work is currently being undertaken by Thieblin and her colleagues [5], who are creating a complex alignment benchmark using the Conference track ontologies within the OAEI [6]. This work is partially completed, and at the time of this writing it covers three of the seven ontologies. The reference alignment we describe herein differs from the effort by Thieblin et al. in that the GeoLink ontologies and alignment constitute real-world datasets designed and used in a practical application by geoscientists, rather than being an artificial artifact designed solely for alignment benchmarking. Furthermore, data from seven geoscience repositories have been published according to the GeoLink schema and are available online[6]. This instance data can in the future be used by alignment systems that employ extensional matching techniques [7]. In contrast to this, significant instance data is currently not readily available for most of the OAEI conference track ontologies. With the increasing requirement of more complex ontology alignment and growing interest in generating complex correspondences in real-world datasets [3, 8],

---

[4]http://oaei.ontologymatching.org
[5]http://www.ontologymatching.org/
[6]http://data.geolink.org

3

the first version of the complex alignment track was introduced in OAEI 2018 [9]. Our GeoLink benchmark is one of the four benchmarks that contain complex correspondences in this track. The other three complex ontology alignment benchmarks are from different domains: conference, hydrography, and plant taxonomy. In addition, different evaluation strategies were applied in evaluating the performance of complex alignment systems on the different benchmarks. More details of evaluations and results can be accessed on the OAEI 2018 website[7].

While alignment systems capable of generating complex alignments are relatively rare, several approaches have been proposed in the literature. Ritze applied pattern-based [10] and linguistic analysis approaches [11] to detect the complex correspondences in a dataset. Jiang [12] accomplished the task of finding a complex alignment by defining knowledge rules and using a probabilistic framework to integrate a knowledge-based strategy with standard terminology-based and structure-based strategies.

Alignment systems that attempt to identify subsumption relations have sometimes used their own manually developed (and sometimes unpublished) reference alignments [13]. Other subsumption systems have evaluated the precision of their approach by manually validating relations produced by their system, while foregoing an assessment of recall [14]. Other related work has centered on developing a benchmark for compound alignments, which the authors define as mappings between class or property expressions involving more than two ontologies [15]. Their first step in this direction was to create a set of reference alignments containing relations of the form $< X, Y, Z, R, M >$, where X, Y and Z are classes from three different ontologies and R is a relation between Y and Z that results in a class expression that is related to X by the relation M. For example, a DisabledVeteran (X) is equivalent to (M) the intersection (R) of Veteran (Y) and Disabled (Z). This benchmark is based on cross-products among the Open Biomedical Ontologies (OBO) Foundry[8], which have been manually validated by at least two experts. The work presented herein differs from these approaches by considering a wider range of relationship types (beyond subsumption and the type of ternary relation described in [15]), as they naturally arose out of the application from which the reference alignment was taken.

3. The GeoLink Modeling Process

Benchmarks come in at least two varieties. On the one hand, there are artificial benchmarks that provide a kind of laboratory setting for evaluation. On the other hand, there are benchmarks created from data as it is used in realistic use cases or even deployed scenarios. Both of these types are important, and they cover different aspects of the spectrum, and may have different advantages. Artificial benchmarks can be made to be balanced, or to focus on certain aspects of a problem, and sometimes they can be used to test scalability issues more easily as different versions of the same benchmark set may be easily producible. Natural benchmarks, on the other hand, may expose issues arising in practice which may easily be overlooked by designers of artificial benchmarks, in particular in a young field such as complex ontology alignment. Natural benchmarks also may come with an independently verified gold standard baseline, as in our case.

The project that this benchmark arose from is called GeoLink [16] and was funded under the U.S. National Science Foundation's EarthCube initiative. This planned decade-long endeavor is a recognition that oftentimes the most innovative and useful discoveries come at the intersection

---

[7]http://oaei.ontologymatching.org/2018/complex/index.html
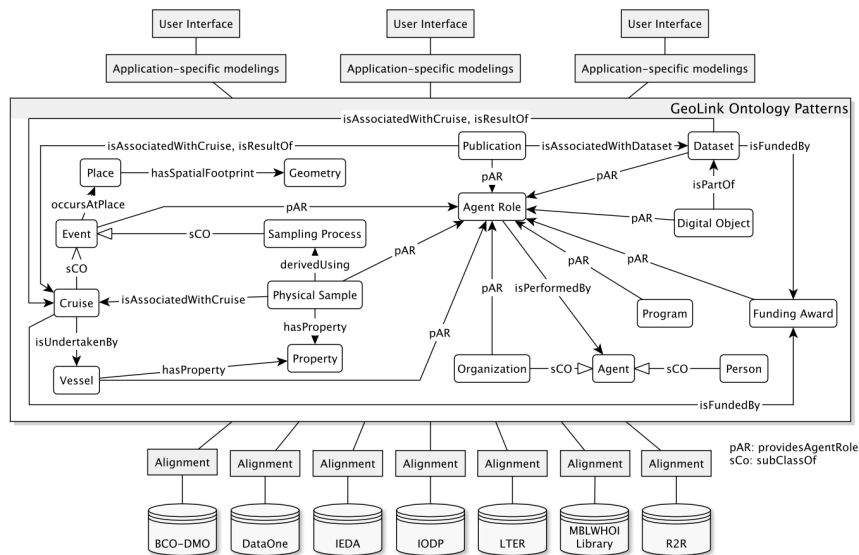[8]http://www.obofoundry.org/

4

Figure 1: Intended usage of the GMO

of traditional fields of research. This is particularly true in the geosciences, which often bring together disparate groups of researchers such as geologists, meteorologists, climatologists, ecologists, archaeologists, and so on. For its part, GeoLink employs semantic web technologies to support data representation, sharing, integration, and discovery [17]. In particular, seven diverse geoscience datasets have been brought together into a single data repository.

At the beginning of the project, some providers' data resided in relational databases while others' had been published as RDF triples and exposed via a SPARQL endpoint. Because each provider had their own schema, the first step in the GeoLink project was to develop a unified schema according to which all data providers could publish their data [17]. Creating a unified schema for independently developed datasets is sometimes difficult, and the final product often ends up requiring providers to shoehorn their data into a schema that does not quite fit. GeoLink uses an approach that relies on ontology design patterns (ODP) in an attempt to avoid this issue [18]. An ODP represents a reusable solution to a recurring modeling problem and generally encodes a specific abstract notion, such as a process, event, agent, etc. These are frequently the small areas of semantic overlap that exist between datasets from different subfields of the same high-level domain. ODPs provide a structured and application-neutral representation of the key concepts within a domain. Throughout the first year of the project, geoscientists, data providers and ontologists worked together to identify and model the important concepts within the geosciences that recurred across two or more datasets. The results of this were what we call ontology modules, based on ODPs, and eventually they were stitched together to form the GeoLink Modular Ontology (GMO) [19].

As shown in Figure 1, the GMO allows data providers to publish only those aspects of their data modeled by the GMO according to that schema. Any data that the provider has that is
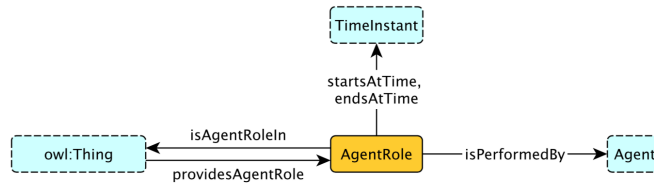
5

Figure 2: The Agent Role pattern

not covered by that schema can be published using the provider's own schema, since no other providers have similar content. For example, in Figure 1, the provider R2R has data related mostly to the cruise and vessel modules in the lower left of the Figure 1, and so it publishes its related data using that terminology. R2R also has data not modeled by the GMO, and so it uses its own terminology when publishing that information. This freedom is intended to make the publishing process easier. However, some problems still remained.

Some of the patterns contain a rather complicated structure, mostly due to reification, which was employed to accommodate different perspectives (e.g., based on granularity) on the data. For example, many of the data providers have information about the sponsor of a project, and R2R has a native relation in their schema called hasSponsor with domain Award and range Organization. However, following best practices, it leads to a more versatile model if being a sponsor is recognized (and thus modeled) as a role which an agent (in this case an organization) can assume. Creating a distinct relation for each type of role on a project (sponsor, chief scientist, research assistant, etc.) is brittle, in the sense that if new roles will be added later, potentially due to the inclusion of a new dataset, then the schema will need to be edited by adding new vocabulary for new roles together with (possibly complex) role relationships. Another issue with using a relation such as hasSponsor is that a more fine-grained data repository may have additional temporal information related to the sponsor role, and then it is not clear how to add this temporal information to the hasSponsor model without punning. Essentially, hasSponsor should better be expressed as a ternary relation between award, organization, and the type of relation (in this case, being a sponsor) expressed using an individual which can be reused in all sponsor relationships. In terms of ODPs, this is realized by reusing the Agent Role pattern, shown in abstract form in Figure 2. This approach both allows new roles to be added easily (by subclassing AgentRole) and supports temporal queries if desired.

Unfortunately, while the data providers recognized the utility of this modeling approach, they found it cumbersome to map their data to it. Looking at their own schemas, they found nothing equivalent to AgentRole, and looking at the GMO, they found no obvious way to model the Sponsor field in their database. Additionally, reification led to the generation of blank nodes and the need to create and maintain many URIs. A simpler interface for the data providers was therefore requested.

To accommodate this, a second ontology, together with a manual alignment between this ontology and the GMO, was created to bridge the gap via an intermediate schema that is "flatter" than the patterns and closer to the data providers' own schemas, but still easy to align to the GMO modules because it has been developed directly out of the GMO. This ontology is referred to as the GeoLink Base Ontology (GBO). The providers publish their data according to the GBO

6

and then SPARQL construct queries which encode the alignment can be used to map data to the GMO. From the very beginning, it was intended that the data integration process would be based on manual, and thus high-quality, mappings between different schemas. As a consequence, ontology alignment systems were not employed to make these mappings, not even to inform human decisions. All mappings were established as a collaborative effort between the data repository providers, the domain experts, and the ontology engineers involved in the modeling and deployment process. Because the GBO was manually engineered directly from the GMO in order to serve this particular purpose, the alignment is guaranteed to be precisely the one intended by the developers. I.e. the alignment is guaranteed to contain all of the relations necessary to solve this real-world alignment problem and no superfluous relations have been included. We argue that this characteristic makes the GeoLink ontologies a good example of a complex ontology alignment problem that can be used as a benchmark for systems that attempt to automate such alignment processes: While it is not a synthetic benchmark, it reflects complex alignment issues encountered in practice.

The example below illustrates the use of the GBO and its alignment to the GMO. In the GBO, there is a relation called hasSponsor with a domain that includes Award and range Organization. This mirrors many of the providers' existing schemas. Providers publish triples either directly according to the GMO schema (e.g., if they have temporal information), or according to the GBO schema.

```
x:award1 a              view:Award ;
         view:hasSponsor x:org1 .
x:org1   a              view:Organization .
```

In the latter case, the GBO-oriented triples are converted into the GMO schema using this SPARQL construct:

```
PREFIX view: <http://schema.geolink.org/dev/view#>
CONSTRUCT {
  ?x    a                 :FundingAward ;
        :providesAgentRole _:bn1 .
  _:bn1 :isPerformedBy     ?y ;
        a                 :SponsorRole .
  ?y    a                 :Organization .
} WHERE {
  ?x    a                 view:Award ;
        view:hasSponsor   ?y .
  ?y    a                 view:Organization
}
```

Let us look at this by means of a schema diagram. In Figure 3, the three nodes and the two solid arrows indicate the graph pattern used to express the sponsoring organization role in the GMO. The dashed arrow is sometimes called a *shortcut* [20]. This shortcut (which is not part of the GMO) "flattens" this part of the GMO, and in the GBO, the :SponsorRole node is removed, but the shortcut is added (and :FundingAward and :Organization have been replaced by the local view:Award and view:Organization, respectively).

Note that there is no doubt here about the intended alignment between the corresponding parts of the GBO and the GMO: view:Award and :FundingAward should be mapped to each other
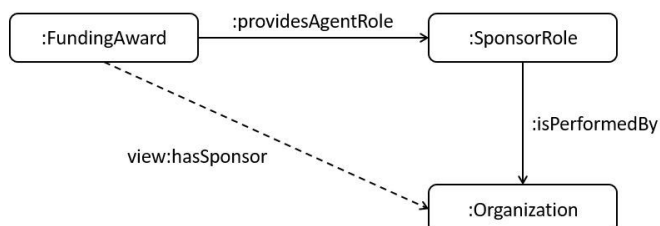
7

Figure 3: A schema diagram to explain an example alignment

(as equivalence), as should `view:Organization` and `:Organization`. It is also clear that the relation `view:hasSponsor` between an `view:Award` and an `view:Organization` should be aligned (as equivalence) to the concatenation of `:providesAgentRole` and `:isPerformedBy`, provided the entity shared by the two relation expressions is of type `:SponsorRole`, and the chain starts at a `:FundingAward` and ends at an `:Organization`. I.e. a complex alignment is required to express this very natural relationship between these two ontology snippets. Below we will give more examples of complex alignments arising from our setting, when we discuss the different alignment patterns we have identified. The example above is a "Typed Property Chain Equivalence" in our classification, and below we discuss this example further.

More information about the GMO and the project is available from [21] and from the project website[9].

4. The GeoLink Complex Alignment Benchmark

4.1. Dataset

In order to prepare the GeoLink ontologies for use as a complex alignment benchmark, some changes to the namespaces were required. As we introduced in the previous section, several ODPs and modules were created to represent the frequently recurring concepts in the GeoLink datasets, and these were stitched together to form the GMO. During this process, the namespace of some entities was changed from one that reflected its originating pattern to the namespace of the GMO, which is `http://gmo#`. For example, the class FundingAward was originally in the fundingaward pattern, with the namespace `http://schema.geolink.org/1.0/pattern/fundingaward#`. After merging these modules, the namespace of the class FundingAward became `http://gmo#`. This has been applied to all entities except those that are imported from other ontologies, which retain their original namespace. For example, the namespace of the class Instant, which is imported from `http://www.w3.org/2006/time#`, remains unchanged. Additionally, the namespace of entities in the GBO has been changed from `http://schema.geolink.org/1.0/base/main#` to `http://gbo#`.

Table 1 shows the number of classes and properties in both ontologies. Both ontologies are comparable in size to ontologies currently used by the OAEI, meaning that they are within the capabilities of most current ontology alignment systems to handle.

---

[9]`http://www.geolink.org/`

Table 1: The number of classes, object properties, and data properties in both GeoLink ontologies

Ontology	Classes	Object Properties	Data Properties
GeoLink Base Ontology	40	149	49
GeoLink Modular Ontology	156	124	46

Table 2: The alignment pattern types found in the GeoLink complex alignment benchmark, along with the number of times each occurs and the type of relation.

Pattern	Occurrences	Category
Class Equivalence	10	1:1
Class Subsumption	2	1:1
Property Equivalence	7	1:1
Property Equivalence Inverse	5	1:1
Class Typecasting Equivalence	4	1:n
Class Typecasting Subsumption	1	1:n
Property Typecasting Subsumption	5	1:n
Property Typecasting Subsumption Inverse	5	1:n
Typed Property Chain Equivalence	26	m:n
Typed Property Chain Equivalence Inverse	17	m:n
Typed Property Chain Subsumption	17	m:n
Typed Property Chain Subsumption Inverse	12	m:n

4.2. Simple and Complex Correspondences

In order to understand the correspondences in the benchmark, we give the formal definition of simple and complex correspondences.

Simple Correspondence. Simple correspondence refers to a basic 1-to-1 simple mapping between two ontologies, in which the entities involved may be either classes or properties. This category not only includes 1-to-1 equivalence relations, but also 1-to-1 subsumption and 1-to-1 disjointness.

Complex Correspondence. Complex correspondence refers to more complex patterns, such as 1-to-n equivalence, 1-to-n subsumption, m-to-n equivalence, m-to-n subsumption, and m-to-n arbitrary relationship.

We have identified 12 different kinds of simple and complex correspondence patterns in the GeoLink complex alignment benchmark. Table 2 presents these different patterns and the corresponding number and category in the whole dataset. As the table shows, the alignment consists predominantly of complex relationships. In the following, we explain these alignment types, from simple 1-to-1 correspondence to complex m-to-n correspondence, with a formal pattern and example for each.

Class Equivalence. The first pattern is just simple 1-to-1 class equivalence. Classes C_1 and C_2 are from ontology O_1 and ontology O_2, respectively.

9

Formal Pattern: $C_1(x) \leftrightarrow C_2(x)$

Example: $\mathsf{Award}(x) \leftrightarrow \mathsf{FundingAward}(x)$

Class Subsumption. This pattern is very similar to the first pattern. But, instead of class equivalence, this pattern describes simple 1-to-1 class subsumption.

Formal Pattern: $C_1(x) \rightarrow C_2(x)$

Example: $\mathsf{GeoFeature}(x) \rightarrow \mathsf{Place}(x)$

Property Equivalence. Property alignment is also an important part of ontology alignment research [20]. This pattern captures simple 1-to-1 property equivalence. Property p_1 and property p_2 are from ontology O_1 and ontology O_2, respectively. The property can be either a data property or an object property.

Formal Pattern: $p_1(x, y) \leftrightarrow p_2(x, y)$

Example: $\mathsf{hasAward}(x, y) \leftrightarrow \mathsf{fundedBy}(x, y)$

Property Equivalence Inverse. This pattern is similar to the previous one, except that the domain and range values of a property are switched when it aligns to a property in another ontology.

Formal Pattern: $p_1(x, y) \leftrightarrow p_2(y, x)$

Example: $\mathsf{isAwardOf}(x, y) \leftrightarrow \mathsf{fundedBy}(y, x)$

Class Typecasting Equivalence. This pattern is more specific than the previous ones. The idea of typecasting, and why it is important in ontology modeling, is formally introduced and discussed in [20]. The pattern indicates that individuals of type C_1 in one ontology are cast into a subclass of C_2 in the other ontology. Note that punning is employed here – x is treated as an individual on the left-hand side of the rule and as a class on the right-hand side. For example, an instance of PlaceType in the GBO might be 'ocean'. This is cast into a subclass of Place in the GMO. The reverse is also true: if the GMO has a subclass of Place called Island, then 'island' is an instance of the class PlaceType in the GBO.

Formal Pattern: $C_1(x) \leftrightarrow \mathsf{rdfs{:}subClassOf}(x, C_2)$

Example: $\mathsf{PlaceType}(x) \leftrightarrow \mathsf{rdfs{:}subClassOf}(x, \mathsf{Place})$

Class Typecasting Subsumption. This pattern is almost identical to the one above, except that the rule only holds in one direction. In the example, a GeoFeatureType (which comes from the General Bathymetric Chart of the Oceans[10] vocabulary) is always a type of Place, but there are

---

[10]https://www.gebco.net

10

types of Places that are not GeoFeatureType.

Formal Pattern: $C_1(x) \rightarrow \mathsf{rdfs{:}subClassOf}(x, C_2)$

Example: $\mathsf{GeoFeatureType}(x) \rightarrow \mathsf{rdfs{:}subClassOf}(x, \mathsf{Place})$

Property Typecasting Subsumption. This pattern is similar in spirit to the Class Typecasting patterns mentioned above. However, in this case, a property is cast into a class assignment statement. In a sense, this alignment drops information, as *y* does not occur on the right-hand side.

Formal Pattern: $p_1(x, y) \rightarrow \mathsf{rdf{:}type}(x, C_2)$

Example: $\mathsf{hasPlaceType}(x, y) \rightarrow \mathsf{rdf{:}type}(x, \mathsf{Place})$

We note here that some rules that fall under this category are not exact translations of the underlying SPARQL queries, due to expressibility constraints in EDOAL (see Section 4.4 below). For instance, instead of the example above, which states that the hasPlaceType object property is subsumed by an rdf:type statement with the range value of Place, we would actually like to state the following, which reflects the SPARQL query:

Formal Pattern: $p_1(x, y) \leftrightarrow \mathsf{rdf{:}type}(x, y) \wedge \mathsf{rdfs{:}subClassOf}(y, C_2)$

Example: $\mathsf{hasPlaceType}(x, y) \leftrightarrow \mathsf{rdf{:}type}(x, y) \wedge \mathsf{rdfs{:}subClassOf}(y, \mathsf{Place})$

For instance, we would like a rule that implies that the GBO statement hasPlaceType(Honolulu,Island) is equivalent to stating that Honolulu is a type of Island *and* that Island is a subclass of Place in the GMO. In other words, one of the individuals occurring as a property filler on the GBO side is cast into a class on the GMO side. At the same time, the other property filler on the GBO side is asserted to be an instance of this class. However, this is not possible because the statement requires a variable (*y*), and that is not supported by the core EDOAL language. The EDOAL specification does mention a pattern language that might enable this type of statement, but it does not appear to be fully supported at this time.

Property Typecasting Subsumption Inverse. This pattern is the same as the one above, except that the property fillers are flipped.

Formal Pattern: $p_1(x, y) \rightarrow \mathsf{rdf{:}type}(y, C_2)$

Example: $\mathsf{isPlaceTypeOf}(x, y) \rightarrow \mathsf{rdf{:}type}(y, \mathsf{Place})$

Again, in some cases we would actually like to state the following, which cannot be fully expressed in EDOAL, to the best of our knowledge:

Formal Pattern: $p_1(x, y) \rightarrow \mathsf{rdf{:}type}(y, x) \wedge \mathsf{rdfs{:}subClassOf}(x, C_2)$

11

Example: isGeoFeatureTypeOf$(x, y) \rightarrow$ rdf:type$(y, x) \land$ rdfs:subClassOf(x, Place)

Typed Property Chain Equivalence. A property chain is a classical complex pattern that was introduced in [10]. This pattern captures the situation related to the hasSponsor property discussed in detail in Section 3. The pattern applies when a property, together with a type restriction on one or both of its fillers, in one ontology has been used to "flatten" the structure of the other ontology by short-cutting a property chain in that ontology. The pattern also ensures that the types of the property fillers involved in the property chain are typed appropriately in the other ontology. The formal pattern and example are shown below. The classes D_i and property r are from ontology O_1, and classes C_i and properties p_i are from ontology O_2.

Formal Pattern:
$$D_1(x_1) \land r(x_1, x_{n+1}) \land D_2(x_{n+1}) \leftrightarrow C_1(x_1) \land p_1(x_1, x_2) \land C_2(x_2)$$
$$\land \cdots \land p_n(x_n, x_{n+1}) \land C_{n+1}(x_{n+1})$$

Example[11]:
$$\text{Award}(x) \land \text{hasSponsor}(x, z) \leftrightarrow \text{FundingAward}(x) \land \text{providesAgentRole}(x, y)$$
$$\land \text{SponsorRole}(y) \land \text{performedBy}(y, z)$$

Note that in this and all following patterns, any of the D_i or C_i may be omitted (in which case they are essentially $\top$). Also, for the left-to-right direction, we assume that $x_2, \ldots x_n$ are existentially quantified variables.

Typed Property Chain Equivalence Inverse. This pattern is the same as the one above, except that the property fillers are flipped.

Formal Pattern:
$$D_1(x_1) \land r(x_1, x_{n+1}) \land D_2(x_{n+1}) \leftrightarrow C_1(x_{n+1}) \land p_1(x_{n+1}, x_n) \land C_2(x_n)$$
$$\land \cdots \land p_n(x_2, x_1) \land C_{n+1}(x_1)$$

Example:
$$\text{Award}(z) \land \text{isSponsorOf}(x, z) \leftrightarrow \text{FundingAward}(z) \land \text{provideAgentRole}(z, y)$$
$$\land \text{SponsorRole}(y) \land \text{performedBy}(y, x)$$

Typed Property Chain Subsumption. This is identical to the Typed Property Chain Equivalence pattern except that the relationship only holds in one direction.

Formal Pattern:
$$D_1(x_1) \land r(x_1, x_{n+1}) \land D_2(x_{n+1}) \rightarrow C_1(x_1) \land p_1(x_1, x_2) \land C_2(x_2)$$
$$\land \cdots \land p_n(x_n, x_{n+1}) \land C_{n+1}(x_{n+1})$$

---

[11]In contrast to the example discussed in Figure 3, we leave out `:Organization` and `view:Organization`, because it is possible, in principle, that a non-organization agent (e.g., an individual) may sponsor.

12

Example:

$\text{Cruise}(x) \wedge \text{hasChiefScientist}(x, z) \rightarrow \text{Cruise}(x) \wedge \text{providesAgentRole}(x, y)$
$$\wedge \text{AgentRole}(y) \wedge \text{performedBy}(y, z)$$

Typed Property Chain Subsumption Inverse. This pattern is the same as the one above, except that the property fillers are flipped.

Formal Pattern:

$D_1(x_1) \wedge r(x_1, x_{n+1}) \wedge D_2(x_{n+1}) \rightarrow C_1(x_{n+1}) \wedge p_1(x_{n+1}, x_n) \wedge C_2(x_n)$
$$\wedge \cdots \wedge p_n(x_2, x_1) \wedge C_{n+1}(x_1)$$

Example:

$\text{Cruise}(z) \wedge \text{isChiefScientistOf}(x, z) \rightarrow \text{Cruise}(z) \wedge \text{providesAgentRole}(z, y)$
$$\wedge \text{AgentRole}(y) \wedge \text{performedBy}(y, x)$$

In [10], four alignment types were identified, some of which are subsumed by ours. We do not at all claim that our classification above is exhaustive, but we consider it a refinement of the ones listed in [10]. We conjecture that there are many additional important types of relevance to other use cases. Mapping out the space of complex alignment types is, in our understanding, helpful for further research into complex alignment algorithms.

4.3. Instance Data Population

Instance-based ontology mapping algorithms have been shown to be effective in several practical use cases [22]. The basic idea of instance-based mapping is to query the instance data of the two entities or constructs in two ontologies and calculate the overlap of the common instances, as assessed by some coreference resolution method. In order to extend the functionality of our benchmark and provide more scalability for researchers to explore algorithms that depend on the instance data, we have included the same instance data published according to both the GBO and the GMO in the GeoLink dataset.

4.3.1. Instance Data Information

The GeoLink knowledge base aims at helping users to query and reason over some of the most prominent geoscience metadata repositories in the United States. These include:

- Rolling Deck to Repository (R2R)[12]

- Biological and Chemical Oceanography Data Management Office (BCO-DMO)[13]

- International Ocean Discovery Program (IODP)[14]

- Marine Biological Laboratory Woods Hole Oceanographic Institution (MBLWHOI) Library[15]

---

[12]http://www.rvdata.us/
[13]https://www.bco-dmo.org/
[14]https://www.iodp.org/
[15]http://www.mblwhoilibrary.org/

13

- System for Earth Sample Registration (SESAR)[16]

- Data Observation Network for Earth (DataONE)[17]

- American Geophysical Union (AGU), the National Geochemical Database (NGDB)[18]

- United States Antarctic Program (USAP)[19]

Owing to these data providers, the GeoLink knowledge base contains over 48 million triples, which are formatted according to the GBO schema. As explained in Section 3, the data providers had difficulty publishing directly to the GMO schema, so the simpler (i.e. "flatter") GBO schema was developed and they published their data according to that. In order to enable instance-based matching systems to utilize our benchmark and evaluate their performance, we have used SPARQL construct queries based on the reference alignment to expand the GeoLink ABox to include the GMO as well as the GBO tags.

4.3.2. Population Approach

As mentioned previously, the Geolink knowledge base contains over 48 million triples. In order to facilitate the convenient storage and distribution of the benchmark, we decided to pare down the size by only populating part of the instance data into the benchmark for future OAEI usage. For each reference mapping between the two ontologies, we randomly selected up to 500 instances from the GBO in the SPARQL construct queries. For usage of OAEI benchmark, currently we only published the instance data that are related to the classes and properties in the reference alignment. If there is increasing demand of other instance data which are not related to the reference alignment in the future, we are also willing to provide more instance data which can be found in the GeoLink Website[20].

As an example, referring to the property equivalence correspondence: $\mathsf{hasAward}(x, y) \leftrightarrow \mathsf{fundedBy}(x, y)$. This mapping means that the property $\mathtt{hasAward}$ in the GBO and the property $\mathtt{fundedBy}$ should be mapped to each other as equivalence. Therefore, instances that are related by the $\mathtt{hasAward}$ property in the GBO should be also related by the $\mathtt{fundedBy}$ in the GMO. The corresponding SPARQL construct query is:

```
PREFIX gbo: <http://gbo#>
PREFIX gmo: <http://gmo#>
CONSTRUCT {?x    gmo:fundedBy    ?y .}
    WHERE {?x    gbo:hasAward    ?y .}
LIMIT 500
```

This SPARQL construct query looks for the triples that have $\mathtt{hasAward}$ as the property in the dataset and creates a new graph corresponding to the $\mathtt{fundedBy}$ relation with the same x and y values. This example illustrates the usage of the SPARQL construct query.

---

[16]http://www.geosamples.org/
[17]https://www.dataone.org/
[18]https://minerals.usgs.gov/science/natl-geochemical-db/
[19]http://www.usap-dc.org/
[20]http://www.geolink.org/

14

```
x:professor1    gbo:hasAward    x:NSF
```

If this is a triple according to the GBO schema (ignoring the namespace "x" of the individual), the SPARQL construct query creates another one for the GMO, which is shown below:

```
x:professor1    gmo:fundedBy    x:NSF
```

Besides this relatively simple mapping, our GeoLink benchmark contains more complex relations that involve reification, which lead to the generation of blank nodes. For an example we refer to the typed property chain equivalence correspondence

$$\mathsf{Award}(x) \wedge \mathsf{hasSponsor}(x, z) \leftrightarrow \mathsf{FundingAward}(x) \wedge \mathsf{providesAgentRole}(x, y)$$
$$\wedge \mathsf{SponsorRole}(y) \wedge \mathsf{performedBy}(y, z)$$

The GBO uses a "flattened" structure for the property `hasSponsor`. Compared to the corresponding structure in the GMO, it is a shortcut of the property chain that involves the property `providesAgentRole` and `performedBy`. The SPARQL construct query for this mapping is shown below. It acquires up to 500 instances that satisfy this relation. A blank node, which is of type `SponsorRole`, is generated to maintain the property refication.

```
PREFIX gbo: <http://gbo#>
PREFIX gmo: <http://gmo#>
CONSTRUCT {
  ?x    a                   gmo:FundingAward ;
        gmo:providesAgentRole _:bn1 .
  _:bn1 a                   gmo:SponsorRole ;
        gmo:performedBy     ?z .
} WHERE {
  ?x    a                   gbo:Award ;
        gbo:hasSponsor      ?z .
}
LIMIT 500
```

We utilize the Jena API [23] to generate the blank node when it is needed by a SPARQL construct query. Then, we leverage the OWL API to insert the assertions into the ontologies and finally finish the population process.

4.3.3. Population Result

After finishing up the population of the instance data into the GBO and the GMO, the total number of individuals in the GBO and GMO are 10897 and 11419, respectively. In addition, the number of axioms in the two ontologies are 18336 and 56318, respectively. However, among 67 reference mappings between the GBO and the GMO, there are 19 mappings that lack any applicable instance data currently, because the data providers do not have any more instance data within their repositories. Therefore, the data providers can not publish the related instance data into the GeoLink knowledge base at this stage. We introduce and discuss some potential methods to rectify this in Section 6. In the meantime, the instance data that is currently present in the knowledge base is sufficient for the detection of most of the complex mappings within the benchmark by automated ontology alignment systems that depend on instances.

15

4.4. Format in Rule syntax and EDOAL format

As mentioned previously, SPARQL construct queries are used to convert data published by the data providers according to the GBO into the schema described in the GMO, because the GMO employs modeling practices that enhance extensibility and facilitate reasoning. However, most ontology alignment benchmarks are not formatted in SPARQL but rather according to the format provided by the Alignment API [13]. The standard alignment format is not expressive enough to capture complex relations. However, the Alignment API also provides a format called EDOAL that can be used to express these types of relations. This format can be read and manipulated programmatically using the Alignment API and is therefore very convenient for ontology alignment researchers. In addition, EDOAL is already accepted by the ontology alignment community. It has been used by others when proposing new alignment benchmarks [15] and [6], and we continue that approach here. Because EDOAL can be difficult for humans to parse quickly, we have also expressed the alignments in using a naive rule syntax. The rule presentation is not intended for programmatic manipulation, but rather to make it easier for humans to read and understand the alignments. Both versions of the alignment, along with the GBO and GMO ontologies, can be downloaded from `http://doi.org/10.6084/m9.figshare.5907172` under a CC-BY License. We applied the HermiT [24] reasoner to the ontologies independently to check satisfiability, since some EDOAL mappings which are part of our benchmark do not seem to be expressible in OWL DL. The GeoLink website[21] contains detailed documentation of the dataset and provides users with more insights about the resource, such as all entities, patterns, and relationships between them in both ontologies.

5. OAEI Evaluation

5.1. Simplified Version of Benchmark

The version of the GeoLink alignment benchmark used for the first version of the complex alignment track in OAEI 2018 was slightly simplified compared to the one discussed in Section 4. Some relatively complex relations involving class typecasting have been removed due to a concern that many automated alignment systems would not consider these as potential mappings. One example is PlaceType(x) ↔ rdfs:subClassOf(x, Place). This mapping expresses that the set of individuals of the class PlaceType in the GBO is equivalent to a subclass of the class Place on the GMO side. This is probably a challenge for current automated alignment systems to detect because it involves entities that are not in either the source or target ontology but are rather a construct of the language (e.g. `rdfs:subClassOf`. In addition, we also removed correspondences that involve the inverse relationship, because at the time the reference alignment was created, an evaluation methodology had not yet been finalized for alignment systems on this task. In particular, our thinking was that if an alignment system managed to find a mapping for either a relation or its inverse (e.g. `isGeoFeatureTypeOf`), but not the other (`hasGeoFeatureType`), then it should not be penalized. Even though using semantic precision and recall [25] as the evaluation metric will probably resolve this issue, GeoLink ontologies fail to be expressed in OWL DL, which makes us decide to leave these mappings that involve inverse relations out of the benchmark for the OAEI 2018. After these two modifications, 67 correspondences including simple and complex relations remained in the simplified version of the benchmark. Table 3 presents the remaining patterns and their corresponding number and category in the simplified the version.

---

[21]`http://schema.geolink.org/`

Table 3: The alignment pattern types found in the GeoLink complex alignment benchmark in OAEI 2018, along with the number of times each occurs and the type of relation.

Pattern	Occurrences	Category
Class Equivalence	10	1:1
Class Subsumption	2	1:1
Property Equivalence	7	1:1
Property Typecasting Subsumption	5	1:n
Typed Property Chain Equivalence	26	m:n
Typed Property Chain Subsumption	17	m:n

5.2. Evaluation Results

There are three subtasks related to the evaluation of complex ontology alignment systems in OAEI 2018:

1. **Entity Identification:** For each entity in the source ontology, the alignment systems will be asked to list all of the entities that are related in some way in the target ontology. For example, referring to the example we used above,

$$\mathsf{Award}(x) \wedge \mathsf{hasSponsor}(x, z) \leftrightarrow \mathsf{FundingAward}(x) \wedge \mathsf{providesAgentRole}(x, y)$$
$$\wedge \mathsf{SponsorRole}(y) \wedge \mathsf{performedBy}(y, z)$$

the expected output from an alignment system is that the property `hasSponsor` in the GBO is related to `FundingAward`, `providesAgentRole`, `SponsorRole`, `performedBy` in the GMO and `Award` in the GBO.

2. **Relationship Identification:** Given a dictionary containing entities from the source ontology paired with all related entities, determine the expression that specifies the nature of the relation. So, in terms of the example above in this subtask, an alignment system needs to eventually determine the relationship between two sides is equivalence.

3. **Full Complex Alignment Identification;** A combination of the two former step to determine the complex alignment that exist between the source and target ontology.

All three subtasks were evaluated based on standard precision, recall and F-measure. There were 16 ontology alignment systems that participated in this year's OAEI. Unfortunately, none of the alignment systems were capable of producing results for subtasks 2 and 3 on the GeoLink benchmark. The Table 4 shows the results of the systems that can produce results on subtask 1. There were seven such systems. The performance of these systems are shown in Table 4. Among the alignments produced by these systems, all correspondences identified between the GBO and the GMO were 1-to-1 equivalences. The precision of most of the systems was relatively high, which means that traditional ontology alignment systems can handle the simple relations in this real-world ontology alignment task fairly well. But, it is not surprising that the low recall reflects that current ontology alignment systems are not capable of identifying more complex relations, a situation that we hope will change in future years.

