Comparison and Evaluation of Ontologies for Units of Measurement

Jan Martin Keil [a,*], Sirko Schindler [a]
[a] *Heinz Nixdorf Chair for Distributed Information Systems,*
Institute for Computer Science,Friedrich Schiller University Jena, Germany
E-mails: jan-martin.keil@uni-jena.de, sirko.schindler@uni-jena.de

Abstract. Measurement units and their relations like conversions or quantity kinds play an important role in many applications. Thus, many ontologies covering this area have been developed. As a consequence, for new projects aiming at reusing one of these ontologies, the process of evaluating them has become more and more time consuming and cumbersome. We evaluated eight major ontologies for units of measurement and the relevant parts of the Wikidata corpus. We automatically collected descriptive statistics about the ontologies and scanned them for potential errors, using an extensible collection of scripts. The computational results were manually reviewed, which uncovered several issues and misconceptions in the examined ontologies. The issues were reported to the ontology authors.

In this paper we will present the evaluation results including statistics as well as an overview of detected issues. We thereby want to enable a well-founded decision upon the unit ontology to use. Further, we hope to prevent errors in future by describing some pitfalls in ontology development—not limited to the domain of units of measurement.

Keywords: measurement unit ontology, ontology comparison, ontology evaluation, ontology quality

1. Introduction

Units of measurement are an essential part in many aspects of modern life: The correct handling of the scale a value is measured in is crucial not only in science, but also in trade, industry and administration. A well documented use of units is especially important, when a project is carried out by different partners with a potentially different background. One of the most prominent examples of neglecting this fact is the crash of the Mars Climate Orbiter in 1999, which the NASA investigation board attributed to a mismatch of used units between two components of its software [1].

Similar integration challenges arise on an even larger scale in the context of Big Data: With the increasing need to integrate datasets of different origins, data annotation—preferably using Semantic Web techniques—gains importance. Using a machine readable annotation is essential for (semi)automatic discovery, verification and integration.

*Corresponding author. E-mail: jan-martin.keil@uni-jena.de.

As a part of these semantic descriptions and to cover the field of units of measurement and related concepts, over the last years several projects were initiated to create respective ontologies [2–5].

Most of these attempts were embedded in bigger research projects and thus catered to their specific needs. This led to a variety of different approaches to model the domain at hand. The created ontologies differ not only in the modeled subset of concepts and relations, but also in the type and amount of units included. Engineers or researchers, who wish to use an existing ontology in their work, are now faced with the choice between several ontologies.

To assist in this decision making process we analyzed eight ontologies in the field of units of measurements and the respective parts of the emerging Wikidata corpus [6]. We used a collection of scripts to extract several kinds of statistics and to detect contradictions between the data sources. The manually review of the results confirmed not only the different emphases of the ontologies, but also revealed the exis-

1570-0844/0-1900/$35.00

tence of several issues in all of them. Contributions of this paper are as follows:

- We provide an extensible collection of scripts, which can be used by unit ontology developers to validate their work against other ontologies.
- We identified multiple issues in existing ontologies, which were reported to the respective authors.
- We identified issue classes and hint towards preventive actions.
- We present the analysis of existing ontologies to support potential users in the decision process of selecting a unit ontology for reuse.

The paper is structured as follows: section 2 will give an overview over the analyzed ontologies and present previous work on ontology evaluation. The terminology used will be specified in section 3. In section 4 we will describe the general approach and some implementation details, before in section 5 the various aspects of the analysis will be presented. Limitations of our approach will be discussed in section 6. We conclude with suggestions to ontology authors to prevent some of the encountered issues in section 7 and final remarks in section 8.

2. Related Work

The discussion of related work is split into two parts: First, we will present the analyzed ontologies and specify the respective version we used. The second part will then describe previous work on ontology evaluation.

2.1. Ontologies for Units of Measurement

Over the last years several projects were initiated to create ontologies modeling the domain of units of measurements. The following selection of eight ontologies focuses on ontologies available in an OWL compatible format. There are, of course, several other unit ontologies within and beyond the OWL world. However, the following subset of ontologies was selected, as they seem to be the most promising candidates regarding the amount of individuals and concepts modeled:

- Measurement Units Ontology (MUO)[1]; result of a project to exploit semantics in mobile environments; the instances were automatically generated from UCUM [7],
- Extensible Observation Ontology (OBOE)[2]; an ontology suite to represent scientific observations,
- Ontology of units of Measure and related concepts (OM 1)[3]; an ontology to model concepts and relations important to scientific research, developed in context of food research [3],
- Ontology of units of Measure (OM 2)[4]; second iteration of the OM ontology,
- Library for Quantity Kinds and Units (QU)[5]; a showcase ontology based on the OMG SysML 1.2 QUDV specifications and the UN/CEFACT Recommendation 20 code list [8],
- Quantities, Units, Dimensions and Data Types Ontologies (QUDT)[6]; developed in context of NASA projects,
- Semantic Web for Earth and Environmental Terminology (SWEET)[7]; also developed in context of NASA projects and
- Units of Measurement Ontology (UO)[8] and Phenotypic Quality Ontology (PATO)[9]; both modules of the OBO family to model units and phenotypic qualities.

We are aware that there are efforts towards a second version of QUDT currently in progress. At the time of writing, however, the current results do not justify an inclusion in this analysis. This is left open for future work. In the following “OM” will refer to both OM 1 and OM 2, when a statement applies to both ontologies alike. If only one ontology is affected, we will use “OM 1” or “OM 2” respectively.

Besides these specialized domain ontologies, linked data initiatives also provide data on units of measurement.

DBpedia extracts structured data from Wikipedia[10] and creates an interlinked knowledge base out of it [9]. Many Wikipedia articles provide structured information in the form of infoboxes. In these infoboxes values

[1] `muo-vocab.owl` and `ucum-instances.owl` dated 2008 from http://idi.fundacionctic.org/muo/

[2] Version 1.0 from https://code.ecoinformatics.org/code/semtools/trunk/dev/oboe/

[3] Version 1.8.3 from http://www.wurvoc.org/vocabularies/om-1.8/

[4] Version 2.0.3 from https://github.com/HajoRijgersberg/OM

[5] `qu.owl` dated 2011-06-20 and `qu-rec20.owl` dated 2010-09-28 from https://www.w3.org/2005/Incubator/ssn/ssnx/qu/

[6] Version 1.1 from http://www.qudt.org/

[7] Version 2.3 from http://sweet.jpl.nasa.gov/

[8] Version 2017-04-24 from http://purl.obolibrary.org/obo/uo.owl

[9] Version 2017-03-22 from http://purl.obolibrary.org/obo/pato.owl

[10] https://www.wikipedia.org/

are oftentime annotated with the respective unit. These values are harmonized by the extraction framework of DBpedia. Unfortunately, the underlying knowledge about units is hard coded into the extraction framework.[11] For articles on units, however, the infoboxes are not yet covered by the DBpedia extraction and mapping process at all. Therefore, it is not contained in the data set dumps, the public SPARQL endpoint and the user interface provided by DBpedia.

On the other hand, Wikidata –a sister project to Wikipedia– aims at collecting the factual knowledge of Wikipedia in a central location [6]. This information can then be linked directly from different Wikipedia articles and thus provide a consistent view even across several language versions, which up to now might provide different values on certain facts like, e.g., population statistics. Besides other facts, Wikidata also includes a rather large number of units and their relations. As this data can also easily be accessed through a SPARQL endpoint, WD was included in the analysis.

– Wikidata (WD); community driven repository of factual data for Wikipedia[12]

2.2. *Ontology Evaluation*

Motivated by the lack of published evaluation processes as well as examination results for well known ontologies an early evaluation approach was developed by Gómez Pérez [10]. This approach consists of two steps: In a first analysis step the ontology will be inspected regarding consistency, completeness, conciseness, expandability and sensitiveness. In a second synthesis step the ontology will be corrected. This approach was applied to the Standard Unit Ontology (SU) by two ontologists and two domain experts. The choice of SU also indicates a particular need for evaluation and improvement in the domain of measurement units.

