Noname manuscript No.
(will be inserted by the editor)

Alignment of Surface Water Ontologies

A comparison of manual and automated approaches

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

Received: date / Accepted: date

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.

Keywords Complex ontology alignment · Schema alignment · Surface water ontologies · Semantic data integration · OAEI

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.

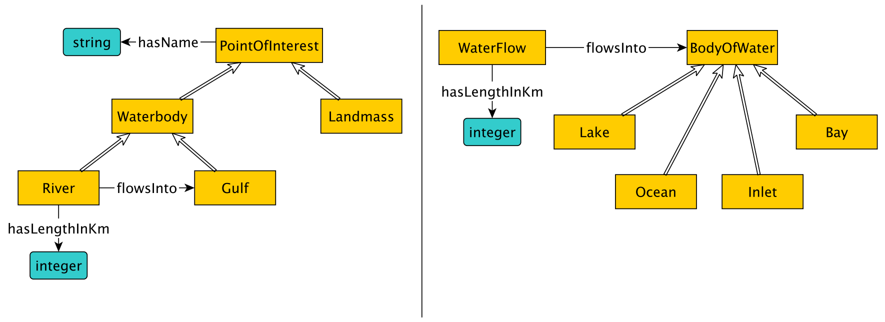

Fig. 1 Two sample ontologies. Differences in scope and granularity create alignment challenges.

```
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

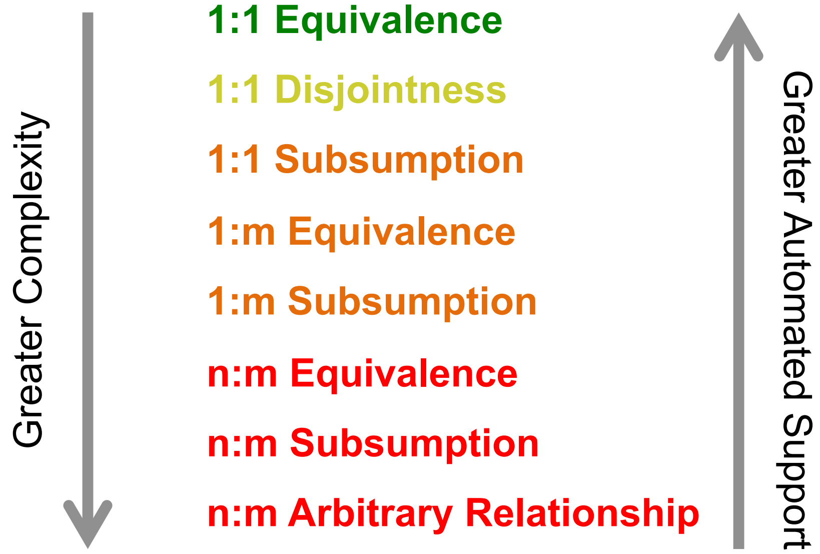

Fig. 2 Spectrum of alignment complexity. The x:y notation refers to the number of entities from the first ontology (x) and the number from the second ontology (y) that are involved in a particular mapping across the ontologies.

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.

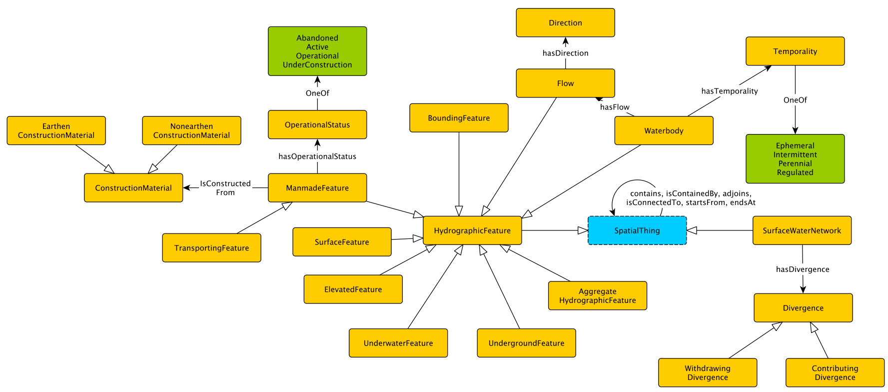

Fig. 3 The upper levels of the SWO class and property hierarchy

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.

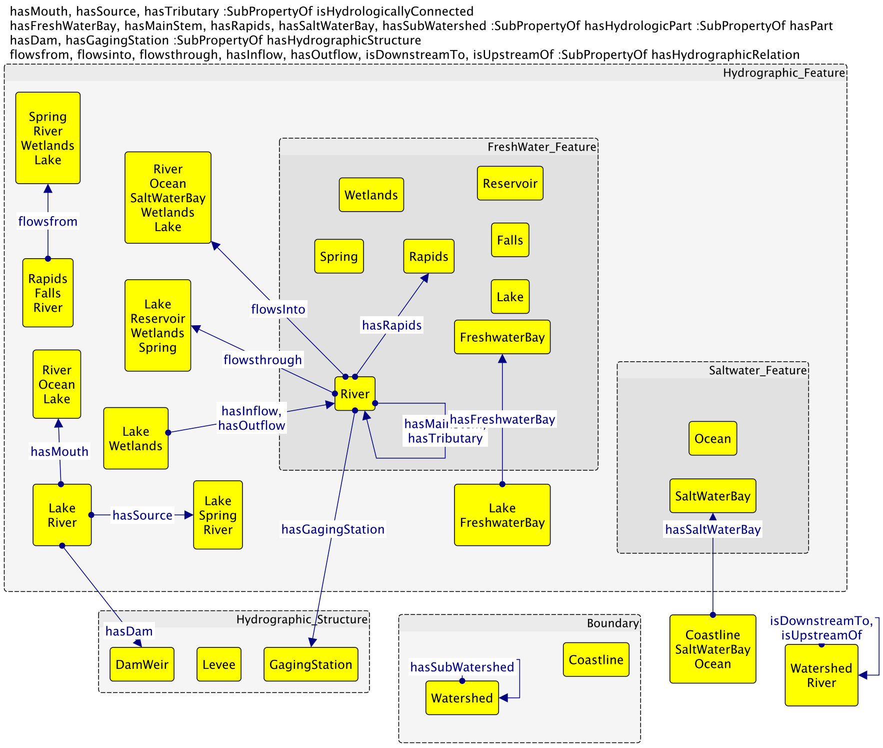

Fig. 4 The Hydro3 module within the HydroGazetteer ontology

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-

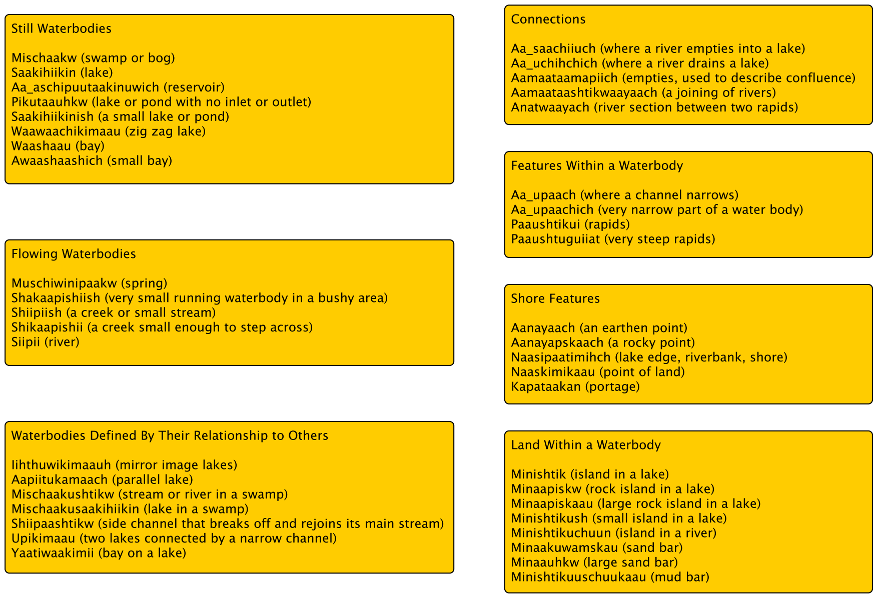