17

Table 4: The Performance of Complex Ontology Alignment Systems in OAEI on Subtask 1

System	Precision	F-measure	Recall
ALOD2Vec	0.78	0.19	0.11
DOME	0.44	0.17	0.11
LogMap	0.85	0.18	0.10
LogMapKG	0.85	0.18	0.10
LogMapLt	0.73	0.19	0.11
POMAP++	0.90	0.17	0.09
Xmap	0.39	0.15	0.09

6. Discussion

This work creates a complex ontology alignment benchmark in real-world ontologies and evaluates the performance of traditional ontology alignment systems. It can be a stepping stone towards deeper understanding and discovery in this area. It is clear that there are still some challenges in the generation, detection, and evaluation of the complex correspondences between real-world ontologies. This section outlines the challenges that we faced during our research and presents some possible methods to solve them.

- **Challenge 1**: The first challenge is how to identify the complex mappings between ontologies, no matter if they are real-world or artificial ontologies, in order to create new benchmarks. So far the process of generating a consensual complex ontology alignment is time-intensive and somewhat tedious, because it requires the ontologists to design or understand the ontologies in the best practice way, and also necessistates that multiple domain experts help the ontologists with the verification of the ground truth manually. This issue could potentially be resolved in the future through creating automated recommendation systems to select and rank the possible entities and relationships from one ontology to another one, which will effectively help people in interactive reference alignment generation. One possible method based on logical RDF compression has been introduced in the paper [8]. We are currently working on this alignment system. And we hypothesize that it will be able to help the researchers to pick the possible mappings between two ontologies effectively.

- **Challenge 2**: The second challenge is how to generate and populate the instance data for the entities in the source and target ontologies. In our GeoLink benchmark, even though there are over 48 million triples provided by the data providers, some entities, like the object property "hasContact" in the GMO, still lacks any corresponding individuals because none of the GeoLink data providers currently use this property. (Note: the GeoLink ontologies were also designed for possible future extension. Therefore, some entities will not be used until the data providers acquire the corresponding dataset in the future.) But the alignment exists between the two ontologies no matter whether the instance data exists or not. Therefore, we still decide to keep these alignment in our reference alignment. The lack of instance data may have a negative impact on the performance of automated complex ontology alignment systems that require instance data to support their algorithms. Similarly, significant instance data is not readily available for most of the artificial benchmarks

18

in OAEI. It is a challenge to supply a large amount of instance data for these benchmarks. One potential method to solve this issue is to first locate useful real-world datasets online based on the domain of the benchmark and then populate the most suitable instance data into the ontologies. For example, it might be possible to incorporate additional geoscience data repositories into the GeoLink Knowledge base to enrich our instance data. However, the amount of real-world instance data may be limited due to a lack of datasets relevant to the domain. In such cases, an artificial population process may be needed to enrich the first step, because the performance of some instance-based ontology alignment systems relies on statistical analysis and computational similarity measures that require a large number of instances [7, 22]. One possible approach might be to use the techniques described in [26].

- **Challenge 3**: The third challenge that we experienced in our research was presenting the complex alignment in EDOAL format and converting between EDOAL and OWL DL. Referring to the example of property typecasting subsumption correspondence in Section 4, we were actually trying to state the following mapping, as expressed in the rule: $hasPlaceType(x, y) \leftrightarrow rdf:type(x, y) \wedge rdfs:subClassOf(y, Place)$. This is currently not supported by the core EDOAL language, because EDOAL is not good at dealing with mapping individuals. Instead of calling this as a mapping, we would probably rather say it as a mapping rule that describes the context of converting datasets from one ontology to another one. But, it seems that it falls outside of the capability of the current automated matching algorithms to detect it directly, as it defines a transformation between entities that are not listed in the ontologies.

  A related problem stems from the inexpressibility of some mappings from the reference alignment in OWL DL. This came up because we originally planned to apply semantic precision and recall [25] as the evaluation metrics to compute the performance of ontology alignment systems on this benchmark, which require a reasoner to test the entailed axioms and therefore need the alignment present in OWL DL. Unfortunately, only 24 of 67 EDOAL expressions in the GeoLink alignment can be expressed in this language. In particular, many mappings that involve typed property chains are valid in EDOAL but not in OWL DL. For example, $Award(x) \wedge hasEndDate(x, z) \leftrightarrow FundingAward(x) \wedge endsOnDate(x, y) \wedge time:Instant(y) \wedge time:inXSDDate(y, z)$. This means that $hasEndDate$ in the GBO is equivalent to the concatenation of $endsOnDate$ and $inXSDDate$ with some additional domain and range restrictions. While this type of concatenation should be unproblematic in terms of semantics, it involves concatenation of an object property with a datatype property, which is not allowed in OWL DL. We are not aware of any good solution to these two issues that we describe here.

- **Challenge 4**: The last and most difficult challenge is how to correctly and accurately evaluate the performance of complex ontology alignment systems. By far, classical precision and recall are the most widely used evaluation metrics to assess performance in the majority of existing work on ontology alignment. However, several complications arise in the use of these metrics when the alignments contain complex relations due to their all-or-nothing syntactic comparisons of individual mappings, which do not distinguish between correspondences that are formally incorrect but closely related to the correct correspon-

19

dences and those that are completely incorrect. For example, this is a mapping in the GeoLink reference alignment:

$$\text{Award}(x) \land \text{hasSponsor}(x, z) \leftrightarrow \text{FundingAward}(x) \land \text{providesAgentRole}(x, y)$$
$$\land \text{SponsorRole}(y) \land \text{performedBy}(y, z)$$

And here are two different mappings that were generated by two different hypothetical complex ontology alignment systems. The first mapping is:

$$\text{Award}(x) \land \text{hasSponsor}(x, z) \leftrightarrow \text{FundingAward}(x) \land \text{providesAgentRole}(x, y)$$
$$\land \text{AgentRole}(y) \land \text{performedBy}(y, z)$$

and the second one is:

$$\text{Award}(x) \land \text{hasSponsor}(x, z) \rightarrow \text{Program}(x) \land \text{providesAgentRole}(x, y)$$
$$\land \text{DataManagerRole}(y) \land \text{performedBy}(y, z)$$

The first mapping is formally incorrect compared to the reference alignment, but it is very closely related to it because `SponsorRole` is a subclass of `AgentRole`. Conversely, the second one is completely incorrect, as it contains incorrect domain and range restrictions of `providesAgentRole` and the relationship between the two sides indicates a subsumption rather than an equivalence relation. Some variations of the traditional precision and recall metrics have been proposed to mitigate the limitations of the basic approach, but these do not resolve all of the issues. For instance, semantic precision and recall [25] compare correspondences based on their semantic meaning rather than their syntactic representation. This is done by applying a reasoner to determine when one mapping is logically equivalent to another. Even though the semantic approaches solve an important problem for evaluating alignments with complex correspondences, they still have several limitations. One is that the reasoning takes a significant amount of time, particularly for large ontologies. Furthermore, such reasoning is not possible at all if the merged ontology is not in OWL DL, like the example introduced in Challenge 3 in our GeoLink benchmark. Therefore, a new evaluation metric will need to be designed to conquer this challenge. This new evaluation metric will need to have more detailed and accurate penalties for different kinds of closeness of entities and relationship comparisons to avoid the all-or-nothing problem in order to provide more nuanced results that can assist researchers in improving their algorithms.

7. Conclusion and Future Work

Complex ontology alignment has been discussed for a long time, but relatively little work has been done to advance the state of the art. The lack of an available complex ontology alignment benchmark may be a primary reason for the slow speed of the development. In addition, most current ontology alignment benchmarks have been created by humans for the sole purpose of evaluating ontology alignment systems, and they may not always represent real-world cases. In this paper, we have proposed a complex ontology alignment benchmark based on the real-world GeoLink knowledge base. The two ontologies and the reference alignment were designed and created by ontologists and geoscience domain experts to support data representation, sharing,

20

integration and discovery. We take advantage of these ontologies to create a complex ontology alignment benchmark. In our benchmark, the alignments not only cover 1:1 simple correspondences, but also contain 1:n and m:n complex relations. All correspondences required to convert between the two ontologies (a key goal of ontology alignment) are guaranteed to be present, because one ontology was consciously created from the other, with SPARQL queries to mitigate each change. In addition, the alignment has been evaluated by domain experts from different organizations to ensure the high quality. Moreover, instance data has been published according to both ontologies, which is important in order to support use of the benchmark by extensional alignment systems. Furthermore, the ontologies and alignments in both rule syntax and EDOAL format have been published in FigShare with an open access license for reusability and can be accessed in OAEI 2018 website as well[22]. The evaluation results of the automated ontology alignment systems that participated in OAEI 2018 are also presented in this paper.

We discuss four challenges in this paper, which we plan to explore in our future work on this topic. Besides this, with respect to the maintenance of the benchmark, our intention is to remain actively involved for years to come in the OAEI complex alignment benchmarking track, and to also develop corresponding alignment methods. We thus have an intrinsic interest in keeping the benchmark maintained and usable, which would, e.g., mean that we are prepared to transfer it to a new benchmarking framework if required in the future. At the same time, based on participants' feedback, we will modify the reference alignment if necessary to perfect the benchmark by making it more convenient to use. This may involve, for example, making the alignment available in additional formats.

Acknowledgment

References

[1] P. Shvaiko, J. Euzenat, Ontology matching: State of the art and future challenges, IEEE Trans. Knowl. Data Eng. 25 (2013) 158–176.

[2] M. Cheatham, P. Hitzler, The properties of property alignment, in: Proceedings of the 9th International Workshop on Ontology Matching collocated with the 13th International Semantic Web Conference (ISWC 2014), Riva del Garda, Trentino, Italy, October 20, 2014., 2014, pp. 13–24. URL: `http://ceur-ws.org/Vol-1317/om2014\_Tpaper2.pdf`.

[3] L. Zhou, M. Cheatham, A. Krisnadhi, P. Hitzler, A complex alignment benchmark: Geolink dataset, in: The Semantic Web - ISWC 2018 - 17th International Semantic Web Conference, Monterey, CA, USA, October 8-12, 2018, Proceedings, Part II, 2018, pp. 273–288. URL: `https://doi.org/10.1007/978-3-030-00668-6\_17`. doi:10.1007/978-3-030-00668-6_17.

[4] J. David, AROMA results for OAEI 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, 2009. URL: `http://ceur-ws.org/Vol-551/oaei09\_paper3.pdf`.

---

[22]`http://oaei.ontologymatching.org/2018/complex/index.html#geolink`
[23]`http://oaei.ontologymatching.org/2018/seals.html` used in the OAEI

21

[5] É. Thiéblin, O. Haemmerlé, N. Hernandez, C. Trojahn, Task-oriented complex ontology alignment: Two alignment evaluation sets, in: The Semantic Web - 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, June 3-7, 2018, Proceedings, 2018, pp. 655–670. URL: `https://doi.org/10.1007/978-3-319-93417-4\_42`. doi:`10.1007/978-3-319-93417-4\_42`.

[6] É. Thiéblin, O. Haemmerlé, N. Hernandez, C. T. dos Santos, Towards a complex alignment evaluation dataset, 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, 2017., 2017, pp. 217–218. URL: `http://ceur-ws.org/Vol-2032/om2017\_poster6.pdf`.

[7] É. Thiéblin, O. Haemmerlé, C. Trojahn, CANARD complex matching system: results of the 2018 OAEI evaluation campaign, 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., 2018, pp. 138–143. URL: `http://ceur-ws.org/Vol-2288/oaei18\_paper4.pdf`.

[8] L. Zhou, A journey from simple to complex alignment on real-world ontologies, in: Proceedings of the Doctoral Consortium at ISWC 2018 co-located with 17th International Semantic Web Conference (ISWC 2018), Monterey, USA, October 8th - to - 12th, 2018., 2018, pp. 93–101. URL: `http://ceur-ws.org/Vol-2181/paper-11.pdf`.

[9] É. Thiéblin, M. Cheatham, C. T. dos Santos, O. Zamazal, L. Zhou, The first version of the OAEI complex alignment benchmark, in: Proceedings of the ISWC 2018 Posters & Demonstrations, Industry and Blue Sky Ideas Tracks co-located with 17th International Semantic Web Conference (ISWC 2018), Monterey, USA, October 8th - to - 12th, 2018., 2018. URL: `http://ceur-ws.org/Vol-2180/paper-67.pdf`.

[10] D. Ritze, C. Meilicke, O. Sváb-Zamazal, H. Stuckenschmidt, A pattern-based ontology matching approach for detecting complex correspondences, 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, 2009. URL: `http://ceur-ws.org/Vol-551/om2009\_Tpaper3.pdf`.

[11] D. Ritze, J. Völker, C. Meilicke, O. Sváb-Zamazal, Linguistic analysis for complex ontology matching, in: Proceedings of the 5th International Workshop on Ontology Matching (OM-2010), Shanghai, China, November 7, 2010, 2010. URL: `http://ceur-ws.org/Vol-689/om2010\_Tpaper1.pdf`.

[12] S. Jiang, D. Lowd, S. Kafle, D. Dou, Ontology matching with knowledge rules, T. Large-Scale Data- and Knowledge-Centered Systems 28 (2016) 75–95.

[13] P. Jain, P. Hitzler, A. P. Sheth, K. Verma, P. Z. Yeh, Ontology alignment for linked open data, in: The Semantic Web - ISWC 2010 - 9th International Semantic Web Conference, ISWC 2010, Shanghai, China, November 7-11, 2010, Revised Selected Papers, Part I, 2010, pp. 402–417. URL: `https://doi.org/10.1007/978-3-642-17746-0\_26`. doi:`10.1007/978-3-642-17746-0\_26`.

[14] F. M. Suchanek, S. Abiteboul, P. Senellart, PARIS: probabilistic alignment of relations, instances, and schema, PVLDB 5 (2011) 157–168.

[15] C. Pesquita, M. Cheatham, D. Faria, J. Barros, E. Santos, F. M. Couto, Building reference alignments for compound matching of multiple ontologies using OBO cross-products, in: Proceedings of the 9th International Workshop on Ontology Matching collocated with the 13th International Semantic Web Conference (ISWC 2014), Riva del Garda, Trentino, Italy, October 20, 2014., 2014, pp. 172–173. URL: `http://ceur-ws.org/Vol-1317/om2014\_poster2.pdf`.

[16] J. You, Geoscientists aim to magnify specialized web searching, 2015.

[17] A. A. Krisnadhi, Y. Hu, K. Janowicz, P. Hitzler, R. A. Arko, S. Carbotte, C. Chandler, M. Cheatham, D. Fils, T. Finin, P. Ji, M. B. Jones, N. Karima, K. A. Lehnert, A. Mickle, T. Narock, M. O'Brien, L. Raymond, A. Shepherd, M. Schildhauer, P. Wiebe, The geolink framework for pattern-based linked data integration, 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., 2015. URL: `http://ceur-ws.org/Vol-1486/paper\_99.pdf`.

[18] P. Hitzler, A. Gangemi, K. Janowicz, A. Krisnadhi, V. Presutti (Eds.), Ontology Engineering with Ontology Design Patterns - Foundations and Applications, volume 25 of *Studies on the Semantic Web*, IOS Press, 2016.

[19] A. Krisnadhi, Y. Hu, K. Janowicz, P. Hitzler, R. A. Arko, S. Carbotte, C. Chandler, M. Cheatham, D. Fils, T. W. Finin, P. Ji, M. B. Jones, N. Karima, K. A. Lehnert, A. Mickle, T. W. Narock, M. O'Brien, L. Raymond, A. Shepherd, M. Schildhauer, P. Wiebe, The geolink modular oceanography ontology, in: The Semantic Web - ISWC 2015 - 14th International Semantic Web Conference, Bethlehem, PA, USA, October 11-15, 2015, Proceedings, Part II, 2015, pp. 301–309. URL: `https://doi.org/10.1007/978-3-319-25010-6\_19`. doi:`10.1007/978-3-319-25010-6\_19`.

[20] A. A. Krisnadhi, P. Hitzler, K. Janowicz, On the capabilities and limitations of OWL regarding typecasting and ontology design pattern views, in: Ontology Engineering - 12th International Experiences and Directions Workshop on OWL, OWLED 2015, co-located with ISWC 2015, Bethlehem, PA, USA, October 9-10, 2015, Revised Selected Papers, 2015, pp. 105–116. URL: `https://doi.org/10.1007/978-3-319-33245-1\_11`.

22

doi:10.1007/978-3-319-33245-1_11.

[21] A. Krisnadhi, Ontology Pattern-Based Data Integration, Ph.D. thesis, Wright State University, 2015.

[22] A. Isaac, L. van der Meij, S. Schlobach, S. Wang, An empirical study of instance-based ontology matching, in: The Semantic Web, 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007, Busan, Korea, November 11-15, 2007., 2007, pp. 253–266. URL: https://doi.org/10.1007/978-3-540-76298-0_19. doi:10.1007/978-3-540-76298-0_19.

[23] J. J. Carroll, I. Dickinson, C. Dollin, D. Reynolds, A. Seaborne, K. Wilkinson, Jena: implementing the semantic web recommendations, in: Proceedings of the 13th international conference on World Wide Web - Alternate Track Papers & Posters, WWW 2004, New York, NY, USA, May 17-20, 2004, 2004, pp. 74–83. URL: https://doi.org/10.1145/1013367.1013381. doi:10.1145/1013367.1013381.

[24] B. Glimm, I. Horrocks, B. Motik, G. Stoilos, Z. Wang, Hermit: An OWL 2 reasoner, J. Autom. Reasoning 53 (2014) 245–269.

[25] J. Euzenat, Semantic precision and recall for ontology alignment evaluation, in: IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007, 2007, pp. 348–353. URL: http://ijcai.org/Proceedings/07/Papers/054.pdf.

[26] B. A. C. Schopman, S. Wang, A. Isaac, S. Schlobach, Instance-based ontology matching by instance enrichment, J. Data Semantics 1 (2012) 219–236.

23

The First Version of the OAEI Complex Alignment Benchmark

Elodie Thiéblin[1], Michelle Cheatham[2], Cassia Trojahn[1], Ondřej Zamazal[3] and
Lu Zhou[2]

[1] IRIT & Université de Toulouse 2 Jean Jaurès, Toulouse, France
[2] Wright State University, Dayton, USA
[3] University of Economics, Prague, Czech Republic
`elodie.thieblin@irit.fr, michelle.cheatham@gmail.com,`
`cassia.trojahn@irit.fr, ondrej.zamazal@vse.cz, zhou.34@wright.edu`

Abstract. We present the first version of the complex benchmark of the
Ontology Alignment Evaluation Initiative campaigns. This benchmark is
composed of four datasets from different domains (conference, hydrology,
geoscience and agronomy) and covers different evaluation strategies.

Keywords: complex ontology alignments, evaluation dataset, OAEI

1   Introduction

Complex correspondences involve transformation functions of literal values or
logical constructors (e.g. $\forall x,\ ekaw{:}AcceptedPaper(x) \equiv \exists y,\ cmt{:}acceptedBy(x,y)$),
which make them more expressive than simple correspondences. Complex align-
ments, composed of at least one complex correspondence, are therefore a com-
plement to simple alignments. Different approaches for complex matching have
emerged in the literature [2,4,5,8]. Most of them, however, have been evaluated
on tailored datasets (e.g., targeting a specific correspondence pattern). Most ef-
forts on systematic evaluation, in the context of the OAEI campaigns[1], are still
dedicated to simple matchers.

   This paper presents the first version of the OAEI complex track, composed of
four datasets (Table 1) from different domains. This domain and correspondence
variety allows for better covering different kinds of heterogeneity between ontolo-
gies. Different evaluation strategies aim at evaluating complex matchers under
different perspectives. The evaluation will be supported by the SEALS platform
and the output alignments must be in EDOAL. The detail of each dataset and
evaluation process can be found on the OAEI's 2018 complex track webpage[2],
and are introduced in the following.

---

[1] `http://oaei.ontologymatching.org/`
[2] `http://oaei.ontologymatching.org/2018/complex/index.html`

Dataset	Ontologies	(1:1)	(1:n)	(m:n)
Conference consensus	3	78	79	0
Hydrography	4	113	69	15
GeoLink	2	24	15	72
Taxon	4	6	17	3

Table 1. Number of ontologies and correspondences by kind in each dataset. (1:1) are simple correspondences, (1:n) and (m:n) are complex correspondences.

2 Conference consensual dataset

This dataset is based on the OntoFarm dataset [9], which is composed of 16 ontologies on the conference organisation domain and simple reference alignments between 7 of them. Here, we consider 3 out of the 7 ontologies from the reference alignments (*cmt*, *conference* and *ekaw*), resulting in 3 alignment pairs. The alignments involve both logical constructors (76 correspondences) and transformations (3 correspondences). Examples are given in the following :

1. $\forall x$, *ekaw:AcceptedPaper(x)* $\equiv \exists y$, *cmt:acceptedBy(x,y)* is a correspondence with the *existential* constructor.
2. $\forall x,y$, *cmt:name(x,y)* $\equiv \exists y_1, y_2$, *conference:has_the_first_name(x,y$_1$)* $\wedge$ *conference:has_the_last_name(x,y$_2$)* $\wedge$ *concatenation(y,y$_1$," ", y$_2$)*, where *concatenation(a,b$_1$, b$_2$, ...)* is a predicate ensuring that its first parameter a is equal to the string concatenation of the others $\{b_1, b_2, ...\}$. It uses a transformation function of the literal values.

The alignments have been manually created by three experts in the domain, following the methodology in [7]. Four experts assessed the generated correspondences to reach a consensus. The systems will be manually evaluated on their output alignments to produce precision and recall scores. Only the complex equivalence correspondences will be assessed. The systems can use a simple reference alignment as input. Confidence scores of correspondences will not be taken into account in the evaluation.

3 Hydrography dataset

The hydrography dataset is composed of 4 source ontologies (Hydro3, hydrOntology_native, hydrOntology_translated and Cree) that each should be aligned to a single target Surface Water Ontology (*swo*). The source ontologies vary in their similarity to the target ontology – Hydro3 is similar in both language and structure, hydrOntology_native and hydrOntology_translated are similar in structure but hydrOntology_translated is in Spanish rather than English, and Cree is very different in terms of both language and structure. The alignments were created by a geologist and an ontologist, in consultation with a native Spanish speaker regarding the hydrOntology_translated, and consist of logical relations such as the one shown below.

1. $\forall x$, *hydrOntology_translated:Aguas_Corrientes(x)* $\leq$ *swo:SurfaceFeature(x)* $\wedge$ *swo:Waterbody(x)* $\wedge \exists y$, *swo:hasFlow(x,y)* $\wedge$ *swo:Flow(y)*

Performance on this dataset will be evaluated on three sub-tasks: 1) identifying the atoms (classes and properties) from the target ontology involved in the relations (e.g., *swo:SurfaceFeature*, *swo:Waterbody*, *swo:hasFlow* and *swo:Flow* from the correspondence above), 2) when given the atoms, identifying the logical relations that hold between them and 3) the full complex alignment task. Evaluation of the first sub-task will use traditional F-measure, while the remaining two subtasks will be evaluated on semantic F-measure [1].

4   GeoLink dataset

This dataset is from the GeoLink project[3], which was funded under the U.S. National Science Foundation's EarthCube initiative. It is composed of 2 populated ontologies: the GeoLink base ontology (*gbo*) and the GeoLink modular ontology (*gmo*). The GeoLink project is a real-world use case of ontologies. The alignment between the ontologies was developed in consultation with domain experts from several Geoscience research institutions. The complex correspondences include not only class and property subsumption and property chains (described in [5]), but also some that involve typecasting (c.f. [3]), for example:

1. Property Chain: $\forall x,z,$ *gbo:Award(x)* $\land$ *gbo:hasSponsor(x,z)* $\equiv$ $\exists y,$ *gmo:FundingAward(x)* $\land$ *gmo:providesAgentRole(x,y)* $\land$ *gmo:SponsorRole(y)* $\land$ *gmo :performedBy(y,z)*
2. Class Typecasting: $\forall x,$ *gbo:PlaceType(x)* $\equiv$ *rdfs:subClassOf(x, gmo:Place)*

More information about this dataset can be found in [10] and the benchmark and alignment can be downloaded here[4]. The performance of alignment systems on this dataset will be evaluated in the same way as the hydrography dataset.

5   Taxon dataset

This dataset is composed of 4 populated ontologies whose common scope is plant taxonomy: AgronomicTaxon (*agtx*), Agrovoc (*agv* and *agronto*), DBpedia (*dbo*) and TaxRef-LD (*txr*). This dataset extends the one proposed in [6] by adding the TaxRef-LD ontology. The alignments were manually created with the help of one expert and involve only logical constructors, as for example:

1. $\forall x,$ *agtx:GenusRank(x)* $\equiv$ *agronto:hasTaxonomicRank(x,agv:c_11125)*
2. $\forall x,$ *agtx:GenusRank(x)* $\equiv \exists y,$ *dbo:Species(y)* $\land$ *dbo:genus(y,x)* $\land$ *dbo:Species(x)*

The evaluation of this dataset is task-oriented. We will evaluate the generated correspondences using a SPARQL query rewriting system and manually measure their ability of answering a set of queries over each dataset. For example, a competency question could be "Retrieve all the genus taxa". For AgronomicTaxon, as source ontology, the corresponding SPARQL query is `SELECT ?x WHERE {?x a agtx:GenusRank.}` and the correspondences output by the systems with Agrovoc as target ontology, should be able to translate the query into: `SELECT ?x WHERE {?x agronto:hasTaxonomicRank agv:c_11125.}`

---

6 Conclusions

This paper has presented the first OAEI complex evaluation track, covering different kinds of complex correspondences, domains and evaluation strategies. For most datasets, the evaluation is still manually performed, opening directions on how complex alignments can be automatically generated and evaluated.

References

1. Euzenat, J.: Semantic precision and recall for ontology alignment evaluation. In: IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007. pp. 348–353 (2007)
2. Jiang, S., Lowd, D., Kafle, S., Dou, D.: Ontology matching with knowledge rules. In: Transactions on Large-Scale Data-and Knowledge-Centered Systems, pp. 75–95 (2016)
3. Krisnadhi, A.A., Hitzler, P., Janowicz, K.: On the capabilities and limitations of OWL regarding typecasting and ontology design pattern views. In: Ontology Engineering - 12th International Experiences and Directions Workshop on OWL, OWLED 2015, co-located with ISWC 2015, Bethlehem, PA, USA, October 9-10, 2015, Revised Selected Papers. pp. 105–116 (2015)
4. Parundekar, R., Knoblock, C.A., Ambite, J.L.: Discovering concept coverings in ontologies of linked data sources. In: ISWC. pp. 427–443 (2012)
5. Ritze, D., Meilicke, C., Šváb Zamazal, O., Stuckenschmidt, H.: A pattern-based ontology matching approach for detecting complex correspondences. In: 4th OM workshop. pp. 25–36 (2009)
6. Thiéblin, E., Amarger, F., Hernandez, N., Roussey, C., Trojahn, C.: Cross-querying lod datasets using complex alignments: an application to agronomic taxa. In: MTSR. pp. 25–37 (2017)
7. Thiéblin, E., Haemmerlé, O., Hernandez, N., Trojahn, C.: Task-oriented complex ontology alignment – two alignment evaluation sets. In: ESWC (2018), (to appear)
8. Walshe, B., Brennan, R., O'Sullivan, D.: Bayes-recce: A bayesian model for detecting restriction class correspondences in linked open data knowledge bases. International Journal on Semantic Web and Information Systems 12(2), 25–52 (2016)
9. Zamazal, O., Svátek, V.: The Ten-Year OntoFarm and its Fertilization within the Onto-Sphere. Journal of Web Semantics 43, 46–53 (2017)
10. Zhou, L., Cheatham, M., Krisnadhi, A., Hitzler, P.: A complex alignment benchmark: Geolink dataset. In: ISWC. Springer (2018), (to appear)

The Enslaved Dataset: A Real-world Complex Ontology Alignment Benchmark using Wikibase

Lu Zhou
Cogan Shimizu
Pascal Hitzler
luzhou@ksu.edu
coganmshimizu@ksu.edu
hitzler@ksu.edu
Kansas State University
Manhattan, Kansas, USA

Alicia M. Sheill
Seila Gonzalez Estrecha
Catherine Foley
Duncan Tarr
Dean Rehberger
asheill@msu.edu
seilage@msu.edu
foleyc@msu.edu
tarrdunc@msu.edu
rehberge@msu.edu
Michigan State University
East Lansing, Michigan, USA

ABSTRACT

Ontology alignment has taken a critical place for helping heterogeneous resources to interoperate. It has been studied for over a decade, and over that time many alignment systems and methods have been developed by researchers to find simple 1:1 equivalence matches between two ontologies. However, very few alignment systems focus on finding complex correspondences. Even if the complex alignment systems are developed, the performance of finding complex relations still has a lot of room for improvement. One reason for this limitation may be that there are still few applicable alignment benchmarks that contain such complex relationships that can raise researchers' interests. In this paper, we propose a real-world dataset from the Enslaved project as a potential complex alignment benchmark. The benchmark consists of two resources, the Enslaved Ontology along with a Wikibase repository holding a large number of instance data from the Enslaved project, as well as a manually created reference alignment between them. The alignment was developed in consultation with domain experts in the digital humanities. The alignment not only includes simple 1:1 equivalence correspondences, but also more complex m:n equivalence and subsumption correspondences and are provided in both Expressive and Declarative Ontology Alignment Language (EDOAL) format and rule syntax. The Enslaved benchmark has been incorporated into the Ontology Alignment Evaluation Initiative (OAEI) 2020 and is completely free for public use to assist the researchers in developing and evaluating their complex alignment algorithms.

CCS CONCEPTS

• **Information systems** → **Information integration**; **Resource Description Framework (RDF)**; **Web Ontology Language (OWL)**; **Ontologies**; • **Computing methodologies** → **Ontology engineering**.

KEYWORDS

Ontology Alignment; Knowledge Graph; Wikibase; Benchmark

1 INTRODUCTION

Ontology alignment is an important step in enabling computers to query and reason across the immense amounts of linked data on the semantic web. It has been considered a "silver bullet" for the semantic heterogeneity problem faced by computer systems. Ontology alignment is a difficult challenge as the ontologies, which are used as knowledge graph schemas, that underlie different linked data can vary significantly in terms of subject area coverage, level of abstraction, ontology modeling philosophy, and language. Due to the importance and difficulty of the ontology alignment problem, it has been an active area of research for over a decade [21].

Ideally, alignment systems should be able to uncover any entity relationships across two ontologies that can exist within a single ontology. Such relationships have a wide range of complexity, from basic 1:1 (1-to-1) equivalence, such as a Person in one ontology being equivalent to a Human in another ontology, to arbitrary m:n (m-to-n) relationships, such as a Professor with a hasRank property value of "Assistant" in one ontology being a subclass of the union of the Faculty and TenureTrack classes in another. Unfortunately,

the majority of the research activities in the field of ontology alignment remain focused on the simplest end of this scale – finding 1:1 equivalence relations between ontologies. Part of the reason for this may be that there are still few widely used and accepted ontology alignment benchmarks that involve complex relations. Even though some benchmarks containing complex relations were proposed in Ontology Alignment Evaluation Initiative (OAEI) in 2018 [22], the performance of the alignment systems is still relatively poor when detecting these complex correspondences between two ontologies [1].

Wikibase is the powerful knowledge base software that drives Wikidata [24]. Wikidata is an immense, crowdsourced knowledge base with persistent data that is available for public use and consumption. It would be very difficult to have an ontology of everything, but Wikidata is probably close enough for this purpose. It contains millions of pieces of knowledge from many different domains in the world. In addition, Wikidata is crowdsourced and can act as a "common resource".[1] People can export data to Wikidata so that it is publicly persistent in an open and transparent manner. Wikidata is an instance of Wikibase. Any organization can adapt it to their own needs, including setting up their own Wikibase repositories to host their data under different licenses, so that the other instances of Wikibase can be linked with the data on Wikidata. Therefore, it is crucial to be able to find alignments between domain or proprietary ontologies and this common resource. It is also a fact that some organizations have their own internal and proprietary knowledge graphs. They can apply their alignments to this public resource as an important tool to augment or induce new information into their own knowledge graph.

This paper seeks to take a step in that direction by proposing a complex alignment benchmark based on two knowledge graphs: the Enslaved knowledge graph, that was developed by ontology engineers and domain experts together for the Enslaved project, and the Wikibase repository storing historical slaved trade data collected from different provenances. The Enslaved benchmark, including the reference alignment, can be considered to be objective as it was created for deployment and not for benchmarking. It is realistic, since it captures an application use case developed for the historical slave trade, and it is a valid ground truth alignment, as the reference alignment was developed together by historian domain experts and ontology engineers. Therefore, it is rather unique in nature and will inform complex ontology alignment research from a practical and applied perspective, rather than an artificial one. The main contributions of this paper are therefore the following:

- Introduction of two knowledge graphs to support data representation, sharing, integration, and discovery for the Enslaved project;
- Creation of alignment between these two knowledge graphs that include 1:1 and m:n correspondences. Given the creation steps and usage of the alignment, it is fair to say that the alignment will constitute a gold-standard reference;
- Publication of the benchmark alignment in both rule syntax and EDOAL format[2] at a persistent URL[3] under a CC-BY 4.0

license, and it is also incorporated into the complex ontology alignment track in OAEI 2020;[4]

- Evaluation of the quality and validity of this benchmark by using a complex alignment system from OAEI and a discussion of the results in detail.

The rest of this paper is organized as follows. Section 2 discusses the few existing ontology alignment benchmarks that involve relationships other than 1:1 equivalence. Section 3 gives further background on the Enslaved project, including its Wikibase repository, knowledge graph schema, and property reification. Section 4 discusses the alignment between two resources, along with some descriptive statistics, an analysis of the types of correspondences constituting the alignment, and the performance of a complex ontology alignment system tested on the Enslaved benchmark to evaluate the quality of the benchmark. Section 5 concludes with a discussion of potential future work in this area.