Based on Gómez Pérezs criteria Rijgersberg et al. evaluated five ontologies concerned with units of measurement [11]. This analysis determined a lack of an ontology containing all important concepts of this domain. In consequence they developed the Ontology of units of Measure and related concepts (OM 1). Subsequently, Rijgersberg et al. published a detailed comparison of OM 1 with the Quantities, Units, Dimensions and Data Types Ontologies (QUDT) in terms of modeled concepts [3].

Another domain specific comparison approach was introduced by Marcus P. Foster. He suggests a ranking of unit ontologies in five levels regarding the supported concepts and use cases [12]. This ranking provides a fast overview of scope and level of development of ontologies. The order of requirements for each ranking-level, however, seems biased by the authors background in parts. An ontology might, e.g., provide concepts of quantity and system of quantities, required by the third level, without supporting unit conversion, required by the second level.

Chau Do and Eric J. Pauwels applied MathML to map unit ontologies [13]. The evaluation of their mapping approach discovered several “incorrect information in the ontologies”. But as the work was not focused on ontology evaluation, they did not continue to systematically search for errors. Their approach requires relations between units themselves like, e.g., unit composition to be present in the ontologies, which unfortunately is not guaranteed for all ontologies.

Samadian et al. provided a comparison tailored to the specific requirements of a medical data integration project [14]. It is focused on the conversion of units and the modeling of their mathematical relation. Due to the narrow focus the benefit for other projects is limited.

In [15] a comprehensive list of use cases in the field of unit ontologies is presented. A simple metric was developed to measure the suitability of a given ontology with respect to those use cases. Finally, a selection of seven ontologies was analyzed and compared using this metric.

Zhang et al. tried to compare four unit ontologies with a focus on the OWL standard vocabulary used for modeling and possible reuse of the ontologies [16]. However, their approach lacks a general methodology, which could be extended to include other ontologies in the comparison. Furthermore, the used vocabulary in an ontology does not enable any conclusions regarding its usefulness in other projects.

A domain independent approach of ontology comparison is the usage of general quality criteria as implemented by, e.g., OntOlogy Pitfall Scanner! (OOPS!) [17]. This tool (semi)-automatically checks for many common pitfalls in ontology development thus giving initial impression of the quality. A detailed analysis of domain specific issues is, however, out of the scope here.

[11] https://github.com/dbpedia/extraction-framework/blob/master/core/src/main/scala/org/dbpedia/extraction/ontology/OntologyDatatypes.scala

[12] Status as of 2017-06-27; using https://query.wikidata.org/

Although the comparison in [3] includes some statistics, all mentioned evaluation efforts are manual assessments almost exclusively remaining on a conceptual level or only cover general quality issues. Thus, there is a lack of evaluation concerning the instances of unit ontologies.

3. Terminology

The unit ontologies differ in the terms they use to describe the same or closely related concepts. For the purpose of this paper these terms have to be unified. In the following the subsequent terms are used as defined in the International vocabulary of metrology [18]:

Definition 1 (Quantity). *"property of a phenomenon, body, or substance, where the property has a magnitude that can be expressed as a number and a reference"* (e.g., width of a particular street or mass of a particular apple)

Definition 2 (Kind of quantity). *"aspect common to mutually comparable **quantities**"* (e.g., diameter, width or mass)

Definition 3 (Dimension). *"expression of the dependence of a **quantity** on the **base quantities** of a **system of quantities** as a product of powers of factors corresponding to the base quantities, omitting any numerical factor"* (e.g., $\dim \textit{diameter} = \mathrm{L}$, $\dim \textit{width} = \mathrm{L}$ or $\dim \textit{force} = \mathrm{LMT}^{-2}$)

Definition 4 (Measurement unit). *"real scalar **quantity**, defined and adopted by convention, with which any other quantity of the same **kind** can be compared to express the ratio of the two quantities as a number"* (e.g., meter or gram)

Definition 5 (System of units). *"set of **base units** and **derived units**, together with their multiples and submultiples, defined in accordance with given rules, for a given **system of quantities**"* (e.g., International System of Units (SI) or Centimetre–Gram–Second System of Units (CGS))

Furthermore, the terms *quantity-value scale* and *conversion factor between units* defined in [18] are summarized as follows:

Definition 6 (Conversion). *tuple of an offset specified by the respective **quantity-value scales** and a **conversion factor** used to convert between two given units* (e.g., the offset 459.67 and the factor 1.8 for the conversion between degree Fahrenheit and Kelvin)

Additionally the following terms, which are not defined but utilized by [18], will be used in the subsequent sections:

Definition 7 (Field of application). *scientific discipline, domain or area of life, where a **kind of quantity** is used* (e.g., physics or navigation)

Definition 8 (Prefix). *named factor defined in a **system of units** to obtain the **multiple or submultiple of a unit*** (e.g., kilo or micro)

Moreover *compound unit* will be used according to the following definition:

Definition 9 (Compound unit). *a composition of **measurement units** using the mathematical operations multiplication, division or exponentiation [11]* (e.g., meter per second or kilogram per meter cubed)

4. Methods

This section presents our approach to evaluate and compare different unit ontologies. First, we describe the efforts to compile our reference data. Following, we give a general description of our approach and implementation. We conclude with a discussion of challenges in the evaluation process and the decisions made.

4.1. Reference Data

When evaluating a unit ontology on the instance level, there are two major aspects: One is *coverage*, which evolves around what concepts are modeled as well as how many and what kind of individuals are included. The other one is *correctness*, which is concerned with how reliable the included values are in an application's context.

To cover both aspects in a meaningful way, there is a need for a comprehensive reference dataset for comparison. In the field of units of measurement, we are not aware of such a dataset. There are some datasets [7, 8, 19, 20], but those either cover only a subset of the units in use or omit some of the concepts described in the previous section. In absence of such a reference corpus, we decided to use the union of all individuals used in the different available ontologies. Although this probably does not reach the same quality level, it can serve as a well enough substitute, especially when the main focus is on comparing the ontologies to one another. To create such a unified set of individuals, one has to extract the individuals and then map them to one another to establish a connection and reduce redundancy.

Having now created some kind of a reference dataset and as a byproduct a mapping between ontologies, one can easily compute all the characteristic statistics for an evaluation of the respective *coverage*. Also by explicitly using this mapping, it is easy to evaluate the overlap of ontologies, that is those objects which have a representation in more than one ontology.

The other major aspect is the *correctness* of the ontologies. In the absence of a reference dataset, using the previously created joint dataset, one can detect deviations between at least two ontologies, but not any issue affecting objects present in just one ontology. In general, it is not possible, however, to automatically decide which of the alternatives is correct and which erroneous. For this, further manual intervention is need along with possibly some literature research.

One of the most obvious issues, that can be checked for, is the correctness of conversions. Most units today are defined with respect to either a physical phenomenon or other units. As those conversions are defined by the respective standard bodies, there is no room for interpretation by the ontologies and hence all have to return the same values for a given set of (matching) units. Another class of issues, which became apparent during the analysis, is duplicate units. Here for different reasons multiple individuals describing the same concept were included in one ontology.

Many of the other issue classes detected with this approach and described later in detail also became apparent while implementing the mapping and gathering the statistical characteristics previously described. To reduce bias, each time a new class of issues was found in one ontology, if possible, its detection was automatized and added to the overall evaluation, so all ontologies are checked in the same way.

A different way to categorize the results is, whether they can be computed for each ontology individually or need the comparison with the reference dataset. Hence, in the following we will distinguish between inter-ontology facts which denote results that need multiple ontologies to be computed and intra-ontology results which rely on a single ontology, only. Simple examples for intra-ontology results are the mere number of individuals per concept and ontology or the completeness of an ontology regarding certain relations between the individuals. Issue detection most of the time is inter-ontology like, e.g., wrong conversions as well as all statistics concerned with actual comparing ontologies, like the overlap in terms of units.

