Quality Assurance of Ontology Content Reuse

Michael Halper[1], Christopher Ochs[2], Yehoshua Perl[1], Sivaram Arabandi[3], Mark A. Musen[4]

[1]NJIT
Newark, NJ 07102 USA
{michael.halper, yehoshua.perl}@njit.edu

[2]Nokia Bell Labs
Murray Hill, NJ 07974 USA
christopher.ochs@nokia-bell-labs.com

[3]Ontopro LLC
Houston, TX 77025 USA
sivaram.arabandi@gmail.com

[4]Stanford University
Stanford, CA 94305 USA
musen@stanford.edu

***Abstract*—*Building ontologies is difficult and time-consuming. As such, content reuse has been promoted as an important guiding principle in ontology development. Reusing content from other ontologies can reduce the overall effort involved in new ontology construction and provide better alignment with existing knowledge modeling. However, reuse is not a panacea, and it comes with its own attendant difficulties. In this paper, we investigate some common quality assurance issues associated with reuse, such as duplicated content and versioning problems. Some heuristic-based approaches are proposed for analyzing ontologies for these kinds of quality assurance issues. An analysis is carried out on a sample of the large collection of BioPortal-hosted ontologies, many of which employ reuse. The findings indicate that curators and authors, particularly those new to the reuse process, should be on the alert when developing an ontology with reused content to avoid introducing problems into their own ontologies.*

***Keywords*—*ontology; modeling; ontology reuse; ontology quality assurance; BioPortal*

I. Introduction

Ontology reuse is a well-established design pattern. An ontology author may reuse content to save on development time and effort, promote interoperability with other ontologies, and ensure that a consistent representation of a domain is included in their ontology. Support for importing and reusing ontology content is included in the Web Ontology Language (OWL) (through the use of *owl:imports* axioms) [1], and the paradigm is supported by the Protégé ontology editing environment [2]. Top-level ontologies such as the Basic Formal Ontology (BFO) [3] were designed specifically to support content reuse and alignment of ontologies. Top-domain ontologies, like the Ontology for General Medical Sciences (OGMS) [4] and BioTop [5], extend the BFO and add general domain knowledge that can also be reused by an ontology author.

While there are enormous benefits to reuse, an ontology author also needs to be keenly aware of potential issues that can affect the quality of the resulting ontology. There may be unintended consequences if reused content is not incorporated correctly or not maintained properly. In previous studies [6], we investigated the issue of quality assurance (QA) in the context of the Sleep Domain Ontology (SDO) [9], the Ontology for Drug Discovery Investigations (DDI) [10], and the Cancer Chemoprevention Ontology (CanCo) [11]. These ontologies all reused content from other ontologies (e.g., BFO), and in the context of these ontologies, some of the QA problems we encountered related to content reuse.

In this paper, we focus strictly on such ontology QA problems and investigate a broader collection of ontologies that reuse content. The main purpose for this study is to alert curators and authors, especially those new to the process, of the pitfalls of reuse in terms of the errors that they are likely to encounter. This awareness will help in avoiding the errors in the first place and enhancing the content of their own ontologies. Let us note that ontology errors can come in a wide range of severity and causes, such as with unsatisfiability, incoherence, and inconsistency of concepts. Even so, we will use the term "error" throughout this paper, though one may argue in certain circumstances whether an irregular modeling issue truly warrants that designation.

Moreover, let us state at the outset that ontology development is intrinsically difficult, and the findings that we present are in no way meant as indictments of anyone's work. In fact, some of the errors reported arose from the work of one of the co-authors (SA), who took great care in the construction of the SDO. Ontology developers have the best intentions to do a good job and take great pains to review their work. Even with that being the case, the inherent complexity of ontology design and the reuse of content makes the appearance of errors almost inescapable. It is our intention to alert ontology maintenance personnel to this fact through the results of our study. Additionally, we are not criticizing reuse in ontology design, with its numerous advantages. We just wish to caution ontology designers to be careful about the potential disadvantages and pitfalls of reuse.