2 RELATED WORK

Most work associated with evaluating the performance of ontology alignment systems has been done in conjunction with the Ontology Alignment Evaluation Initiative (OAEI).[5] These yearly events allow developers to test their alignment systems on various tracks that evaluate performance on different facets of the problem such as instance matching, knowledge graph matching, and interactive matching, among others. Currently, most of these tracks involve the identification of 1:1 equivalence relationships, such as Person being equivalent to Human. A discussion at the last two Ontology Matching workshops[6] made it clear that the community is interested in complex ontology alignment, but that lack of applicable benchmarks is hindering progress. In OAEI 2018, the complex ontology alignment track was proposed and organized for the first time [22]. The first version of the complex track is comprised of four benchmarks containing complex relationships from the conference, hydrography, ocean science, and plant taxonomy domains respectively. In OAEI 2019, in order to extend the functionality of the benchmarks and provide more scalability for researchers to explore algorithms that depend on the instance data, Thieblin et al. populated the Conference benchmark with some instances collected from the Extended Semantic Web Conference (ESWC), along with some synthetic data.[7], and Zhou et al. also populated a large number of real-world instances that are currently used in the GeoLink Project[8] as part of the GeoLink benchmark [28]. In addition, different evaluation strategies were applied in evaluating the performance of complex alignment systems on different benchmarks. More details of evaluations and results can be accessed on the OAEI website.[9][10]

Wikidata is a free and open knowledge base that covers many interesting topics, with similar coverage to Wikipedia. There are several ways to access Wikidata; there are built-in tools, external

[1]https://en.wikipedia.org/wiki/Commons
[2]http://alignapi.gforge.inria.fr/edoal.html
[3]https://doi.org/10.6084/m9.figshare.12400976

[4]http://oaei.ontologymatching.org/2020/complex/index.html#popenslaved
[5]http://oaei.ontologymatching.org
[6]http://www.ontologymatching.org/
[7]https://framagit.org/IRIT_UT2J/conference-dataset-population
[8]https://www.geolink.org/
[9]http://oaei.ontologymatching.org/2018/complex/index.html
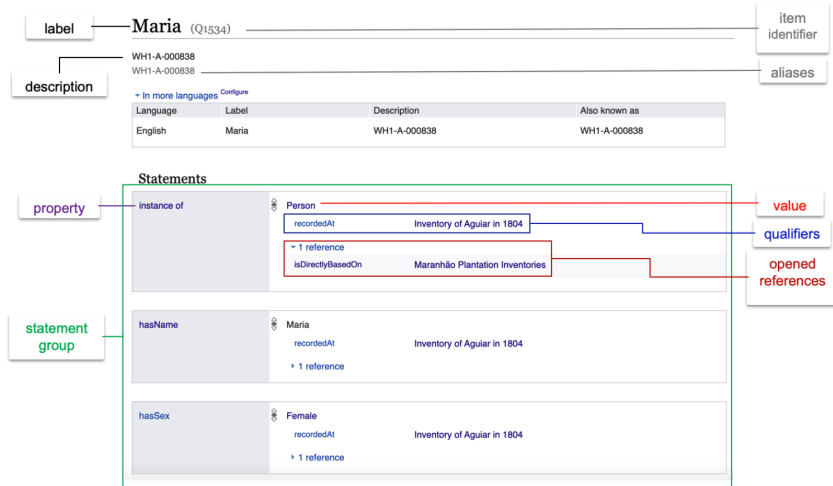[10]http://oaei.ontologymatching.org/2019/complex/index.html

Figure 1: An Excerpt of the Wikibase Repository for the Enslaved Project

tools, or programming interfaces, such as Wikidata Query[11] and Reasonator[12] for searching and examining Wikidata items. Therefore, it has also been considered as a useful external knowledge base for many alignment tasks, particularly for instance matching or entity resolution. For example, Geiß et al. utilized the information on locations and places extracted from Wikidata as ground truth for their entity resolution task [8]. However, the Wikidata knowledge graph hasn't been widely used for ontology alignment tasks due to its young age relative to Wikipedia and DBpedia [16]. The knowledge graphs track in OAEI [12] executes the DBpedia Extraction Framework on several different Wikis from Fandom,[13] which is one of the most popular Wiki Farms and generates several knowledge graphs for both the instance matching (i.e., entities derived from pages about the same real-world entity in different Wikis) and schema matching (i.e., classes and properties derived from different constructs in different Wikis).

The Enslaved benchmark we describe herein differs from current benchmarks in OAEI in some aspects. First, the Enslaved benchmark is a good reflection of real-life data since the Enslaved project comprises over 33 million triples currently from real-world datasets shared by different researchers and contributors from different institutions, while the conference benchmark only consists of synthetic instance data. Second, the Enslaved benchmark utilizes a Wikibase repository as a central storage repository to represent the knowledge in the historical slave trade domain. Anyone with any level of expertise can access the content in the same way they access Wikidata and make use of the knowledge graph. So, it greatly improves the availability of the benchmark. Furthermore, it is useful and important to align the domain ontology to the Wikidata schema, in order to further enrich other external knowledge graphs. Third,

the two knowledge graph schemas in the Enslaved benchmark are completely designed and modeled independently, while the two ontologies in the GeoLink benchmark were developed together for the same project, which may not be common occurrences. Therefore, it further improves the quality and generality of the benchmark and can be considered as a potential good benchmark for complex alignment research.

3 BACKGROUND

3.1 The Enslaved Project and Ontology

The Enslaved Ontology was developed as part of an ongoing project entitled Enslaved: Peoples of the Historical Slave Trade[14] and funded by The Andrew W. Mellon Foundation where the focus is on tracking the movements and details of peoples in the historical slave trade. It further acts as an interchange format between a number of heterogeneous data formats among other projects in the digital humanities, because it has built a proof of concept for a slave data hub. At the heart of the project, the Enslaved Hub allows students, researchers, and the general public to search over numerous databases to reconstruct the lives of individuals who were part of the historical slave trade. The Enslaved project leverages Linked Open Data (LOD) techniques, including the use of Wikibase and a graph database, to create an innovative and compelling centralized Hub for engaging with historical slave trade data from a variety of sources. LOD is a method of exposing, sharing, and connecting data on the semantic web. Data from the different sources is standardized, aggregated, and formatted in such a way that it is machine-readable and is predicated on the relationship between data as developed with the Enslaved Ontology. The central notion of the Enslaved Ontology models *records* of historical agents [19].[15]

---

[11]https://query.wikidata.org/
[12]https://tools.wmflabs.org/reasonator/
[13]https://www.fandom.com/

[14]https://enslaved.org/
[15]Comprehensive documentation can be found in [20]

The key observation is that the ontology is necessarily a secondary (or further) source and thus cannot purport to state ontological truth. As such, it models, instead, the observations that historians or record keepers have made over time.

The development of the ontology was a collaborative effort and was carried out using a modular ontology modeling approach based on ontology design patterns [4, 7, 13]. Such a methodology is designed to ensure high quality and reusability of the ontology, as well as cater to future expansions, both in terms of scope and in terms of granularity. This allows the Enslaved Ontology to adapt as needs evolve and the number of researchers and contributors increases. The modular ontology modeling approach and its rationale have been described in [15], and it is closely related to the eXtreme Design approach [3]. The modeling team included domain experts, data experts, software developers, and ontology engineers.

The primary purpose of the formal axiomatization is to disambiguate the model, i.e., we were striving for as complete an axiomatization as possible while avoiding ontological over-commitments. Each axiom was discussed in detail between the ontology engineers and the historians on the team. The axiomatization is expressed using the OWL 2 DL profile. The primary goal was not to do formal reasoning over the ontology, but it was authored in such a way as to not rule out such goals in the future (e.g. the use of reasoning for consistency checking) [14].

3.2 The Wikibase Repository and Wikidata Knowledge Graph Schema

Wikibase is a powerful, flexible, and customizable knowledge base software. Its primary components are the Wikibase Repository, an extension for storing and managing data. Wikibase makes collaboration easy for humans and machines alike, and its data model prioritizes language independence and knowledge diversity.

Wikidata is the largest website that is powered by Wikibase. It is an open knowledge base that was launched in 2012. Similar to all the other projects of Wikimedia, anyone can freely edit it.[16] The main goal of Wikidata is to act as central storage for the structured data to provide support for Wikipedia. However, it has grown out of that, since it provides structured linked data about lots of interesting topics in the world, and it is licensed under Creative Commons CC-Zero, which is very close to the public domain and anyone can use it for any purpose. Wikibase is the software that Wikidata has utilized for such success. The Enslaved project uses its own installation of the Wikibase platform to a similar purpose, creating the Enslaved Hub, as mentioned in the previous section. For brevity, we will use the acronym, EKG for the Enslaved Wikibase Knowledge Graph Schema, and EWI for the Enslaved Project's Wikibase installation.

Figure 1 shows the Enslaved Wikibase page for a Person named Maria that appears in the Enslaved benchmark. In the center, we can see the language and label of the entire descriptions of what it means. The important thing is that Maria could be ambiguous because multiple person records may have the same name or there could be other items which are called Maria. To make this item uniquely identified, an item identifier is used as a Q followed by a number, such as Q1534 in this case.
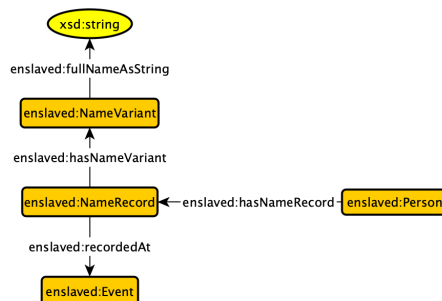
Figure 2: The Example of Enslaved Knowledge Graph for the Enslaved Project

The main part of any Enslaved Wikibase page is the statements section that can be seen in the center of Figure 1. For example, there is an object property in the EKG called *instance of* with the value of Class *Person.* It can be interpreted that an entity *Maria* connects to an entity *Person* by an edge. The edge is labeled as *instance of.* Properties in Wikibase have a P prefix followed by a number, such as instance of (P1), hasName (P20), and hasSex (P31). The references are used to point to specific sources that back up the data provided in a statement. For instance, the statement, "Maria is an instance of Person" which "is directly based on" Maranhão Plantation Inventories [11]; the latter statement allows an interested user to track the provenance of the information contained in the previous statement. This single example is just a small excerpt of EKG.

We mapped the OWL classes and properties in the Enslaved ontology with the items and properties in the Enslaved Wikibase knowledge graph. To the authors' knowledge, this is the first time an OWL ontology had been mapped onto a Wikibase installation. The Enslaved project team found that Wikibase was especially useful for organizing the historical slave trade data, as it had built-in tools that, for example, add qualifiers and references to every statement about the Enslaved data. Such features helped to connect time-bound statements to specific events and connect provenance information to each data point. The EWI stores the instance data, including all of the controlled vocabularies and multiple examples of people, events, and places. Through this process the exact manner in which people are connected to events, events are connected to places at specific periods of time, and how every piece of data is attached to provenance information can be examined. The work mapping the raw data onto the Enslaved ontology via Wikibase has proven that the fields developed for the Enslaved Hub can in fact represent diverse datasets.

3.3 Property Reification

Property reification is a classic strategy for adding context to a property. We mention this here, in particular, as it is frequently utilized in the Enslaved Ontology and Wikibase repository. Two such examples can be found in Figures 2 and 3.
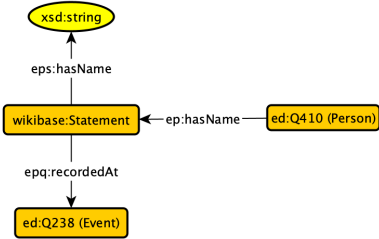
Figure 3: The Example of Wikibase Repository for the Enslaved Project

In the abstract, property reification is an addition of a node between two other nodes, in a knowledge graph. For example, in Figure 2, the property of hasName is reified. That is, there is not a direct link from the enslaved:Person to the xsd:string. This is done in order to add further context. Within the confines and purposes of the Enslaved Ontology, a name is not a constant thing. Enslaved people frequently went by aliases, different family names, incorrect spellings in documents, and so on. By reifying the property of having a name to account for both historical records and the temporality of a name, we have a more accurate, historical model of that particular enslaved person.

In any Wikibase repository, reification is also very common. In order to track provenance in the background, many items are reified.

4 THE ENSLAVED COMPLEX ONTOLOGY ALIGNMENT BENCHMARK

In this section, we present the details of the Enslaved benchmark, including the process of dataset preparation, some descriptive statistics of the benchmark, the types of alignment rules and correspondence patterns with examples, the expression format of reference alignment, and the evaluation of the quality of the benchmark.

4.1 Dataset

The Enslaved benchmark consists of two knowledge graphs. The first knowledge graph utilizes the Enslaved ontology generated by the ontology engineers in consultation with domain experts from different institutions [19]. The Enslaved ontology serves as the underlying schema for the Enslaved knowledge graph in order to enable the historical slave trade data sharing and integration for different communities. The other one is the Enslaved Wikibase repository that is currently being built and used on Wikibase, which employs a completely different schema.

The Enslaved knowledge base currently contains over 33 million triples which are formatted according to the Enslaved Wikibase repository schema, and the number of triples is still growing as additional researchers and contributors start sharing their data for the integration. In order to utilize these two Enslaved knowledge graphs to establish a complex ontology alignment benchmark and facilitate the convenient storage and distribution for OAEI, we decided to pare down the size by only populating part of the instance

data into the benchmark. For each reference alignment between two knowledge graph schemas, we randomly selected up to 500 instances and populated them into both knowledge graphs with their underlying schema. If there is increasing demand of more instance data which are not related to reference alignment in the future, we can provide more data which can be found in the Enslaved Wiki pages.[17]

After finishing up the population of the instance data into the Enslaved knowledge graph and the Enslaved Wikibase knowledge graph, Table 1 shows the number of classes, properties, axioms, and instances in both resources respectively. Both of the knowledge graphs are comparable in size to the benchmarks currently used by the OAEI, which means that they are within the capabilities that most current ontology alignment systems to handle.

Table 1: The number of classes, properties, axioms and instances in two knowledge graph schema

Ontology	Classes	Properties	Axioms	Instances
Enslaved Knowledge Graph	43	75	67,613	13,763
Enslaved Wikibase Knowledge Graph	20	50	83,512	18,415

4.2 Simple and Complex Correspondences

There are two different types of correspondences, which are simple correspondence and complex correspondence [27]. Simple correspondence refers to basic 1:1 simple alignment between two ontologies, such as 1:1 class equivalence, property equivalence, and 1:1 class subsumption, property subsumption. Complex correspondence usually consists of more complex patterns compared to simple correspondence. It may comprise more than one class or property in both ontologies, such as 1:n equivalence, m:n equivalence, and m:n arbitrary relationship. With respect to the correspondence patterns, Zhou et al. list roughly 12 different types of simple and complex correspondence patterns [27]. In the Enslaved benchmark, there are three different types that emerge most frequently in ontology matching tasks, which are listed in Table 2. In the following, we explain the alignment types with a formal pattern and example for each. Some namespaces that are frequently used in the following examples are listed below.

```
@prefix ed:<https://lod.enslaved.org/entity/> .
@prefix ep:<https://lod.enslaved.org/prop/> .
@prefix eps:<https://lod.enslaved.org/prop/statement/> .
@prefix epq:<https://lod.enslaved.org/prop/qualifier/> .
@prefix wikibase:<http://wikiba.se/ontology#> .
@prefix enslaved:<https://enslaved.org/ontology/> .
```

- **Class Equivalence.** is simple 1:1 class equivalence. Classes C_1 and C_2 are from ontology O_1 and ontology O_2, respectively.

  Formal Pattern: $C_1(x) \leftrightarrow C_2(x)$
  Example: enslaved:Person$(x) \leftrightarrow$ ed:Q410(Person)(x)

  Note that ed:Q410 has the label of Person in Wikidata.[18] In order to better understand the example, we use both the

---

[17]https://lod.enslaved.org/wiki/Meta:Main_Page
[18]https://lod.enslaved.org/wiki/Q410

Table 2: The alignment pattern types found in the Enslaved complex alignment benchmark, along with the number of times each occurs and the type of relation.

Pattern	Occurrences	Category
Class Equivalence	15	1:1
Typed Property Chain Equivalence	67	m:n
Typed Property Chain Subsumption	16	m:n

unique identifier, which is presented by a Q prefix followed by a number and the label in the alignment rules. However, in the real alignment rules, only the identifier is kept. This also applies to all the following examples.

- **Typed Property Chain Equivalence.** A property chain is a classical complex pattern that was introduced by Ritze et al. [18]. The pattern applies when a property, together with a type restriction on one or both of its fillers, in one ontology has been used to "flatten" the structure of the other ontology by short-cutting a property chain in that ontology. The pattern also ensures that the types of property fillers involved in the property chain are typed appropriately in the other ontology. The formal pattern and example are shown below. The classes D_i and property r are from ontology O_1, and classes C_i and properties p_i are from ontology O_2.

Formal Pattern:

$$D_1(x_1) \wedge r(x_1, x_{n+1}) \wedge D_2(x_{n+1}) \leftrightarrow C_1(x_1) \wedge p_1(x_1, x_2) \wedge C_2(x_2)$$
$$\wedge \cdots \wedge p_n(x_n, x_{n+1}) \wedge C_{n+1}(x_{n+1})$$

Example:

$$\text{enslaved:Person}(x) \wedge \text{enslaved:hasSexRecord}(x, y) \wedge$$
$$\text{enslaved:SexRecord}(x, y) \wedge \text{enslaved:hasValue}(y, z) \wedge$$
$$\text{enslaved:SexTypes}(z) \leftrightarrow \text{ed:Q410(Person)}(x) \wedge$$
$$\text{ep:P31(hasSex)}(x, y) \wedge \text{wikibase:Statement}(y) \wedge$$
$$\text{eps:P31(hasSex)}(y, z) \wedge \text{ed:Q291(Sex)}(z)$$

Note that in this and all following patterns, any of the D_i or C_i may be omitted (in which case they are essential ⊤). Also, for the left-to-right direction, we assume that $x_2, \ldots x_n$ are existentially quantified variables.

- **Typed Property Chain Subsumption.** This is identical to the Typed Property Chain Equivalence pattern except that the relationship only holds in one direction.

Formal Pattern:

$$D_1(x_1) \wedge r(x_1, x_{n+1}) \wedge D_2(x_{n+1}) \leftarrow / \rightarrow C_1(x_1) \wedge p_1(x_1, x_2) \wedge C_2(x_2)$$
$$\wedge \cdots \wedge p_n(x_n, x_{n+1}) \wedge C_{n+1}(x_{n+1})$$

Example:

$$\text{enslaved:Person}(x) \wedge \text{enslaved:hasNameRecord}(x, y) \wedge$$
$$\text{enslaved:NameRecord}(x, y) \leftarrow \text{ed:Q410(Person)}(x) \wedge$$
$$\text{ep:P20(hasName)}(x, y) \wedge \text{wikibase:Statement}(y)$$

4.3 Format in EDOAL and Rule Syntax

Most ontology alignment benchmarks are formatted according to the format provided by the Alignment API [5]. The standard alignment format is not expressive enough to capture complex relations. Fortunately, the Alignment API also provides a format called EDOAL that can be used to express these types of complex relations. This format can be read and manipulated programmatically using the Alignment API and is therefore very convenient for ontology alignment researchers. In addition, EDOAL is already accepted by the ontology alignment community. It has been used by others when proposing new alignment benchmarks (e.g. [23, 27]) and we continue that approach here. Because EDOAL can be difficult for humans to parse quickly, we have also expressed the alignments in using a naive rule syntax. The rule presentation is not intended for programmatic manipulation, but rather to make it easier for humans to read and understand the alignments. Both versions of the alignment, along with the ontologies, can be downloaded from http://doi.org/10.6084/m9.figshare.12400976 under the Creative Commons CC-BY 4.0 license. We apply HermiT [9] reasoning to the ontologies independently to check satisfiability, since some EDOAL mappings which are part of our benchmark do not seem to be expressible in OWL DL. The Enslaved project website[19] and Enslaved data in Wikibase repository website[20] contains more detailed information, and corresponding documentation of the project which provides users with more insights about the resource, such as all entities and relationships between them. The complex ontology alignment track in OAEI 2020[21] also introduces the detailed information of the Enslaved benchmark, including the benchmark download link and method to evaluate the performance.

4.4 Evaluation using Complex Alignment Systems

In order to examine the quality of this benchmark to see if it is within the capability of current complex ontology alignment systems to handle in OAEI, we apply the Association Rule-based Ontology Alignment System (AROA) [25, 26] on this benchmark since AROA participated in the evaluation of OAEI 2019 and achieved the best performance in terms of F-measure [2]. Table 3 lists the relaxed precision, recall, and F-measure [6] with different thresholds of minimum support and minimum confidence in association rule mining [10, 17]. Minimum support refers to an indication of how frequently the itemset appears in the dataset, while minimum confidence refers to an indication of how often the rule has been found to be true. From Table 3, we can find that the best precision is 0.94 when the minimum support with a value of 0.03 and the minimum confidence with a value of 0.5. And it reaches the best recall of 0.39 when the minimum support value is 0.01, and the minimum confidence value is 0.5. The best F-measure is 0.51, which is achieved when the minimum support and minimum confidence are 0.01 and 1.0 respectively. Overall, the higher the minimum support and minimum confidence, the higher the precision. The lower the minimum support and minimum confidence, the higher the recall. In terms of F-measure, for the same minimum support, the

---

Table 3: The Performance of AROA system on Enslaved Benchmark

MinSupport	MinConfidence	Relaxed Precision	Relaxed Recall	Relaxed F-measure
0.01	0.5	0.63	**0.39**	0.49
0.01	0.6	0.64	0.39	0.49
0.01	0.7	0.65	0.38	0.49
0.01	0.8	0.71	0.38	0.49
0.01	0.9	0.72	0.38	0.50
0.01	1.0	0.80	0.38	**0.51**
0.02	0.5	0.66	0.38	0.48
0.02	0.6	0.66	0.37	0.48
0.02	0.7	0.67	0.37	0.48
0.02	0.8	0.72	0.36	0.48
0.02	0.9	0.72	0.36	0.48
0.02	1.0	0.82	0.36	0.50
0.03	0.5	**0.94**	0.27	0.42
0.03	0.6	0.94	0.27	0.42
0.03	0.7	0.94	0.27	0.42
0.03	0.8	0.94	0.27	0.42
0.03	0.9	0.94	0.27	0.42
0.03	1.0	0.94	0.27	0.43
0.04	0.5	0.94	0.26	0.41
0.04	0.6	0.94	0.26	0.41
0.04	0.7	0.94	0.26	0.41
0.04	0.8	0.94	0.26	0.41
0.04	0.9	0.94	0.26	0.41
0.04	1.0	0.94	0.26	0.42
0.05	0.5	0.94	0.26	0.41
0.05	0.6	0.94	0.26	0.41
0.05	0.7	0.94	0.26	0.41
0.05	0.8	0.94	0.26	0.41
0.05	0.9	0.94	0.26	0.41
0.05	1.0	0.94	0.26	0.41

best F-measure is usually achieved when the value of minimum confidence is 1.0. Figure 4 demonstrates the trend of the performance when the minimum confidence is set to 1.0. We can find that the variation of performance tends to be flat and steady after the minimum confidence with a value of 0.03. The reason is that the number of alignment rules generated is getting smaller, which it is reasonable to explain the higher precision, but with the lower recall. The results of more alignment systems will be available in the coming OAEI 2020. In this paper, we have not intended to focus on the improvement of the alignment algorithm. Instead, we would only like to prove that the Enslaved benchmark is within the capability of the current complex ontology alignment systems in OAEI. And based on the results, it also indicates that there is still much space for the improvement of the current alignment systems to detect more complex correspondences and solve the challenge of the knowledge graph and ontology integration problem. Thereby, the Enslaved benchmark can be considered as a useful potential resource to advance the development of the research in the complex ontology alignment field.

5 CONCLUSION

Complex alignment has been discussed for a long time, but relatively little work has been done to advance the state of the art of complex ontology alignment. The lack of applicable complex alignment benchmarks may be a primary reason for the slow speed of development. In addition, most current alignment benchmarks have been created by humans for the sole purpose of evaluating alignment systems, and they may not always represent real-world cases. In this paper, we have proposed a complex alignment benchmark based on the real-world Enslaved project. The two knowledge graphs and the reference alignment were designed and created by ontologists and historians to support data representation, sharing, integration, and discovery. Additionally, we take advantage
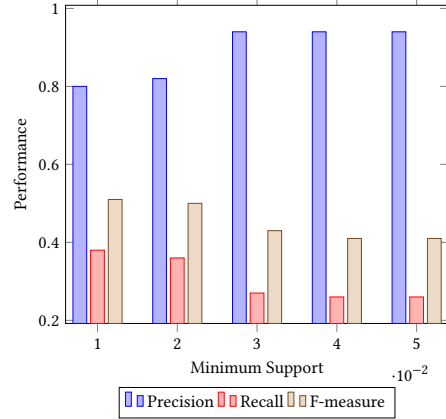

Figure 4: The Trend of the Performance When Minimum Confidence is 1.0

of Wikibase as a tool to represent the data, which is convenient for users with any level of expertise to use. Detecting alignments between ontologies and Wikibase knowledge graphs are helpful to solve many practical problems and enrich knowledge graphs by aligning common resources in Wikidata. In our benchmark, the alignments not only cover 1:1 simple correspondences but also contain m:n complex relations. All correspondences required to convert between the two ontologies (a key goal of ontology alignment) are guaranteed to be present. In addition, the alignment has been evaluated by domain experts from different organizations, and we also test the complex alignment systems on the benchmark to ensure high quality. Moreover, the alignments in both rule and EDOAL syntax have been published in FigShare and OAEI with an open-access license for reusability.

As future work in this area, we have put forth this benchmark into the complex track within the OAEI. We intent to remain actively involved for years to come in the OAEI complex alignment benchmarking track and to also develop corresponding alignment methods. We thus have an intrinsic interest in keeping the benchmark maintained and usable, which would, e.g., mean that we are prepared to transfer it to a new benchmarking framework if required in the future. At the same time, based on participants' feedback, we will modify the reference alignment if necessary to perfect the benchmark by making it more convenient to use. This may involve, for example, making the alignment available in additional formats. Furthermore, we also plan to make use of Wikidata to generate more benchmarks for Multilingual ontology matching, instance matching, and knowledge graph matching tasks. We plan to generate and improve an automated alignment system to tackle the alignment problem set forth by this benchmark.

6 ACKNOWLEDGMENTS

REFERENCES

[1] Alsayed Algergawy, Michelle Cheatham, Daniel Faria, Alfio Ferrara, Irini Fundulaki, Ian Harrow, Sven Hertling, Ernesto Jiménez-Ruiz, Naouel Karam, Abderrahmane Khiat, Patrick Lambrix, Huanyu Li, Stefano Montanelli, Heiko Paulheim, Catia Pesquita, Tzanina Saveta, Daniela Schmidt, Pavel Shvaiko, Andrea Splendiani, Élodie Thiéblin, Cássia Trojahn, Jana Vatascinová, Ondrej Zamazal, and Lu Zhou. 2018. Results of the Ontology Alignment Evaluation Initiative 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 (CEUR Workshop Proceedings)*, Pavel Shvaiko, Jérôme Euzenat, Ernesto Jiménez-Ruiz, Michelle Cheatham, and Oktie Hassanzadeh (Eds.), Vol. 2288. CEUR-WS.org, 76–116. http://ceur-ws.org/Vol-2288/oaei18_paper0.pdf

[2] Alsayed Algergawy, Daniel Faria, Alfio Ferrara, Irini Fundulaki, Ian Harrow, Sven Hertling, Ernesto Jiménez-Ruiz, Naouel Karam, Abderrahmane Khiat, Patrick Lambrix, Huanyu Li, Stefano Montanelli, Heiko Paulheim, Catia Pesquita, Tzanina Saveta, Pavel Shvaiko, Andrea Splendiani, Élodie Thiéblin, Cássia Trojahn, Jana Vatascinová, Ondrej Zamazal, and Lu Zhou. 2019. Results of the Ontology Alignment Evaluation Initiative 2019. In *Proceedings of the 14th International Workshop on Ontology Matching co-located with the 18th International Semantic Web Conference (ISWC 2019), Auckland, New Zealand, October 26, 2019 (CEUR Workshop Proceedings)*, Pavel Shvaiko, Jérôme Euzenat, Ernesto Jiménez-Ruiz, Oktie Hassanzadeh, and Cássia Trojahn (Eds.), Vol. 2536. CEUR-WS.org, 46–85. http://ceur-ws.org/Vol-2536/oaei19_paper0.pdf

[3] Eva Blomqvist, Karl Hammar, and Valentina Presutti. 2016. Engineering Ontologies with Patterns – The eXtreme Design Methodology. In *Ontology Engineering with Ontology Design Patterns – Foundations and Applications*, Pascal Hitzler, Aldo Gangemi, Krzysztof Janowicz, Adila Krisnadhi, and Valentina Presutti (Eds.). Studies on the Semantic Web, Vol. 25. IOS Press, 23–50.

[4] Eva Blomqvist and Kurt Sandkuhl. 2005. Patterns in Ontology Engineering: Classification of Ontology Patterns. In *ICEIS 2005, Proceedings of the Seventh International Conference on Enterprise Information Systems, Miami, USA, May 25-28, 2005*, Chin-Sheng Chen, Joaquim Filipe, Isabel Seruca, and José Cordeiro (Eds.). 413–416.

[5] Jérôme David, Jérôme Euzenat, François Scharffe, and Cássia Trojahn dos Santos. 2011. The Alignment API 4.0. *Semantic Web* 2, 1 (2011), 3–10. https://doi.org/10.3233/SW-2011-0028

[6] Marc Ehrig and Jérôme Euzenat. 2005. Relaxed Precision and Recall for Ontology Matching. In *Integrating Ontologies '05, Proceedings of the K-CAP 2005 Workshop on Integrating Ontologies, Banff, Canada, October 2, 2005 (CEUR Workshop Proceedings)*, Benjamin Ashpole, Marc Ehrig, Jérôme Euzenat, and Heiner Stuckenschmidt (Eds.), Vol. 156. CEUR-WS.org. http://ceur-ws.org/Vol-156/paper5.pdf

[7] Aldo Gangemi. 2005. Ontology Design Patterns for Semantic Web Content. In *The Semantic Web – ISWC 2005, 4th International Semantic Web Conference, ISWC 2005, Galway, Ireland, November 6-10, 2005, Proceedings (Lecture Notes in Computer Science)*, Yolanda Gil, Enrico Motta, V. Richard Benjamins, and Mark A. Musen (Eds.), Vol. 3729. Springer, 262–276.

[8] Johanna Geiß, Andreas Spitz, Jannik Strötgen, and Michael Gertz. 2015. The Wikipedia location network: overcoming borders and oceans. In *Proceedings of the 9th Workshop on Geographic Information Retrieval, GIR 2015, Paris, France, November 26-27, 2015*, Ross S. Purves and Christopher B. Jones (Eds.). ACM, 2:1–2:3. https://doi.org/10.1145/2837689.2837694

[9] Birte Glimm, Ian Horrocks, Boris Motik, Giorgos Stoilos, and Zhe Wang. 2014. HermiT: An OWL 2 Reasoner. *J. Autom. Reasoning* 53, 3 (2014), 245–269. https://doi.org/10.1007/s10817-014-9305-1

[10] Jiawei Han, Jian Pei, Yiwen Yin, and Runying Mao. 2004. Mining Frequent Patterns without Candidate Generation: A Frequent-Pattern Tree Approach. *Data Min. Knowl. Discov.* 8, 1 (2004), 53–87. https://doi.org/10.1023/B:DAMI.0000005258.31418.83

[11] Walter Hawthorne. 2020. Maranhão Inventories Slave Database. *Journal of Slavery and Data Preservation* 1, 1 (2020). https://jsdp.enslaved.org/fullDataArticle/40-59-4

[12] Sven Hertling and Heiko Paulheim. 2020. The Knowledge Graph Track at OAEI - Gold Standards, Baselines, and the Golden Hammer Bias. In *The Semantic Web - 17th International Conference, ESWC 2020, Heraklion, Crete, Greece, May 31-June 4, 2020, Proceedings (Lecture Notes in Computer Science)*, Andreas Harth, Sabrina Kirrane, Axel-Cyrille Ngonga Ngomo, Heiko Paulheim, Anisa Rula, Anna Lisa Gentile, Peter Haase, and Michael Cochez (Eds.), Vol. 12123. Springer, 343–359. https://doi.org/10.1007/978-3-030-49461-2_20

[13] Pascal Hitzler, Aldo Gangemi, Krzysztof Janowicz, Adila Krisnadhi, and Valentina Presutti (Eds.). 2016. *Ontology Engineering with Ontology Design Patterns – Foundations and Applications*. Studies on the Semantic Web, Vol. 25. IOS Press.

[14] Pascal Hitzler and Adila Krisnadhi. 2016. On the Roles of Logical Axiomatizations for Ontologies. In *Ontology Engineering with Ontology Design Patterns – Foundations and Applications*, Pascal Hitzler, Aldo Gangemi, Krzysztof Janowicz, Adila Krisnadhi, and Valentina Presutti (Eds.). Studies on the Semantic Web, Vol. 25.

IOS Press, 73–80.

[15] Adila Krisnadhi and Pascal Hitzler. 2016. Modeling With Ontology Design Patterns: Chess Games As a Worked Example. In *Ontology Engineering with Ontology Design Patterns – Foundations and Applications*, Pascal Hitzler, Aldo Gangemi, Krzysztof Janowicz, Adila Krisnadhi, and Valentina Presutti (Eds.). Studies on the Semantic Web, Vol. 25. IOS Press, 3–21.

[16] Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N. Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick van Kleef, Sören Auer, and Christian Bizer. 2015. DBpedia - A large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web* 6, 2 (2015), 167–195. https://doi.org/10.3233/SW-140134

[17] Gregory Piatetsky-Shapiro. 1991. Discovery, Analysis, and Presentation of Strong Rules. In *Knowledge Discovery in Databases*, Gregory Piatetsky-Shapiro and William J. Frawley (Eds.). AAAI/MIT Press, 229–248.

[18] Dominique Ritze, Christian Meilicke, Ondrej Sváb-Zamazal, and Heiner Stuckenschmidt. 2009. A Pattern-based Ontology Matching Approach for Detecting Complex Correspondences. 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 (CEUR Workshop Proceedings)*, Pavel Shvaiko, Jérôme Euzenat, Fausto Giunchiglia, Heiner Stuckenschmidt, Natalya Fridman Noy, and Arnon Rosenthal (Eds.), Vol. 551. CEUR-WS.org. http://ceur-ws.org/Vol-551/om2009_Tpaper3.pdf

[19] Cogan Shimizu, Pascal Hitzler, Quinn Hirt, Dean Rehberger, Seila Gonzalez Estrecha, Catherine Foley, Alicia M Sheill, Walter Hawthorne, Jeff Mixter, Ethan Watrall, et al. 2020. The enslaved ontology: Peoples of the historic slave trade. *Journal of Web Semantics* (2020), 100567.

[20] Cogan Shimizu, Pascal Hitzler, Quinn Hirt, Alicia Sheill, Seila Gonzalez, Catherine Foley, Dean Rehberger, Ethan Watrall, Walter Hawthorne, Duncan Tarr, Ryan Carty, and Jeff Mixter. 2019. *The Enslaved Ontology 1.0: Peoples of the Historic Slave Trade*. Technical Report. Enslaved: Peoples of the Historic Slave Trade. Available from http://docs.enslaved.org.

[21] Pavel Shvaiko and Jérôme Euzenat. 2013. Ontology Matching: State of the Art and Future Challenges. *IEEE Trans. Knowl. Data Eng.* 25, 1 (2013), 158–176. https://doi.org/10.1109/TKDE.2011.253

[22] Élodie Thiéblin, Michelle Cheatham, Cássia Trojahn dos Santos, Ondrej Zamazal, and Lu Zhou. 2018. The First Version of the OAEI Complex Alignment Benchmark. In *Proceedings of the ISWC 2018 Posters & Demonstrations, Industry and Blue Sky Ideas Tracks co-located with 17th International Semantic Web Conference (ISWC 2018), Monterey, USA, October 8th - to - 12th, 2018 (CEUR Workshop Proceedings)*, Marieke van Erp, Medha Atre, Vanessa López, Kavitha Srinivas, and Carolina Fortuna (Eds.), Vol. 2180. CEUR-WS.org. http://ceur-ws.org/Vol-2180/paper-67.pdf

[23] Élodie Thiéblin, Ollivier Haemmerlé, Nathalie Hernandez, and Cássia Trojahn dos Santos. 2017. Towards a complex alignment evaluation dataset. 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, 2017 (CEUR Workshop Proceedings)*, Pavel Shvaiko, Jérôme Euzenat, Ernesto Jiménez-Ruiz, Michelle Cheatham, and Oktie Hassanzadeh (Eds.), Vol. 2032. CEUR-WS.org, 217–218. http://ceur-ws.org/Vol-2032/om2017_poster6.pdf

[24] Denny Vrandecic and Markus Krötzsch. 2014. Wikidata: a free collaborative knowledgebase. *Commun. ACM* 57, 10 (2014), 78–85. https://doi.org/10.1145/2629489

[25] Lu Zhou, Michelle Cheatham, and Pascal Hitzler. 2019. AROA Results for 2019 OAEI. In *Proceedings of the 14th International Workshop on Ontology Matching co-located with the 18th International Semantic Web Conference (ISWC 2019), Auckland, New Zealand, October 26, 2019 (CEUR Workshop Proceedings)*, Pavel Shvaiko, Jérôme Euzenat, Ernesto Jiménez-Ruiz, Oktie Hassanzadeh, and Cássia Trojahn (Eds.), Vol. 2536. CEUR-WS.org, 107–113. http://ceur-ws.org/Vol-2536/oaei19_paper4.pdf

[26] Lu Zhou, Michelle Cheatham, and Pascal Hitzler. 2019. Towards Association Rule-Based Complex Ontology Alignment. In *Semantic Technology - 9th Joint International Conference, JIST 2019, Hangzhou, China, November 25-27, 2019, Proceedings (Lecture Notes in Computer Science)*, Xin Wang, Francesca Alessandra Lisi, Guohui Xiao, and Elena Botoeva (Eds.), Vol. 12032. Springer, 287–303. https://doi.org/10.1007/978-3-030-41407-8_19

[27] Lu Zhou, Michelle Cheatham, Adila Krisnadhi, and Pascal Hitzler. 2018. A Complex Alignment Benchmark: GeoLink Dataset. In *The Semantic Web - ISWC 2018 - 17th International Semantic Web Conference, Monterey, CA, USA, October 8-12, 2018, Proceedings, Part II (Lecture Notes in Computer Science)*, Denny Vrandecic, Kalina Bontcheva, Mari Carmen Suárez-Figueroa, Valentina Presutti, Irene Celino, Marta Sabou, Lucie-Aimée Kaffee, and Elena Simperl (Eds.), Vol. 11137. Springer, 273–288. https://doi.org/10.1007/978-3-030-00668-6_17

[28] Lu Zhou, Michelle Cheatham, Adila Krisnadhi, and Pascal Hitzler. 2019. GeoLink Data Set: A Complex Alignment Benchmark from Real-world Ontology. *Data Intelligence* (2019), 1–26.