4.2. Approach

A strictly manual assessment of the given ontologies is quite cumbersome and oftentimes lacks the advantages of an automated, systematic approach. Also over time new ontologies will emerge, which would require a reevaluation for the already existing ones. These considerations led us to the semi-automatic approach shown in Figure 1, that will be described in the following. A running example is given in Figure 2 to Figure 6.

In a first step, the individuals of the different ontologies for each concept as well as their relations are extracted. As the different ontologies use different modeling techniques, this ensures a common relational format for further analysis. As a side effect, the approach may also be used to compare against other datasets not necessarily in the format of an ontology. The result is a unified, relational representation of the data contained in each ontology (Figure 2).

Afterwards, in a second step the intra-ontology checks can be applied, which can also validate the extraction results (Figure 3).

In the third step the individuals for each concept and relation are mapped to one another. As the individuals were extracted separately for each concept in the first step, the mapping can work on each concept individually. Besides reducing the computational effort, this allows for employing different mapping strategies. Whereas for some concepts a mere comparison by name might be sufficient, others like the units themselves need a more sophisticated approach, which also takes the ordering of terms into account (Figure 4).

In the fourth step the inter-ontology statistics and checks over the previously gathered results are computed like, for example, the amount of instances per

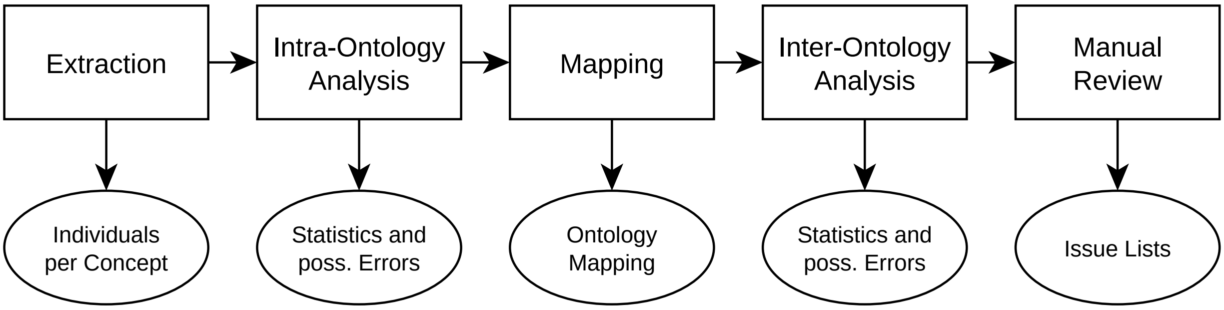

Fig. 1. Schematic of evaluation approach and the respective results

Unit	Label
ex1:meter	"meter"
ex1:kilometer	null
ex2:metre	"metre"
ex2:kilogram	"kilogram"
...	...

Prefix	Factor
ex1:kilo	1000
ex1:centi	0.01
ex2:kilo	0.001
...	...

Unit	Prefix
ex1:kilometer	ex1:kilo
ex2:kilogram	ex2:kilo
ex2:milligram	ex2:milli
...	...

Fig. 2. Exemplary results for step one: Extraction.

Missing Label
ex1:kilometer

Missing Prefix
ex2:kilometer

Individuals per Concept
ex1 has 100 units
ex2 has 200 units
...

Fig. 3. Exemplary results for step two: Intra-Ontology Checks and Statistics.

meter (unit)	kilogram (unit)	kilo (prefix)
ex1:meter	ex2:kilogram	ex1:kilo
ex2:metre		ex2:kilo

Fig. 4. Exemplary results for step three: Mapping.

type[13] or the number of prefixed and non-prefixed units per ontology (Figure 5).

It is important to note, that the results just contain lists of potential errors or a representation, which makes it easy to spot them. Therefore, in a final phase, the output has to be reviewed manually to extract actual issues (Figure 6). On the other hand, that manual evaluation process and the search for the reasons of the respective issues sometimes leads to completely new insights, which in turn led to new checks.

4.3. Implementation

To apply the described approach to existing unit ontologies an implementation in the form of a collection of scripts was created. It is publicly available for free use and extension [21]. To easily be able to run just a subset of scripts, but yet preserve dependencies, each script will store its results into a set of JSON files, which will then be used by its successors.

The first phase requires the extraction of individuals. For each ontology we formulated SPARQL queries to extract the individuals per concept as well as the respective relations. Those ontologies that are given in an OWL-compatible format are added to a Sesame triple store[14], before the corresponding set of queries is executed. To simplify the extraction queries a RDF-schema reasoner was used. For Wikidata on the other hand an online SPARQL endpoint is provided[15], which we used for extraction. A missing query for a specific concept in an ontology is treated as though that concept is not present in that particular ontology.

While the second phase is straight forward in terms of implementation, the subsequent mapping of individuals in the third phase required more care as the labels had to be harmonized regarding certain aspects.

For measurement units, this included harmonizing British and American notation (e.g., “metre” and “meter”), differences in phrasing (e.g., “square meter” and “meter squared”) and the position of a potential system of units (e.g., “US Survey Foot” and “Foot US Survey”). Additionally the transitive closure of conversions per ontology is computed. Thus the amount

[13]There are cases, where computing these values over the mapped units differ from computing them in the second phase. Some ontologies contain the same individual under different URIs, which would have been counted twice in Phase two. For more details see section 5.

[14]http://rdf4j.org/

[15]https://query.wikidata.org/

wrong prefix factor	overlap	unique mappings
ex1:kilo ↔ ex2:kilo	units: ex1 80% in ex2 units: ex2 40% in ex1	units: 220 prefixes: 20

Fig. 5. Exemplary results for step four: Inter-Ontology Checks and Statistics.

wrong prefix factor	Missing Label	Missing Prefix
ex2:kilo	ex1:kilometer	ex2:kilometer

Fig. 6. Exemplary results for step five: Manual Review.

of available conversions for the later review is significantly expanded: If, for example, the conversions *milligram to gram* and *gram to kilogram* exist, the missing, but inferable conversion *milligram to kilogram* is added).

4.4. Challenges

The wide range of languages in Wikidata posed a particular challenge for the mapping. We found examples of units where the label in some languages was plain wrong: See, e.g., `WD→Q25343`, which represents the unit "square metre". The Hindi label घन मीटर, however, translates to "cubic meter".

On the other hand some individuals have the same label although they depict different units of measurement (units): See, e.g., `WD→Q16830567` and `WD→Q103510` both sharing the same label in German "Bar", but while one denotes the common unit for pressure the other one represents an old Persian unit of mass. As the other ontologies besides OM and WD only use English labels or omit the language tag altogether (see Table 5), we decided just to map units based on three types of labels: English labels, labels without a given language and an artificial label created from the local name of the URI. The published code, however, also includes a mapping implementation that includes all languages and can be activated in the configuration.

Another challenge in WD are duplicates with no overlap in the languages of their labels. See `WD→Q2649540` and `WD→Q10778756`, which appear to reference to the same unit "Alma", but currently can not be matched as `WD→Q2649540` provides only a German label while `WD→Q10778756` only has a Czech one. Finding such duplicates is out of the scope of this paper.

Beside the mapping by name we utilized mappings already defined by the ontologies using `owl:sameAs` and other similar relations. This includes mappings within a single ontology as well as to other ontologies.

The manual review of the mappings uncovered several missed mappings. As adding specialized code for particular instances would have made the codebase unnecessarily complex, one can predefine certain mappings in separate files, which will then be used in the mapping process.