Our focus is on the collection of ontologies hosted in BioPortal [12]. The specific QA issues that we wish to examine are duplicated content (including duplicated classes and properties), versioning problems with respect to source ontologies of reuse, and mechanical import errors. The heuristic methods that were used in our analyses are described, and our findings from among the BioPortal ontologies are reported.

II. Background

A. Prior Reuse QA

Ochs *et al*. [6] performed a QA review of the SDO and discovered several significant issues related to the import of content from other ontologies. For example, pairs of duplicated classes (e.g., two *Clinical finding* classes and two *Organism* classes), originating from different ontologies, were found and corrected. However, on revisiting the SDO using a change analysis methodology called a *diff partial-area taxonomy* [13], which visually summarizes the differences between two releases of a given ontology, several additional QA issues related to the reuse of content were uncovered.

These preliminary studies, along with further discussions with ontology authors and maintainers, motivated the research described herein. The reuse design pattern, and the way it is applied, can have serious, unintended impacts on an ontology. The advantages of reusing content often come with a cost to the quality of the overall ontology.

B. Prior Analysis of Ontology Reuse

Previous studies have reviewed the existence and prevalence of ontology reuse. Kamdar *et al*. [14] analyzed term reuse among ontologies and noted several error patterns with ontology reuse. Ghazvinian *et al*. [15] reviewed the orthogonality of the OBO Library [16] ontologies. Ochs *et al*. [18] investigated how reused content is utilized in a sample of 355 ontologies in BioPortal.

Among the ontologies in BioPortal, reuse of the BFO, an upper-level ontology, is somewhat common. This is expected given the principle of a "commitment to collaboration" espoused by the OBO Foundry [16]. Content reuse from top-domain ontologies, like the OGMS [4], and domain-specific ontologies, like GO [19] and ChEBI [20], is also fairly common.

C. Methods of Ontology Reuse

There are several ways an author of an ontology *O* can reuse content from a source ontology *S*. Each method of reuse has several advantages and disadvantages, particularly in relation to maintaining and updating reused content. Content included from another ontology may be updated periodically at its source. Corrections of errors and inconsistencies performed during maintenance of the source ontology *S* will need to be propagated to *O*. While an ontology like the BFO may be updated only once every several years (e.g., BFO 1.1 was released in 2009 and BFO 2.0 was released in 2015), other ontologies are updated much more frequently. ChEBI and GO, which Ochs *et al*. [18] found to be reused by 37 and 33 ontologies, respectively, are updated quite frequently: ChEBI, almost every month, and GO, on a daily basis (though a new version may only be published monthly).

An ontology author may include content via the *owl:imports* mechanism defined in OWL syntax, and implemented in the OWL API [1]. This approach includes the entire contents of *S* into *O* "on the fly," which allows updates in *S* to be included in *O* without work from the author of *O*. However, there may be unexpected consequences downstream, especially after classification.

Alternatively, the author of *O* may reuse a fixed version of *S*'s content (either the complete contents of the ontology or a selected subset of the ontology, extracted using, e.g., the MIREOT approach [21]). Reusing a fixed version of the content provides the author of *O* with greater control over when reused content is updated, at the expense of making it labor intensive to align changes from *S* into *O*.

III. Methods

In this study, we reviewed a collection of ontologies from BioPortal, looking for errors and inconsistencies arising from the reuse of content from other ontologies. In analyzing the collection, we employed several heuristic-based methodologies to determine the prevalence of duplicate content, versioning problems, and any import issues. The collection that was examined was extracted from the 355 ontologies studied by Ochs *et al*. [18], which were obtained from BioPortal in April 2015. Specifically, the collection consisted of the 197 ontologies (55.5%) that were found to reuse content.

We define a *source ontology* as an ontology that has content included in another ontology *O*. As in Ochs *et al*. [18], we define reuse according to the URIs of the entities in an ontology. For each ontology in this study, we identified its *base URI* (e.g., the base URI of the BFO is *http://purl.obolibrary.org/obo/bfo*). Similarly, the base URI of the SDO is *http://mimi.case.edu/ontologies/2009/1/SDO.owl*. In general, all of the entities in an ontology have a URI that starts with the ontology's base URI. Different versions of an ontology may have different base URIs. For this study, an entity (i.e., class or property) was considered reused if it had a different base URI from the ontology it is residing in (e.g., a BFO class in SDO will have the BFO base URI). In this study, we did not distinguish between content imported directly and content imported by transitivity.