Alignment of Surface Water Ontologies
A comparison of manual and automated approaches

**Michelle Cheatham · Dalia Varanka ·
Fatima Arauz · Lu Zhou**

Abstract Studying the surface water systems of the earth is important for many fields, from biology to agriculture to tourism. Much of the data relevant to surface water systems is stored in isolated repositories that interface with different ontologies, such as the US Geological Survey's Surface Water Ontology or the Environment Ontology (ENVO). Effectively using this data requires integrating the ontologies so that the data can be seamlessly queried and analyzed. Automated alignment algorithms exist to facilitate this data integration challenge. In this paper we examine the utility of two leading automated alignment systems to integrate four pairs of ontologies from the surface water domain. We show that the performance of such systems in this domain lags behind their results on popular benchmarks, and therefore incorporate the alignment task described here into the set of benchmarks used by the alignment community. We also show that, with minor modifications, existing alignment algorithms can be used effectively within a semi-automated system for the surface water domain. In addition, we analyze the unique challenges of this domain with respect to data integration and discuss possible solutions to pursue in order to address these challenges.

M. Cheatham
Wright State University 3640 Colonel Glenn Hwy. Dayton, OH 45435 USA
Tel.: +01-937-343-1429
Fax: +01-937-775-5009
E-mail: michelle.cheatham@wright.edu

D. Varanka
U.S. Geological Survey 1400 Independence Road Rolla, MO 65401 USA

F. Arauz
Wright State University 3640 Colonel Glenn Hwy. Dayton, OH 45435 USA

L. Zhou
Wright State University 3640 Colonel Glenn Hwy. Dayton, OH 45435 USA

1 Introduction

Much of the earth's surface is covered by water, and so various research organizations around the globe have data related to surface water features stored in data repositories. Of course, the earth's water is an inherently interconnected system, and more powerful analyses of this data could be conducted if these individual data repositories could be queried or otherwise accessed in a uniform manner. Two things standing in the way of integrating these data silos are syntactic differences between different datasets and semantic heterogeneity. The syntactic differences are currently being addressed through the application of semantic web protocols, such as using HTTP uniform resource identifiers (URIs) to represent entities, encoding information about those entities in the Resource Description Framework (RDF), and providing links between related entities [5]. These technologies are moving towards widespread acceptance, as evidenced by the growth of the Linked Data Cloud [1]. However, semantic heterogeneity remains a challenge.

Data within repositories is organized according to some vocabulary, or schema. In the case of the Semantic Web, these schemas generally take the form of an ontology. There are many ontologies related to the surface water domain [11, 28]. Four of these: the US Geological Survey's Surface Water Ontology, the Hydro3 module from the University of Maine's HydroGazetteer, the Cree surface water ontology, and the Spanish National Geographic Institute's hydrOntology, are discussed in detail in Section 3. Other ontologies contain some entities that are related to surface water features but are overall more general in scope, such as ENVO [6] and SWEET [26].

Engineering ontologies is not a deterministic process – many design decisions must be made, and the designers' backgrounds and the application they are targeting will influence their decisions in different ways. The end result is that even two ontologies that represent the same domain will not be identical. They may use synonyms for the same concept or the same word for different concepts, they may be at different levels of abstraction, they may not include all of the same concepts, and they may not even be in the same language. As a specific example, the United States Geological Survey (USGS) considers surface water features from the perspective of the Earth's terrain and the water bodies and flows between them that the geography induces. On the other hand, the conceptualization of surface water features by the indigenous Cree-speaking people of Northern Canada is based on their utility for transportation via canoe and is therefore largely focused on water bodies' locations relative to one another. These different viewpoints mean that these two ontologies have many low-level classes in common (e.g. River, Pond, Swamp), but the class hierarchies look very different because water bodies are considered "similiar" for different reasons.

Semantic heterogeneity can sometimes be resolved by *aligning* the different ontologies. The goal of ontology alignment is to determine when an entity in one ontology is semantically related to an entity in another ontology. Ontology alignment is an important part of realizing the potential of the Semantic Web.

Alignments between two ontologies can be used to browse a combined data set according to either ontology's vocabulary, to federate search queries, to perform logical reasoning across multiple domains, and other important tasks. While some of these applications require high-quality alignments that must be created manually, which often takes weeks even for small ontologies, some uses can benefit from automated alignment that sacrifice some accuracy in favor of timely results. Examples include identifying other data repositories that are related to an existing one and finding linking points for modular ontology development [17].

The overall goal of this paper is to assess the utility of automated alignment systems on real-world ontology alignment tasks from the surface water domain. The paper makes the following contributions:

- A revised version of the USGS Surface Water Ontology (SWO) is presented.
- Manual alignments between three existing surface water ontologies and the new version of the SWO have been created. These alignments constitute a new benchmark within the annual Ontology Alignment Evaluation Initiative as of 2018.
- The performance of two state of the art ontology alignment systems on this benchmark is examined in detail, with a focus on how aspects relevant to the surface water domain pose unique challenges.
- A modified version of an existing alignment system that performs significantly better than the original in this domain is presented.
- Potential avenues to address the alignment challenges raised by surface water ontologies are discussed.

2 Background and Related Work

An ontology is a way to model the semantics of a domain of study. An ontology is typically expressed in a formal language, such as the Web Ontology Language (OWL). It contains classes to represent types of things in the domain of interest, individuals that are specific things, and properties, which are relationships that hold between two things, or between a thing and a value. For example, in the ontology on the left in Figure 1, the items in the yellow squares, including PointOfInterest, Waterbody, and Gulf, are classes. The arrow labeled flowsInto represents an object property (i.e. a relation that holds between two individuals that both belong to a class, in this class River is the domain of the relation and Gulf is the range) and the arrows labeled hasName and hasLengthInKm represent data properties, which hold between an individual and a literal value.

The information shown in Figure 1 comprises the schema, or T-box, of the ontologies. In addition, an ontology often contains instance data. For example, the following statements indicate that there are instances called Mississippi_River and Gulf_Of_Mexico that are of type River and Gulf, respectively, that the Mississippi_River flowsInto the Gulf_of_Mexico, and that it hasLengthInKm 3730.
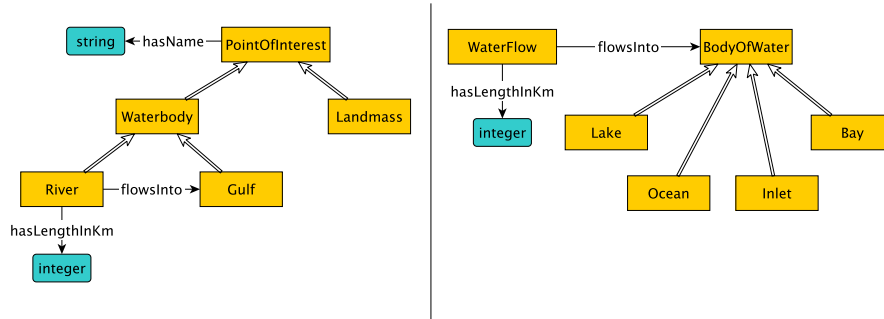
Fig. 1 Two sample ontologies. Differences in scope and granularity create alignment challenges.

```
ont1:Mississippi_River rdf:type ont1:River
ont1:Gulf_Of_Mexico rdf:type ont1:Gulf
ont1:Mississippi_River ont1:flowsInto ont1:Gulf_Of_Mexico
ont1:Mississippi_River ont1:hasLengthInKm 3730
```

As mentioned previously, the goal of ontology alignment is to determine when an entity in one ontology is semantically related to an entity in another ontology. Continuing the example from Figure 1, the ontology on the left has a class called Gulf that does not exist in the one on the right. However, a Gulf is defined by Wikipedia and other general knowledge sources as a large inlet or bay. An alignment could indicate this relationship by stating that Gulf is a subclass of the union of Inlet and Bay. A comprehensive discussion on ontology alignment is outside the scope of this paper, but a more detailed explanation can be found in [14].

An automated alignment algorithm takes as input two ontologies and produces a set of matches consisting of a URI specifying one entity from each ontology, a relationship, and an optional confidence value that is generally in the range of 0 to 1, inclusive. In order to produce this output, automated alignment systems generally employ one or more similarity metrics that determine the type and strength of relationship between two or more entities. These similarity metrics typically fall into one of three groups: syntactic, semantic, and structural. Syntactic metrics compare entities from each of the ontologies to be aligned based on strings associated with the entities. The strings are generally the entity label, but can also include comments or other annotations of the entity. Referring to the alignment problem in Figure 1, a syntactic metric would likely align the classes Waterbody and BodyOfWater, in addition to the properties, which have syntactically identical labels. Semantic similarity metrics attempt to use the meanings of entity labels rather than their spellings. External resources such as thesauri, dictionaries, encyclopedias, and web search engines are often used to calculate semantic similarity.

A semantic metric might be able to detect that a Gulf is related to an Inlet or Bay by looking up the term Gulf in an electronic dictionary. Structural techniques consider the neighborhoods of two entities when determining their similarity. For instance, two entities with the same superclass that share some common instances are considered more similar than entities that do not have these things in common. Graph matching techniques are often used for this. In our example, River might be scored as fairly similar to Waterflow because it is the domain of two properties that have already been matched via a syntactic metric. An alignment system may use zero or more of each type of similarity metric. The values from multiple approaches may be combined to form a single measure of similarity, or they may be used in a serial fashion to filter potential matches down to the most likely candidates. At some point, a final list of related entities is generated, frequently by including any matches with a confidence (similarity) value higher than some threshold. Additionally, alignment systems may use some form of inconsistency checking and repair after the matching process in order to ensure a merged ontology produced using the alignment is logically consistent. More detail about ontology matching systems can be found in Euzenat and Shvaiko's book on the subject [14].

Ontology alignment is a well established field. There are dozens of automated alignment systems (see [14] and [24] for surveys), and an annual ontology alignment evaluation initiative (OAEI) for these systems to compare their performance on benchmark alignment tasks.[1] Ideally, alignment systems should be able to uncover any entity relationships across two ontologies that can exist within a single ontology. Such relationships have a wide range of complexity, as shown in Figure 2. Nearly all existing alignment systems fall at the simplest end of the scale. A few systems, including ASMOV [19], RiMoM [22], BLOOMS [18] and PARIS [30], attempt to identify subsumption relationships across ontologies. CSR [29] and TaxoMap [16] attempt to find 1-to-many equivalence and subsumption relationships. In general though, most research activity in the field of ontology alignment remains focused on finding 1-to-1 equivalence relations. This limitation was mentioned in 2013 [27] and again in 2017 [9] as a challenge for the field. One reason for the lack of systems that attempt to find more complex matches may be that current benchmarks have not historically contained any complex relations. This is changing, however – the surface water alignment task described in this paper has been accepted as part of a new OAEI complex alignment track as of 2018.[2]

In this work we analyze the performance of two of the best performing automated alignment systems from the OAEI on the task of aligning surface water ontologies. While many ontologies exist to model surface water features (these are surveyed in Section 3), this is to the best of our knowledge the first time that the performance of automated alignment systems has been evaluated on this domain.

---

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

[2] http://oaei.ontologymatching.org/2018/complex/index.html

1:1 Equivalence

1:1 Disjointness

1:1 Subsumption

1:m Equivalence

1:m Subsumption

n:m Equivalence

n:m Subsumption

n:m Arbitrary Relationship

Greater Complexity

Greater Automated Support
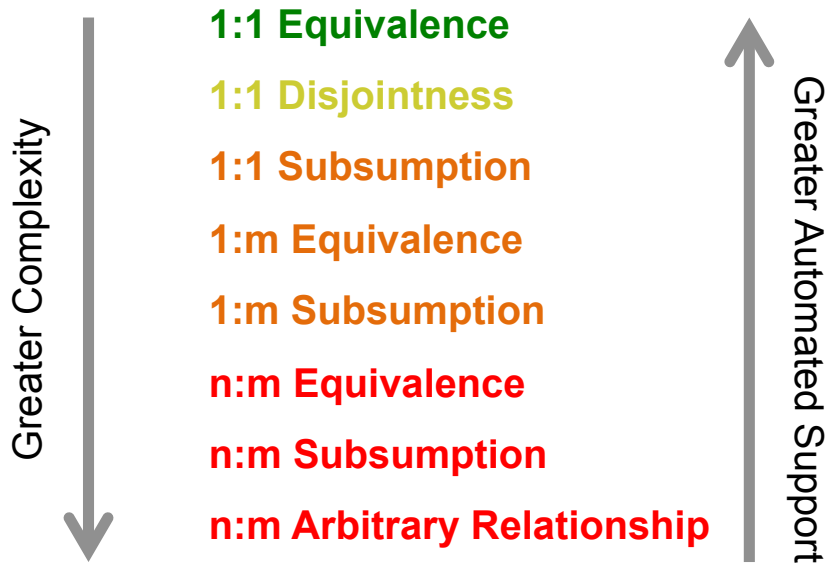
Fig. 2 Spectrum of alignment complexity. The x:y notation refers to the number of entities from the first ontology (x) and the number from the second ontology (y) that are involved in a particular mapping across the ontologies.

If we look more broadly at aligning geographical ontologies in general, we find more related work. Much of this involves using semi-automated approaches to create alignments at the bottom of the complexity spectrum shown in Figure 2. For example, the Geodint project aligned several ontologies related to geographic points of interest, including Facebook Places, Foursquare, and DBpedia using COMA++, a visually oriented semi-automatic alignment system [23]. Sunna and Cruz focus on structure-based similarity metrics to align ontologies related to wetlands [31]. A paper describing the G-MAP alignment system mentions the ability to identify complex relations, but it defines complex as relations between properties rather than classes [4]. These relations are still 1-to-1 and would therefore not be considered complex as defined in this paper. As we will see in Section 4, most of the relations of interest between the surface water ontologies used in this study are 1-to-many in nature.

Another common theme in research related to alignment of geospatial ontologies is extensional matching [10,12,13,7]. Extensional matchers begin by trying to determine when two instances represent the same spatial feature. For example, they may try to determine that Mississippi_River in one ontology is equivalent to Greater_Mississippi_Rvr in another ontology, often based on the coordinates associated with each entity. They can then use these instance-level matches to find schema-level relations, for example by using inductive inferencing. The work presented here differs in that it does not assume instance level data exists in both ontologies being aligned.

3 Surface Water Ontologies

As mentioned previously, there are many existing ontologies relevant to the surface water domain. The USGS SWO was chosen as a focal point for this work because it is the domain ontology with which the authors have the most familiarity. The other three ontologies discussed in this section were chosen to create a spectrum of difficulty level regarding the alignment task: Hydro3 is similar to the SWO in terms of both organization and language, the hydrOntology has a similar organization but is in a different language, and the Cree ontology differs greatly from the SWO in terms of both organization and language. In this work each of these ontologies (Hydro3, hydrOntology and Cree) will be aligned to the SWO. These nature of these ontology pairs allows us to evaluate the performance of automated alignment systems on a range of real-world hydrographic ontology alignment tasks.

3.1 USGS Surface Water Ontology

In 2001, as part of its National Map project, the US Geological Survey (USGS) began development of the National Hydrography Dataset. The dataset consists of surveys conducted both in the field and from aerial photographs of surface water features across the United States and is maintained via edits and additions submitted by the individual states. The NHD was originally stored in a relational database, but in 2014 the data was also made available as an RDF triplestore. As part of this process, the USGS developed the SWO, which was originally presented in [33]. The SWO was initially designed to closely follow the underlying relational database. Our goal with this revision was to make it more broadly applicable to other hydrographic datasets.
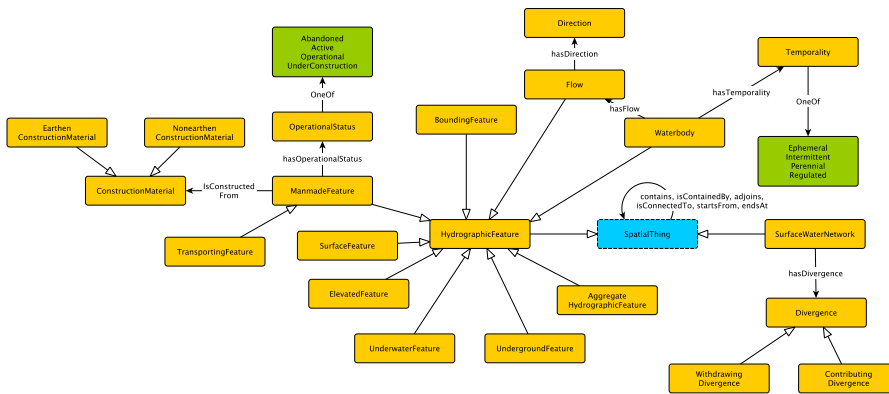

Fig. 3 The upper levels of the SWO class and property hierarchy

The changes made to the ontology in this revision fall into three general categories. First, the modeling of hydrographic measurements was fleshed out. The original ontology represents the taking of a hydrographic measurement via a class called Gaging, which is a subclass of Event. There are also classes to represent what was measured (e.g. SpatialQuality and its subclasses such as Area and Length) and the value and unit of the measurement (e.g. SpatialMeasurement). This version retains that basic model, but adds new object properties such as isMeasurementOfFeature, measuresSpatialQuality, producedMeasurement, and takenAtStage to fully relate a measurement to the SpatialQuality and HydrographicFeature being measured and to capture the corresponding provenance information, including the WaterStage of the HydrographicFeature during the measurement (represented as a controlled vocabulary using the OWLOneOf construct, in order to force consensus on this aspect of the measurement, which is key to understanding the context of the data). The second group of changes involved the creation of an abstract layer in the ontology (shown in Figure 3), which contains the upper levels of the class and property hierarchy. This layer is important both because it enables the SWO to apply in many more applications involving surface water features and because the more concrete features in the ontology are often defined in terms of this layer. Finally, specific hydrographic features, such as seas, rivers, dams, and shorelines are now defined using axioms that relate a concept to others within the ontology, often from the abstract layer. These axioms range from relatively simple, such as that a shore is something that bounds a body of water

```
SubClassOf(swo:Shore swo:BoundingFeature)
```

or that a sea or ocean is a perennial waterbody

```
SubClassOf(swo:SeaOrOcean swo:Waterbody
ObjectHasValue(swo:hasTemporality swo:Perennial))
```

to more complex, such as that an estuary must adjoin both a sea or ocean and a shore.

```
SubClassOf(swo:Estuary OWLIntersectionOf(swo:Waterbody
ObjectSomeValuesFrom(swo:adjoins swo:SeaOrOcean)
ObjectSomeValuesFrom(swo:adjoins swo:Shore))
```

3.2 Hydro3

An ontology called HydroGazetteer was developed by individuals at the University of Maine in order to support expanded gazetteer functions using topology and semantic inference [34]. The Hydro3 module of this ontology, shown in Figure 4, overlaps significantly with the SWO.
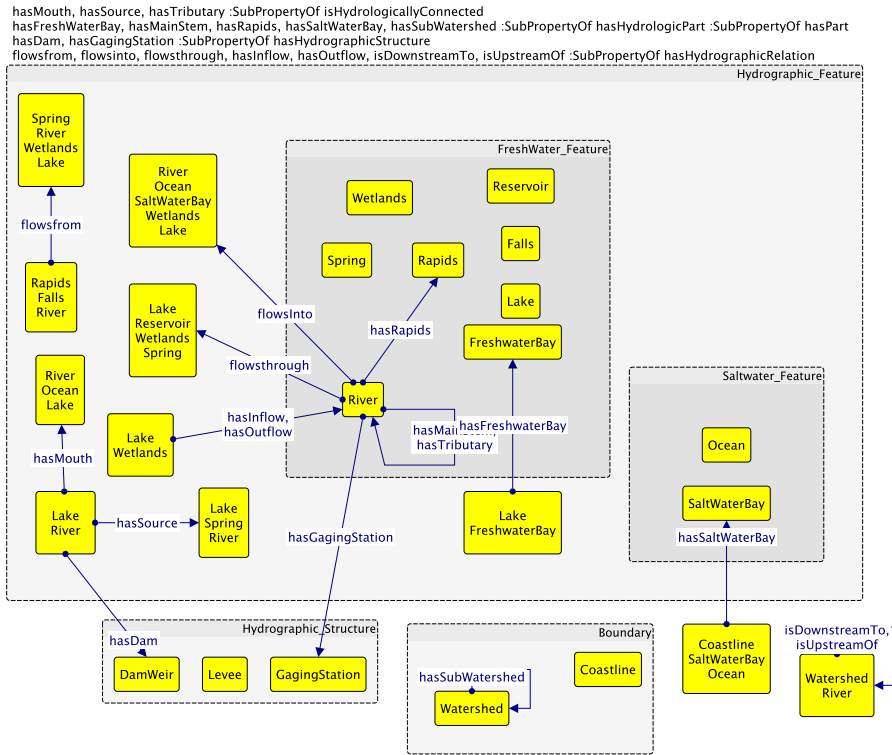
Fig. 4 The Hydro3 module within the HydroGazetteer ontology

3.3 Cree

The Cree surface water ontology is described in [36]. Cree is a language spoken by some of the native inhabitants of northern Canada. This region is densely covered with surface water features, and Cree speakers have a very rich vocabulary to describe them. Many Cree terms do not have a direct translation into English. The authors of [36] worked with native speakers in order to establish an ontology of Cree surface water features, along with English descriptions. The classes in this ontology are shown in Figure 5. The Cree speakers do not have a hierarchical view of different types of water bodies, so the ontology is very flat and does not contain any abstract notions.

3.4 hydrOntology

Another non-English hydrography ontology is the hydrOntology, which was developed by the Spanish National Geographic Institute (IGN) [35]. The hydrOntology was originally created to assist Spanish cartographers in coordinating their products, and has since been expanded into a complete hydro-
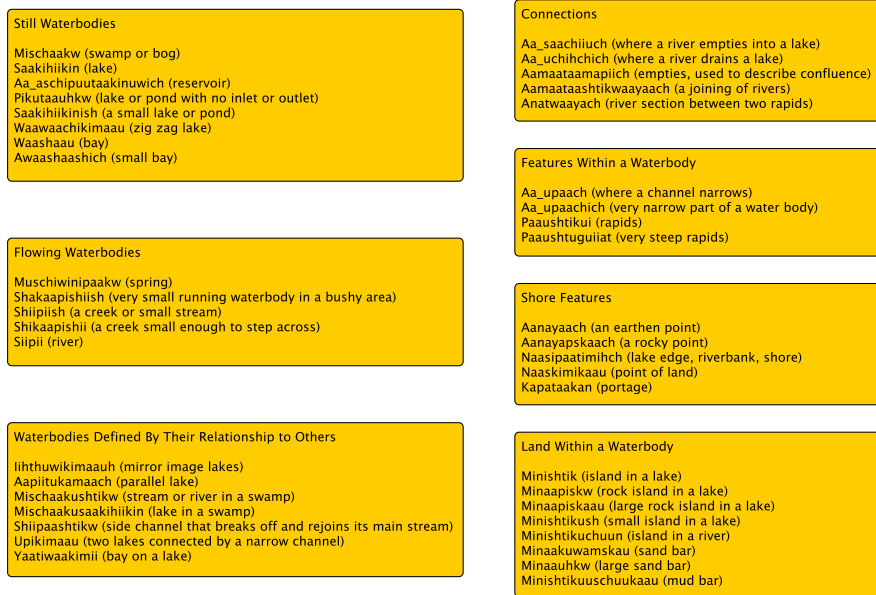
Fig. 5 The classes within the Cree ontology. Class groupings (e.g. Still Waterbodies, Connections) have been added for convenience and are not part of the ontology.

graphic domain ontology. The ontology's design was informed by numerous feature catalogs, including those of the IGN, the European Water Framework Directive, and the Alexandria Digital Library, as well as by several geographic data repositories owned by the IGN. Like the SWO, the hydrOntology can be thought of as two layers: one describing the relationships among abstract hydrographic concepts and the other containing concrete hydrographic features that are generally defined in terms of their relation to one or more of the abstract concepts. The upper layer is shown in Figure 6. The properties in the hydrOntology are largely similar to that of the SWO, but they have extensive domain and range restrictions involving classes in the concrete layer of the ontology, while the SWO has few of these.

Table 1 presents some basic characteristics of the ontologies described in this section. In comparison to other ontology alignment benchmarks, the surface water ontologies presented here have some characteristics that pose different challenges and possibilities for automated alignment systems. Existing benchmarks primarily involve ontologies related to either conference organization or the life sciences (e.g. anatomy, diseases, biodiversity, and ecology).[3] In comparison to the ontologies that make up those alignment tasks, the surface water ontologies presented here vary more in their level of granularity. For example, the SWO has a single class that represents a lake while the Cree ontology has classes to represent nine different types of lakes, and these classes do not all

---

[3] http://oaei.ontologymatching.org/2018/
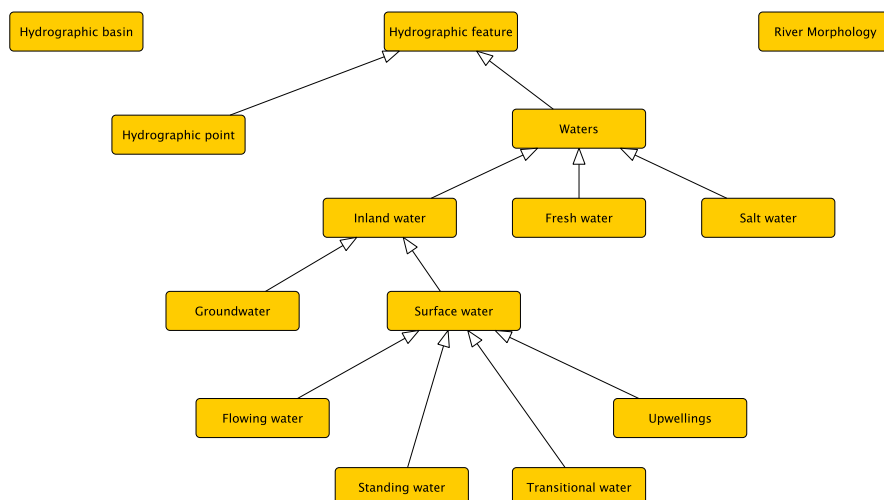
Fig. 6 The upper levels of the hydrOntology class hierarchy

Table 1 Number of entities within the chosen ontologies

	SWO	Hydro3	hydrOntology	Cree
Classes	85	22	154	83
Object Properties	20	34	47	21
Data Properties	1	0	75	7

share a common ancestor in the class hierarchy. This is likely to cause trouble for alignment systems that heavily employ structural similarity metrics or on identifying a fairly large set of anchor mappings based on lexical equivalence of entity labels. In addition, many concepts within the surface water domain are defined in terms of other concepts within the domain. For instance, an estuary can be defined as "a semi-enclosed coastal body of water which has a free connection with the open sea and within which sea water is measurably diluted with fresh water derived from land drainage". While this information is often available in natural language comments, most surface water ontologies other than the SWO lack formal axioms to express these relations, which makes it difficult for automated alignment systems to make use of them. Furthermore, spatial relationships are particularly important within the surface water domain. This is true of both classes, such as the Cree class Iihthuwikimaauh, defined as "mirror image lakes", and properties, such as parallelTo. Few alignment systems are currently able to consider these types of relationships.

4 Manual Alignments

In order to evaluate the performance of automated systems on ontology align-
ment tasks within the hydrography domain, we must first establish reference
alignments that serve as the gold standards for these tasks. Each reference
alignment was manually created by an ontologist. Semantic correctness was
verified by an earth scientist working for the US Geological Survey, and log-
ical consistency was verified using the HermiT reasoner. In the case of the
hydrOntology, labels and comments were translated from Spanish to English
by a native Spanish speaker so that the other (non-Spanish-speaking) team
members could accomplish their work.

There were some instances in which there appeared to be mistakes in the
surface water ontologies used for this study. For example, in some cases sub-
class axioms within the Cree ontology appear to be missing. For instance,
Aanayapskaach (a rocky point) is not a subclass of Naaskimikaau (a point of
land). Likewise, in Hydro3 there may be a mistake related to property domain
and range restrictions. The property hasHydrographicPart is a subproperty of
hasPart, and the domain of the hasPart property is River and the range is the
union of Rapids and Falls. On the other hand, the hasHydrographicPart rela-
tion has no restrictions on its domain or range. It seems that the domain and
range restrictions should be on hasHydrographicPart rather than on hasPart.
Our view is that in real-world use, ontology alignment systems will often be
presented with arguably imperfect ontologies. Because of this, no changes were
made to the underlying ontologies when developing the reference alignments.
The only exceptions to this were changes necessary to enable the ontologies
to open in Protégé. For example, the hydrOntology had some < characters in
comments that caused parse errors and were therefore removed. In addition,
the hydrOntology had a cardinality restriction involving the parte_de prop-
erty on the Aguas_de_Transición class, while in the Cree ontology the HasPart
property, which has an inverse called PartOf, was involved in a cardinality re-
striction. Cardinality restrictions on these properties push the ontology from
OWL DL into OWL Full and render it undecidable by a reasoner, so they were
removed.

When developing the reference alignments, we attempted to find the *sim-
plest* relation that holds between classes and properties in the source ontology
(e.g. Hydro3, hydrOntology, and Cree) and those in the target ontology (i.e.
the SWO). For example, relations that involved an atom (i.e. a single class or
property) were given preference over those that involved an expression (e.g.
union, intersection, cardinality or value restrictions, etc.) and equivalence re-
lations were given preference over subsumption and disjointness. This is the
same approach followed by the developers of the reference alignments discussed
in [32].

A typical atom-to-atom relation is:

```
<EquivalentClasses>
    <Class abbreviatedIRI="hydrOnt:Wetlands"/>
```

```
    <Class abbreviatedIRI="swo:SwampOrMarsh"/>
</EquivalentClasses>
```

An example of an expression-to-expression relation is shown below. (Note that *origen del agua* translates to *origin of the water*.)

```
<SubClassOf>
    <DataSomeValuesFrom>
        <DataProperty abbreviatedIRI="hydrOnt:origen_del_agua"/>
        <Datatype abbreviatedIRI="rdfs:Literal"/>
    </DataSomeValuesFrom>
    <ObjectSomeValuesFrom>
        <ObjectInverseOf>
            <ObjectProperty abbreviatedIRI="swo:startsFrom"/>
        </ObjectInverseOf>
        <Class abbreviatedIRI="owl:Thing"/>
    </ObjectSomeValuesFrom>
</SubClassOf>
```

There are 106 entities in the SWO. The Hydro3-to-SWO alignment involved 24 unique SWO entities, while the Cree-SWO alignment involved 42, and the hydrOntology-SWO alignment referred to 84 unique SWO entities. Table 2 shows the number of relations between classes and properties, organized by complexity type (i.e. those that involve atomic entities versus expressions). From this it is evident that the Hydro3-to-SWO alignment is the most straightforward – 24 out of 27 relations involve atomic entities from both ontologies. On the other end of the spectrum, the hydrOntology-to-SWO alignment can be considered the most complex in terms of number of expressions, because it is the only case in which a majority of the relations involve expressions rather than atoms.

Table 2 also shows that there are no equivalent property relationships in any of the reference alignments. This is because most of the properties in the source ontologies have domain and range restrictions, whereas the SWO does not place these restrictions on most of its properties. Because of this, even very related properties cannot be declared equivalent; instead, most of the source ontology properties must be represented as subproperties of things in the SWO. It is possible to represent the domain and range restrictions on the SWO properties in terms of classes from that ontology, but this would complicate the relations, and our approach is to identify the simplest correspondences between the two ontologies.

Table 3 shows how often different OWL constructs appear when a relation involves an expression. Note that a single relation can involve multiple OWL constructs. By far the most frequently appearing constructs are intersection (OWLIntesectionOf) and object value restrictions (OWLSomeValuesFrom and OWLAllValuesFrom). A typical relation using these constructs is shown below. (Aquas corrientes translates to running water.)

Table 2 Complexity of the reference alignments. Alignments are between the indicated ontology and the SWO.

		Hydro3	hydrOntology	Cree	Total
Class Equivalence	atom-atom	6	11	2	19
	atom-expr	1	2	1	4
	expr-atom	0	0	0	0
	expr-expr	0	1	0	1
Class Subsumption	atom-atom	6	32	10	48
	atom-expr	1	42	14	57
	expr-atom	1	8	0	9
	expr-expr	0	5	0	5
Class Disjointness	atom-atom	0	4	3	7
	atom-expr	0	3	0	3
	expr-atom	0	0	0	0
	expr-expr	0	0	0	0
Property Equivalence	atom-atom	0	0	0	0
	atom-expr	0	0	0	0
	expr-atom	0	0	0	0
	expr-expr	0	0	0	0
Property Subsumption	atom-atom	12	12	11	35
	atom-expr	1	4	0	5
	expr-atom	0	0	0	0
	expr-expr	0	0	0	0
Property Disjointness	atom-atom	0	0	4	4
	atom-expr	0	0	0	0
	expr-atom	0	0	0	0
	expr-expr	0	0	0	0
Total		28	123	45	196

```
<SubClassOf>
    <Class abbreviatedIRI="hydrOnt:Aguas_Corrientes"/>
    <ObjectIntersectionOf>
        <Class abbreviatedIRI="swo:SurfaceFeature"/>
        <Class abbreviatedIRI="swo:Waterbody"/>
        <ObjectSomeValuesFrom>
            <ObjectProperty abbreviatedIRI="swo:hasFlow"/>
            <Class abbreviatedIRI="swo:Flow"/>
        </ObjectSomeValuesFrom>
    </ObjectIntersectionOf>
</SubClassOf>
```

As mentioned previously, these ontology matching tasks have been incorporated into a new complex alignment track within the annual Ontology Alignment Evaluation Initiative. Links to download all of the ontologies and the reference alignments are available from the OAEI website.[4]

---

[4] http://oaei.ontologymatching.org/2018/complex/index.html#hydrography

Table 3 OWL constructs in the reference alignments. Alignments are between the indicated ontology and the SWO.

	Hydro3	hydrOntology	Cree
Union	3	6	1
Intersection	0	29	15
Complement	0	1	1
Property Inverse	0	10	0
Object Value Restriction	0	62	15
Data Value Restriction	1	9	0
Cardinality Restriction	0	3	9

5 Automated Alignments

In order to evaluate the performance of automated systems in this domain, we used two automated alignment systems, AgreementMakerLight (AML) [15] and LogMap [20], to perform the same alignment tasks described in the previous section. AML allows users to select a set of different matchers to run (or does so automatically based on a profile of the ontologies to be matched) and runs each matcher individually. The resulting alignments are combined using a greedy selection strategy and any logical inconsistencies are removed. Matchers available within AML include lexical and structural algorithms, as well as approaches that leverage background information such as from Word-Net or domain-specific lexicons. On the other hand, LogMap compares two entities based on their ISUB (i.e. string) similarity and scope (i.e. the degree of overlap of their neighborhoods). Additionally, LogMap's approach to ontology alignment heavily involves consideration of whether or not a relation would conflict with another relation that has a higher confidence value. For example, the system either filters out or more carefully scrutinizes what it calls "dangerous" and logically inconsistent relations.