Not all ontologies model the relation between units and their potential prefixes. To have a more conclusive comparison, a heuristic was used to identify those prefixed units, that the respective ontologies do not flag as such. This heuristic checks the unit label for appearances of prefixes at the start of each word. That way, e.g., the prefix *kilo* will be detected in the units *kilometer* or *meter per kilogram* and they will be marked as prefixed. In addition after the mapping the heuristic will use the information from the other ontologies to find unlabeled prefixed units.

The next two phases are again straight forward to implement. In an additional phase before the manual review, however, all final results were presented in HTML files. In some cases this included the integration of different results, if they were closely related. Also on other occasions simply reordering the results by different criteria eases the review process.

5. Results

The analysis of the examined ontologies using the collection of scripts provides a deeper insight into the ontologies with respect to broadness and completeness. Furthermore, it revealed several issues of different types and some pitfalls specific to unit ontologies. The raw results are publicly available [22]. In this section first a statistical overview of the ontologies will be discussed, followed by a description of discovered issue classes and an overview of issues per ontology.

Note that Wikidata is constantly evolving and no specific version-numbers are issued. The numbers stated in the following for Wikidata will probably already have changed by the time of reading. They can, however, give a first impression about the amount of data managed by the project.

5.1. Statistics

For an initial overview the supported concepts and relations in the ontologies as well as the respective number of individuals are given in Table 1. Also included are the number of *non-prefixed* and *prefixed* units, which incorporate the results of the aforementioned heuristic to identify prefixed units.

The first notable observation is the absence of a connection between *prefixed units* and *prefixes* in MUO, QUDT and UO, even though both exist. First of all, despite containing both prefixed units and prefixes MUO, QUDT and UO do not model the connection between both. This hinders attempts to retrieve multiples or submultiples of a given unit. Furthermore, individuals for the concepts of *field of application*, *dimension* and *system of units* are just included in small subset of the ontologies.

OM is the only ontology with two versions present in this analysis. Most improvements between both OM versions are concerned with the simplification of the ontology structure, which is outside of the scope of this paper, but is documented in the SPARQL queries provided [21]. An important observation in the comparison of both versions is the absence of almost 40% of the individuals for kinds of quantities (qks) and systems of units (systems) in OM 2. This also has a major impact on the number of app⋈qk and qk⋈unit relations. Remarkably, the numbers for system⋈unit remain almost the same. This suggests, that the removed individuals for systems were poorly connected anyways. In contrast to the absence of many kinds of quantities and systems of units, there is a considerable increase of app⋈unit relations, whose number has more than doubled.

Just looking for the ontology with the most units WD takes first place with almost as much units as all other ontologies combined. Manual inspection reveals a large number of historic units to be present in WD, which might explain this large number. Another reason are possible duplicates that were not recognized due to the reasons mentioned in subsection 4.4. The second place is taken by OM, which does not suffer from these limitations. If the prefixed units are discarded, however, the situation changes in favor of QUDT. A closer look at the actual individuals in the ontologies suggests, that a systematic way has been used to create the prefixed pendants for all SI units in OM, which vastly increased the total number of units.

In the next step, the individuals for each concept were mapped to one another. Table 2 compares the amount of individuals present in the ontologies and the number of mapped distinct units overall. The number of found mappings especially for units is surprisingly low. To eliminate shortcomings of the mapping process itself the list of mapped units was carefully examined by hand, which led to a number of mappings added manually, but to no substantial overall change. Due to the wide range of languages we were not able to perform this manual duplicate search for WD to the same extent as for other ontologies. Therefore in some of the further analysis the values reported for WD are not as reliable as the ones reported for the other ontologies.

Following up on differences in the amount of units included and the lack of mappings, we analyzed the frequency in which units appear in the different ontologies. Table 3 lists the amount of prefixed and non-prefixed units, which are present in a certain number of ontologies. The actual number of widely used units was less than anticipated as over 85 %[17] of all units appeared in just one ontology.

Also notable is the fact that there seems to be little consent which units are essential to an ontology. Only 17 units are present in all ontologies.[18] This includes, e.g., no unit for *mass* as both *gram* (SWEET) and *kilogram* (MUO) are missing in one ontology.

To explore this even more, the overlap between the ontologies with respect to the included units is given in Table 4. It displays the fraction of both prefixed and non-prefixed units from one ontology, that are included in another.[19]

As can be seen, for any two given ontologies less than three quarters of the units of one ontology are part of the other. Even the smaller ontologies with respect to the number of units include units, that are not covered by the larger ones. There are different possible reasons for this fact. The ontologies were created for different projects or domains and hence needed to meet different demands regarding the included units. This is especially noticeable in the different sets of compound

[16]Information is included in the ontology, but not explicitly modeled using a specific concept or relation.

[17]Excluding WD lowers this number to 70 %.

[18]Units present in all ontologies: *ampere*, *candela*, *day*, *degree*, *degree Celsius*, *hour*, *kelvin*, *lumen*, *meter*, *minute*, *mole*, *newton*, *pascal*, *radian*, *second*, *siemens* and *volt*.

[19]A similar table showing the same relation but just for non-prefixed units can be created using the provided scripts. Although the exact numbers differ of course, the general trend remains the same.

Table 1

Number of individuals per concept and number of relations between two concepts per ontology. (A ⋈ B denotes a relation between concept A and B.)

	MUO	OBOE	OM 1	OM 2	QU	QUDT	SWEET	UO	WD
app	-	-	18	17	12	-[16]	49	-	-
conversion	-	64	827	832	674	55	37	-	603
dim	-	-	68	76	-	130	-	-	-
prefix	24	-	28	28	20	28	12	20	-
qk	78	67	879	549	598	235	847	324	1419
system	-	-	13	8	-	11	-	-	35
unit	267	138	1367	1384	801	768	140	221	4404
non-prefixed unit	264	71	371	388	433	652	111	111	4143
prefixed unit	3	67	996	996	368	116	29	110	261
app⋈qk	-	-	625	335	653	-	489	-	-
app⋈unit	-	-	410	992	-	-	-	-	-
dim⋈unit	-	-	1144	1157	-	491	-	-	-
prefix⋈unit	-	-	989	866	163	-	9	-	-
qk⋈unit	267	140	5157	4027	129	1093	1028	944	4177
system⋈unit	-	-	819	821	-	4124	-	-	166

Table 2

Comparison of the number of individuals as present in the ontologies with the number of distinct objects after the mapping.

	app	dim	prefix	qk	system	unit
extracted	96	274	160	4996	67	9490
mapped	76	140	29	3644	51	6413

Table 3

Frequency of unit occurrence in the tested ontologies.

	present in ontologies	1	2	3	4	5	6	7	8	9	total
WD incl.	#units	4771	1037	275	136	73	50	33	21	17	6413
	#non-prefix units	4448	299	111	80	29	34	25	15	17	5058
WD excl.	#units	1086	1008	257	90	65	33	24	17	-	2580
	#non-prefix units	786	198	137	41	46	24	17	17	-	1266

Table 4

Percentage of units from one ontology (row) that are also present in another ontology (column).

	MUO	OBOE	OM 1	OM 2	QU	QUDT	SWEET	UO	WD
MUO	100.0	10.3	40.7	41.1	28.1	31.6	12.9	17.5	32.7
OBOE	19.7	100.0	67.2	67.9	67.9	46.7	27.7	39.4	68.6
OM 1	7.8	6.7	100.0	99.9	21.6	18.4	6.2	9.6	25.9
OM 2	7.8	6.7	98.6	100.0	21.8	18.8	6.2	9.5	25.9
QU	9.3	11.7	36.9	37.8	100.0	28.9	10.3	13.0	26.3
QUDT	10.9	8.4	32.9	34.1	30.2	100.0	10.8	9.8	42.0
SWEET	25.2	28.1	62.2	63.7	60.7	60.7	100.0	43.0	52.6
UO	20.8	24.4	59.3	59.3	47.1	33.9	26.2	100.0	52.0
WD	2.0	2.2	8.1	8.2	4.8	7.3	1.6	2.6	100.0

units included.[20] Another reason might be the inclusion of prefixed units. While on the one hand ontologies will include next to no prefixed units (MUO), others seem to have systematically added all prefixed variations for a given unit (OM). Most ontologies, however, choose a more relaxed approach adding some prefixed units, but not all of them.