In the following, we describe the kinds of errors that were sought and the approaches to finding them. Examples from the SDO are used to illustrate how each type of error may manifest itself during the ontology editing process. Additionally, for each kind of error, we describe the heuristic-based approach that we utilized to determine the prevalence of the error among the set of 197 BioPortal ontologies.

A. Duplicated Content

An author may reuse content from multiple ontologies. If content from two ontologies is reused, and the ontologies cover a similar domain, the potential exists for the inclusion of duplicate classes (i.e., the author could inadvertently include two classes, from two different ontologies, that represent the same concept). This kind of duplicate information is not desired. As mentioned previously, we identified several pairs of duplicated classes in the SDO [6].

Class duplication can cause significant issues. For example, the abovementioned duplicate *Clinical finding* classes in SDO had the same name and represented the same entity but were not set equivalent, and their restrictions were not the same.

This issue will cause problems for both users and authors alike, as they will typically not suspect a duplicate and will likely not suspect that classes representing the same entity will have different modeling within a single ontology.

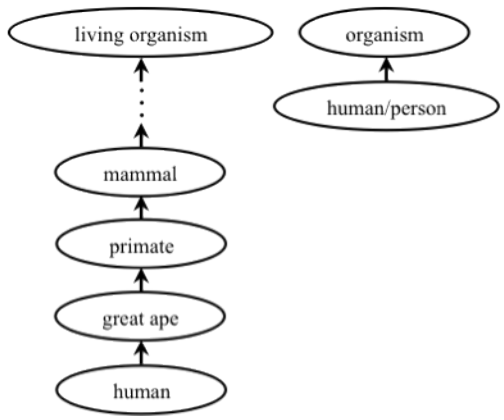

Fig. 1. Hierarchical paths between classes *human* and *organism* in BioTop (left) and CPRO (right).

Beyond individual classes being duplicated, two source ontologies may have subhierarchies of duplicate (or very similar) classes, often modeled with different levels of granularity. For example, in the SDO, we found duplicated classes for *organism* and *human*, originating from BioTop and CPRO [22]. (Actually, the terms are slightly different in each: *living organism* vs. *organism* and *human* vs. *human/person*, respectively.) In BioTop, *living organism* is a distant ancestor of *human*; there are seven other ancestor classes on the ancestry path between them (e.g., *great ape, primate,* and *mammal*). In CPRO, *human/person* is a direct subclass of *organism*. See Fig. 1. The two versions of *human* have different relationship structures. The one on the left has a defined *participates in* relationship. The one of the right does not, though its two children, *patient* and *physician*, do have have the relationship. The use of one *human* version alone may lead to deficient modeling in an application.

Duplicate classes can also be unknowingly included. The duplicate content may be imported by transitivity, i.e., an ontology was reused by another reused ontology and the author may or may not have been aware of this. Different versions of the same ontology may be reused. For example, as we report below, we identified several ontologies that appear to import content from multiple versions of the BFO.

Duplicate properties can also be introduced into an ontology via reuse. Let us point out that the presence of duplicate properties in itself is not necessarily an error. It is the inconsistent use of such properties that constitutes an error. This situation is analogous to the software engineering scenario where multiple libraries are imported; in such a case, there is a high potential for similar functions to be present.

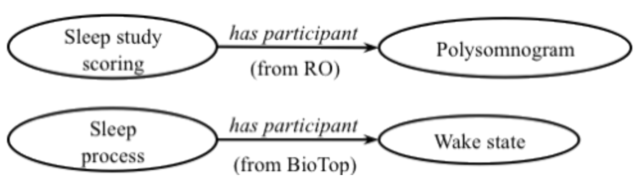

Fig. 2. Two examples of SDO classes using the *has participant* property.