Fig. 5 The classes within the Cree ontology. Class groupings (e.g. Still Waterbodies, Connections) have been added for convenience and are not part of the ontology.

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/

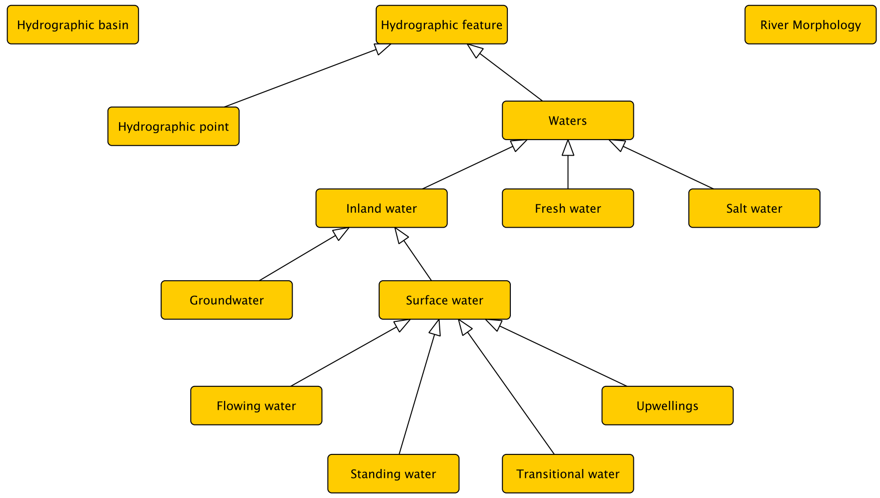

Fig. 6 The upper levels of the hydrOntology class hierarchy

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 alignment 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 logical 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 subclass 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 relation 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 property 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 restriction. 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 *simplest* 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 relations 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 WordNet 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 RiMoM 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.

Acknowledgements This work was funded by the United States Geological Survey under Agreement Number G16AC00120.

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 Enrichment for Spatial Data, chap. An Architecture for Enhancing Semantic Interoperability 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 semantics **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: International 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)