WD provides a wealth of languages for its labels – totaling over 300 different language tags including multiple regions for some of the languages. Following this observation we were interested in the frequency of each individual language. An excerpt of this is given in Table 5. Note that just the five most frequent languages for each type are shown for WD, while all present languages are listed for the other ontologies. Most ontologies choose to only add English labels. When the language tags were missing, the label turned out to be in English in all cases examined (QU, QUDT, UO). The only exception for domain ontologies is OM which also added Dutch and Chinese labels for a subset of individuals. In WD English is not as dominant as in other ontologies: Considering just the subset of analyzed individuals English is surpassed by German and rivaled at least by Russian in numbers. On the other end of the spectrum are OBOE and SWEET which omit labels altogether for most of their individuals.

The final examined aspect is the coverage of relations. It is important not only to include individuals for each type, but also link them together in an appropriate manner. All units, e.g., could be linked to at least one kind of quantity to represent what physical property can be measured using a particular unit. That statement, however, does not hold true for all relations: Not all units can be linked to a matching prefix as there is no prefix for non-prefixed units. This also shows that the both types within a single relation have to be considered individually in terms of coverage. Hence, in the example for each prefix there should at least one unit be linked to it. The current state of the ontologies is shown in Table 6.

Note that a frequency of 1.00 does not suggest that all possible links are included, but only that each individual is part of at least one link. To illustrate this restriction assume an ontology includes just two units `EX→meter` and `EX→foot` as well as two kinds of quantities `EX→Breadth` and `EX→Height`. To achieve a frequency value of 1.00 this ontology only needs to include two relations like, e.g., (`EX→meter`,`EX→Breadth`) and (`EX→foot`,`EX→Height`). The other two possible relations would not be need and their absence would not be detected by this metric.

A complete coverage in the discussed way is rarely achieved by any ontology. Noteworthy are especially those entries, that suggest no link at all despite both types being modeled. MUO, QUDT and UO although having both prefixes and units did chose not to link them at all. Similarly QU and SWEET did not link field of application and units. This begs the question why prefixes or fields of application were included in these ontologies in the first place. Other relations are usually given for just a subset of individuals with just few exceptions of a full coverage.

5.2. Encountered Issue Classes

The analysis revealed several classes of issues affecting the ontologies. Some are relevant to all kinds of ontologies, others are specific to unit ontologies. Following we describe common issue classes encountered, except the most obvious classes like missing, wrong or duplicated values, assignments or entities. An ontology specific discussion will follow in the next section.

Entity IRIs contain ontology version. The IRIs of entities contain the version number of the ontology. This is in contrast to the idea of Linked Data [23] as an update of the ontology will break references by other ontologies or systems. The version number should be contained in an additional version-specific IRI of the ontology [24].

Label concatenation. A single label contains the concatenation of multiple alternative names of a single entity, e.g., the label “tonne (metric ton)” actually represents two labels: “tonne” and “metric ton”. This concatenation might become a problem in automation and should be replaced by multiple or alternative labels.

Mix-up of equivalent and identical units. A conversion of factor one and offset zero does not imply two unit individuals to be the same. A unit also consist of other properties like the use within a system of units or the assignment to a kind of quantity. Consider the following:

```
Individual: :cubicDecimetre
 SameAs:
  :liter
 Facts:
  :unitOfSystem :SI
```

[20] In the absence of a respective modeling in most ontologies, exact numbers for the share of compound units could not reliably be extracted.

Table 5

Labels per language as present in the ontologies.(Σ ... total number of individuals; ∅ ... missing labels; ? ... labels without language tag)

	MUO	OBOE	OM 1	OM 2	QU	QUDT	SWEET	UO	WD
app			Σ: 18 en: 18 nl: 18	Σ: 17 en: 17 nl: 17	Σ: 12 ∅: 12		Σ: 49 ∅: 49		
dim			Σ: 68 en: 68 nl: 15	Σ: 76 en: 76 nl: 15		Σ: 130 ?: 130			
prefix	Σ: 24 en: 24		Σ: 28 en: 28 nl: 28 zh: 15	Σ: 28 en: 28 nl: 28 zh: 15	Σ: 20 ?: 20	Σ: 28 ?: 28	Σ: 12 ∅: 12	Σ: 20 ?: 20	
qk	Σ: 78 en: 78	Σ: 67 ∅: 62 en: 5	Σ: 879 en: 879 nl: 256 zh: 24	Σ: 549 en: 549 nl: 228 zh: 26	Σ: 598 ?: 598	Σ: 235 ?: 235	Σ: 847 ∅: 839 ?: 8	Σ: 324 ?: 324	Σ: 1419 en: 1313 de: 1073 ru: 1014 fr: 1011 nl: 929
system			Σ: 13 en: 13 nl: 13	Σ: 8 en: 8 nl: 8		Σ: 11 ?: 11			Σ: 35 en: 31 es: 28 de: 28 ru: 24 ja: 23
unit	Σ: 267 en: 267	Σ: 138 ∅: 138	Σ: 1367 en: 1367 nl: 1214 zh: 104	Σ: 1384 en: 1383 nl: 1223 zh: 106 ∅: 1	Σ: 801 ?: 801	Σ: 768 ?: 768	Σ: 140 ∅: 140	Σ: 221 ?: 221	Σ: 4404 de: 3227 en: 1824 ru: 1423 es: 1151 nl: 1142

This would erroneously imply that liter is a SI unit, which is wrong [20]. Similarly, a fraction is not identical to its simplified form (e.g., $\frac{m^3}{m^2}$ is not identical to m) as discussed in [25].

Namespace proliferation. An ontology or ontology family contains multiple namespace definitions for the same or similar namespaces or the same prefix is used for multiple namespaces. Assume the following two ontology files:

```
a.owl
Prefix ref: <http://example.org/b#>

b.owl
Prefix a:   <http://example.org/a>
Prefix ref: <http://example.org/a#>
```

The namespaces in `b.owl` differ only in whether or not to include the tailing `#`. Furthermore `ref` is used for two completely different namespaces.

SPARQL incompatible namespace. The namespace may be separated from the local name of an entity by using `#`, but the namespace definition does not include the `#`. Thus the namespace can not be reused in SPARQL as `#` is reserved for comments in SPARQL and has to be escaped each time a entity is referred using the namespace prefix.

Vague designation of units. The existence of several units of the same name in different systems of units has been ignored. For example, *acre* might refer to *acre (international)* as well as *acre (U.S. survey)*, which represent different sizes of an area [7]. The respective systems of units should be explicitly declared or alternatively, the conversions of a unit should be defined with respect to a specific system of units as done in WD. The disadvantage of the second option would be the lack of an IRI for a unit of a certain system of units and therefore impedes referring them.

Table 6

Individuals present in at least one relation.
Total represents the number of individuals of that type, while the respective frequencies are given as fraction of that total number.