As with duplicate classes, the introduction of duplicate properties may be due to the fact that two ontologies cover a similar domain. For example, in the SDO, there are *has participant* object properties included from the Relations Ontology (RO) [23] and BioTop, both of which represent the same kind of relationship. Both properties are utilized in the modeling of the SDO. Some SDO classes have restrictions using the RO version of *has participant*, while other SDO classes have the BioTop version. See Fig. 2 for some examples of this. These properties were not defined as equivalent in the SDO.

To identify ontologies with duplicated classes, we can utilize two heuristic-based methods. (These methods can also be used to identify duplicate properties). First, an ontology may have duplicate classes if it reuses classes from two source ontologies that cover a similar, or identical, domain. For example, if an ontology reuses classes from FMA [24] and Uberon [25], ontologies that model the domain of anatomy, then there is a greater chance of finding duplicate classes than in ontologies that reuse content from only one ontology.

Second, if an ontology reuses two classes with the same label, but those classes originate from different source ontologies, then they may be duplicates. One can search for all pairs (or, in general, sets) of classes where the label is the same but the URIs of the classes are different. While this method potentially returns many false positives—e.g., "cold," as in temperature, and "cold," as in the disease, which are expected to have different URIs and may be modeled in different domains—it provides an indicator for a potential problem.

B. Versioning Problems

There is also the potential of versioning problems when reusing properties. In general, ontologies should reuse content consistently from a single version of a source ontology. However, an ontology may inadvertently include content from multiple versions of the same source ontology. This may occur due to import by transitivity. For example, the SDO includes multiple versions of the *part of* property: one from an old version of RO included via FMA and another from a more recent version of RO via CPRO. Again, both of these properties represent the same relationship. See Fig. 3 for illustrations. Below, we identify several ontologies that reuse content from multiple versions of the BFO.

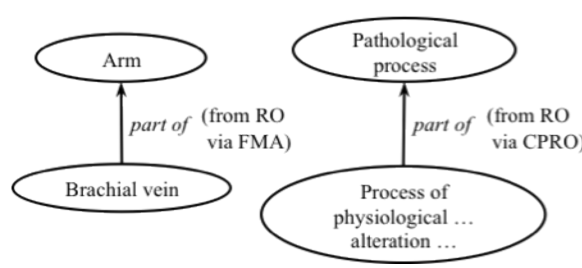

Fig. 3. Property *part of* from two versions of RO in the SDO.

To analyze inconsistent versioning, we can identify the base URI of every ontology in our data set, under the heuristic that a different base URI indicates a significantly different version of the same ontology. A number of source ontologies, such as the BFO, GO, FMA, and others, were found to have multiple base URIs among the BioPortal ontologies. For example, in Ochs *et al*. [18], we identified six base URIs for FMA, and below we describe several base URIs for BFO. We mapped each source ontology to its set of base URIs. If an ontology *O* included entities from the same source ontology *S*, but the entities had different base URIs, the ontology is considered to have a reuse versioning problem.

C. *Owl:imports Errors*

OWL's *owl:imports* mechanism enables an ontology author to include external ontologies without defining the classes and properties in their own ontology. The entire external ontology will be included when the importing ontology is opened (e.g., in the OWL API). However, if the URI for the source ontology is not correct, or the ontology is no longer available at the specified URI, then the source ontology cannot be loaded.

To investigate issues related to *owl:imports* errors, we opened every ontology with the OWL API and logged which ontologies encountered an error related to a missing *owl:imports* file(s).

IV. RESULTS

Our analysis of the various kinds of errors resulting from reuse was carried out on the 197 ontologies in BioPortal that were found to reuse content by Ochs *et al*. [18]. See [12] for more information pertaining to the individual ontologies referred to in this section.