AML and LogMap were chosen based on their strong performance in the OAEI over several years. In addition, we endeavored to explore the performance of the systems mentioned in Section 2 that attempt to identify subsumption relations between ontologies (i.e. the class Document subsumes the class Book). Unfortunately, BLOOMS, CSR and ASMOV could not be located and the authors could not provide us with those systems. TaxoMap and Ri-MoM have executable versions available online, but they had errors when run on the surface water ontologies that could not be fixed without the source code. PARIS requires instance data from both ontologies, which is not available for this alignment task.

5.1 1-to-1 Class Equivalence

Because AML and LogMap focus on identifying 1-to-1 class equivalences, we first analyzed their performance on just this aspect of the surface water ontology alignments (i.e. the topmost section of Table 2). The results are shown in

Table 4 Atom-to-atom class equivalence

		AML	LogMap
Hydro3	True Positives	5	4
	False Positives	1	2
	False Negatives	1	2
	Precision	0.833	0.667
	Recall	0.833	0.667
	F-measure	0.833	0.667
hydrOnt (translated)	True Positives	3	4
	False Positives	6	5
	False Negatives	8	7
	Precision	0.333	0.444
	Recall	0.273	0.364
	F-measure	0.300	0.400
hydrOnt (native)	True Positives	0	0
	False Positives	0	0
	False Negatives	11	11
	Precision	0	0
	Recall	0	0
	F-measure	0	0
Cree	True Positives	0	0
	False Positives	0	0
	False Negatives	2	2
	Precision	0	0
	Recall	0	0
	F-measure	0	0

Table 4. Precision reflects the percentage of mappings found by the system that were correct, while recall is related to the number of correct mappings that the system found. F-measure is the harmonic mean of precision and recall. AML was able to identify five of the six 1-to-1 class equivalences between the SWO and Hydro3 ontologies with one false positive, while LogMap found four with two false positives. The performance on the version of the hydrOntology that was translated into English was significantly worse, with AML and LogMap correctly identifying three and four relations out of 11, respectively. Neither system was capable of producing any results on the non-English ontologies.

Even though both AML and LogMap were designed to identify 1-to-1 class equivalences, their performance on these ontologies from the surface water domain are significantly below what they have achieved on the OAEI benchmarks. For example, in 2017 AML had an F-measure of 0.76 on ontologies from the domain of conference organization and 0.94 when matching ontologies about human and mouse anatomy. The corresponding values for LogMap were 0.73 and 0.88 [2]. A detailed analysis of the results of AML and LogMap on finding the 1-to-1 class equivalences among the surface water ontologies considered here shows these systems' reliance on lexical similarity among entity labels. For example, all of AML's correct results on the Hydro3-to-SWO alignment task involve either exact matches of entity labels (e.g. Levee to

Levee) or significant lexical similarity (e.g. Falls to Waterfall). This approach can sometimes lead to incorrect results, such as AML's treatment of Hydrographic_Feature in Hydro3 as equivalent to HydrographicFeature in the SWO, when the correct mapping is between the union of Hydrographic_Feature, Hydrographic_Structure and Boundary of Hydro3 and the SWO HydrographicFeature class. A heavy reliance on lexical metrics causes AML and LogMap to miss some fairly clear mappings, such as Wetlands to SwampOrMarsh in the Hydro3-to-SWO task, and to fail completely when the ontologies are not in the same language. This issue has been noted previously, as in [8].

The dependence of automated alignment systems on syntactic similarity between entity labels is not unique to AML and LogMap: the results from the complex alignment track of the OAEI show the same pattern. Within that track were several different data sets, including the one presented here, one based on the domain of academic conference organization, and one from the geosciences.[5] The conference and geosciences alignments both involve more syntactically similar entity labels than the surface water ontology alignments. The average normalized Levenstein distance between related source and target entity labels in the conference ontologies is .28. For the geosciences it is .24, while for the surface water ontologies the corresponding value is .16. Unsurprisingly, the participating alignment systems performed better on the conference and geosciences tasks than on the surface water case, in terms of the number of systems that could generate meaningful results. For the conference case, two alignment systems were able to identify complete complex mappings. No systems were capable of this for the geosciences and surface water tests, so instead, systems were evaluated based on their ability to determine which target entities were related to a given source entity. The average F-measure for the surface water ontologies was .10, versus .18 for the geosciences. More detail about the performance of alignment systems on the 1-to-1 class equivalence task for these ontologies can be found in [3].

5.2 Identification of related entities

As shown in Table 2, the majority of relations between these surface water ontologies are not 1-to-1 class equivalences, but rather relations in which an entity in one ontology is related in some way (equivalence, subsumption, or disjointness) to an expression involving multiple entities from the other ontology. As discussed previously, most current automated alignment systems, including AML and LogMap, cannot directly identify these types of relations. However, these systems do contain a set of similarity metrics that is used to assess the degree of relevance of one entity to another. In this section we explore the ability of these alignment systems to effectively rank target ontology entities for each entity in the source ontology.

---

[5] There was also a fourth data set from the plan taxonomy domain, but we could not include it in our analysis because the reference alignments are not public.

We evaluate the performance against the reference alignments in terms of mean reciprocal rank. This is a standard evaluation metric in situations in which results are ordered according to how well they apply to the current search or query, such as search results or auto-completion suggestions. In this case, the "query" is the given entity from the source ontology. As an example, consider the relation below, which appears in the alignment between the hydrOntology and the SWO:

```
<SubClassOf>
    <Class abbreviatedIRI="hydrOnt:Aguas_Corrientes"/>
    <ObjectIntersectionOf>
        <Class abbreviatedIRI="swo:SurfaceFeature"/>
        <Class abbreviatedIRI="swo:Waterbody"/>
        <ObjectSomeValuesFrom>
            <ObjectProperty abbreviatedIRI="swo:hasFlow"/>
            <Class abbreviatedIRI="swo:Flow"/>
        </ObjectSomeValuesFrom>
    </ObjectIntersectionOf>
</SubClassOf>
```

Assume an alignment system produced the following ordered set of SWO entities and similarity values for the hydrOntology entity Aguas_Corrientes:

```
River 0.97
Rapids 0.96
Waterbody 0.94
Flow 0.91
Waterfall 0.89
hasFlow 0.88
...
```

We calculate the reciprocal rank by summing the inverses of the ranks of each correct answer and dividing by the number of answers. An entity's rank is its place in the ordered list minus the number of entities involved in the relationship (3 in this case). Continuing with the example, Waterbody has a rank of 1 because it is among the first three entities. Flow has a rank of 2 and hasFlow has a rank of 4. The reciprocal rank for this relation is therefore $(1/1 + 1/2 + 1/4) / 3 = 0.58$. The mean reciprocal rank is then just the average of the reciprocal ranks over all relations in the reference alignment. A value of 1.0 means that the alignment system always ranks the entities involved in the relation most highly, while a value of 0.0 occurs if the system consistently ranks the related entities last in its list. This metric was chosen because it can differentiate between the two system's performance even when neither one produces the correct answer, in effect recognizing one as "closer" than the other based on how high the related target entities are in its list.

In order to use AML and LogMap in this way, a few changes needed to be made. In particular, we changed AML so that the system would display

aggregate similarity values for every entity in the target ontology when considering an entity from the source ontology. This involved commenting out code that forced an alignment to be 1-to-1.[6] When we ran the system, we did not enable the filtering and repair functionality. As for LogMap, its approach to ontology alignment more heavily involves consideration of whether or not a relation would conflict with another relation that has a higher confidence value. This approach is not conducive to a complete ranking of all possible relations, so in order to generate such a ranking while keeping the spirit of the LogMap approach, we modified LogMap so that rather than filtering out inconsistent or dangerous mappings, it allows them but assesses a penalty on their confidence value.

The results of this effort are shown in the first two data columns of Table 5. The table clearly shows the increasing difficulty level of the alignment tasks. Additionally, we see that this version of AML outperforms the modified version of LogMap on this task, which is not surprising given that much of the underlying principle of that system assumes that the goal is to generate 1-to-1 relations. As with the 1-to-1 results, upon detailed analysis of the results in this section we again see that lexical similarity explains the vast majority of the performance. Both AML and LogMap tend to rank syntactically similar target entities highly, so if these are the ones involved in the complex mappings, the mean reciprocal rank benefits. This tendency is more important than any other factor, such as the number or types of entities involved in the mapping.

One thing of note is that neither AML nor LogMap make use of comments encoded within the ontology. This may be because most of the ontologies involved in the OAEI benchmarks do not contain comments. However, except for Hydro3, all of the ontologies from the surface water domain covered here make extensive use of comments. The comments in the Cree ontology are particularly helpful given the challenges of the language. We therefore added a new matcher to AML that leverages these comments and evaluated its performance in the same way as the other systems. To do this, we modified the AML Lexicon to store comments in addition to entity labels. We then created a CommentMatcher class. The match method in this class iterates through all of the comments in the source ontology and identifies entities in the target ontology whose names are mentioned in the comment. Relationship strength is based on the number of words in common between the comment and the entity name, divided by the number of words in the comment. We have made this system publicly available on GitHub.[7] The results of this approach, shown in the last column of Table 5, show a large increase in performance when English comments are available.

---

[6] This code is in the string matcher and the neighborhood matcher within AML

[7] https://github.com/mcheatham/aml-comments

Table 5 Related entity recommendation (assessed by mean reciprocal rank)

	AML	LogMap	AML with comments
Hydro3	0.91	0.69	0.91
hydrOnt (translated)	0.50	0.36	0.79
hydrOnt (native)	0.15	0.10	0.19
Cree	0.05	0.06	0.98

6 Discussion

Results like those described in Section 5 are useful because they highlight the performance of top performing existing automated alignment systems in the surface water domain and raise new challenges that can be addressed in the future. We see that identifying complex relationships between two ontologies is a very challenging task. This is particularly true in the surface water domain, because such ontologies frequently have less syntactic and structural (due to differing levels of abstraction) similarity than ontologies in other domains that have been a focus for alignment system developers. Here we present some possible research threads to improve the performance of automated alignment systems in this domain.

The relative success of the AML with comments system in identifying related entities is an important first step that could be leveraged in a more complete complex alignment system. Its performance is good enough that it can already be of some utility in a semi-automated approach to complex ontology alignment in this domain. For example, we have developed a web application called WorldView that assists a domain expert (for example, a native Cree speaker) and an ontologist in building a complex alignment between an ontology familiar to the domain expert and an unfamiliar one from the same domain for which instance data (e.g. coordinates) are available. A screenshot is shown in Figure 7. The user clicks on a word from the familiar (source) ontology in the upper left quadrant (Area 1) and an automated alignment system such as AML with comments ranks entities from the target ontology in terms of relevance. The user can then click on these ranked entities to be shown pictures of them in the map view on the right (Area 3). The domain expert and ontologist can then work together using the axiom authoring tool in the bottom left (Area 2) to refine the relation until the domain expert is satisfied that the things highlighted in the map view match his or her definition of the surface water feature. This tool differs from the systems discussed in Section 2 in that it requires instance data only for the target rather than both ontologies. The source code is available on GitHub.[8]

While including comments in the alignment process significantly improves performance, further gains will require more advanced techniques. In the situation discussed in this paper, instance data is available for only one of the
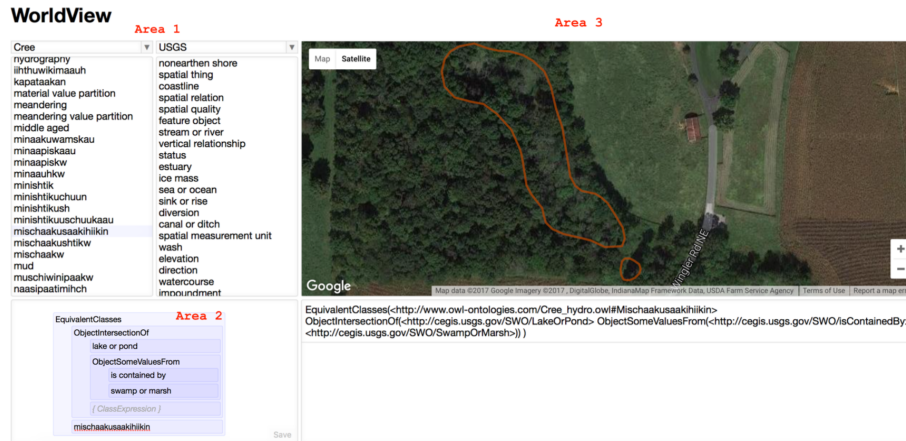
---

[8] https://github.com/mcheatham/worldview

Fig. 7 The WorldView semi-automated alignment system

ontologies (the SWO). However, in cases where instance data is available for both the source and target ontologies being matched, extensional alignment approaches could be considered. Extensional alignment systems consider the overlap between instances when aligning entities at the schema level. In our future work on this topic, we plan to create an extensional matcher that leverages logical RDF compression [21]. Logical RDF compression uses the FP-Growth data mining algorithm to generate rules that can be stored in lieu of the triples they are based on. While logical RDF compression seeks to find any rules that can be used to shrink the dataset, it is possible that some of these rules represent meaningful semantic relations that hold between entities. Because the FP-Growth algorithm underlying logical RDF compression can generate a very large number of rules, some mechanism must be put in place to choose the more semantically meaningful rules rather than the ones that result in the most compression. Our planned approach for this is to choose rules that involve the entities suggested by traditional alignment systems. Another possibility, when coordinates of surface water features is available in both the source and target ontologies, is to take advantage of the spatial nature of this domain by extending the semi-automated approach of WorldView to a fully automated system.

In order to deal with the challenges presented by the varied vocabulary used to describe surface water features and the interrelated nature of their definitions, alignment systems would likely benefit from incorporating external resources, similar to the way AML leverages upper level life sciences ontologies as a source of background knowledge when aligning ontologies from that domain [15]. Unfortunately, the surface water domain is currently somewhat lacking in these resources. Another approach might be to leverage more general purposes knowledge sources, such as Wikipedia. Working with unstructured text in this context is difficult, but relatively recent advances in word embeddings ([25]) might make such an approach feasible.

7 Conclusion

This paper explored the nature of the relationships that exist across a set of
ontologies from the surface water domain and examined the performance of
current automated alignment systems in this domain. Characteristics common
to surface water ontologies, such as lack of syntactic similarity of entity labels,
differences in modeling granularity, and the tendency for surface water features
to be defined in terms of other features pose particular challenges for current
systems. Our results show that existing alignment systems do not perform as
well in this domain as they do on standard ontology alignment benchmarks.
In addition, no current systems were able to find relations other than 1-to-
1 equivalences. The reference alignments presented here have therefore been
introduced as part of a new track within the Ontology Alignment Evaluation
Initiative, in an effort to spur researchers to improve performance on this
domain and to develop alignment systems capable of identifying the complex
relationship types present among surface water ontologies. This paper provides
background knowledge and baseline results for system developers interested in
participating in that track. In addition, a discussion of possible next steps to
improve performance in this domain is included in order to provide ideas for
future work on this topic.

References

1. Abele, A., McCrae, J., Buitelaar, P., Jentzsch, A., Cyganiak, R.: Linking open data
   cloud diagram 2017 (2017). URL http://lod-cloud.net/
2. Achichi, M., Cheatham, M., Dragisic, Z., Euzenat, J., Faria, D., Ferrara, A., Flouris,
   G., Fundulaki, I., Harrow, I., Ivanova, V., et al.: Results of the ontology alignment
   evaluation initiative 2017. In: OM 2017-12th ISWC workshop on ontology matching,
   pp. 61–113. No commercial editor. (2017)
3. Algergawy, A., Cheatham, M., Faria, D., Ferrara, A., Fundulaki, I., Harrow, I., Hertling,
   S., Jiménez-Ruiz, E., Karam, N., Khiat, A., et al.: Results of the ontology alignment
   evaluation initiative 2018. In: 13th International Workshop on Ontology Matching co-
   located with the 17th ISWC (OM 2018), vol. 2288, pp. 76–116 (2018)
4. Bakillah, M., Mostafavi, M.A.: Universal Ontology of Geographic Space: Semantic En-
   richment for Spatial Data, chap. An Architecture for Enhancing Semantic Interoper-
   ability Based on Enrichment of Semantics of Geospatial Data. IGI Global (2012)
5. Berners-Lee, T., Hendler, J., Lassila, O., et al.: The Semantic Web. Scientific American
   284(5), 28–37 (2001)
6. Buttigieg, P.L., Morrison, N., Smith, B., Mungall, C.J., Lewis, S.E.: The environment
   ontology: contextualising biological and biomedical entities. Journal of biomedical se-
   mantics **4**(1), 43 (2013)
7. Chaabane, S., Jaziri, W.: A novel algorithm for fully automated mapping of geospatial
   ontologies. Journal of Geographical Systems **20**(1), 85–105 (2018)
8. Cheatham, M., Hitzler, P.: String similarity metrics for ontology alignment. In: Inter-
   national Semantic Web Conference, pp. 294–309. Springer (2013)
9. Cheatham, M., Pesquita, C.: Semantic data integration. In: Handbook of Big Data
   Technologies, pp. 263–305. Springer (2017)

10. Cruz, I.F., Xiao, H.: Data integration for querying geospatial sources. (2008)
11. Dornblut, I., Atkinson, R.: Ogc hy_features: a common hydrologic feature model. Open Geospatial Consortium Technical Report OGC 11-039r3, 55pp (2014)
12. Du, H., Alechina, N., Jackson, M., Hart, G.: Matching formal and informal geospatial ontologies. In: Geographic information science at the heart of europe, pp. 155–171. Springer (2013)
13. Duckham, M., Worboys, M.: Automated geographical information fusion and ontology alignment. In: Spatial Data on the Web, pp. 109–132. Springer (2007)
14. Euzenat, J., Shvaiko, P., et al.: Ontology matching, vol. 18. Springer (2007)
15. Faria, D., Pesquita, C., Santos, E., Palmonari, M., Cruz, I.F., Couto, F.M.: The agreementmakerlight ontology matching system. In: OTM Confederated International Conferences" On the Move to Meaningful Internet Systems", pp. 527–541. Springer (2013)
16. Hamdi, F., Safar, B., Niraula, N.B., Reynaud, C.: Taxomap alignment and refinement modules: Results for oaei 2010. In: Proceedings of the 5th International Conference on Ontology Matching-Volume 689, pp. 212–219. CEUR-WS. org (2010)
17. Hitzler, P., Gangemi, A., Janowicz, K.: Ontology Engineering with Ontology Design Patterns: Foundations and Applications, vol. 25. IOS Press (2016)
18. Jain, P., Hitzler, P., Sheth, A.P., Verma, K., Yeh, P.Z.: Ontology alignment for linked open data. In: International Semantic Web Conference, pp. 402–417. Springer (2010)
19. Jean-Mary, Y.R., Shironoshita, E.P., Kabuka, M.R.: Ontology matching with semantic verification. Web Semantics: Science, Services and Agents on the World Wide Web 7(3), 235–251 (2009)
20. Jiménez-Ruiz, E., Grau, B.C.: Logmap: Logic-based and scalable ontology matching. In: International Semantic Web Conference, pp. 273–288. Springer (2011)
21. Joshi, A.K., Hitzler, P., Dong, G.: Logical linked data compression. In: Extended Semantic Web Conference, pp. 170–184. Springer (2013)
22. Li, J., Tang, J., Li, Y., Luo, Q.: Rimom: A dynamic multistrategy ontology alignment framework. IEEE Transactions on Knowledge and data Engineering 21(8), 1218–1232 (2009)
23. Matuszka, T., Kiss, A.: Geodint: towards semantic web-based geographic data integration. In: Asian Conference on Intelligent Information and Database Systems, pp. 191–200. Springer (2014)
24. Otero-Cerdeira, L., Rodríguez-Martínez, F.J., Gómez-Rodríguez, A.: Ontology matching: A literature review. Expert Systems with Applications 42(2), 949–971 (2015)
25. Pesaranghader, A., Muthaiyah, S.: Definition-based information content vectors for semantic similarity measurement. In: Soft Computing Applications and Intelligent Systems, pp. 268–282. Springer (2013)
26. Raskin, R., Pan, M.: Semantic Web for earth and environmental terminology (SWEET). In: Proc. of the Workshop on Semantic Web Technologies for Searching and Retrieving Scientific Data, vol. 25 (2003)
27. Shvaiko, P., Euzenat, J.: Ontology matching: state of the art and future challenges. IEEE Transactions on knowledge and data engineering 25(1), 158–176 (2013)
28. Sinha, G., Mark, D., Kolas, D., Varanka, D., Romero, B.E., Feng, C.C., Usery, E.L., Liebermann, J., Sorokine, A.: An ontology design pattern for surface water features. In: International Conference on Geographic Information Science, pp. 187–203. Springer (2014)
29. Spiliopoulos, V., Vouros, G.A., Karkaletsis, V.: On the discovery of subsumption relations for the alignment of ontologies. Web Semantics: Science, Services and Agents on the World Wide Web 8(1), 69–88 (2010)
30. Suchanek, F.M., Abiteboul, S., Senellart, P.: Paris: Probabilistic alignment of relations, instances, and schema. Proceedings of the VLDB Endowment 5(3), 157–168 (2011)
31. Sunna, W., Cruz, I.F.: Structure-based methods to enhance geospatial ontology alignment. In: International Conference on GeoSpatial Sematics, pp. 82–97. Springer (2007)
32. Thiéblin, É., Haemmerlé, O., Hernandez, N., Trojahn, C.: Towards a complex alignment evaluation dataset
33. Varanka, D.E., Usery, E.L.: An applied ontology for semantics associated with surface water features. Land Use and Land Cover Semantics: Principles, Best Practices, and Prospects p. 145 (2015)

34. Vijayasankaran, N.: Enhanced place name search using semantic gazetteers (2015)
35. Vilches-Blázquez, L., Ramos, J., López-Pellicer, F.J., Corcho, Ó., Nogueras-Iso, J.: An approach to comparing different ontologies in the context of hydrographical information. In: Information fusion and geographic information systems, pp. 193–207. Springer (2009)
36. Wellen, C.C., Sieber, R.: Toward an inclusive semantic interoperability: the case of cree hydrographic features. International Journal of Geographical Information Science **27**(1), 168–191 (2013)

A Journey From Simple to Complex Alignment on Real-World Ontologies

Lu Zhou

DaSe Lab, Wright State University, Dayton OH 45435, USA
`zhou.34@wright.edu`

Abstract. Ontology alignment has been an active research topic for over a decade. Over that time, many developers have focused on creating alignment systems and methods to find simple 1-to-1 equivalence matches between two ontologies. However, very few alignment systems focus on finding complex correspondences. There are several reasons for this limitation. First, there are no widely accepted alignment benchmarks that contain such complex relationships. Second, the traditional evaluation metrics like precision, recall, and f-measure are not accurate enough to evaluate the performance of a complex alignment system. And third, the approaches most commonly used to find simple equivalences do not handle the increased computational complexity of finding complex equivalences well. Therefore, it becomes a big challenge for many developers to create and evaluate the systems. In this paper, in order to advance the development of ontology matching, we seek to address the problem by first developing potential complex alignment benchmarks from real-world ontologies. In addition, we utilize traditional automated alignment systems to suggest complex correspondences, and finally plan to achieve our ultimate goal of creating and evaluating our own complex alignment system based on logical RDF data compression.

1 Problem Statement

Similar to database integration, ontology alignment is an important process in enabling computers to query and reason across the many linked open datasets on the semantic web. This is a difficult challenge because the ontologies underlying different linked datasets can vary in terms of subject area coverage, level of abstraction, ontology modeling philosophy, and even language.

In order to solve this problem, an ontology alignment system aims to identify the entity relationships between two and more ontologies. Such relationships have a wide range of complexity, from basic 1-to-1 equivalence to arbitrary m-to-n relationships. However, over a decade, the majority of the studies in the field still focuses on the simplest end of this scale – finding 1-to-1 equivalence relations between ontologies. Very few ontology alignment systems and methods are developed to uncover complex relations. The reasons for this limitation may lie in the following. First, there are no widely used and accepted ontology alignment benchmarks that involve complex relations. Without these benchmarks,

even if there were a complex alignment system, it is not only very hard to evaluate if this system could correctly detect the complex correspondences, but also it is a challenge to evaluate whether the system is comprehensive enough to find all different kinds of complex patterns. Second, the traditional approach of precision, recall, and f-measure does not seem fine-grained enough to evaluate complex correspondences. A better version of precision and recall is needed [2].

This work seeks to progress in the direction of fostering the development of research activities in the field of complex ontology alignment. We firstly show that real-world ontologies involve many complex relations. And based on these real-world datasets, we develop high quality complex alignment benchmarks, including creating complex alignments and categorizing them into complex patterns. In addition, to decrease the complexity of detecting complex relations, we leverage automated alignment systems to uncover and suggest possible complex relations. Moreover, we plan to apply logical RDF compression with the results that are generated by traditional automated alignment systems to create a new complex alignment system.

2   Relevancy

Ontology alignment seeks to address the conceptual heterogeneities between ontologies. Over a decade, the community of this field remains on creating and improving the algorithms of finding simple alignments. The reason that the researchers do not really dig into the complex alignment for such a long time, is that the community is still discovering and analyzing how to reach the goal. Nowadays, the research related to simple alignment has been well studied. It is actually a good timing to move on to complex ontology alignment, because more and more good alignment systems and algorithms have been published, and also more and more data are populated into ontologies and published as linked open data, the applications that utilize these LODs are required to involve ontology matching and data integration processes [3]. In addition, due to the complexity of the alignments between ontologies, only identifying traditional simple 1-to-1 alignment is not enough to fulfill the growing high demand of most of these applications. Therefore, it is necessary to create complex alignment systems and methods to uncover complex relations in real-world use cases.

This work focuses on addressing the problem from several different perspectives. We prepared benchmarks that involve complex relations from real-world ontologies and will try to distribute them as a new track in the Ontology Alignment Evaluation Initiative (OAEI), which was started in 2005 with the intent to allow researchers in the field to compare the performance of their approaches on a consistent set of benchmarks over time. Since then, it has been more convenient for researchers from different organizations to test their methods. Second, as we've seen, it is a difficult challenge to detect complex relations. This work seeks to narrow down this issue by leveraging traditional alignment systems to suggest possible complex candidates. This would be a valuable starting point for determining the exact relation. Moreover, our ultimate goal in this work is to

create a complex alignment system that has better performance than traditional automated alignment systems.

3   Related work

Regarding the creation of complex alignment benchmarks, Thieblin [11] is creating a complex alignment benchmark using the Conference track ontologies within the OAEI. This work is partially completed, and at the time of this writing it covers three of the seven ontologies. In addition, we are collaborating with them (under their direction) to complete the dataset and prepare a new task in OAEI to evaluate complex alignment systems.

However, even though there are no widely accepted and used benchmarks that involve complex relations, some researchers still tried to create alignment systems and evaluate them using their ow manually developed reference alignment. BLOOMS [3] is an alignment system based on Wikipedia to detect the subsumption relations. Other subsumption systems have evaluated the precision of their approach by manually validating relations produced by their system, while foregoing an assessment of recall [9]. There are some more general approaches based on complex patterns to detect complex correspondences. Ritze *et al* [7, 8] proposed several complex correspondences patterns. Such as: Class by Attribute Value, Class by Attribute Type, Class by Inverse Attribute Type, Inverse Properties, and Property Chain. In addition, Ritze also utilized linguistic analysis techniques, like detection of antonymy, active form, etc to help detect these complex patterns. Other similar work was done by Šváb-Zamazal and Svátek [10]. It firstly detected N-ary relations in the source ontology. And then, it matched the detected N-ary relations to an object property in the target ontology.

Our work differs from the above methods in several aspects. First, we focus on real-world ontologies, which we found that these datasets are not only used by academic researchers, but also the industries and governments to develop applications for the usage of normal human life. There are some interesting relations that have not yet been mentioned in the current benchmark from OAEI. In addition, the instance data of these real-world ontologies are ready to be used as additional information to help improve the performance of alignment process. In contrast to this, significant instance data is not readily available for most of the OAEI Conference Track ontologies. Moreover, regarding the creation of a complex alignment, instead of comparing each entity in the source ontology to each entity in the target ontology, we apply logical RDF compression to list a set of available rules, and narrow down them based on the suggestion generated by traditional alignment systems to finally output the complex relation. More details are discussed in Section 5.

4   Research Questions and Hypotheses

The research questions that we plan to address are listed as follows:

1. Do real-world ontology alignments contain complex relations?
2. How well do traditional automated alignment systems work on real-world matching tasks that contain complex relations?
3. Can we create an automated alignment system that performs better than traditional alignment systems on finding complex relations that exist between ontologies?

Our hypotheses associated to the above research questions are the following:

1. Most real-world ontologies contain many complex relations.
2. Traditional automated alignment systems may not be able to identify complex relationships directly, but they may be able to suggest the atomic entities involved in such relations.
3. A complex alignment system that leverages logical RDF compression can effectively identify complex relations between ontologies.

5   Approach

Hypothesis 1 Our previous work with the NSF EarthCube Initiative and the US Geological Survey involved the time consuming task of manually aligning several real-world ontologies. These alignments have been discussed and evaluated by domain experts and ontology engineers to guarantee that they are of high quality. We will inventory these alignments, along with any other real-world alignments we can acquire, to answer the first research question: Do real-world ontology alignments contain complex relations?

Hypothesis 2 To answer the second research question, "How well do traditional automated alignment systems perform on real-world matching tasks that contain complex relations?", we plan to first evaluate several state of the art alignment systems on the alignment tasks mentioned above. Since traditional alignment systems only attempt to identify simple relations between ontologies, their performance will be limited to the percentage of the alignments that involve these types of relations. However, it is possible that these systems, while they cannot identify the precise relationship that holds between an entity in the source ontology and two or more entities in the target ontology, they can at least identify the entities involved in the relationship. For example, in the relation below, the class $Mischaakusaakihiikin$ in Cree ontology is equivalent to the intersection of instances of $LakeOrPond$ and entities that $isContainedBy$ a $SwampOrMarsh$ in the SWO ontology. While a traditional alignment system cannot identify things like intersection or value restrictions, it may be able to determine that $LakeOrPond$, $isContainedBy$, and $SwampOrMarsh$ are related in some way to $Mischaakusaakihiikin$. To check this, for each entity e_s

in source ontology O_s, we will use the automated alignment systems to give a list of candidates e_t in the target ontology O_t, ordered by the similarity assigned to them by the alignment system. We will evaluate the performance against the benchmark using mean reciprocal rank [6].

```
EquivalentClasses(cree:Mischaakusaakihiikin
        ObjectIntersectionOf(swo:LakeOrPond
                ObjectSomeValuesFrom(swo:isContainedBy
                        swo:SwampOrMarsh)))
```

Hypothesis 3 As we mentioned, the ultimate goal of this work is to see if we can create an automated alignment system that effectively identifies complex relationships that exist between two ontologies. Our planned approach is to create an extensional matcher (i.e. one that relies upon instance data) that leverages logical RDF compression [5]. Logical RDF compression uses the FP-Growth data mining algorithm to generate rules that can be stored in lieu of the triples they are based on. For example, say that a linked dataset contains triples about university students. There might be many triples of the form <ind1 hasMajor ComputerScience> and many corresponding triples of the form <ind1 isEnrolledIn CollegeOfEngineering> because, according to this dataset, all Computer Science majors are enrolled in the College of Engineering. Logical RDF compression would replace the second set of triples with a single rule: if x is hasMajor ComputerScience then x isEnrolledIn CollegeOfEngineering, and these triples could then be generated on-the-fly in response to queries, thereby saving space in the linked dataset. While logical RDF compression seeks to find any rules that can be used to shrink the dataset, it is possible that some of these rules represent meaningful semantic relations that hold between entities. For example, if hasMajor ComputerScience exists in one ontology and isEnrolledIn CollegeOfEngineering exists in another ontology, then it may be possible to infer the relation below.

```
SubClassOf(ObjectSomeValuesFrom(ont1:hasMajor ont1:ComputerScience)
    ObjectSomeValuesFrom(ont2:isEnrolledIn ont2:CollegeOfEngineering))
```

Because the FP-Growth algorithm underlying logical RDF compression can generate a very large number of rules, some mechanism must be put in place to choose the more semantically meaningful rules rather than the ones that result in the most compression. Our planned approach for this is to choose rules that involve the entities suggested by traditional alignment systems.

The overall work flow is shown in Figure 1. We first apply the traditional alignment systems to suggest the candidates as we described above. And then, we use RDF compression [5] on the source ontology to list a set of compression rules. Based on the suggested candidates from traditional alignment systems, we can create a filter to pick up the compression rules, and finally output the complex relations.
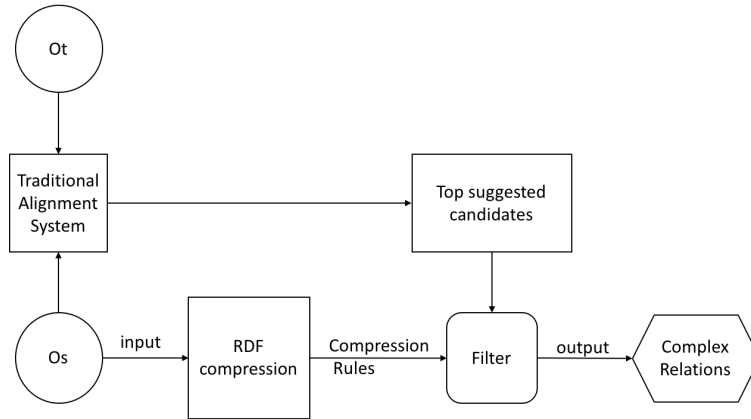
Fig. 1. The Work Flow of the Proposed Complex Alignment System

6   Preliminary Results

We have some preliminary results in terms of finding complex relations from real-world ontologies. There are two different datasets that we are currently working on. One dataset is from GeoLink project[1] that was funded under the U.S. National Science Foundation's EarthCube initiative. Another dataset is a set of ontologies from surface water domain. Based on these two sets of ontologies, we have developed the alignments in consultation with domain experts from different institutions. In GeoLink datasets, the number of classes, object properties, and data properties in GeoLink base ontology (GBO) and GeoLink modular ontology (GMO) are showed in Table 1. We found that the 87 out of 111 relations in them are complex relations, including not only class subsumption, property subsumption, property chain equivalence, and property chain subsumption that were introduced in [7], but also some typecasting relations. The idea of typecasting, and why it is important in ontology modeling, is formally introduced and discussed in [4]. Moreover, the alignment is also available in both EDOAL and rules syntax for the purpose of manipulating and reading respectively. The full dataset has been uploaded to the FigShare[2]. We also wrote a paper to describe the creation and submitted to ISWC2018, which is currently under-review.

In hydrography dataset, it consists of four ontologies. (a) The Surface Water Ontology (SWO), which was originally presented in [12], was developed by the US Geological Survey (USGS). (b) The Hydro3 ontology was developed by individuals at the University of Maine in order to support expanded gazetteer functions using topology and semantic inference [13]. (c) The HydrOntology is a non-English ontology, which was developed by the Spanish National Geographic Institute (IGN) [1]. (d) Cree surface water ontology is in a language, Cree, which

---

[1] https://www.geolink.org/
[2] https://doi.org/10.6084/m9.figshare.5907172

Table 1. The Number of Classes, Object Properties, and Data Properties in Both GeoLink Ontologies

Ontology	Classes	Object Properties	Data Properties
GBO	40	149	49
GMO	156	124	46

is spoken by some of the native inhabitants of northern Canada. The reason of choosing these four ontologies is that Hydro3, HydrOntology and Cree have a large degree of overlap with SWO. Therefore, we utilize these four ontologies to create a reference alignment manually. Table 2 shows the number of classes, and object properties, and data properties. And we found that the 84 out of 197 relations in them are complex relations. We are currently writing a paper about evaluating the performance of traditional automated alignment systems on this dataset. The traditional automated alignment systems are able to find the simple relationships. But, they may not be able to identify these complex relationships directly. We hypothesize that they may be able to suggest some possible complex relations partially. Moreover, it is able to greatly narrow down the entities and mitigates the high complexity of computation.