		MUO	OBOE	OM 1	OM 2	QU	QUDT	SWEET	UO	WD
app	total	-	-	18	17	12	-	49	-	-
	⋈qk	-	-	1.00	1.00	1.00	-	0.14	-	-
	⋈unit	-	-	0.28	0.94	0	-	0	-	-
dim	total	-	-	68	76	-	130	-	-	-
	⋈unit	-	-	0.99	1.00	-	0.95	-	-	-
prefix	total	24	-	28	28	20	28	12	20	-
	⋈unit	0	-	1.00	1.00	0.75	0	0.67	0	-
qk	total	78	67	879	549	598	235	847	324	1419
	⋈app	-	-	0.67	0.58	0.98	-	0.46	-	-
	⋈unit	1.00	0.51	0.62	0.95	0.21	0.74	0.55	0.30	0.10
system	total	-	-	13	8	-	11	-	-	35
	⋈unit	-	-	0.62	1.00	-	0.73	-	-	0.63
unit	total	267	138	1367	1384	801	768	140	221	4404
	⋈app	-	-	0.25	0.65	0	-	0	-	-
	⋈dim	-	-	0.84	0.84	-	0.64	-	-	-
	⋈qk	1.00	1.00	0.77	0.77	0.16	0.98	0.61	0.76	0.90
	⋈system	-	-	0.56	0.55	-	0.73	-	-	0.03

Unspecified language. An ontology provides literals containing natural language without identifying the used language with a language tag. In automated processing this may cause language mixups if ignoring those literals is not an option.

Depending on the usage context of an ontology or the specific task it was created for, some of the listed issues might not be considered problems at all. For example, concatenated labels might not do harm or even be an advantage, if the ontology is just used for manual labeling. However, if the same ontology is used for automatic annotation, it is hard to work around that issue.

5.3. Ontology Assessment

Following we discuss each ontology with regard to encountered issues. A full list of issues found was sent to the authors of the respective ontologies. They were also asked about the status of the development, the intended handling of the issues and asked for permission to report on their answer. Helpful responses were provided by authors of OM and UO. They affirmed the ongoing development of both ontologies. The authors of OM fixed the reported issues immediately. All the other ontologies at the time of writing still contain the reported issues. As WD is a collaborative project we did not report discovered issues to specific authors. However, a more detailed verification and correction will be subject of future work.

MUO is one of the smaller ontologies and almost exclusively includes non-prefixed units. Since the data was imported from UCUM [7], most of the issues seem to be shortcomings of the import process or UCUM itself:

In six labels the Unicode symbol #160 ("no break space") was used instead of a spaces (e.g. in "calorie at 15␣°C"). This might cause problems when comparing labels to different inputs. We found two most likely unintended unit duplicates (`MUO→liter-2` and `MUO→the-number-ten-for-arbitrary-powers-2`). Although UCUM defines both `MUO→calorie` and `MUO→British-thermal-unit` just as aliases to other units, the ontology includes them as independent units. A typing error has been recognized in the IRI and the label of `MUO→pound-per-sqare-inch`. The entity `MUO→horsepower` is a case of vague designation.

OBOE is the smallest unit ontology included in the evaluation. Nevertheless we detected a substantial amount of issues. Most critical, ten conversions contained wrong factor or offset values. In eight of these cases the direction of the conversion has been mixed up resulting in an inverse conversion factor. In case of `OBOE→NanogramToKilogram` this would, e.g., result in a calculation error of factor 10^{24}. Less critical, one conversion (`OBOE→HectopascalToPascal`) has been defined without any conversion values. A typing error occurred in the IRI of `OBOE→MilliliterPerLIter`. Further, we found 14 units with a vague designation (e.g., `OBOE→Acre` and `OBOE→Gallon`). One conversion occurred twice (`OBOE→AcreToMeterSquared` and `OBOE→AcreToSquareMeter`). Three times equivalent and identical units have been mixed up, e.g. `OBOE→MeterCubed` and `OBOE→Kiloliter` have been connected using `owl:sameAs` but should only have a conversion with factor one and offset zero.

OM is the largest unit ontology. Thus a larger amount of issues is to be expected. However, OM stands out as most reported issues below have been fixed immediately. Due to further development and the resolving of reported issues updated releases have been published by the authors (1.8.3 and 2.0.3 respectively).

While we started our analysis with version 1.8, the latest version 2.0.3 is still affected by the following issues:

All IRIs contain the version number of the ontology. Four entities classified as units should be scales instead. For example, `OM→_1-10` should by a `OM→Ordinal_scale`. Five temperature scale individuals were connected to the according temperature scale class by `owl:equivalentClass` which makes them instances of themselves. Six temperature quantity classes erroneously contained the according temperature scale classes in their definition. For example, the definition of `OM→CelsiusTemperature` contained the following statement:

```
Class: om:CelsiusTemperature
 SubClassOf:
  om:hasValue only (om:usesUnit only
   (om:CelsiusTemperatureUnit or
    om:CelsiusTemperatureScale))
```

That caused the temperature scale classes to be a subclass of `OM→Unit`.

Further, we found the following issues that have been fixed in the current version (2.0.3):

OM 2 was inconsistent due to missing datatype declarations of property values of `OM2→hasFactor`. Therefore, OM 2 was not usable with a reasoner. Two conversions contained wrong factors. For example, the factor between `OM1→metre_of_mercury` and `OM1→pascal` was `1.33322e2` instead of `1.33322e5`. Three dimension individuals had wrong dimension vector values. For example `OM1→thermal_resistance-dimension` used the exponent -2 instead of 0 for the length component. Nine typing errors in labels, comments, symbols and alternative symbols have been found like “reciprocal metre kilogram” instead of “kilogram second to the power -2” in `OM1→kilogram_second-time_to_the_power_-2`. Seven entity IRIs contained typing errors like in `OM1→SI_luminous_intensity_dimension_dimension_exponent` a duplicated “dimension” breaking references to the entity. Four duplicated units were included like `OM1→metre_to_the_power_-2`, which is a duplicate of `OM1→reciprocal_square_metre`. We recognized 18 units with a vague designation like `OM1→foot`. The unit `OM1→`$g \cdot m^{-2} \cdot m^{-1}$ was missing m^{-1} in its composition definition. Two units contained a wrong dimension like `OM1→reciprocal_kelvin` with dimension `OM1→thermodynamic_temperature-dimension`. Nine labels were falsely tagged English instead of Dutch like “Gaussische eenhedenstelsel” on `OM1→Gaussian_system_of_units`. The property `OM2→hasFactor` contained the class `OM2→Prefix` as an additional domain that caused the most units to be a prefix to. The property `OM2→hasUnit` instead of `OM2→commonlyHasUnit` was used to assigne the unit `OM2→are` and the quantity kind `OM2→Area`.

QU is one of the medium sized ontologies. We found in `QU→xo15degCCalorie` one wrong conversion factor. We also recognized six units with concatenated labels, e.g., in the label “tonne (metric ton)” of `QU→tonneMetricTon`. The entity `QU→micrometreMicron` contains a mix-up of equivalence and identity of units and should be separated, since “micrometre” is a SI unit, but “micron” is not [19]. Three units (`QU→dryPintUs` and `QU→pintUsDry`, `QU→dryQuartUs` and `QU→quartUsDry`, `QU→perSquareMetre` and `QU→reciprocalSquareMetre`) occurred twice in QU. Further we found 14 units with vague designation.

QUDT, which is also one of the medium sized ontologies, contained the following issues: Most important, the ontology is inconsistent. For example, it contains the following definitions:

```
Class: qudt:Unit
 SubClassOf:
  qudt:typePrefix exactly 1 owl:Thing,
  qudt:typePrefix value "U"^^xsd:string
Class: qudt:ResourceUnit
 SubClassOf:
  qudt:Unit
Class: qudt:FinancialUnit
 SubClassOf:
```

```
qudt:ResourceUnit,
qudt:typePrefix value "UF"^^xsd:string
```

This means that every individual of the class `FinancialUnit` has the type prefix “UF”. Moreover, because it is also a individual of the class `Unit`, it has the type prefix “U”. As the number of type prefixes is, however, limited to one, this results in a conflict. As a result QUDT can not be used with a reasoner. Also critical, three conversions and two prefixes contained wrong factors. For example the factor of the binary prefix `QUDT→Pebi` was `125899906842624` instead of `1125899906842624` (2^{50}). We found for duplicates of units, like `QUDT→AtomicMassUnit` and `QUDT→Dalton2` for `QUDT→Dalton` [19]. Further `QUDT` $\xrightarrow{ind.}$ `AtomicMassUnit` was classified as `QUDT→AtomicPhysicsUnit` but should be a `QUDT` $\xrightarrow{class}$ `AtomicMassUnit`[21]. Typing errors occurred in `QUDT→RadianPerMinute`, which was labeled “radians per second” and in the namespace of `QUDT→schema/quantity#Mass` which should be `QUDT→vocab/quantity#Mass`. We found 38 units with vague designation.