A. *Duplicate Classes and Properties*

Reuse of classes from multiple sources is common, with an average of more than five sources [18]. But we found that it is relatively uncommon for an ontology to reuse classes from two or more ontologies that cover a similar domain. However, when we investigated cases where ontologies did reuse such content, there were several potential errors. For example, the Cell Line Ontology (CLO) reuses content from the FMA and Uberon. In it, we found several potential duplicate class pairs. For example, there is *Scalp* from Uberon and *Scalp* from FMA. There were also duplicate *Pelvis* classes from EFO and Uberon. Many such classes are related using class equivalence axioms (e.g., *Amnion, Colon,* and *Intestine*). However, other duplicate classes are not related in this way (e.g., *Scalp, Aorta,* and *Liver*). Analyzing these examples, one can see that CLO includes the *Anatomical structure* subhierarchy from Uberon and the *Organism part* subhierarchy from EFO. In such a case, the potential exists for additional duplicate classes.

When looking for pairs of classes with the same label but different base URIs, we found that class duplication does not occur frequently. In total, 149 ontologies were found to reuse at least one class from another ontology. Among the 149 ontologies, 46 ontologies (30.1%) contain at least one potential duplicate pair based on our criteria. In general, we found very few such pairs in a given ontology. Most of the 46 ontologies either have just a single pair or between two and ten pairs. We did find several ontologies (e.g., CLO, CSEO, and SYN) that have many such pairs, and these ontologies reuse content from multiple ontologies that cover the same—or similar—domains.

For example, in the Synapse Ontology (SYN), there are many (apparently) duplicated classes reused from NCIt and CL (e.g., pairs of *acinar cell* classes). Within SYN, we found three separate *Cell* subhierarchies. One subhierarchy, from GRO, consists of two classes. The other two *Cell* subhierarchies, from NCIt and CL, are much larger. There are no equivalences set between the classes in these subhierarchies. For the use case of SYN, this might be an intentional design decision, but from an ontology design perspective, it is not typical compared to other ontologies that reuse NCIt, CL, etc.

The CSEO contains over 200 potential duplicate class pairs. It includes a large portion of the *Disease* subhierarchy from NCIt and defines its own *Finding* subhierarchy. In these two subhierarchies, there are many similar classes (e.g., *Abscess*) that represent diagnoses. Looking more closely, we found additional pairs of duplicate diagnoses. Similarly, many classes related to various kinds of anatomical structures and tissues (e.g., *Tongue* and *Uterus*) are included from NCIt and added in CSEO. In all of these cases, there are no connections (e.g., equivalences or restrictions) to indicate that these pairs of classes are related to one another. On the other hand, CSEO does define equivalences between classes reused from NCIt and classes reused from UO (e.g., *Lux* and *Liter*).

For duplicated properties, we found 31 ontologies with properties that have the same label and different base URIs. Twenty of these (64.5%) were found to contain one or more pairs of duplicated properties. For instance, ENM contains several pairs of duplicated properties from BAO, RO, and NPO (e.g., properties named *derives from* and *has part*).

B. *Versioning Problems*

The large majority of cases of reuse that appear to have versioning problems, based on different base URIs, were found among ontologies that reuse the BFO and RO. We identified eleven BioPortal ontologies (3.1% of all the ontologies in the BioPortal at the time) that included classes from multiple versions of the BFO. For example, the DDI uses all 39 classes from an OWL release of BFO and one class from a version of the BFO with an OBO URI. Fig. 4 shows eight examples of ontologies that include classes from multiple versions of the BFO. Two of the ontologies, CHEMINF and TEO, include all of the content from two versions of the BFO.