Table 2. The Number of Classes, Object Properties, and Data Properties in Hydrography Ontologies

Ontology	Classes	Object Properties	Data Properties
SWO	85	20	1
Hydro3	22	34	0
HydrOntology	154	47	75
Cree	83	21	7

7   Evaluation Plan

In this section, we introduce the evaluation plan for each research question. For research question 1, as we showed in Section 6, it is considered successful that we have found many complex alignments in real-world ontologies. In addition, we are also preparing to incorporate the dataset into OAEI as a new track for other researchers accessing it. For research question 2, after achieving the list of entities involved in a complex relation, we will evaluate the performance against the benchmark using mean reciprocal rank as we discussed in Section 5. For research question 3, it is a challenge to evaluate the performance of a complex alignment system. The traditional precision, recall, and f-measure metrics do not

seem fine-grained enough. For example, there is a relation between Hydro3 and SWO, if one alignment system identified this:

```
EquivalentClasses(
        ObjectIntersectionOf(
                hydro3:Hydrographic_Feature
                hydro3:Hydrographic_Structure
                hydro3:Boundary)
        swo:HydrographicFeature))
```

and another identified this:

```
SubClassOf(
        ObjectUnionOf(
                hydro3:Hydrographic_Feature
                hydro3:Island
                hydro3:Shore)
        swo:HydrographicFeature))
```

Based on the reference alignment, we need a metric to consider the first system "more correct" than the second. We plan to develop a performance metric that more accurately reflects the performance of a complex alignment system. Another challenge is that, to the best of our knowledge, there are no existing complex alignment systems against which to compare our approach. Therefore, we might consider evaluating the performance based on our manually created reference alignment.

8   Reflections

It is primarily difficult to identify complex relationships between ontologies because of computational complexity. A naive approach would need to compare every entity in the source ontology to every possible *combination* of entities in the target ontology, which is not feasible. Instead of doing this, our proposed approach has a good chance of success because it is based on a logical RDF compression method that has already been shown to be applicable to large datasets, and we also can further limit the search space by using the output from traditional alignment systems to narrow the focus. There are some reflections. The performance of using logical RDF compression in our alignment system is primarily based on the *Abox* information in the ontology. It is still not clear that how to apply our alignment algorithm to a more generalized scenario. However, our approach is feasible, and can be a good starting point to achieve the ultimate goal in the future.

References

1. Blázquez, L.M.V., Gargantilla, J.Á.R., López-Pellicer, F.J., Corcho, Ó., Nogueras-Iso, J.: An approach to comparing different ontologies in the context of hydrographical information. In: Information Fusion and Geographic Information Systems, Proceedings of the Fourth International Workshop, IF&GIS 2009, 17-20 May 2009, St. Petersburg, Russia. pp. 193–207 (2009)

2. Euzenat, J.: Semantic precision and recall for ontology alignment evaluation. In: IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007. pp. 348–353 (2007)

3. Jain, P., Hitzler, P., Sheth, A.P., Verma, K., Yeh, P.Z.: Ontology alignment for linked open data. In: The Semantic Web - ISWC 2010 - 9th International Semantic Web Conference, ISWC 2010, Shanghai, China, November 7-11, 2010, Revised Selected Papers, Part I. pp. 402–417 (2010)

4. Krisnadhi, A.A., Hitzler, P., Janowicz, K.: On the capabilities and limitations of OWL regarding typecasting and ontology design pattern views. In: Ontology Engineering - 12th International Experiences and Directions Workshop on OWL, OWLED 2015, co-located with ISWC 2015, Bethlehem, PA, USA, October 9-10, 2015, Revised Selected Papers. pp. 105–116 (2015)

5. Pan, J.Z., Gómez-Pérez, J.M., Ren, Y., Wu, H., Wang, H., Zhu, M.: Graph pattern based RDF data compression. In: Semantic Technology - 4th Joint International Conference, JIST 2014, Chiang Mai, Thailand, November 9-11, 2014. Revised Selected Papers. pp. 239–256 (2014)

6. Radev, D.R., Qi, H., Wu, H., Fan, W.: Evaluating web-based question answering systems. In: Proceedings of the Third International Conference on Language Resources and Evaluation, LREC 2002, May 29-31, 2002, Las Palmas, Canary Islands, Spain (2002)

7. Ritze, D., Meilicke, C., Sváb-Zamazal, O., Stuckenschmidt, H.: A pattern-based ontology matching approach for detecting complex correspondences. 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 (2009)

8. Ritze, D., Völker, J., Meilicke, C., Sváb-Zamazal, O.: Linguistic analysis for complex ontology matching. In: Proceedings of the 5th International Workshop on Ontology Matching (OM-2010), Shanghai, China, November 7, 2010 (2010)

9. Suchanek, F.M., Abiteboul, S., Senellart, P.: Paris: Probabilistic alignment of relations, instances, and schema. Proceedings of the VLDB Endowment **5**(3), 157–168 (2011)

10. Šváb-Zamazal, O., Svátek, V.: Towards ontology matching via pattern-based detection of semantic structures in owl ontologies. In: Proceedings of the Znalosti Czecho-Slovak Knowledge Technology conference (2009)

11. Thieblin, E., Haemmerle, O., Hernandez, N., Trojahn, C.: Towards a complex alignment evaluation dataset. In: Proceedings of the 16th International Conference on Ontology Matching-Volume 2032. CEUR-WS. org (2017)

12. Varanka, D.E., Usery, E.L.: An applied ontology for semantics associated with surface water features. Land Use and Land Cover Semantics: Principles, Best Practices, and Prospects p. 145 (2015)

13. Vijayasankaran, N.: Enhanced place name search using semantic gazetteers (2015)

Towards Association Rule-Based Complex Ontology Alignment

Lu Zhou[1], Michelle Cheatham[2], and Pascal Hitzler[1]

[1] DaSe Lab, Kansas State University, Manhattan KS 66506, USA
{luzhou, hitzler}@ksu.edu
[2] Wright State University, Dayton OH 45435, USA
michelle.cheatham@wright.edu

Abstract. Ontology alignment has been studied for over a decade, and over that time many alignment systems have been developed by researchers in order to find simple 1-to-1 equivalence alignments between ontologies. However, finding complex alignments, i.e., alignments that are not simple class or property equivalences, is a topic largely unexplored but with growing significance. Currently, establishing a complex alignment requires domain experts to work together to manually generate the alignment, which is extremely time-consuming and labor-intensive. In this paper, we propose an automated method based on association rule mining to detect not only simple alignments, but also more complex alignments between ontologies. Our algorithm can also be used in a semi-automated fashion to effectively assist users in finding potential complex alignments which they can then validate or edit. In addition, we evaluate the performance of our algorithm on the complex alignment benchmark of the Ontology Alignment Evaluation Initiative (OAEI).

1 Introduction

Ontology alignment is an important step in enabling computers to query and reason across the many linked datasets on the semantic web. This is a difficult challenge because the ontologies underlying different linked datasets can vary in terms of subject area coverage, level of abstraction, ontology modeling philosophy, and even language. Due to the importance and difficulty of the ontology alignment problem, it has been an active area of research for over a decade [21].

Ideally, alignment systems should be able to uncover any entity relationship across two ontologies that can exist within a single ontology. Such relationships have a wide range of complexity, from simple 1-to-1 equivalence, such as a Person in one ontology being equivalent to a Human in another ontology, to arbitrary m-to-n complex relationships, such as a Professor with a hasRank property value of "Assistant" in one ontology being a subclass of the union of the Faculty and TenureTrack classes in another. Unfortunately, the majority of the research activities in the field of ontology alignment remains focused on the simplest end of this scale – finding 1-to-1 equivalence alignments between ontologies. Indeed, identifying arbitrarily complex alignment is known to be significantly

harder than finding 1-to-1 equivalences. In the latter case, a naive approach can compare every entity from the source ontology against every entity in the target ontology, which is feasible for small- and medium-sized ontologies. However, a complex alignment can potentially involve many entities from both ontologies, so pair-wise comparison is insufficient, and the search space become very large even for small ontologies. It is indeed very difficult for either a human expert or an automated system to evaluate all possible combinations [2, 19].

In this paper, we propose a complex alignment algorithm based on association rule mining. Our algorithm automatically discovers potential complex correspondences which can then be presented to human experts in order to effectively generate complex alignment between two ontologies with populated common instance data. We evaluate the performance of our system on one of the benchmarks from the complex alignment track of the OAEI 2018,[3] the GeoLink benchmark, which contains around 74k instances from real-world datasets. Significant instance data, which is required for the association rule mining approach, is not available for the remaining benchmarks.[4] The main contributions of this paper are the following:

 – The association rule-based algorithm automatically detects not only 1-to-1 equivalences, but also more complex alignment between two ontologies.
 – A detailed analysis of the results provides a good understanding of the efficacy of this approach and identifies further directions for advancement.

There is a side contribution when we analyze the results, which is that our algorithm shows that shared instance data between two ontologies can be a good resource to improve the performance of ontology alignment.

The rest of the paper is organized as follows. Section 2 discusses related work in ontology alignment using association rule mining and instance data and complex ontology alignment, including existing alignment algorithms and relevant benchmarks. Section 3 gives background on the FP-growth association rule mining algorithm. Section 4 illustrates the association rule-based alignment algorithm in detail, along with the alignment patterns used to generate the alignment between ontologies. The analysis of the performance of the system is discussed in Section 5. Section 6 concludes with a discussion of potential future work in this area.

2   Related Work

Association rule mining has already been used for finding 1:1 simple alignments. AROMA [4] is a hybrid, extensional and asymmetric ontology alignment method that makes use of association rules and a statistical measure. It relies on the idea that "An entity A will be more specific than or equivalent to an entity B if the vocabulary used to describe A and its instances tends to be included in that of B and its instances." In addition, association rule mining is also used in discovering rules in ontological knowledge bases [10] and logical linked data compression [15].

---

[3] http://oaei.ontologymatching.org/2018/complex/index.html
[4] It might be available for OAEI 2019.

There are also some instance-based ontology alignment systems that utilize Abox information to generate 1:1 simple alignments between ontologies. GLUE [6] uses joint probability distributions to describe the similarity of concepts in two ontologies. For example, $p(A, B)$ is the probability that an instance in the domain belongs to both concept A and concept B. And then, if the instances of concept A and concept B are in isolation, GLUE uses the instances of A to learn a classifier for A, and then classifies instances of B according to that classifier, and vice-versa. FCA_MERGE also utilizes common instances between ontologies [22]. FCA_MERGE extracts instances from a given set of domain-specific text documents by applying nature language processing techniques. Based on the extracted instances, FCA_MERGE applies mathematical techniques to derive a lattice of concepts as a structural result of FCA_MERGE. More instance-based alignment systems have been discussed in the survey [26].

There are some related studies on creating algorithms to find complex alignment between ontologies. Early work on generating complex alignment is [19, 20]. Therein, three complex alignment patterns were described, which are Class by Attribute Type (CAT), Class by Attribute Value (CAV), and Property Chain (PC). Based on these patterns, the authors generated complex alignments on the Conference and Benchmark datasets from the OAEI. [13] identified complex alignments by defining knowledge rules and using a probabilistic framework to integrate a knowledge-based strategy with standard terminology-based and structure-based strategies. More recent related work is currently being undertaken by Thieblin et al. [24]. They propose a complex alignment approach that relies on the notion of Competency Question for Alignment (CQA). The approach translates a CQA into a SPARQL query and extracts a set of instance data from the source ontology. Then the matching is performed by finding the lexically similar surroundings between the set of instance data and the instances in the target ontology. This approach resulted in the CANARD system [23]. However, the current version of the system is limited to finding complex correspondences that only involve classes. More complex correspondences containing properties are still not taken into account [23]. Another alignment system that works on the detection of the complex alignment is the complex version of AgreementMakerLight (AMLC) [9]. This system focuses on the complex Conference benchmark to find alignments that follow the CAT and CAV patterns.

In OAEI 2018, the first version of the complex alignment track [25] opened new perspectives in the field of ontology matching. It comprised four different benchmarks containing complex relations. However, the results from the first year were rather poor. Only 2 out of 15 systems, AMLC and CANARD, were able to generate any correct complex correspondences on the complex Conference and Taxon benchmarks, and the correct number of mappings found was quite limited. The very limited performance of the two systems of course shows avenues for improvement in the future. More details of evaluations and results can be accessed on the OAEI 2018 website.[5]

---

[5] http://oaei.ontologymatching.org/2018/complex/index.html

Our algorithm differs from the above methods in several aspects. First, [9], [13], and [19] focus on computing lexical or terminological similarity to decide on complex alignments, while our system takes advantage of instance data to generate association rules between ontologies. While the CANARD system also relies on the instance data, we use it in completely different ways. In addition, the current version of CANARD is limited to finding complex correspondences that involve only classes, while our algorithm does not have this limitation. Second, our evaluation of results is more detailed, in order to provide insight into how to improve the performance of complex alignment algorithms. Specifically, we break the evaluation process down into two subtasks: entity identification and relationship identification. We utilize a variation of traditional evaluation metrics called relaxed precision, recall, and f-measure [7] to present the final results of the full complex alignment.

3   Background

In order to help the reader understand how we apply association rule mining and the FP-growth algorithm on the ontology alignment task, we introduce here some concepts that we frequently mention in the rest of the paper.

Association Rule Mining. Our alignment system mainly depends on a data mining algorithm called association rule mining, which is a rule-based machine learning method for discovering interesting relations between variables in large databases [17]. Over the years, association rule mining has played an important role in many data mining tasks, such as market basket analysis, web usage mining, and bioinformatics. Many algorithms for generating association rules have been proposed, like Apriori [1] and FP-growth algorithm [11]. In this paper, we use FP-growth to generate association rules between ontologies, since the FP-growth algorithm has been proven superior to other algorithms [11] and will improve the algorithm in terms of run-time.

Transaction Database. Let $I = \{i_1, i_2, \ldots, i_n\}$ be a set of distinct attributes called items. Let $D = \{t_1, t_2, \ldots, t_m\}$ be a set of transactions where each transaction in D has a unique transaction ID and contains a subset of the items in I. Table 1 shows a list of transactions corresponding to a list of triples. The data in an ontology can be displayed as a set of triples, each consisting of subject, predicate, and object. Here, subjects represent the identifiers and the set of corresponding properties with the objects represent transactions, which are separated by the symbol "|". I.e., a transaction is a set $T = (s, Z)$ such that s is a subject, and each member of Z is a pair (p, o) of a property and an object such that (s, p, o) is a triple.

FP-growth. The FP stands for frequent pattern. The FP-growth algorithm is run on the transaction database in order to determine which combinations of items co-occur frequently. The algorithm first counts the number of occurrences of all individual items in the database. Next, it builds an FP-tree structure by inserting these instances. Items in each instance are sorted by descending order of their frequency in the dataset, so that the tree can be processed quickly. Items

Table 1. Triples and Corresponding Transactions

| s_1 p_1 o_1 |
| s_1 p_2 o_2 |
| s_1 p_4 o_4 |
| s_2 p_1 o_1 |
| s_2 p_2 o_2 |
| s_2 p_3 o_3 |
| s_2 p_4 o_4 |
| s_3 p_1 o_1 |
| s_3 p_2 o_2 |

TID	Itemsets
s_1	$p_1|o_1$, $p_2|o_2$, $p_4|o_4$
s_2	$p_1|o_1$, $p_2|o_2$, $p_3|o_3$, $p_4|o_4$
s_3	$p_1|o_1$, $p_2|o_2$

Table 2. Examples of Association Rules

Antecedent	Consequent
$p_4|o_4$, $p_1|o_1$	$p_2|o_2$
$p_2|o_2$	$p_1|o_1$
$p_4|o_4$	$p_1|o_1$

in each instance that do not meet the predefined thresholds, such as minimum support and minimum confidence (see below for these terms), are discarded. Once all large itemsets have been found, the association rule creation begins.

Association Rule. Every association rule is composed of two sides. The left-hand-side is called the antecedent, and the right-hand-side is the consequent. These rules indicate that whenever the antecedent is present, the consequent is likely to be as well. Table 2 shows some examples of association rules generated from the transaction database in Table 1.

Support. Support indicates how frequently an itemset appears in the dataset. The FP-growth algorithm finds the frequent itemsets from the dataset based on the minimum support threshold. In our alignment system, the minimum support value is examined and set to 0.001 to guarantee the best performance.

Confidence. Confidence is an indication of how often an association rule has been found to be true, i.e. how often the presence of the antecedent is associated with the presence of the consequent. The minimum confidence can be tuned to find relatively accurate rules. In this paper, we use the minimum confidence of 0.3 as default value. And we tune the value to 1 when we mine the association rules that may contain complex relations, because our algorithm would focus on precision-oriented results.

Lift. Lift is the ratio of the observed support to that expected if the antecedent and consequent were independent. If the lift is greater than 1, it means that the two items are dependent on one another, which indicates that the association rule useful. In our approach, lift is used to choose between otherwise equal options when detecting simple mappings. When the confidence values of two association rules are the same, the one with higher lift value is selected as the basis for the mapping.

4   Association Rule-Based Alignment Algorithm

In this section, we introduce the proposed ontology alignment algorithm based on association rule mining in detail. Figure 1 illustrates the overview of our proposed algorithm.
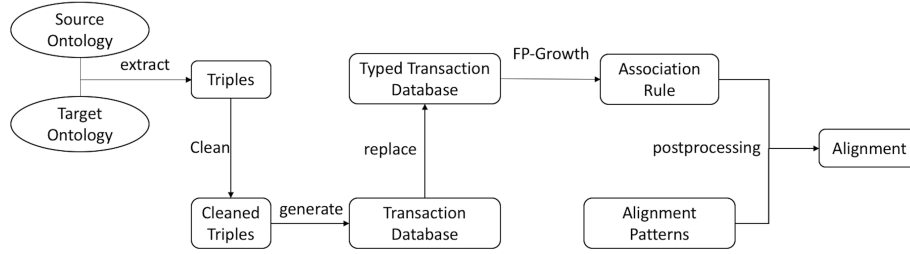
Fig. 1. Overview of The Proposed Alignment Algorithm

4.1   Data Preparation

We first extract all triples ⟨Subject, Predicate, Object⟩ from the source and target ontologies. Each item in a triple is expressed as a web URI. After collecting all of the triples, we prepare the data as follows: we only keep the triples that contain at least one entity under the source or the target ontology namespace and also the triples that contain rdf:type information, since our algorithm relies on this information. After this, there are still some triples containing less useful information for association rule mining, which follow this format: x rdf:type owl:NamedIndividual. This triple is not very informative except stating the subject x is an individual. But, it frequently occurs in the dataset and may lead to noises when applying the FP-growth algorithm, since the frequency of occurrence impacts the results of FP-growth. So, we filter out such noise from the dataset as well.

After this filtering process, we generate the transaction database for the FP-growth algorithm based on all of the remaining triples. The subjects serve as the transaction IDs, and the predicates with the objects separated by the symbol "|" are the items for each transaction. Then we replace the object in the triples with its rdf:type,[6] because we focus on generating schema-level (rather than instance-level) mapping rules between two ontologies, and the type information of the object is more meaningful than the original URI. If an object in a triple has rdf:type of a class in the ontology, we replace the URI of the object with its class. If the object is a data value, the URI of the object is replaced with the datatype. If the object already is a class in the ontology, it remains unchanged. Tables 3 and 4 show some examples of the conversion.

4.2   Association Rule and Alignment Generation

We run the FP-growth algorithm on the transaction database and generate a set of association rules. Since we are trying to find the mappings *between* two ontologies, we focus on mining the rules whose antecedent only contains entities

---

[6] Our evaluation data has only single type. If there are multiple types of the object, it can also combine the subject and predicate as additional information to determine the correct type, or keep both types as two triples.

Table 3. Original Transaction Database

TID	Itemsets
x_1	gbo:hasAward|y_1, gmo:fundedBy|y_2
x_2	gbo:hasFullName|y_3, gmo:hasPersonName|y_4
x_3	rdf:type|gbo:Cruise, rdf:type|gmo:Cruise

Table 4. Typed Transaction Database

TID	Itemsets
x_1	gbo:hasAward|gbo:Award, gmo:fundedBy|gmo:FundingAward
x_2	gbo:hasFullName|xsd:string, gmo:hasPersonName|gmo:PersonName
x_3	rdf:type|gbo:Cruise, rdf:type|gmo:Cruise

from the source ontology and whose consequent only contains entities from the target ontology. The association rules tell us which source entities are related to which target entities, but they do not give us information on *how* those entities are related. In order to determine this, we analyze the output of the association rule mining step in light of the common alignment patterns introduced in [19, 27]. In the following, we introduce how we leverage these alignment patterns to filter the association rules and generate the corresponding alignment. The following examples that we use in this paper are from the GeoLink benchmark [27]. gbo: is the prefix of the namespace of the GeoLink Base Ontology (GBO), and gmo: is the prefix of the namespace of the GeoLink Modular Ontology (GMO). The alignment between the two ontologies contains both simple and complex correspondences. To deal with the redundancy of generated association rules, we always keep the simpler rule as the result. For example, there are two association rules generated by our system. Cruise in the GBO is equivalent to the domain of fundedBy with it range of FundingAward in the GMO. And Cruise in the GBO is also equivalent to Cruise in the GMO, which is the domain of fundedBy. Therefore, the two mapping rules are semantically equivalent. And we only keep the second rule which is the simpler one as our result.

Simple Alignment. Simple alignment is a set of simple correspondences that refer to basic 1-to-1 simple mappings between two ontologies, in which the entities involved may be either classes or properties.

1:1 Class Alignment. The first pattern is simple 1-to-1 class relationships. Classes C_1 and C_2 are from ontology O_1 and ontology O_2, respectively. So, we target the association rules with the following format:

Association Rule format: rdf:type|C_1 → rdf:type|C_2
Example: rdf:type|gbo:Award → rdf:type|gmo:FundingAward
Generated Alignment: gbo:Award(x) → gmo:FundingAward(x)

The left and right hand side of the arrow represent the antecedent and consequent in the association rules, respectively. In the example, the associa-

tion rule implies that if an individual x has rdf:type of gbo:Award, then x also has rdf:type of gmo:FundingAward. This means that gbo:Award is a subclass of gmo:FundingAward. If there is another association rule containing the reverse information, which means that gmo:FundingAward is also a subclass of gbo:Award then we can generate an alignment based on the two association rules stating that gbo:Award is equivalent to gmo:FundingAward. This method of choosing between subsumption and equivalence relationships is used for all of the following types of correspondences as well.

1:1 Property Alignment. This pattern captures simple 1-to-1 property mappings. The property can be either an object property or a data property.

(1) Object Property Alignment. Since we have the information of the type of the object in the association rule, we can use the type information to filter the mapping candidates. When we align two object properties, the range types of the properties are usually either equivalent to each other or compatible (because they are in a subclass or superclass relationship). In this paper, our algorithm is precision-oriented. Therefore, we require the object properties in the two ontologies to have equivalent (rather than compatible) ranges in order to be considered equivalent. Range equivalence is determined through the results of the simple class alignment introduced above. Object Property op_1 with its range type t_1 and object property op_2 with its range type t_2 are from ontology O_1 and ontology O_2, respectively. In order to find this alignment, we select the association rules with the following format:

Association Rule format: $op_1|t_1 \rightarrow op_2|t_2$
Example: gbo:hasAward|gbo:Award $\rightarrow$ gmo:fundedBy|gmo:FundingAward
Generated Alignment: gbo:hasAward$(x, y) \rightarrow$ gmo:fundedBy(x, y)

We know from the results of the simple class alignment that gbo:Award is equivalent to gmo:FundingAward. This association rule says that gbo:hasAward is subsumed by gmo:fundedBy. If there is another association rule containing the reverse relationship, we can generate the mapping that gbo:hasAward is equivalent to gmo:fundedBy.

(2) Data Property Alignment. Similar to aligning object properties, when aligning two data properties, the range values of the two properties should be of a compatible datatype. In this paper, we only investigate equivalent datatypes. Data Property dp_1 with its range value t_1 and property dp_2 with its range value t_2 are from ontology O_1 and ontology O_2, respectively.

Association Rule format: $dp_1|t_1 \rightarrow dp_2|t_2$
Example:
       gbo:hasIdentifierValue|xsd:string $\rightarrow$ gmo:hasIdentifierValue|xsd:string
Generated Alignment:
       gbo:hasIdentifierValue$(x, y) \rightarrow$ gmo:hasIdentifierValue(x, y)

(3) Data/Object to Object/Data Property Alignment. It is possible that two ontologists may model the same property differently – e.g., there is an example in the OAEI GeoLink complex alignment benchmark [27]. The entity

hasIdentifierScheme is modeled as an object property in the GBO with a range of class IdentifierScheme. But, this entity is modeled as a data property in the GMO with a range of the string datatype. In this case, we calculate the Levenshtein string similarity between the labels of the two properties and keep the pairs within a predefined threshold (0.9 is examined to get the best performance). The association rule should have the following format:

Association Rule format: $op_1/dp_1|t_1 \rightarrow dp_2/op_2|t_2$
Example:
      gbo:hasIdentifierScheme|gbo:IdentifierScheme $\rightarrow$
                      gmo:hasIdentifierScheme|xsd:string
Generated Alignment:
      gbo:hasIdentifierScheme$(x, y) \rightarrow$ gmo:hasIdentifierScheme(x, y)

Complex Alignment. Complex alignment is a set of Complex correspondences that refer to more complex patterns, such as 1-to-n equivalence, 1-to-n subsumption, m-to-n equivalence, m-to-n subsumption, and m-to-n arbitrary relationship.

1:n Class Alignment. This type of pattern was first introduced in [19]. It contains two different patterns: the Class by Attribute Type pattern (CAT) and the Class by Attribute Value pattern (CAV). In addition, [27] introduced another pattern called Class Typecasting.

(4) Class by Attribute Type. This pattern states that a class in the source ontology is in some relationship to a complex construction in the target ontology. This complex construction may comprise an object property and its range type. Class C_1 is from ontology O_1, and object property op_1 and its range type t_1 are from ontology O_2.

Association Rule format: rdf:type$|C_1 \rightarrow$ op$_1|$t$_1$
Example: rdf:type|gbo:PortCall $\rightarrow$ gmo:atPort|gmo:Place
Generated Alignment: gbo:PortCall$(x) \rightarrow$ gmo:atPort$(x, y) \wedge$ gmo:Place(y)

In this example, this association rule implies that if the subject x is an individual of class gbo:PortCall, then x is subsumed by the domain of gmo:atPort with the range type of gmo:Place. The equivalence relationship can be generated by combining another association rule holding the reverse information.

(5) Class by Attribute Value. This pattern is similar to the previous one. It just replaces the object property with a data property. Class C_1 is from ontology O_1, and data property dp_1 and its datatype of the range value t_1 are from ontology O_2.

Association Rule format: rdf:type$|C_1 \rightarrow$ dp$_1|$t$_1$
Example: rdf:type|gbo:Identifier $\rightarrow$ gmo:hasIdentifierScheme|xsd:string
Generated Alignment: gbo:Identifier$(x) \rightarrow$ gmo:hasIdentifierScheme(x, y)

(6) Class Typecasting. This pattern indicates that an individual x of type C_1 in one ontology O_1 is cast into a subclass of C_2 in the other ontology O_2.

Association Rule format: rdf:type$|C_1 \rightarrow$ rdfs:subClassOf$|C_2$
Example: gbo:PlaceType $\rightarrow$ rdfs:subClassOf|gmo:Place
Generated Alignment: gbo:PlaceType $\rightarrow$ rdfs:subClassOf$(x,$ gmo:Place$)$

1:n Property Alignment This pattern represents a Property Typecasting relationship that is defined in [27].

(7) 1:n Property Typecasting. This pattern is similar in spirit to the Class Typecasting patterns mentioned above. However, in this case, a property from one ontology is cast into a class assignment statement in the other ontology.

Association Rule format: $p_1|t_1 \rightarrow \mathsf{rdf{:}type}|C_2$

Example: $\mathsf{gbo{:}hasPlaceType}|\mathsf{gbo{:}PlaceType} \rightarrow \mathsf{rdf{:}type}|\mathsf{gmo{:}Place}$

Generated Alignment:

$\mathsf{gbo{:}hasPlaceType(x,y)} \land \mathsf{gbo{:}PlaceType(y)} \rightarrow \mathsf{gmo{:}Place(x)}$

m:n Complex Alignment. This group contains the most complex correspondences.

(8) m:n Property Chain. This pattern applies, for example, when a property, together with type restrictions on one or both of its fillers, in one ontology, has been used to "flatten" the structure of the other ontology by short-cutting a property chain in that ontology. The pattern also ensures that the types of the property fillers involved in the property chain are typed appropriately in the other ontology. The class C_1 and property r_1 with its range restriction t_1 are from ontology O_1, and classes B_i and properties p_i with its range restriction d_i are from ontology O_2.

Association Rule format:

$\mathsf{rdf{:}type}|C_1, r_1|t_1 \rightarrow \mathsf{rdf{:}type}|B_1, p_1|d_1, \dots, \mathsf{rdf{:}type}|B_i, p_i|d_i$

Example:

$\mathsf{gbo{:}Award}, \mathsf{gbo{:}hasSponsor}|\mathsf{gbo{:}Organization}$
$\rightarrow \mathsf{rdf{:}type}|\mathsf{gmo{:}FundingAward},$
$\mathsf{gmo{:}providesAgentRole}|\mathsf{gmo{:}SponsorRole},$
$\mathsf{gmo{:}performedBy}|\mathsf{gmo{:}Organization}$

Generated Alignment:

$\mathsf{gbo{:}Award(x)} \land \mathsf{gbo{:}hasSponsor(x,z)} \land \mathsf{gbo{:}Organization(z)}$
$\rightarrow \mathsf{rdf{:}type}|\mathsf{gmo{:}FundingAward(x)} \land$
$\mathsf{gmo{:}providesAgentRole(x,y)} \land \mathsf{gmo{:}SponsorRole(y)} \land$
$\mathsf{gmo{:}performedBy(y,z)} \land \mathsf{gmo{:}Organization(z)}$

In this example, the association rule implies that in the GBO, the property gbo:hasSponsor with the domain type of gbo:Award and the range type of gbo:Organization has been used to "flatten" the complex structure in the GMO by short-cutting a property chain. Note that in this pattern, C_1 and any of the B_i may be omitted (in which case they are essentially $\top$).

5   Evaluation

In this section, we show the experimental results of our proposed alignment algorithm on the OAEI GeoLink benchmark and analyze the results in detail. The GeoLink benchmark [27] is composed of two ontologies in the geosciences domain. These two ontologies are both populated with 100% shared instance data collected from the real-world GeoLink knowledge base [3], in order to help

the evaluation of alignment algorithms depending on instance data.[7] The subset used for this study contains around 74k triples, which is suitable for applying association rule mining.

We originally planned to compare the performance of our system against pattern based system in [19], CANARD, and AMLC. However, the GeoLink benchmark is a property-oriented dataset which involves many object or data properties in the complex correspondences. As we discussed in Section 2, CANARD is currently limited to finding complex mappings that only involve classes. Even though pattern based system in [19] can generate property-based complex correspondences, like property chains, there are several rules that the system follows that largely limit its results, and it ends without finding any complex alignment on the GeoLink ontology pair. AMLC currently only works for the complex Conference benchmark [2, 9]. Therefore, there are no complex alignment systems against which we could compare the performance of our system. So in this paper we are limited to reporting the performance of our system against the reference alignment when it comes to the identification of complex alignment. Performance on the identification of simple alignment is compared against that of systems that participated in the OAEI 2018.

Because the systems we compare against are only capable of identifying simple correspondences, we present the results on the simple and complex portions of the overall alignment separately.[8] For simple correspondences, we use the traditional precision, recall and F-measure metrics, in order to compare against other simple alignment systems. However, in order to provide more insight into the underlying nature of the performance on complex correspondences, we take a slightly different approach. Semantic precision and recall, which compare correspondences based on their semantic meaning rather than their syntactic representation [8]. This is done by applying a reasoner to determine when one mapping is logically equivalent to another. Even though the semantic approaches solve an important problem for evaluating alignments with complex correspondences, they still have several limitations. One is that the reasoning takes a significant amount of time, particularly for large ontologies. Furthermore, such reasoning is not possible if the merged ontology is not in OWL DL. The GeoLink benchmark is one example of this case, since there are many correspondences involving an object property on one side and a data property on another side, which is not permissible in OWL DL. Instead, we utilize relaxed precision and recall [7]. More specifically, a correspondence consists of two aspects: the entities involved, and the relationship between them (e.g. equivalence, subsumption, disjunction). In order to assess performance on both of these aspects, we evaluate them separately. This roughly corresponds to the first and second subtasks described for some of the test sets within the complex track of the OAEI.[9] However, the types

Table 5. The Performance Comparison of Matchers on the Simple Alignment

Matcher	# of 1:1 Class Equiv.	# of 1:1 Class Subsum.	# of 1:1 Property Equiv.	Precision	Recall	F-measure
Reference Alignment	10	2	7	-	-	-
Our Results	10	0	5	**0.94**	**0.79**	**0.86**
CANARD [18]	9	0	3	0.67	0.63	0.64
DOME [12]	9	0	4	0.41	0.68	0.51
LogMap [14]	9	0	1	0.77	0.53	0.63
LogMapKG [14]	9	0	1	0.77	0.53	0.63
LogMapLt [14]	9	0	5	0.63	0.73	0.68
POMAP++ [16]	9	0	0	0.89	0.47	0.62
XMap [5]	9	0	0	0.39	0.47	0.43

of relationships we consider are limited to equivalence and subsumption rather than the arbitrary OWL constructs considered there.

5.1   Simple Alignment Evaluation

In the GeoLink benchmark, there are 19 simple mappings, including 10 class equivalences, 2 class subsumptions, and 7 property equivalences. Table 5 shows the simple mapping comparison between our algorithm and the matchers that participated in the OAEI 2018. We list the numbers of correctly identified mappings for each matcher and calculate the precision, recall, and f-measure. The confidence value for picking association rules is set to 0.3, since we find it generates the best performance for simple alignments.

Based on the results, our algorithm outperforms other systems on finding the simple mappings in this benchmark. We can argue that leveraging the instance data is a contributing factor, since our algorithm takes advantages of the instance data, while the other alignment systems do not use it. In addition, most traditional alignment systems focus on accurate detection only of 1:1 class equivalences, which limits their performance on this benchmark. The only 1:1 class equivalence that other alignment systems do not find, but our algorithm does, is gbo:Award$(x) \leftrightarrow$ gmo:FundingAward(x). This may also own to the populated instance data. The reason that our algorithm does not achieve 100% precision is that we mistakenly identify that gbo:PortCall is equivalent to gmo:Fix. The correct relationship should be subsumption. This relation can be easily refined by a semi-automated approach in the future.

5.2   Complex Alignment Evaluation

We set the confidence threshold to 1 when running the association rule mining algorithm in order to generate the results described in this section. This is a precision-oriented approach. However, these values can be tuned to fulfill various purposes of alignment systems.

As mentioned previously, in order to assess the quality of a mapping, there are two dimensions that we can look into. First, we can evaluate if the mapping contains the correct entities that should be involved based on the reference alignment. Another dimension is the relationship between the entities, like equivalence and subsumption. Based on this, we break the evaluation procedure down into two subtasks.

Table 6. Similarity for Relationship Identification

Found Relation	Correct Relation	Similarity	Comment
$=$	$=$	1	correct relation
$\subset$	$\subset$	1	correct relation
$\supset$	$\supset$	1	correct relation
$\subset$	$=$	0.8	return less information, but correct
$=$	$\supset$	0.8	return less information, but correct
$\supset$	$=$	0.6	return more information, but incorrect
$=$	$\subset$	0.6	return more information, but incorrect
$\subset$	$\supset$	0.3	incorrect relation
$\supset$	$\subset$	0.3	incorrect relation

(1) Entity Identification: For each entity in the source ontology, the alignment systems will be asked to list all of the entities that are related in some way in the target ontology. For example, referring to the example we used above,

$$\mathsf{Award}(x) \wedge \mathsf{hasSponsor}(x,z) \leftrightarrow \mathsf{FundingAward}(x) \wedge \mathsf{providesAgentRole}(x,y)$$
$$\wedge \, \mathsf{SponsorRole}(y) \wedge \mathsf{performedBy}(y,z),$$

the expected output from an alignment system is that `hasSponsor` in the GBO is related to `FundingAward`, `providesAgentRole`, `SponsorRole` and `performedBy` in the GMO and `Award` in the GBO. Based on the two lists of entities from the reference alignment and the matcher, precision, recall, and f-measure can be calculated.