SWEET contains only a few units more than the smallest ontology in our test, but likewise contains a lot of issues. Most adversely, the SWEET unit module file `reprSciUnits.owl` does not contain valid RDF[22], due to a duplicated ID. For the further evaluation, we removed one of the duplicates (line 635–638) to get a valid file. A second general issue is the namespace proliferation and the SPARQL incompatibility of some namespace definitions. Beside this general issues we found several entity specific issues:

The prefix `SWEET→micro` has no symbol. There are two duplicates like `SWEET→FTU` for `SWEET→FormazinTurbidityUnit`. Three conversions contain a wrong factor or a wrong offset. Two of them used the inverse values. The units `SWEET→FTU` and `SWEET→FNU` are marked as identical but are not. Further, a mix-up of equivalence and identity occurred in case of `SWEET→hertz` and `SWEET→perSecond` as well as `SWEET→micrometer` and `SWEET→micron`.

UO contained the following issues: The most important issue regarding UO is the lack of a stringent version management. Although most of the issues have been fixed or the erroneous concepts have been removed in the development repository[23] the updates were never made available at the location referenced by the ontology’s IRI. The development version for example contains most missing links between unit and prefixes, as described in subsection 5.1, whereas they are missing in the released version. Our analysis is based on the released version as we believe most users would resort to this one. Four wrong subclass assignments have been found, like e.g. `UO→UO_0000325` (“megaHertz”) is subclass of `UO→UO_0000106` (“hertz”). We found nine violations of the UO naming convention as applied to all other entities of the same type. E.g. the class `UO→UO_0000219` containing units related to electric charge is named “electric charge” instead of “electric charge unit”. Further UO uses a special kind of annotation properties[24] to categorize classes as units or groups of units. We found 23 cases of missing or wrong assignments of this entity classification. We also found a heterogeneous usage of the property `UO→hasExactSynonym` for alternative names and unit symbols. Though this is compliant to the definition of the property, it might become challenging in applications.

WD Please note that our analysis of WD was not as detailed as the analysis of the other ontologies. This will be part of future work. However we found the following issues:

WD contains links to entities from OM 1 and QUDT. The links to QUDT, however, point to the URLs of an HTML representation of QUDT, but do not refer the entities themselves. The references to OM 1 contain one wrong entry: The unit `WD→Q734439` (“hour angle”) points towards `OM→Hour_angle`, which is a kind of quantity, but should point to `OM→hour-hour_angle`. The conversion of “1 E-16 s” to “second” contained a wrong precision. Eight conversions contained a wrong factor like from `WD→Q21014455` (“metre per minute”) to `WD→Q182429` (“meter per second”) with the factor 0.166666667 instead of 0.0166666667. Further the conversion offset in the conversion of “degrees Celsius” to “Kelvin” was missing.

---

[21]QUDT contains two entities with this name: an individual `http://qudt.org/vocab/unit#AtomicMassUnit` and a class `http://qudt.org/schema/qudt#AtomicMassUnit`

[22]RDF Validation: https://www.w3.org/RDF/Validator

[23]https://github.com/bio-ontology-research-group/unit-ontology/blob/master/uo.owl

[24]The approach, which is also used in the Gene Ontology, is explained on http://geneontology.org/page/go-slim-and-subset-guide.

6. Limitations

This approach contains unavoidable limitations: Most important, any particular individual or relation, which is only part of one ontology, can not be compared and hence be validated. Thus possible errors in that case can not be detected. This situation may, however, improve in the future, if new ontologies emerge or existing ones get extended.

Another issue is manual involvement. Currently only mismatches between the ontologies are highlighted. Afterwards, the mismatches have to be manually examined to decide upon them. This, of course, opens up the possibility of human error, which might lead to missing out on some errors or false positives.

One occasion, where deciding upon the correct value can be difficult are conversion factors. Through it is usually easy to identify a number that is way off, certain differences in rounded values are harder to spot and verify. Some stem from minor deviations in the definitions for units of the same name, but belonging to different systems of units. Others, however, can be traced back to varying degrees of precision used by the ontology authors. Furthermore, as for some conversion no precise value can be given using floating point arithmetics[25], it remains an open issue, whether those values are correct in a strict sense.

As previously noted, the coverage of concepts and individuals by the different ontologies is diverse. Therefore, the overlap for some concepts is limited. As a consequence, some characteristics, that might be of interest in the ontology selection process, could not be acquired to a meaningful degree and hence are left out for the time being. Examples include *dimensions*, the distribution of units over the various *fields of application* or the amount of compound units. In conjunction with the relatively small overlap in units as shown in Table 4, this would not result in any reliable comparison of the different ontologies.

In the development of the implementation some errors were discovered, which did not lead to a separate automatic script. So, e.g., some natural language descriptions were attached to the wrong individual or contained typos and wrong information. To automatically detect this natural language processing would be required, which is outside the scope of this work.

[25]E.g., the conversion factor between *degree Fahrenheit* and *Kelvin*, which equals $\frac{5}{9}$ or between *degree* and *radian*, which is $\frac{\pi}{180}$.

7. Suggestions

The evaluation of the nine ontologies or knowledge bases uncovered several modeling pitfalls in the measurement unit domain. In this section we try to give suggestions on how to avoid them.

Careful design and documentation of conversion properties helps both, users and developers, to prevent misinterpretations of conversions. Most ontologies employ two properties for the factor and the offset in relation to an implicit or explicit reference unit to define a conversions. Unfortunately, four reasonable interpretations of those values are possible, as shown in Figure 7.

Given that in most cases the offset value is zero, the problem is usually reduced to two reasonable interpretations. However, two ontologies (OBOE and SWEET) used more than one interpretation within their conversion definitions. This highlights the importance of a proper definition and its documentation for error prevention and later usage. Ideally the modeling approach allows only one reasonable interpretation. The authors of OM 1 employed definition individuals to model conversion that allow only one interpretation:

```
Individual:  :litre
 Facts:
  :definition  :_1.0e-3_cubic_metre
Individual:  :_1.0e-3_cubic_metre
 Facts:
  :unit_of_measure_or_measurement_scale
                            :cubic_metre,
  :numerical_value  "1.0e-3"
```

However, in favor of a leaner structure the additional definition individuals have been removed in OM 2.

Precision and numerical stability of conversions should also be concerned. In some cases a simple change of the direction of the modeled conversion can improve the accuracy (e.g., 1.8 instead of $0.\overline{5}$ for the conversion between *degree Fahrenheit* and *Kelvin*). If that is not possible, one could think about representing problematic values with fractions given by numerator and denominator instead of a single floating point number. Besides this issue, almost none of the considered ontologies provides information about the precision of conversion values, which again makes it hard for a potential user to evaluate the adequacy for the given use case. Only WD allows to define the precision of values by stating a lower and an upper bound.

Naming conventions, particularly for the URI, can prevent duplicates and simplify usage. This corre-

$$value\ stated\ in\ a = value\ stated\ in\ b \times factor + offset$$

$$value\ stated\ in\ a = (value\ stated\ in\ b + offset) \times factor$$

$$value\ stated\ in\ a \times factor + offset = value\ stated\ in\ b$$

$$(value\ stated\ in\ a + offset) \times factor = value\ stated\ in\ b$$

Fig. 7. Reasonable interpretations of conversion factor and offset.

sponds to [10]. The denominator of a fraction for a derived unit can be named in several ways: Using "per", "reciprocal", "inverse" or "to the power of -1". Furthermore, there might be one modifier per part of the divisor or just one modifier for the whole divisor. Similarly, for exponents the terms "squared" and "cubed" or the general phrase "to the power x" can be used. Moreover the order of base and exponent, as well as the order of all factors or the position of an addition (e.g., the system of units in "U.S. survey mile" or the kind of quantity in "second (angle)") can be ambiguous. This highlights the need for a well defined and enforced naming scheme to prevent multiple individuals for the same unit. We do not suggest to assign only one label per entity, but provide different labels and adhere to a systematic approach. Multiple names should never be assigned using only one (concatenated) label.

8. Conclusion

We presented the first in-depth evaluation of unit ontologies considering instances and included facts. We evaluated eight major ontologies for units of measurement and the relevant parts of the Wikidata corpus. The gathered information is important for projects that intend to use a unit ontology as it accelerates the decision making process in selecting a suitable ontology.

The evaluation is based on a extensible collection of scripts. Due to the modular nature of the implementation, it can easily be extended by adding further scripts for new checks and comparisons or by adding other ontologies to be evaluated. The implementation [21] and the presented results [22] are publicly available. Every interested party is welcome to use, review or improve them. This way both developers as well as users of unit ontologies now and in the future can use these results to choose or improve their ontologies.

The analysis provided not only statistical information about the ontologies, but also highlighted several issues within. The respective ontology authors were contacted and in at least one ontology this led to a new release, which fixed the reported issues. As a byproduct, a mapping between the evaluated ontologies was created, that can be used to integrate or migrate between the involved ontologies.

During the analysis we identified several classes of issues concerning the development of ontologies—not limited to the domain of units of measurement. We hope that the characterization of this issue classes will help ontology authors to prevent them in future.

Our future work will focus on a more detailed analysis of Wikidata. Major challenges are the detection of wrong labels in the wealth of more than 300 different languages and the mapping of identical entities with disjoint label language sets.

Acknowledgments

Part of this work was funded by DFG in the scope of the LakeBase project within the Scientific Library Services and Information Systems (LIS) program. We specially thank the ontology author Hajo Rijgersberg (OM) for the feedback on our work.

References

[1] Mars Climate Orbiter Mishap Investigation Board Phase I Report, Technical Report, NASA, 1999.

[2] J. Madin, S. Bowers, M. Schildhauer, S. Krivov, D. Pennington and F. Villa, An ontology for describing and synthesizing ecological observation data, *Ecological Informatics* **2**(3) (2007). doi:10.1016/j.ecoinf.2007.05.004.

[3] H. Rijgersberg, M. van Assem and J. Top, Ontology of Units of Measure and Related Concepts, *Semantic Web* **4**(1) (2013), 3–13. doi:10.3233/SW-2012-0069.

[4] R. Raskin, Semantic Web for Earth and Environmental Terminology (SWEET), in: *Earth Science Technology Conference*, 2003.

[5] G.V. Gkoutos, P.N. Schofield and R. Hoehndorf, The Units Ontology: a tool for integrating units of measurement in science, *Database* **2012** (2012). doi:10.1093/database/bas033.

[6] D. Vrandečić and M. Krötzsch, Wikidata: a free collaborative knowledgebase, *Communications of the ACM* **57**(10) (2014), 78–85. doi:10.1145/2629489.

[7] G. Schadow and C.J. McDonald, The Unified Code for Units of Measure, 2014, visited on 26 Apr. 2016. http://unitsofmeasure.org/ucum.html.

[8] United Nations Economic Commission for Europe (UNECE), Codes for Units of Measure Used in International Trade, Recommendation, 20, UN/CEFACT Information Content Management Group, 2009, Revision 6.

[9] J. Lehmann, R. Isele, M. Jakob, A. Jentzsch, D. Kontokostas, P.N. Mendes, S. Hellmann, M. Morsey, P. Van Kleef, S. Auer et al., DBpedia–a large-scale, multilingual knowledge base extracted from Wikipedia, *Semantic Web* **6**(2) (2015), 167–195. doi:10.3233/SW-140134.

[10] A. Gómez-Pérez, Evaluation of ontologies, *International Journal of Intelligent Systems* **16**(3) (2001), 391–409. doi:10.1002/1098-111X(200103)16:3<391::AID-INT1014>3.0.CO;2-2.

[11] H. Rijgersberg, M. Wigham and J. Top, How semantics can improve engineering processes: A case of units of measure and quantities, *Advanced Engineering Informatics* **25**(2) (2011), 276–287. doi:10.1016/j.aei.2010.07.008.

[12] M.P. Foster, Quantities, Units and computing, *Computer Standards & Interfaces* **35**(5) (2013), 529–535. doi:10.1016/j.csi.2013.02.001.

[13] C. Do and E.J. Pauwels, Using MathML to Represent Units of Measurement for Improved Ontology Alignment, in: *Intelligent Computer Mathematics - MKM, Calculemus, DML, and Systems and Projects 2013, Held as Part of CICM 2013, Bath, UK, July 8-12, 2013. Proceedings*, Lecture Notes in Computer Science, Vol. 7961, Springer, 2013, pp. 310–325. doi:10.1007/978-3-642-39320-4_21.

[14] S. Samadian, B. McManus and M.D. Wilkinson, Automatic detection and resolution of measurement-unit conflicts in aggregated data, *BMC Medical Genomics* **7**(1) (2014), 12. doi:10.1186/1755-8794-7-S1-S12.

[15] M.D. Steinberg, S. Schindler and J.M. Keil, Use Cases and Suitability Metrics for Unit Ontologies, in: *OWL: Experiences and Directions - Reasoner Evaluation - 13th International Workshop, OWLED 2016, and 5th International Workshop, ORE 2016, Bologna, Italy, November 20, 2016, Revised Selected Papers*, 2016, pp. 40–54. doi:10.1007/978-3-319-54627-8_4.

[16] X. Zhang, K. Li, C. Zhao and D. Pan, A survey on units ontologies: architecture, comparison and reuse, *Program* **51**(2) (2017), 193–213. doi:10.1108/PROG-08-2015-0056.

[17] M. Poveda-Villalón, A. Gómez-Pérez and M.C. Suárez-Figueroa, OOPS! (OntOlogy Pitfall Scanner!): An On-line Tool for Ontology Evaluation, *International Journal on Semantic Web and Information Systems* **10**(2) (2014), 7–34. doi:10.4018/ijswis.2014040102.

[18] Joint Committee for Guides in Metrology, *International vocabulary of metrology*, 3rd edn, 2012.

[19] A. Thompson and B.N. Taylor, Guide for the Use of the International System of Units (SI), Technical Report, National Institute of Standards and Technology, 2008.

[20] B.N. Taylor and A. Thompson, *The International System of Units (SI). National Institute of Standards and Technology Special Publication 330 (2008 Edition)*, National Institute of Standards and Technology / U.S. Department of Commerce, 2008.

[21] S. Schindler and J.M. Keil, Unit Ontology Review v1.3.0, Zenodo, 2017. doi:10.5281/zenodo.823686.

[22] S. Schindler and J.M. Keil, Unit Ontology Review v1.3.0 Results, Zenodo, 2017. doi:10.5281/zenodo.823675.

[23] C. Bizer, T. Heath and T. Berners-Lee, Linked Data - The Story So Far, *Int. J. Semantic Web Inf. Syst.* **5**(3) (2009), 1–22. doi:10.4018/jswis.2009081901.

[24] C. Bock, A. Fokoue, P. Haase, R. Hoekstra, I. Horrocks, A. Ruttenberg, U. Sattler and M. Smith, OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax, Recommendation, 27, W3C, 2009. https://www.w3.org/TR/2009/REC-owl2-syntax-20091027/.

[25] I. Johansson, Metrological thinking needs the notions of parametric quantities, units and dimensions, *Metrologia* **47**(3) (2010), 219–230. doi:10.1088/0026-1394/47/3/012.