(2) Relationship Identification: In terms of the example above, an alignment system needs to eventually determine that the relationship between the two sides is equivalence. Based on our algorithm, if there is only one association rule holding the information, we consider the relationship to be subsumption. If there are two association rules containing the information for both directions, an equivalence relationship is generated. At this stage, we do not further assess other complex relationships. Table 6 shows the different similarities for different situations. We slightly penalize differently for the situations in finding less information, but all the information returned is correct, and finding more information, but part of the information is incorrect. We do not penalize the incorrect relationship by giving a ZERO value because that would completely neglect the entity identification outputs without considering whether it is a reasonable result or a completely incorrect one. In order to generate the final results, we multiply the results from the entity identification by the penalty of the relations.[10] The formulas for computing the final results are as follows:

```
Relaxed_precision = Precision_entity × Relation_similarity
Relaxed_recall = Recall_entity × Relation_similarity
Relaxed_f-measure = F-measure_entity × Relation_similarity
```

---

[10] To be accurate, it could also have been better aggregated with other aggregation functions rather than multiplication [7]. But we would not focus on this question in this paper.

Table 7. Comparative Performance of Generating Complex Alignment

Matcher	1:n Property subsum.	m:n Complex equiv.	m:n Complex subsum.
Reference Alignment	5	26	17
Our Algorithm	3	15	7
Relaxed_Precision	0.60	0.90	0.53
Relaxed_Recall	0.36	0.36	0.16
Relaxed_F-measure	0.45	0.51	0.24

Table 7 shows the results of our algorithm. In total there are 48 complex mappings in the reference alignment. For 1:n property subsumption, our algorithm finds 3 mappings that fall into this category. For example, we find that the domain of gbo:hasSampleType is equivalent to gmo:PhysicalSample. However, the correct relationship should be subsumption. So, the final result should be penalized based on Table 6. For m:n complex equivalence, since our default confidence value for complex alignment is 1, the alignment that we found may miss some entities that should exist in the alignment. For example, referring to the example we use in the entity identification, the expected output from the alignment system is that the property hasSponsor in the GBO is related to FundingAward, providesAgentRole, SponsorRole, performedBy in the GMO and Award in the GBO. However, our algorithm misses one entity which is performedBy in the GMO. Errors such as this may of course be easily corrected by human interaction. For m:n complex subsumption, our algorithm does not generate the correct relationships for all the mappings we found. However, overall, our association rule-based algorithm can effectively come up with rather high quality simple and complex alignment automatically.[11]

6   Conclusion

Complex ontology alignment has been discussed for a long time, but relatively little work has been done to advance the state of the art in this field. In this paper, we proposed a complex ontology alignment algorithm based on association rule mining. Our algorithm takes advantage of instance data to mine frequent patterns, which show us which entities in one ontology are related to which entities in the other. Then we apply common simple and complex patterns to arrange these related entities into the formal alignment. We evaluated our system on the complex alignment benchmark from the OAEI and analyzed the results in detail to provide a better understanding of the challenges related to complex ontology alignment research.

There are still some limitations of our algorithm. First, our system relies on instance data for mining the association rules, which is not available for all ontology pairs. However, this could possibly be resolved with automated

---

[11] All the data and alignment that we use and generate can be accessed via the link http://tiny.cc/rojy4y. We utilize the Apache Spark frequent pattern mining library to generate association rules.

instance data generation to populate common instances into the ontologies or instance matching techniques. Second, we incorporate some common patterns that have been widely accepted in the ontology alignment community in this paper. This could be another limitation, since the set of mapping patterns in our system is likely not comprehensive. However, our algorithm is extensible, more patterns can be easily added in the future as the need arises. Third, it is possible that there are situations that the association rule would fail in term of finding simple property alignment. For example, if there are two properties *livesIn* and *bornIn* in source and target ontologies respectively, and the association rules would say if livesIn|Place, then bornIn|Place if they occur frequently. *livesIn* and *bornIn* would be considered as equivalent. In this case, there are many different methods that could be applied to improve the performance, like using lexical-based comparison or utilizing external knowledge base to annotate these entities. Fourth, we are collaborating with other benchmark and system developers to enable the comparison and prepare our alignment system to participate in the complex alignment track of the OAEI.

References

1. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules in large databases. In: VLDB'94, Proc. of 20th International Conference on Very Large Data Bases, September 12-15, 1994, Santiago de Chile, Chile. pp. 487–499 (1994)
2. Algergawy, A., et al.: Results of the ontology alignment evaluation initiative 2018. In: Proceedings of the 13th International Workshop on Ontology Matching, OM@ISWC 2018, Monterey, CA, USA, October 8, 2018. pp. 76–116 (2018)
3. Cheatham, M., et al.: The GeoLink knowledge graph. Big Earth Data **2**(2), 131–143 (2018)
4. David, J., et al.: Association rule ontology matching approach. Int. J. Semantic Web Inf. Syst. **3**(2), 27–49 (2007)
5. Djeddi, W.E., et al.: XMap: results for OAEI 2018. In: Proceedings of the 13th International Workshop on Ontology Matching, OM@ISWC 2018, Monterey, CA, USA, October 8, 2018. pp. 210–215 (2018)
6. Doan, A., et al.: Ontology matching: A machine learning approach. In: Handbook on Ontologies, pp. 385–404 (2004)
7. Ehrig, M., Euzenat, J.: Relaxed precision and recall for ontology matching. In: Integrating Ontologies '05, Proceedings of the K-CAP 2005 Workshop on Integrating Ontologies, Banff, Canada, October 2, 2005 (2005)
8. Euzenat, J.: Semantic precision and recall for ontology alignment evaluation. In: IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007. pp. 348–353 (2007)
9. Faria, D., et al.: Results of AML participation in OAEI 2018. In: Proceedings of the 13th International Workshop on Ontology Matching, OM@ISWC 2018, Monterey, CA, USA, October 8, 2018. pp. 125–131 (2018)

---

[12] https://daselab.cs.ksu.edu/projects/spex

10. Galárraga, L.A., et al.: AMIE: association rule mining under incomplete evidence in ontological knowledge bases. In: 22nd International World Wide Web Conference, WWW '13, Rio de Janeiro, Brazil, May 13-17, 2013. pp. 413–422 (2013)
11. Han, J., et al.: Mining frequent patterns without candidate generation: A frequent-pattern tree approach. Data Min. Knowl. Discov. **8**(1), 53–87 (2004)
12. Hertling, S., Paulheim, H.: DOME results for OAEI 2018. In: Proceedings of the 13th International Workshop on Ontology Matching, OM@ISWC 2018, Monterey, CA, USA, October 8, 2018. pp. 144–151 (2018)
13. Jiang, S., et al.: Ontology matching with knowledge rules. T. Large-Scale Data- and Knowledge-Centered Systems **28**, 75–95 (2016)
14. Jiménez-Ruiz, E., Grau, B.C., Cross, V.: LogMap family participation in the OAEI 2018. In: Proceedings of the 13th International Workshop on Ontology Matching, OM@ISWC 2018, Monterey, CA, USA, October 8, 2018. pp. 187–191 (2018)
15. Joshi, A.K., et al.: Logical linked data compression. In: The Semantic Web: Semantics and Big Data, 10th International Conference, ESWC 2013, Montpellier, France, May 26-30, 2013. Proceedings. pp. 170–184 (2013)
16. Laadhar, A., et al.: OAEI 2018 results of POMap++. In: Proceedings of the 13th International Workshop on Ontology Matching, OM@ISWC 2018, Monterey, CA, USA, October 8, 2018. pp. 192–196 (2018)
17. Piatetsky-Shapiro, G.: Discovery, analysis, and presentation of strong rules. In: Knowledge Discovery in Databases, pp. 229–248. AAAI/MIT Press (1991)
18. Portisch, J., Paulheim, H.: ALOD2Vec matcher. In: Proceedings of the 13th International Workshop on Ontology Matching, OM@ISWC 2018, Monterey, CA, USA, October 8, 2018. pp. 132–137 (2018)
19. Ritze, D., et al.: A pattern-based ontology matching approach for detecting complex correspondences. In: Proceedings of the 4th International Workshop on Ontology Matching (OM-2009), Chantilly, USA, October 25, 2009 (2009)
20. Ritze, D., et al.: Linguistic analysis for complex ontology matching. In: Proceedings of the 5th International Workshop on Ontology Matching (OM-2010), Shanghai, China, November 7, 2010 (2010)
21. Shvaiko, P., Euzenat, J.: Ontology matching: State of the art and future challenges. IEEE Trans. Knowl. Data Eng. **25**(1), 158–176 (2013)
22. Stumme, G., Maedche, A.: FCA-MERGE: bottom-up merging of ontologies. In: Proc. of the Seventeenth International Joint Conference on Artificial Intelligence, IJCAI 2001, Seattle, Washington, USA, August 4-10, 2001. pp. 225–234 (2001)
23. Thiéblin, É., et al.: CANARD complex matching system: results of the 2018 OAEI evaluation campaign. In: Proc. of the 13th Int. Workshop on Ontology Matching, OM@ISWC 2018, Monterey, CA, USA, Oct. 8, 2018. pp. 138–143 (2018)
24. Thiéblin, É., et al.: Complex matching based on competency questions for alignment: a first sketch. In: Proc. of the 13th International Workshop on Ontology Matching, OM@ISWC 2018, Monterey, CA, USA, Oct. 8, 2018. pp. 66–70 (2018)
25. Thiéblin, É., et al.: The first version of the OAEI complex alignment benchmark. In: Proc. of the ISWC 2018 Posters & Demonstrations, Industry and Blue Sky Ideas Tracks at (ISWC 2018), Monterey, USA, October 8th - to - 12th, 2018. (2018)
26. Thiéblin, É., et al.: Survey on complex ontology alignment. Semantic Web Journal (2019), to appear
27. Zhou, L., et al.: A complex alignment benchmark: Geolink dataset. In: The Semantic Web - ISWC 2018 - 17th International Semantic Web Conference, Monterey, CA, USA, October 8-12, 2018, Proceedings, Part II. pp. 273–288 (2018)

Towards Evaluating Complex Ontology Alignments

LU ZHOU[1], ELODIE THIÉBLIN[2], MICHELLE CHEATHAM[3], DANIEL FARIA[4],
CATIA PESQUITA[5], CASSIA TROJAHN[2], ONDŘEJ ZAMAZAL[6]

[1] *Data Semantics Laboratory, Kansas State University, Manhattan, USA*
e-mail:luzhou@ksu.edu
[2] *IRIT & Université de Toulouse 2 Jean Jaurès, Toulouse, France*
e-mails:elodie.thieblin@irit.fr, cassia.trojahn@irit.fr
[3] *Wright State University, Dayton, USA*
e-mail:michelle.cheatham@wright.edu
[4] *Instituto Gulbenkian de Ciência, Oeiras, Portugal*
e-mail:dfaria@igc.gulbenkian.pt
[5] *Lasige, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal*
e-mail:clpesquita@fc.ul.pt
[6] *Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic*
e-mail:ondrej.zamazal@vse.cz

Abstract

The development of semi-automated and automated ontology alignment techniques is an
important part of realizing the potential of the Semantic Web. Until very recently, most
existing work in this area was focused on finding simple (1:1) equivalence correspondences
between two ontologies. However, many real-world ontology pairs involve correspondences that
contain multiple entities from each ontology. These "complex" alignments pose a challenge for
existing evaluation approaches, which hinders the development of new systems capable of finding
such correspondences. This position paper surveys and analyzes the requirements for effective
evaluation of complex ontology alignments and assesses the degree to which these requirements
are met by existing approaches. It also provides a roadmap for future work on this topic taking into
consideration emerging community initiatives and major challenges that need to be addressed.

1   Introduction

Ontology alignments specify the relations that hold between entities in two or more ontologies.
Identifying these relations is critical for integrating data across the Semantic Web. The
development of automated and semi-automated techniques to establish alignments between
ontologies has been an active area of research since at least 2004; however, the vast majority
of existing alignment systems seek to identify relatively simple (1:1) equivalence and (more
rarely) subsumption relationships. While simple (1:1) relationships are limited in expressiveness
by linking single entities, complex matching approaches are able to generate correspondences
which better express the relationships between entities of different ontologies. Earlier works have
introduced the need for complex alignments (Maedche et al., 2002; Visser et al., 1997).

   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. These relations may
involve set operations such as union, intersection, disjunction, cardinality restrictions, and other
constraints. For example, two ontologies representing the domain of conference organization may
have the following relationship between their entities, which states that the class *AcceptedPaper*
in the source ontology is equivalent to the intersection of the class *Paper* with entities that
appear in the domain of the *acceptedBy* property: ⟨o1:AcceptedPaper, *intersectionOf*(o2:Paper,
minCardinality(1, o2:acceptedBy)), ≡, 1.0⟩.

These more complex relationships often make up half or more of the relations within an alignment, as discussed in Zhou et al. (2018). It is therefore an important research area for developers of alignment systems to consider. Unfortunately, the topic of complex ontology alignment has received relatively little attention thus far. Different complex matching approaches have emerged in the literature (Ritze et al., 2009, 2010; Jiang et al., 2016; Parundekar et al., 2010, 2012; Walshe et al., 2016), however most efforts on evaluation are still dedicated to the matching approaches dealing with simple alignments.

We posit that part of the reason for the lack of research on complex alignment systems is a lack of benchmarks that contain complex relations and a lack of appropriate metrics with which to evaluate the performance of systems on such benchmarks. The issue of the lack of ontology alignment benchmarks involving complex relationships is being addressed with the introduction of a new complex alignment track within the Ontology Alignment Evaluation Initiative[1] (OAEI), as described in (Thiéblin et al., 2018a). This paper begins work on the second issue.

The most common evaluation approach for ontology alignments is to perform an exact match comparison between the correspondences suggested by an alignment system and those in a reference alignment and to compute precision and recall based on this. This is a somewhat unforgiving approach. For example, in the case of the aforementioned conference ontologies, if an alignment system identified a relation between the ontologies of the form $\langle$o1:AcceptedPaper, $unionOf$(o2:Paper, $minCardinality$(1, o2:acceptedBy)), $\equiv$, 1.0$\rangle$, i.e. with union instead of intersection, it would be considered a false positive, and the correct relation would be considered a false negative. While the system should clearly be penalized for not producing the correct relation, considering this as completely incorrect lacks important nuance. For instance, this relation could be relatively easily corrected by a user in a semi-automated alignment system. Moreover, the alignment system developer would likely benefit from knowing how close the system came to generating the correct output in this case.

The primary goal of this position paper is to survey and analyze the requirements for effective evaluation of complex ontology alignments, assess the degree to which these requirements are met by existing approaches, and provide a roadmap for future work on this topic. We begin by discussing related work in Section 2. Section 3 presents the relevant background information, including a formal definition of complex alignments and their representation formats. A generic model of the ontology alignment evaluation process that highlights the choices implicit in implementing a complete evaluation strategy based on reference alignments is then presented in Section 4. The paper then surveys existing ontology alignment evaluation metrics and analyzes their strengths and weaknesses with respect to evaluation of complex alignments when a reference alignment is available. Section 5 overviews the alternative evaluation measures applicable in the absence of reference alignments. The paper continues with a discussion of the gaps that exist between the current state of the art and what is needed for effective evaluation of complex alignments following with its feasibility analysis in Section 6 and then argues about necessary future work to fill in those gaps (Section 7).

2   Related Work

Early studies have introduced the need for complex ontology alignments (Visser et al., 1997; Maedche et al., 2002) and different approaches for generating such alignments have been proposed in the literature since. These approaches rely on diverse methods such as correspondence patterns (Ritze et al., 2009, 2010), knowledge-rules (Jiang et al., 2016), statistical methods (Parundekar et al., 2010, 2012; Walshe et al., 2016), or genetic programming Nunes et al. (2011) and path-finding algorithms (Qin et al., 2007). While most work on complex ontology matching has been dedicated to the development of complex matching approaches, automatic support for evaluating complex approaches has still not been extensively addressed in the literature.

---

[1] http://oaei.ontologymatching.org/2018/complex/

The evaluation of most existing approaches has been done by manually calculating the precision of the alignments generated by the systems (Ritze et al., 2009, 2010; Parundekar et al., 2012; Walshe et al., 2016). In order to be able to measure recall, specific datasets have been constructed. The approach of Parundekar et al. (2012) estimated their recall based on the recurring pattern between DBpedia and Geonames:$\exists dbpedia{:}country.\{theCountryInstance\} \equiv \exists geonames{:}countryCode.\{theCountryCode\}$ where $theCountryInstance$ is a country instance of DBpedia such as $dbpedia{:}Spain$ and $theCountryCode$ is a country code such as "ES". They estimated the number of occurrences of this pattern between these ontologies and calculated the recall based on this estimation. 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 50 classes with known *Class-by-attribute-value* (a correspondence pattern) correspondences with DBpedia and 50 classes with no known correspondences with DBpedia. Both ontologies are populated with the same instances.

As described by Thiéblin et al. (2018b), the metrics of *accuracy* and *top−x accuracy* have been also applied in evaluation settings in which the number of correspondences is predefined, e.g., there is one correspondence for each entity of the target schema/ontology. The accuracy is then the percentage of predefined questions having a correct answer. A "question" in this context could be a source entity to be matched and the "answers" the correspondences having this entity as source member. Some approaches output various answers for each question, e.g., a ranked list of correspondences for each source entity. In this case the top-x accuracy is the percentage of questions whose correct answer is in the top-x answers to the question. For example, top-3 accuracy is the fraction of source entities for which the correct correspondence is in the three best correspondences output by the system. Alternatively, the approach in Thiéblin et al. (2017), to evaluate complex correspondences between agronomic ontologies is based on manually comparing the results of the reference queries and queries automatically rewritten with the help of the complex alignments.

More recently, complex evaluation was introduced in the 2018 Ontology Alignment Evaluation Initiative (Thiéblin et al., 2018a). The track consisted of four datasets from a variety of domains: conference organization, hydrography, geoscience, and plant taxonomies. Each dataset was evaluated in a different way. For the conference dataset, precision and recall of the system's alignment were manually calculated based on exact match with respect to the reference alignment. For the plant taxonomy 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. Finally, for the hydrography and geoscience datasets, the evaluation plan was to divide the alignment task into three subtasks and assess performance on each one separately: 1) given an entity from the source ontology, identify all related entities in the source and target ontology; 2) given an entity in the source ontology and the set of related entities, identify the logical relation that holds between them; 3) identify the full complex correspondences. The first subtask was evaluated based on precision and recall with respect to exact match against the reference alignment and the latter two were evaluated using semantic precision and recall.

The evaluation plan for the hydrography and geoscience datasets was not really put to the test in 2018, however, because no alignment systems were capable of finding complex correspondences across these ontologies. The manual nature of the evaluation for the conference organization and plant taxonomy datasets was feasible because only two alignment systems, AMLC and CANARD, were able to generate any complex relations for those datasets; however, there are obvious limitations to a manual approach, both during system development (system developers

cannot quickly test modifications to their system to assess whether or not they improve the performance) and evaluation (the time taken is prohibitive for the OAEI track organizers if many systems participate). Additionally, manual evaluation might introduce bias or inconsistencies into the performance assessment.

3   Background

The examples of both simple and complex correspondences provided throughout this paper are based on the OntoFarm ontologies from the conference domain (Šváb et al., 2005; Zamazal and Svátek, 2017). Complex examples are based on the complex version of this dataset, which consists of alignments between all combinations of three of the OntoFarm ontologies, *ekaw*, *cmt* and *conference*, created by domain experts from three universities who were all familiar with ontology alignment (Thiéblin et al., 2018a).

3.1   Complex ontology alignment

We define ontology matching as the process of generating an alignment A between two ontologies: a source ontology O and a target ontology O', as in (Euzenat and Shvaiko, 2013). A is directional, denoted $A_{O \to O'}$, and is a set of correspondences $\langle e, e', r, s \rangle$. Each correspondence contains a relation r (e.g., equivalence ($\equiv$), subsumption ($\leq$, $\geq$)) between two members e and e', and s expresses the strength or confidence (in [0;1]) of this correspondence. Each member can be a single ontology entity (class, object property, data property, individual, value) of respectively O and O' or a more complex construction that is composed of some entities using constructors or transformation functions.

We consider two types of correspondences depending on the type of their members (Thiéblin et al., 2018; Zhou et al., 2018).

- a correspondence is **simple** if both e and e' are single entities (represented as IRIs):
  $\langle$ *ekaw:Paper, cmt:Paper, $\equiv$, 1* $\rangle$

- a correspondence is **complex** if at least one of e or e' involves a constructor or a transformation function:
  $\langle$ *ekaw:AcceptedPaper, someValuesFrom(cmt:hasDecision, cmt:Acceptance), $\equiv$, 1.0* $\rangle$
  $\langle$ *concatenation(edas:hasFirstName, " ", edas:hasLastName), cmt:name, $\to$, 1* $\rangle$

A simple correspondence is usually noted (1:1), and a complex correspondence can be (1:n) if its source member is a single entity, (m:1) if its target member is a single entity or (m:n) if neither of the members are single entities. Note that these cardinalities refer to the number of entities from the source and target ontologies *in a single correspondence*, not across all correspondences within the alignment. For example, a (1:n) correspondence means that one source entity is related to n target entities via a relationship expressed in a single complex correspondence, not that the same source entity is mapped in a (1:1) manner to n different target entities.

Because relations between instances are generally (1:1) in nature (e.g. *sameAs*, *differentFrom*), complex correspondences predominantly involve entities from the TBox of the ontologies rather than the ABox.

3.2   Representation formats

A general understanding of formats used to express complex correspondences between entities is necessary to comprehend some of the metrics designed to measure the similarity between such correspondences. This section provides an overview of common approaches.

The $\langle e, e', r, n \rangle$ tuples making up a simple alignment are most often encoded using RDF in a representation format commonly referred to as the Alignment API format, which was introduced in Euzenat (2004). This API is used by the OAEI and has wide adoption within the

ontology alignment research community. Version 4 of the Alignment API, described in David et al. (2011), also contains a representation format for complex correspondences, known as the Expressive and Declarative Ontology Alignment Language (EDOAL) (see Euzenat et al. (2007)). While in the simple alignment format e and e' are single IRIs, in EDOAL these are expressions involving classes and properties that can be combined using intersection, union, disjunction and composition operators and/or restricted using constraints on attributes, such as domain, range, cardinality or value restrictions. The EDOAL representation for the correspondence $\langle$cmt:ProgramCommitteeMember, $someValuesFrom$(conference:was_a_member_of, conference:Program_committee), $\equiv$, 1.0$\rangle$ is shown below.

```
<map>
    <Cell>
        <entity1>
            <edoal:Class rdf:about="&cmt;ProgramCommitteeMember"/>
        </entity1>
        <entity2>
            <edoal:AttributeDomainRestriction>
                <edoal:onAttribute>
                    <edoal:Relation rdf:about="&conference;was_a_member_of"/>
                </edoal:onAttribute>
                <edoal:exists>
                    <edoal:Class rdf:about="&conference;Program_committee"/>
                </edoal:exists>
            </edoal:AttributeDomainRestriction>
        </entity2>
        <measure rdf:datatype="&xsd;float">1.</measure>
        <relation>Equivalence</relation>
    </Cell>
</map>
```

While its general acceptance and associated toolset make EDOAL a convenient choice for representing complex relationships between ontologies, there are some limitations to this approach. For instance, while EDOAL supports a limited set of transformations, this aspect of the language is somewhat immature. Another issue is that in some cases a concept that is represented as a class in one ontology is modeled as an instance in another ontology (or, one may need to restrict a set of possible instance values involved in a relationship based on their type). This is similar to the OWL concept known as punning, but it is not currently possible in EDOAL. Finally, some relations may be modelled as object properties in one ontology and data properties in another. This occurs frequently when one ontology author has used a "strings as things" approach while the other has instead created instances. EDOAL does not allow one to specify relationships between object and data properties. Indeed, this is not possible in OWL DL either, though it is permissible in OWL Full.

EDOAL is the most common representation format for complex alignments, but they can be represented in a variety of different ways. For example, OWL can be used directly. This has the benefit of existing tool support for creating, modifying, and reasoning with the alignment, as well as merging ontologies based on it, but it limits the possible complex correspondences to those expressible in OWL (or OWL DL if reasoning is desired), which in particular makes it difficult to encode relationships that involve transformation functions. Another option is to use logical rules following one of a range of different syntaxes, which has the benefit of being generally easier for humans to parse from text than either EDOAL or OWL, but there is a lack of tool support for direct use of alignments expressed in this way. Other possibilities for complex correspondence representation include using a dedicated vocabulary or representing them as queries. As described

in (Xiao et al., 2018), in the area of OBDA (Ontology-Based Data Access) the R2RML format, a W3C standard, has been extended in many different ways, including for this purpose. For a more complete survey on the representation of ontology alignments, we refer the reader to the one presented in (Scharffe, 2009).

4 Evaluation with a reference alignment

The evaluation of ontology alignments is often performed with respect to a reference alignment, as is the case in most of the OAEI tracks. Usually, this evaluation relies on the traditional information retrieval evaluation metrics of precision and recall, and only contemplates correspondences that are exactly equal between the evaluated and reference alignments. However, as we will overview in this section, several alternative approaches to score inexact matches between the evaluated and reference alignments have been proposed.

4.1 Generic evaluation process

The generic process of evaluating an ontology alignment A_{eval} using a reference alignment A_{ref} can be decomposed into four steps, as schematized in Figure 1: anchor selection, correspondence comparison, scoring, and aggregation. Note that these steps are not independent, and in fact, much existing work on the topic of ontology alignment evaluation conflates the latter three steps (Ehrig and Euzenat, 2005; Euzenat, 2007). In practice the correspondence comparison approach selected and corresponding scoring scheme have ramifications throughout the evaluation process.
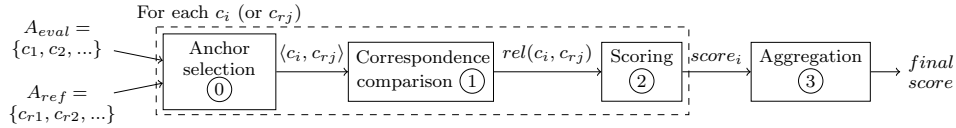

Figure 1 Evaluation process of the alignment A_{eval} with the reference alignment A_{ref}.

In the **anchor selection** step, the set of correspondences c_{rj} from the reference alignment A_{ref} that have to be compared with each correspondence c_i from the evaluated alignment A_{eval} (or vice versa) is computed. This selection depends on the correspondence comparison approach adopted. In the traditional evaluation where only exactly matching correspondences are to be scored, only these need be selected in this step. But if related correspondences are also contemplated, then each evaluated correspondence may have several such correspondences in the reference alignment, and all of them will need to be compared unless it is evident *a priori* which is the most similar (e.g. if there is an equivalent correspondence).

In the **correspondence comparison** step, for each pair of correspondences $\langle c_i, c_{rj} \rangle$, where $c_i = \langle e_i, e_i', r_i, s_i \rangle$ and $c_{rj} = \langle e_{rj}, e_{rj}', r_{rj}, s_{rj} \rangle$, a relation $rel(c_i, c_{rj})$ between c_i and c_{rj} is computed. $rel(c_i, c_{rj})$ can be decomposed into the relations between the elements of c_i and c_{rj}:

$$rel(c_i, c_{rj}) = \begin{cases} rel(e_i, e_{rj}) \\ rel(e_i', e_{rj}') \\ rel(r_i, r_{rj}) \\ rel(s_i, s_{rj}) \end{cases}$$

As we will overview in Section 4.2, relations between entities include syntactic equivalence, semantic equivalence, and semantic relatedness; the relation between the correspondence relations includes equivalence and relatedness; and the relation between the confidence scores, when considered, is typically numerical similarity. In the traditional evaluation, the relation between

the confidence scores is ignored, and correspondences are considered equivalent if both the entities and the correspondence relation are syntactically equivalent.

In the **scoring** step, a scoring function is applied to the relation $rel(c_i, c_{rj})$ between c_i and c_{rj}. This is usually done by applying the scoring scheme associated with the correspondence comparison approach to score the relations between each element in the correspondences, then multiplying these scores. $score_i$ is the result of this scoring function. In the traditional evaluation, equivalent correspondences are treated as true positives and scored 1, and no other correspondences are scored.

In the **aggregation** step, the scores are aggregated over the whole alignment to produce the $final\ score$. In the traditional evaluation, this aggregation means computing precision and recall by tallying the true positives and dividing by the number of correspondences in the evaluated and reference alignments, respectively. Correspondences in the evaluated alignment that are not in the reference alignment are false positives, and those in the latter and not in the former are false negatives. In cases where inexact correspondence matches are contemplated, then the aggregation must also include the selection of which correspondence pairs to score, as each evaluated correspondence may have a non-zero score when compared with several reference correspondences. Intuitively, it makes sense to select only the most similar reference correspondence for each evaluated correspondence (which in the trivial case would be an exact match). However, the fact that multiple evaluated correspondences may have the same reference correspondence as the most similar makes this selection less straightforward. There is some argument to enforcing that each correspondence from both the reference and evaluated alignments be considered only once in the aggregation, but this may not make sense when neither evaluated correspondence is related to any other reference correspondence.

While this generic evaluation workflow is valid for both simple and complex alignments, it is challenging to apply it to complex alignments due to the fact that complex correspondences feature expressions of arbitrary complexity with a wide range of constructs, rather than singular entities. Thus, one cannot simply compare URIs of the mapped entities between two correspondences and check for identity or a semantic relation between them, as there are additional layers to contemplate when comparing correspondences. This affects the anchor selection step as it may not be trivial to determine that two complex correspondences are related in a manner that is computationally more efficient than the worse-case scenario of skipping anchor selection and making the full pairwise comparison of all correspondences in the subsequent step. It also affects the correspondence comparison step, as determining the relation between complex entities requires comparing all the singular entities they list, as well as the expressions in which they are listed, likely in recursive fashion, as there is no theoretical limit to the nesting of expressions within expressions. Furthermore, there are cases where one might want to consider making a joint evaluation of two or more correspondences against a single reference correspondence, which complicates both the correspondence comparison and the aggregation step. For example, consider this reference correspondence from the Conference test set in the complex alignment track of the OAEI:

$\langle intersectionOf$(ekaw:Paper_Author, $complementOf(someValuesFrom$(ekaw:reviewerOfPaper, ekaw:Paper))), $intersectionOf$(conference:Regular_Author, $complementOf$(conference:Reviewer)), $\equiv$, 1.0 $\rangle$

Consider also the following two correspondences produced by an alignment system:

$\langle$ekaw:Paper_Author, conference:Regular_Author, $\equiv$, 1.0 $\rangle$

$\langle someValuesFrom$(ekaw:reviewerOfPaper, ekaw:Paper), conference:Reviewer, $\equiv$, 1.0 $\rangle$

In this scenario, if neither of the system correspondences were in the reference alignment, it is arguable that both should be scored against the reference correspondence together, as the latter can be logically derived from them ({A $\equiv$ A'; B $\equiv$ B'} $\Rightarrow$ A $\cap$!B $\equiv$ A' $\cap$!B').

An additional challenge to the evaluation of complex alignments is that, in practice, there is a greater variety of correspondence relationships, since most simple ontology alignment

benchmarks consist entirely of equivalence relations. This aggravates the difficulty in comparing correspondences, as the relation may factor into how two correspondences are related. Picking up on our example above, consider the following correspondence produced by a matching system:
$\langle intersectionOf$(ekaw:Paper_Author, $complementOf(someValuesFrom$(ekaw:reviewerOfPaper, ekaw:Paper))), conference:Regular_Author, $\leq$, 1.0 $\rangle$

This correspondence is logically derived from the reference correspondence, and thus formally correct (if less specific than desired) whereas it would not be correct if the relation were equivalence.

Finally, the several layers involved in comparing complex correspondences make it desirable to use comparison approaches that generate more nuanced similarity scores than the simple all-or-nothing approach traditionally used in alignment evaluation. This means that there will likely be more correspondence comparisons involved in evaluating complex alignments, and the aggregation step will be less straightforward.

4.2 Existing approaches for correspondence comparison

As we detailed in the previous section, the correspondence comparison approach affects the whole evaluation workflow, as it determines which correspondences are selected as anchors, how they are compared and how they should be scored, as well as how they can be aggregated. Due to this central importance, and to the fact that they are the characterizing factor of different forms of alignment evaluation, this section is devoted to surveying existing approaches for correspondence comparison and discussing their application to complex alignments. This is not an exhaustive survey, but rather an attempt to provide insights on the strengths and weaknesses of each type of approach when used to evaluate complex ontology alignments.

The following example will be used throughout this section. Correspondences in the reference alignment (R):

1. $\langle$cmt:Author, conference:Regular_Author, $\equiv$, 1.0$\rangle$

2. $\langle$cmt:ProgramCommitteeMember, $someValuesFrom$(conference:was_a_member_of, conference:Program_committee), $\equiv$, 1.0$\rangle$

3. $\langle$cmt:User, $unionOf$(conference:Regular_Author, conference:Reviewer), $\geq$, 1.0$\rangle$

4. $\langle$cmt:AuthorNotReviewer,$intersectionOf$(conference:Regular_Author, $complementOf$(conference:Reviewer)), $\equiv$, 1.0$\rangle$

Correspondences generated by alignment system 1 (S1):

1. $\langle$cmt:Author, conference:Regular_Author, $\leq$, 1.0$\rangle$

2. $\langle$cmt:ProgramCommitteeMember, $minCardinality$(1, conference:was_a_member_of, conference:Program_committee), $\equiv$, 1.0$\rangle$

3. $\langle$cmt:User, conference:Regular_Author, $\geq$, 1.0$\rangle$

4. $\langle$cmt:User, conference:Reviewer, $\geq$, 1.0$\rangle$

5. $\langle$cmt:AuthorNotReviewer, $unionOf$(conference:Regular_Author, conference:Reviewer), $\equiv$, 1.0$\rangle$

Correspondences generated by alignment system 2 (S2):

1. $\langle$cmt:Author, conference:Contribution_1th-Author, $\equiv$, 1.0$\rangle$ (Note that conference:Contribution_1th-Author is a subclass of conference:Regular_Author)

2. $\langle$cmt:AuthorNotReviewer,$intersectionOf$(conference:Conference_participant,
   $complementOf$(conference:Committee_member)), $\equiv$, 1.0$\rangle$

We will use the notation $c_i = <e_i, e_i', r_i, s_i>$ to refer to any correspondence generated by an alignment system, and $c_{rj} = <e_{rj}, e_{rj}', r_{rj}, s_{rj}>$ to refer to any reference correspondence.

4.2.1 Syntactic

Syntactic approaches to alignment evaluation compare the elements of two correspondences based on their syntactic description (i.e., the URIs of entities, or the identifiers of correspondence relations or complex expressions). This includes the traditional evaluation approach of scoring only exact matches, where a correspondence is scored 1 if both of its entities and its relation are syntactically equivalent to the reference correspondence (i.e., $e_i \equiv e_{rj}$, $e_i' \equiv e_{rj}'$, and $r_i \equiv r_{rj}$) and scored 0 otherwise.

This exact match approach is used to compare correspondences in most existing work on ontology alignment, including in most OAEI tracks and in the majority of ontology alignment papers. In fact, if papers do not explicitly state what evaluation approach they are using, it is assumed to be exact match. Thus, this approach has the advantage of being both simple and widely used. It is often possible to compare the results of an alignment system to previous work based on this approach by referring to the original papers rather than re-running the experiments. Furthermore, available computational tools for handling ontology alignments, such as the Alignment API, usually contain evaluation facilities based on exact match and do not require users to write additional code.

However, this approach is unforgiving in that it treats as incorrect correspondences that, while not listed in the reference alignment, can be logically derived from it (or even equivalent to it), and thus are formally correct. Furthermore, it does not distinguish between correspondences that are formally incorrect but closely related to correct correspondences, and those that are completely incorrect. Referring to the example above, the first correspondence in the reference alignment is an equality, but the first alignment system identifies the relation between the same entities as subsumption. This is considered completely incorrect under the exact match approach, even though it is formally correct (if imprecise) and may be a useful result in some applications of the alignment, such as query answering. Meanwhile, the second alignment system correctly identified the equality relationship for cmt:Author, but rather than conference:Regular_Author, it specified conference:Contribution_1th-Author, a subclass of conference:Regular_Author, as the equivalent entity. This is formally incorrect, but the correct correspondence can be inferred from it, so it is only partially incorrect, as the correspondence holds true for a subset of cmt:Author. If the alignment system had specified conference:Chair, which has no relation at all to conference:Regular_Author, then the correspondence would be fully incorrect. The case of the second reference correspondence is even more grave, as the first system identified a correspondence that is syntactically different but logically equivalent and thus formally correct. Under a syntactic approach, this correspondence would result in both a false positive and a false negative (as the syntactically correct correspondence is missing) whereas it should clearly result in a true positive. Regarding the third correspondence from the reference alignment, the first alignment system states that cmt:User is related to both conference:Regular_Author and conference:Reviewer, yet this is also treated as incorrect (specifically, as one false negative and two false positives), because the system specified each relation separately instead of as a union. Finally, both alignment systems generate relations that are somewhat similar to the fourth one from the reference alignment. The first system has the correct entities but incorrect expressions while the second has the expressions correct but incorrect entities. Both of these are treated as completely incorrect.

An alternative to the traditional binary syntactic evaluation is the weighted syntactic 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). In this approach, which is also implemented in the Alignment API, the true positive count is replaced by the sum of the products of confidence scores $s_i * s_{rj}$, and the false positive and false negative counts were replaced by the sum of differences of confidence scores $|s_i - s_{rj}|$ respectively for $s_i < s_{rj}$ and $s_i > s_{rj}$. 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, albeit relying on a vector representation of the ontology alignments, was also proposed by (Sagi and Gal, 2018).

4.2.2 Rule-based semantic and reasoning-based semantic

Semantic approaches compare correspondences based on their semantic meaning rather than their syntactic representation. This is done by looking at the correspondence within the context of the ontologies and determining whether they are semantically related. If they are "closely" related, but not equivalent, they are typically scored in $(0; 1]$, depending on the scoring scheme of the approach.

An example of such an approach is the relaxed precision and recall metric proposed by (Ehrig and Euzenat, 2005), which defines different similarity functions for the various elements of a correspondence, depending on whether precision or recall is to be computed. It scores the similarity between two entities, e_i and e_{rj} according to:

$$entity\ prec\ similarity = \begin{cases} 1 & \text{if } e_i \leq e_{rj} \\ 0.5 & \text{if } e_i > e_{rj} \\ 0 & \text{otherwise} \end{cases} \qquad entity\ rec\ similarity = \begin{cases} 1 & \text{if } e_i \geq e_{rj} \\ 0.5 & \text{if } e_i < e_{rj} \\ 0 & \text{otherwise} \end{cases}$$

where $>$ and $<$ stand for direct sub- or super-classes/properties only. The similarity between two relations is defined only for the case where r_i is $\equiv$ (as most matching systems tend to produce only equivalence correspondences) and depends only on r_{rj} according to:

$$relation\ prec\ similarity = \begin{cases} 1 & \text{if } \equiv \text{ or } < \\ 0.5 & \text{if } > \\ 0 & \text{otherwise} \end{cases} \qquad relation\ rec\ similarity = \begin{cases} 1 & \text{if } \equiv \text{ or } > \\ 0.5 & \text{if } < \\ 0 & \text{otherwise} \end{cases}$$

Finally, the similarity between confidence scores, s_i and s_{rj}, is scored according to:

$$score\ similarity = 1 - |s_i - s_{rj}|$$

This approach thus aims to reward correspondences that are semantically close to the correct correspondence from the perspective of query answering. Namely, in the case of precision, it does not penalize at all correspondences that are narrower than the correct correspondence (but implied by it) since these would result in missing but only correct query results (full precision). Likewise, in the case of recall, it does not penalize correspondences that are broader than the correct correspondence (and imply it) since these would result in no missing results but some incorrect ones (full recall).

Another semantic approach also proposed by (Ehrig and Euzenat, 2005) focuses on the perspective of alignment validation rather than query answering, and seeks to account for the effort it would take a human reviewer to correct an erroneous correspondence that is semantically close to the correct one. This approach can be considered a simple edit-distance approach, as it attributes a cost to each edition necessary for converting an incorrect correspondence to a correct one. Under this approach, the similarity between entities is given by:

$$entity\ effort\ similarity = \begin{cases} 1 & \text{if } e_i \equiv e_{rj} \\ 0.6 & \text{if } e_i < e_{rj} \\ 0.4 & \text{if } e_i > e_{rj} \\ 0 & \text{otherwise} \end{cases}$$

where again, $>$ and $<$ stand for direct sub- or super-classes/properties only. The rationale behind attributing a different similarity to sub- and super-entities is that typically ontology entities are expected to have more sub- than super-entities, and thus correcting to a broader entity requires less effort than correcting to a narrower entity. The similarity between relations is 1 if the relations are the same and 0.5 if they are different, under the rationale that correcting the relation is fairly trivial even if the relation predicted by the matching system is completely off.

Another semantic approach is the semantic precision and recall proposed by (Euzenat, 2007). Under this approach, a reasoner is employed to count the number of correspondences suggested by the alignment system that are entailed by a merged ontology consisting of the source and target ontologies and the reference alignment. This count is then divided by the number of correspondences in the proposed alignment to produce the system's precision. Analogously, recall is computed by counting the number of relations in the reference alignment that are entailed by a merged ontology consisting of the source and target ontologies and the proposed alignment, then dividing this by the number of correspondences of the reference alignment.

Regardless of their scope and implementation, semantic approaches tend to mitigate some of the issues we reported for syntactic approaches, since they account for correspondences that are semantically close to the correct ones. Here we discuss the score produced by each metric for the example alignments presented at the start of Section 4.2. This information is summarized in Table 1. The first alignment system identified the relation as subsumption rather than equivalence. This would score 0 under a syntactic approach, but would score respectively 1 and 0.5 in relaxed precision and recall, 0.5 in effort similarity, and 1 and 0 in semantic precision and recall (as subsumption is entailed by but does not entail equivalence). Likewise, the correspondence proposed by the second system, in which the entity from the target ontology was a subclass of the correct entity, would score 0 under a syntactic approach, but respectively 1 and 0.5 in relaxed precision and recall, and 0.6 in effort similarity. In this case it would also score 0 under semantic precision and recall, as equivalence to a class neither entails nor is entailed by equivalence to its superclass. The contrary happens in the second reference correspondence in the example, in which the first system produced a logically equivalent correspondence. In this case, the system's correspondence would score 1 under semantic precision and recall, but would still be scored 0 under relaxed precision and recall as well as effort similarity, as these rule-based approaches have no provisions for complex alignment expressions and thus cannot detect that these correspondences are logically equivalent. Similarly, the fifth correspondence of the first system, which differs from the fourth correspondence of the reference alignment only in that a union was used instead of an intersection, would score respectively 1 and 0 in semantic precision and recall (as equivalence to the union entails equivalence to the intersection but not the other way around) but also be scored 0 under the other approaches, for the same reason as in the previous case. From the perspective of alignment validation, such a correspondence should be fairly trivial to correct, and thus should have a non-zero score. In the case of the third reference correspondence, for which system one predicts two related correspondences (3 and 4), these would also be scored respectively 1 and 0 under semantic precision and recall (as superclass of the union entails superclass of each element in the union but not the other way around), but again 0 under the other approaches, as they have no provision for comparing correspondences other than on a one-to-one basis.

In summary, the main limitations of rule-based semantic approaches with respect to complex alignments are that no such approach has been proposed that encompasses the range of

Issue	Related Items	Exact Match	Relaxed Prec.	Relaxed Rec.	Effort	Sem. Prec.	Sem. Rec.
Mismatched relation	R1,S1$_1$	0	1	0.5	0.5	1	0
Subclass rather than exact class match	R1,S2$_1$	0	1	0.5	0.6	0	0
Logically equivalent	R2,S1$_2$	0	0	0	0	1	1
Correct but expressed as multiple correspondences	R3, S1$_3$, $S1_4$	0	0	0	0	1	0
Correct entities; incorrect construction	R4,S1$_5$	0	0	0	0	1	0
Correct construction; incorrect entities	R4,S2$_2$	0	0	0	0	0	0

Table 1 Scores of the surveyed metrics on the sample alignments.

expressions possible in these alignments, and that they do not contemplate joint correspondence evaluation in the cases where a correspondence is decomposed into several related ones. Furthermore, proposed approaches are coarse in granularity, and only distinguish between identical entities, direct sub-/super-entities, and all other cases. They do not account for cases of other relations, such as indirect sub-/super-entities, even though (Ehrig and Euzenat, 2005) did suggest that more a granular approach could do so by explicitly taking into the account the edge-distance between entities in the similarity function.

By contrast, reasoning-based semantic approaches do account for all complex expressions that can be encoded in OWL, and also handle cases of correspondence decomposition well. However, they only score correspondences that are logically entailed, ignoring those that are semantically related but not entailed. Thus, in cases that can be handled by both reasoning-based and rule-based approaches, reasoning-based approaches are stricter in their assessment of performance for purposes such as query answering or alignment validation. Furthermore, reasoning-based approaches are computationally more complex than rule-based approaches, and may not be applicable in practice to very large ontologies, as reasoning over these is still a computational challenge. Finally, reasoning is only possible if the merged ontology is in OWL DL, which may not be the case in complex alignments even if the original ontologies are (for example if a correspondence is made between an object property and a datatype property).

4.2.3 Instance-based

Instance-based approaches compare two correspondences between ontology classes based on the overlap between their sets of instances. In (Isaac et al., 2007), instance-based similarity measures are divided into two primary categories: traditional set similarity metrics and information-theoretic measures. Set similarity is most often computed based on the Jaccard index, which is the ratio of instances that belong to both the source and target classes to the number of instances belonging to either the source or target classes. Information-theory measures reflect the degree to which knowledge of an instance's categorization via e of a correspondence provides knowledge about the appropriateness of the e' categorization. Examples include point-wise mutual information, log likelihood ratio, and information gain. More recent work has proposed instance-based metrics based on locality-sensitive hashing (Duan et al., 2012) an on Cohen's kappa coefficient (Kirsten et al., 2007).

Instance-based correspondence comparisons are powerful in that they directly correspond to the underlying definition of ontological entities as sets of instances that are related in some way. However, the applicability of such metrics is limited to the evaluation of class correspondences, and only in cases in which common instances exist in both ontologies. These common instances can either be the same individuals (with identical URIs) or individuals with different URIs that have been declared identical through the use of a co-reference resolution procedure (though this procedure can of course introduce errors that would negatively impact the alignment evaluation). Furthermore, even if dual-typed instance data exists, there may be particular valid complex correspondences for which few or no instances are available, which can compromise the evaluation (even though some, metrics such as the log likelihood ratio and the modified version of the Jaccard metric described in (Isaac et al., 2007), handle sparse data better than others). A solution for handling sparse data is to synthetically generate additional instance data, as described in (Schopman et al., 2012), but this has the potential of biasing the evaluation and no assurance of covering particular complex correspondences better.

5 Evaluation without a reference alignment

Constructing reference alignments is a time-consuming task that requires the involvement of domain experts. In the absence of time, an alternative evaluation strategy can be the manual validation of sample alignments, as detailed in (Van Hage et al., 2007), although this still requires significant involvement of domain experts. Alternative approaches consider the generation of natural language questions to support end-users in the validation task (Abacha and Zweigenbaum, 2014) or validation of correspondences in a semi-automatic way (Serpeloni et al., 2011).

In the absence of both reference alignments and domain experts, there are two families of approaches to ontology alignment evaluation: one that uses quality metrics to assess the logical soundness of the alignment (Meilicke and Stuckenschmidt, 2008; Solimando et al., 2017), and another that focuses on the suitability of the alignment for a specific task or application (Isaac et al., 2008; Hollink et al., 2008; Solimando et al., 2014). In this section, we discuss how complex alignments can be evaluated using these strategies.

5.1 Alignment quality metrics

The union of two ontologies through an alignment can lead to logical errors such as unsatisfiable classes (i.e., classes than can only be interpreted as empty sets) even if both ontologies were originally logically sound. In such cases, the merged ontology is said to be incoherent, and by extension, so is the ontology alignment. Since, for many applications, incoherence would cause problems, there are several approaches to measure ontology incoherence (Qi and Hunter, 2007). Derived from these, (Meilicke and Stuckenschmidt, 2008) proposed two measures to assess an alignment's quality based on its logical coherence: one based on counting unsatisfiable classes; and another, named maximum cardinality measure (degree of incoherence), based on the minimum number of correspondences that must be removed to obtain a coherent merged ontology. Additionally, the authors proposed a variant of the latter measure that considers the confidence scores of the correspondences and measures the minimum loss of total confidence required for coherence, called the maximum trust measure. Interestingly, they reported that the maximum cardinality measure can be used to compute a strict upper bound of precision (Meilicke and Stuckenschmidt, 2008).

Also on the topic of logical soundness, Jiménez-Ruiz et al. (2011) proposed three principles for ontology alignments: consistency, conservativity, and locality. The consistency principle states that correspondences should not lead to unsatisfiable classes in the merged ontology. This is a bit of a misnomer, as the principle pertains to ontology coherence (all classes are realizable) rather than ontology consistency (there are no contradicting axioms). Compliance with this principle can thus be assessed by using the metrics described above. The conservativity principle states that correspondences should not introduce, in the merged ontology, new semantic relationships

between any two entities that were originally from the same input ontology. Compliance with this principle can be assessed by counting the number of violation to it, as proposed by Solimando et al. (2017). Finally, the locality principle states that correspondences tend not to be semantically isolated in the ontologies, which is to say, two semantically related concepts from one of the input ontologies are more likely to be aligned to two concepts from the other input ontology that are themselves semantically related, than to unrelated concepts. This principle is more a guideline for identifying potential false correspondences than a basis for assessing alignment quality, which is why no metric to assess its compliance has been proposed.

We must note that coherence and conservativity may sometimes be at odds with alignment completeness, as ontologies have different modelling views of their domain, which have to be reconciled when two ontologies are merged, possibly leading to new semantic relations between entities of one or both of them, as well as to logical conflicts (Pesquita et al., 2013). Thus, it may very well be that the complete and correct alignment between two ontologies is incoherent and/or unconservative. Nevertheless, alignment coherence is critical for several applications, such as ontology merging and query answering, and therefore is commonly used as an evaluation criterion in the OAEI, in tracks such as *Anatomy*, *Conference*, *Large Biomedical Ontologies*, *Disease and Phenotype*, and *Ontology Alignment for Query Answering (OA4QA)*. Evaluation modalities include binary assessment of coherence, the maximum cardinality measure, and the number or fraction of unsatisfiable classes. The relevance of alignment conservativity is more debatable, as it is not strictly required for any application, but it has also been used as an evaluation criterion in the OAEI *Conference* and *OA4QA* tracks. Note also that neither coherence nor conservativity evaluations are a substitute for an evaluation of alignment completeness and correctness, and they have always been used in complement of the latter in the OAEI. In the extreme case, an empty alignment is fully coherent and conservative, but utterly useless.

In complex alignments, assessing coherence is particularly desirable, as the very interest in making a complex alignment is underpinned by a concern with semantic precision beyond what simple alignments allow. However, assessing coherence requires reasoning and is computationally challenging, particularly for large and/or semantically complex ontologies, and even more so if the alignment itself is large and/or complex. Even more important, assessing coherence requires that the merged ontology be expressible in OWL DL, which may not be the case in complex alignments, even if the input ontologies are. Some complex correspondences are not expressible in OWL at all, while others are expressible in OWL but not OWL DL.

Assessing conservativity of complex alignments makes less sense than doing so for simple alignments, as complex alignments tend to contribute substantially to the semantics of both input ontologies by design (e.g. by defining ontology restrictions) and thus it is not at all unexpected that they lead to conservativity violations. That said, assessing conservativity violations in complex alignments should be little harder than doing so for simple alignments, assuming the correspondences can be encoded in OWL DL.

5.2 Task-based evaluation

The quality of an alignment can also be assessed regarding its suitability for a specific task or application. Considering that ontology alignments are, in practice, constructed for a given application or with a given task in mind, it would be useful to set up experiments that do not stop at the delivery of the alignment but carry on to the application or task for which the alignment was constructed. This is especially true when there is a clear measure of success for the overall task or application, but even when there is not, it can be useful to share corresponding aggregate measures associated with a task or application profile.

With respect to application-oriented evaluation, Isaac et al. (2008) proposed ontology alignment evaluation methods for the specific scenarios of thesaurus merging and data translation. They defined sets of tasks which need an alignment or part of it, then evaluated the alignment

on how well it fulfilled these tasks in terms of quality (for each task how good is the answer) and quantity (how many tasks were fulfilled by the alignment).

Regarding task-oriented evaluation, 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, that have different constraints in terms of coverage and run time. One such task-oriented evaluation approach was introduced in the OAEI in 2015 at the *OA4QA* track[2] (Solimando et al., 2014), 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 ABoxes. A given query, such as $Q(x):=Author(x)$ expressed using the vocabulary of the Cmt ontology, was executed over the merged ontology $Cmt \cup Ekaw \cup A$, where A is an alignment between *Cmt* and *Ekaw*. Precision and recall were calculated with respect to model answer sets, i.e., for each ontology pair and query $Q(x)$, and for each alignment A computed by each matching system. An alternative approach for evaluating query answering without using instances was proposed by David et al. (2018), where queries are compared without instance data, by grounding the evaluation on query containment.

While task-based evaluation is equally valid for both simple and complex alignments, some tasks tend to have higher expressiveness requirements, and thus to more often involve complex alignments, such as query answering/rewriting and ontology merging (Thiéblin et al., 2018). Query answering in particular has already been a subject of focus for complex alignments, with (Makris et al., 2012) presenting a set of complex correspondences used for query rewriting[3] for a few pairs of ontologies. More recently, complex correspondences have been exploited for the task of query rewriting for federating agronomic taxonomy knowledge on the LOD cloud (Thiéblin et al., 2017). This (Taxon) dataset was also used in the *Complex* track of the OAEI 2018 campaign, with the aim of assessing the performance of matching systems over large knowledge bases. The evaluation was performed based on the quality of the generated alignments (in terms of precision) and on the ability to rewrite SPARQL queries using these alignments. In particular, a manual analysis of the number of queries satisfyingly rewritten based on the alignments was carried out. The queries written for the source ontology were rewritten automatically when dealing with (1:1) or (1:n) correspondences, using the system described by Thiéblin et al. (2016), and manually when dealing with (m:n) correspondences.

Given the relevance of complex alignments for query answering, and the fact that this task is one of the main applications of these alignments, evaluation approaches based on this task would be highly relevant. One of the main challenges in implementing such approaches lies in establishing a query rewriting scheme that encompasses the expressivity and cardinality of complex correspondences. In the case of simple alignments, a naive approach for rewriting SPARQL queries can be to simply replace the IRI of an entity of the initial query by the IRI of the corresponding entity in the alignment, as described in David et al. (2011). For complex alignments, such a naive approach is obviously not possible, as the semantics of the alignment itself has to be taken under consideration. Euzenat et al. (2008) proposed an approach for writing specific SPARQL *construct* queries, but most query rewriting systems still rely on simple or (1:n) complex correspondence and fail in covering highly expressive (m:n) complex correspondences.

6 Discussion

The nature of complex ontology alignments presents unique evaluation challenges that were not considered when existing evaluation techniques were developed. This section outlines those challenges and analyzes the areas in which current approaches are lacking.

---

[2] http://www.cs.ox.ac.uk/isg/projects/Optique/oaei/oa4qa/index.html
[3] http://www.music.tuc.gr/projects/sw/sparql-rw/

6.1   Challenges

Regarding evaluation using reference alignments, challenges exist at each stage of the evaluation process:

Anchor Selection: Given the less bounded nature of complex matching, it is to be expected that systems will produce a large number of correspondences.

Challenge 1: Selecting which candidates will be compared to which reference correspondences in order to avoid the necessity of a full pairwise comparison of all candidates in the comparison step

Correspondence Comparison: Complex correspondences can consist of entities of arbitrary complexity and be expressed in a multitude of semantically equivalent or nearly-equivalent ways.

Challenge 2: Determining the relation between a candidate correspondence and a reference correspondence, which requires comparing all of the singular entities and the expressions in which they are listed for both correspondences

Challenge 3: Handling correspondence decomposition, which involves comparing sets of correspondences to a single correspondence, since the combination of several correspondences (simple or complex) can be equivalent or related to a single complex correspondence

Challenge 4: Comparing correspondences whose relation differs (e.g. a subsumption to an equivalence)

Scoring: Complex correspondences contain more axes than simple correspondences, because e and e' are not single entities but rather (potentially nested) combinations of entities, constructors, and transformation functions. This necessitates more nuanced scoring metrics, which can be used to determine how close a correspondence is to a reference correspondence. This allows for measuring the effort required of a human validator or by matching approaches creators to understand the limitations of their approaches and thus drive development.

Challenge 5: Accurately reflecting the quality of a correspondence, especially considering that in complex matching, a correspondence is still useful even if only partially correct.

Aggregation: Existing aggregation approaches for alignment evaluation with a reference alignment were designed with simple alignments in mind. They are tightly coupled to particular correspondence comparison and scoring methods and tend to take an all-or-nothing, or at best all, half, or nothing, approach.

Challenge 6: Factoring correspondences that are partially correct into the scoring process

Challenge 7: Considering a set of candidate correspondences in conjunction as related to a single reference correspondence (and vice-versa)

Challenge 8: Handling the occurrence of multiple correct candidate correspondences that are implied by a single reference correspondence (as is the case in correspondence 3 from the reference alignment and 3 and 4 from the first alignment system)

Evaluation when no reference alignment is available presents an orthogonal set of challenges. Task-based evaluations require a well defined task, for which a quality metric is definable. The quality of the alignment is measured by proxy through the quality of the task results, which results in a narrow scope for the evaluation. Furthermore, for tasks where the output needs to be manually evaluated (e.g., query rewriting) the manual effort required presents an additional challenge.

Challenge 9: Developing generalizable quality metrics for task-based complex alignment evaluation

Challenge 10: Automating the query rewriting process based on a set of complex correspondences

In addition to being able to handle the aforementioned challenges, evaluation metrics for complex alignments should also be fully automated and independent of manual input, even if the alignment is intended to be manually validated post-hoc. This is a crucial feature to further promote the development of complex matching approaches, by shortening the time between development cycles. Consequently, techniques for the evaluation of complex alignments need to be able to handle the computational complexity the challenges pose, both at the correspondence level and at the alignment level.

6.2 Gap analysis

We now turn our attention to assessing the degree to which existing alignment evaluation approaches address the challenges above. This analysis begins with approaches focused on cases in which a reference alignment is available (i.e. those relevant to challenges 1 through 8).

Syntactic approaches are unsuited to address challenges 3 and 4, since they do not employ reasoning and consider correspondences that are logically equivalent or can be derived as incorrect. They are also unable to address challenges 5, 6 and 7, given that they do not consider closely related correspondences. By virtue of their simplicity, they struggle less with challenges 1, 2 and 8, which are related to the computational complexity of the approach.

Rule-based semantic approaches provide strategies that can partially address challenges 5 and 6, since they are able to account for closely related correspondences involving direct super/subclasses. However, they are unable to handle the full gamut of expressions required by complex matching. Furthermore they do not address the remaining challenges. Edit-distance metrics, which assess the number of modifications that must be made to a candidate correspondence in order to arrive at reference correspondence, can be considered a type of rule-based semantic approach in the context of complex alignment evaluation. Examples of edit-distance metrics for strings include Levenstein and Smith-Waterman. These metrics are potentially able to handle challenges 4, 5 and 6, while not specifically addressing the remaining challenges. However, we are not aware of any existing edit-distance metrics for any of the common complex alignment representation languages discussed in Section 3.

Reasoning-based semantic approaches are better suited to answer challenges 2, 3 and 7, since they can cover the semantic complexity of complex expressions, and also handle correspondence decomposition. However, this is restricted to the cases where the merged ontologies are in OWL-DL. Furthermore, they are unable to handle challenges 5 and 6 since they only cover correspondences that can be logically derived, they consider closely related correspondences as incorrect. They also do not offer any specific features to address challenges 1 and 8.

Instance-based approaches circumvent many of the outlined challenges, by simplifying correspondence evaluation to a measure of the overlap between sets of instances. However they are applicable to class correspondences and transformations, but it is not straightforward to apply them to property correspondences. Furthermore, they require that all classes in the alignment be populated with instances.

With respect to evaluation approaches that do not require a reference alignment, the existing work primarily consists of manually intensive evaluation strategies that were uniquely developed

for particular cases. There is significant room for future work on the challenges relevant to these metrics (i.e. challenges 9 and 10).

6.3 Feasibility analysis

The task of evaluating complex correspondences is inherently expensive computationally, due to the syntactical and semantic complexity of these correspondences.

From a syntactic perspective, there is no theoretical limit to the complexity of the expressions that can be constructed through nesting. This is not a challenge for the use of the traditional *syntactic* evaluation metric, which can still be implemented with $\mathcal{O}(n)$ time complexity. It is, however, a substantial challenge for the implementation of more sophisticated and promising approaches, such as *rule-based* and particularly edit-distance metrics, which have to cope with a potentially endless search space of possible combinations of constructions and transformations. This means that, in all likelihood, 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, the more expressive complex correspondences go beyond OWL DL, and thus may not be decidable, while 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.

By contrast, instance-based approaches are largely 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. One approach for this would be to determine, for each correspondence c_i in the evaluated alignment, the relation between the sets of instances I_s and I_t, belonging to the source and target members of the correspondence respectively. Each correspondence could then be classified as *equivalent*, *subsumed*, *overlapping*, or *disjoint*, given the relation between I_s and I_t, or *empty* if $I_s = I_t = \emptyset$ (i.e., if both members are either unsatisfiable or non-populated entities). Having a reference alignment, one would know what are the sets of expected instances to be compared. Different precision scores could then be computed for each type of correspondence member relation: the *equivalent* precision would measure the percentage of correspondences whose members are exactly populated with the same instances, and likewise, the *subsumed*, *overlapping* and *not disjoint* precision would measure the percentage of correspondences whose members subsume one another, overlap, or are not *disjoint*, respectively.

Such a strategy could rely on expressing complex correspondences as SPARQL queries, which would cover also transformation functions. As we discussed previously, it is limited in coverage, since it can be applied to the evaluation of class (expression) correspondences or transformations, but it is not straightforward to apply to the evaluation of property (expression) correspondences. Furthermore, it requires the knowledge bases to be consistently populated (i.e., complete population of all entities the complex correspondences are supposed to cover). However, the cost of creating such a knowledge base (e.g. with artificially populated data) is smaller than the cost of creating reference alignments or applying evaluation strategies such as query rewriting.

7 Conclusions and Future Work

In this paper we have defined complex ontology alignments and shown that the few systems that have attempted to generate such alignments have been evaluated using methods that are difficult to generalize and/or labor intensive. A survey of existing evaluation approaches, which were developed with simple alignments in mind, has shown that they are insufficient in several ways. In particular, the most common evaluation approach, based on exact syntactic match, lacks the nuance necessary to distinguish between completely unhelpful correspondence suggestions and

those that are "almost correct." Other existing evaluation techniques are not scalable to the complex case, can only be used under certain conditions (e.g. when dually-annotated instance data is present or when the alignment is expressible in OWL DL), or have other drawbacks. We have enumerated what we view as the most pressing gaps between current techniques and what are needed for complex alignment evaluation. In the remainder of this section, we propose future work that can potentially bridge these gaps.

Evaluating complex ontology alignments is too broad a challenge to tackle with a single approach, as there are multiple aspects to take into account, and different tasks will likely merit different evaluation paradigms. Considering that the two main applications of complex alignments are ontology/linked-data integration, and query answering/rewriting, it stands to reason to focus our efforts in developing evaluation approaches with these applications in mind.

With respect to ontology/linked-data integration, it is unlikely that the state of the art in ontology alignment ever reaches a point where human validation is unnecessary. This is true even for simple alignments, but particularly so for complex alignments, given the inherent difficulty in generating them automatically with reasonable precision or recall. Under this premise, we believe that the most adequate approach to evaluate complex alignments in the context of this application would be an edit-distance approach that reflected the effort involved in human validation, in the same spirit as the effort similarity approach we reviewed in Section 4.2.2. Therefore, our future work will concentrate on developing an analogous edit-distance approach that encompasses all the requirements and nuances of complex alignment evaluation. Concretely, the approach must explicitly contemplate all complex expressions in use and define costs for inter-converting them, and must adequately handle cases of correspondence decomposition, where a reference correspondence should be compared with two or more system correspondences that cover it partially (or vice versa). Greater granularity with respect to the edit-costs between semantically related classes would also be desirable. Last, but not least, the approach should be scalable, and avoid the need to do all-vs-all correspondence comparisons. Given these constraints, we believe that a deterministic rule-based edit-distance approach that covers all the key complex correspondences constructions explicitly, in a way that reflects the effort required to correct them, would be the best candidate.

With respect to query answering/rewriting, we believe that there are two major hurdles to be tackled: developing an automated converter for transforming any complex alignment into a query rewriting scheme; and developing a query generating algorithm that can automatically generate queries adequate in coverage and scope to the complex alignment to evaluate. The primary focus of our future work will be the first hurdle, as only by overcoming it can we use query-based approaches to fully evaluate complex alignments automatically. Overcoming the second hurdle will be essential to enable the widespread use of query-based evaluation, and will also contribute to make query-based evaluation efforts more comprehensive and comparable, as otherwise queries have to be manually defined for each test case.

We will also explore instance-based evaluation approaches, such as the one delineated in the previous section. This approach can complement or even replace the edit-distance approach in a linked-data integration scenario, and can be a computationally efficient and labor-friendly alternative to query answering.

8   Acknowledgments

References

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

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.

David, J., Euzenat, J., Genevès, P., and Layaïda, N. (2018). Evaluation of query transformations without data: Short paper. In *Companion of the The Web Conference 2018 on The Web Conference 2018, WWW 2018, Lyon , France, April 23-27, 2018*, pages 1599–1602.

David, J., Euzenat, J., Scharffe, F., and Trojahn dos Santos, C. (2011). The alignment api 4.0. *Semantic web*, 2(1):3–10.

Duan, S., Fokoue, A., Hassanzadeh, O., Kementsietsidis, A., Srinivas, K., and Ward, M. J. (2012). Instance-based matching of large ontologies using locality-sensitive hashing. In *International Semantic Web Conference*, pages 49–64. Springer.

Ehrig, M. and Euzenat, J. (2005). Relaxed precision and recall for ontology matching. In *Integrating Ontologies '05, Proceedings of the K-CAP 2005 Workshop on Integrating Ontologies, Banff, Canada, October 2, 2005*.

Euzenat, J. (2004). An api for ontology alignment. In *International Semantic Web Conference*, pages 698–712. Springer.

Euzenat, J. (2007). Semantic precision and recall for ontology alignment evaluation. In *IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007*, pages 348–353.

Euzenat, J., Polleres, A., and Scharffe, F. (2008). Processing ontology alignments with sparql. In *2008 International Conference on Complex, Intelligent and Software Intensive Systems*, pages 913–917.

Euzenat, J., Scharffe, F., and Zimmermann, A. (2007). Expressive alignment language and implementation.

Euzenat, J. and Shvaiko, P. (2013). *Ontology Matching*. Springer Berlin Heidelberg.

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.

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.

Isaac, A., Van Der Meij, L., Schlobach, S., and Wang, S. (2007). An empirical study of instance-based ontology matching. In *The Semantic Web*, pages 253–266. Springer.

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*, pages 75–95. Springer.

Jiménez-Ruiz, E., Grau, B. C., Horrocks, I., and Berlanga, R. (2011). Logic-based assessment of the compatibility of umls ontology sources. *Journal of biomedical semantics*, 2(1):S2.

Kirsten, T., Thor, A., and Rahm, E. (2007). Instance-based matching of large life science ontologies. In *International Conference on Data Integration in the Life Sciences*, pages 172–187. Springer.

Maedche, A., Motik, B., Silva, N., and Volz, R. (2002). Mafra—a mapping framework for distributed ontologies. In *International Conference on Knowledge Engineering and Knowledge Management*, pages 235–250. Springer.

Makris, K., Bikakis, N., Gioldasis, N., and Christodoulakis, S. (2012). SPARQL-RW: transparent query access over mapped RDF data sources. In *15th International Conference on Extending Database Technology*, pages 610–613. ACM.

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.

Nunes, B. P., Mera, A., Casanova, M. A., Breitman, K. K., and Leme, L. A. P. (2011). Complex Matching of RDF Datatype Properties. In *Proceedings of the 6th International Conference on Ontology Matching-Volume 814*, pages 254–255. CEUR-WS. org.

Parundekar, R., Knoblock, C. A., and Ambite, J. L. (2010). Linking and building ontologies of linked data. In *ISWC*, pages 598–614. Springer.

Parundekar, R., Knoblock, C. A., and Ambite, J. L. (2012). Discovering concept coverings in ontologies of linked data sources. In *ISWC*, pages 427–443. Springer.

Pesquita, C., Faria, D., Santos, E., and Couto, F. M. (2013). To repair or not to repair: reconciling correctness and coherence in ontology reference alignments. In *OM*, pages 13–24.

Qi, G. and Hunter, A. (2007). Measuring incoherence in description logic-based ontologies. In *The Semantic Web*, pages 381–394. Springer.

Qin, H., Dou, D., and LePendu, P. (2007). Discovering executable semantic mappings between ontologies. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*, pages 832–849. Springer.

Ritze, D., Meilicke, C., Šváb Zamazal, O., and Stuckenschmidt, H. (2009). A pattern-based ontology matching approach for detecting complex correspondences. In *4th ISWC workshop on ontology matching*, pages 25–36.

Ritze, D., Völker, J., Meilicke, C., and Šváb Zamazal, O. (2010). Linguistic analysis for complex ontology matching. In *5th workshop on ontology matching*, pages 1–12.

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

Scharffe, F. (2009). *Correspondence Patterns Representation*. PhD thesis, Faculty of Mathematics, Computer Science and University of Innsbruck.

Schopman, B., Wang, S., Isaac, A., and Schlobach, S. (2012). Instance-based ontology matching by instance enrichment. *Journal on Data Semantics*, 1(4):219–236.

Serpeloni, F., Moraes, R., and Bonacin, R. (2011). Ontology mapping validation. *International Journal of Web Portals*, 3(3):1–11.

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.

Solimando, A., Jiménez-Ruiz, E., and Pinkel, C. (2014). 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.

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.

Thiéblin, É., Amarger, F., Haemmerlé, O., Hernandez, N., and dos Santos, C. T. (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.

Thiéblin, E., Amarger, F., Hernandez, N., Roussey, C., and Trojahn, C. (2017). Cross-querying lod datasets using complex alignments: An application to agronomic taxa. In *Research Conference on Metadata and Semantics Research*, pages 25–37. Springer.

Thiéblin, E., Haemmerlé, O., Hernandez, N., and Trojahn, C. (2018). Task-oriented complex ontology alignment: Two alignment evaluation sets. In *The Semantic Web - 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, June 3-7, 2018, Proceedings*, pages 655–670.

Thiéblin, E., Cheatham, M., Trojahn, C., Zamazal, O., and Zhou, L. (2018a). The first version of the oaei complex alignment benchmark. In *ISWC Posters and Demos*. Springer.

Thiéblin, E., Haemmerlé, O., and Trojahn, C. (2018b). Complex matching based on competency questions for alignment: a first sketch. In *OM 2018 - 13th ISWC workshop on ontology matching*.

Van Hage, W. R., Isaac, A., and Aleksovski, Z. (2007). Sample evaluation of ontology-matching systems. In *EON*, volume 2007, pages 41–50.

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.

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.

Xiao, G., Calvanese, D., Kontchakov, R., Lembo, D., Poggi, A., Rosati, R., and Zakharyaschev, M. (2018). Ontology-based data access: A survey. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*, pages 5511–5519. International Joint Conferences on Artificial Intelligence Organization.

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.

Zhou, L., Cheatham, M., Krisnadhi, A., and Hitzler, P. (2018). A complex alignment benchmark: Geolink dataset. In *The Semantic Web - ISWC 2018 - 17th International Semantic Web Conference, Monterey, CA, USA, October 8-12, 2018, Proceedings, Part II*, pages 273–288.