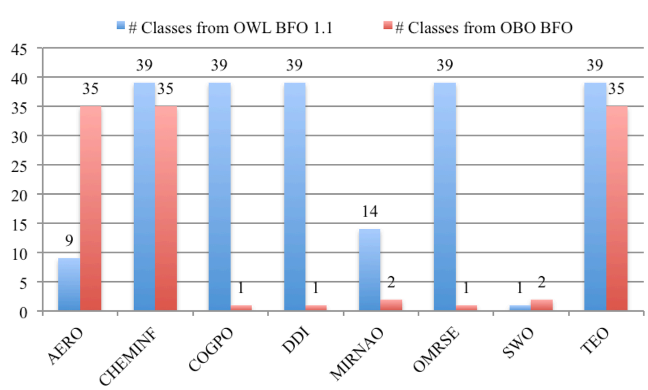

Fig. 4. Example ontologies reusing content from multiple versions of the BFO

The reuse of classes from multiple versions of non-BFO ontologies was relatively uncommon. We identified a few ontologies that included classes from multiple versions of the same ontology. For example, COGPO and DDI include classes from multiple versions of PATO and UO. These classes are not set equivalent. In cases where multiple versions of an ontology appear, the numbers of classes reused from each are typically disproportionate. For example, the Cell Culture Ontology (CCONT) includes classes from multiple versions of EFO (one class, *obsolete normal*, from one version and 4,882 classes from another). Both ENM and EP include multiple versions of PATO. In the case of EP, 48 classes are included from one version and 1,570 classes from another. HUPSON includes one class from one ChEBI version, and 83 classes from another. MF includes classes from multiple versions of NBO; MIRNAO includes several classes from multiple versions of the GO.

We found many different versions of the RO, OBO REL, and BFO properties (e.g., *has part* and *part of*) reused in our data set. Along with SDO, we found several ontologies that reuse properties from multiple versions of these ontologies (often in class restrictions). Consider, for example, the *has part* property. We identified 14 versions of this object property in our dataset (see Table I). Reviewing the ontologies enumerated in Table I, we identified a total of 20 ontologies that include multiple versions of these (and other) RO relationships. Four ontologies, namely, AERO, ONSTR, TAO, and VSO, include object properties from three versions of the RO.

TABLE I. VARIOUS VERSIONS OF THE *HAS PART* PROPERTY FOUND AMONG THE BIOPORTAL ONTOLOGIES

URI	# Ontologies That Reuse *has part* Property
http://purl.obolibrary.org/obo/bfo_0000050	48
http://www.obofoundry.org/ro/ro.owl#part_of	28
http://purl.obolibrary.org/obo/temp#part_of	27
http://purl.obolibrary.org/obo/obo_rel#_part_of	7
http://purl.obolibrary.org/obo/bfo_00000050	5
http://purl.obolibrary.org/obo/obo_rel_part_of	4
http://purl.org/obo/owl/obo#part_of	2
http://purl.org/obo/owl/obo_rel#part_of	2
http://purl.org/obo/owl/ro#part_of	2
http://purl.obolibrary.org/obo/http://www.obofoundry.org/ro/ro.owl#part_of	1
http://purl.org/obo/owlapi/relationship#obo_rel_part_of	1
http://www.ifomis.org/obo/ro/1.0#partof	1
http://obofoundry.org/ro/ro.owl#part_of	1
http://purl.obofoundry.org/ro/ro.owl#part_of	1
Total:	**130**

C. *owl:imports Errors*

A total of 44 ontologies could not be loaded by the OWL API due to errors caused by missing imported ontologies. There were several reasons for these errors; however, the large majority were caused by URIs being no longer valid web addresses. For example, the DDI ontology includes *http://www.obofoundry.org/ro/ro.owl*, but no ontology file exists at that location. In a similar manner, RoleO includes *http://purl.obolibrary.org/obo/RoleO/external/bfo_import.owl*. Similarly, SDO includes its custom-built Units Ontology via an *owl:imports* statement, but the ontology no longer exists at the specified location. Five of the 44 ontologies (11.4%) were previously hosted on Google Code (which is no longer available, as of January 2016).

There were relatively few errors caused by other types of invalid import statements. For example, various Psychology (APA thesaurus) ontologies on BioPortal all have *owl:imports* statements that reference local files. We note that not all instances of an ontology using *owl:imports* are instances of reuse, since the *owl:imports* mechanism is also frequently used to include modules from the same ontology.

V. DISCUSSION

We note that when an author is designing an ontology, it is often with the intention of supporting a specific set of use-cases, or some specific application. Thus, some of the issues we identified in this paper may not be problematic for the intended purposes. However, once it has been discovered that an ontology appears to contain inconsistencies due to reuse, the issue should be brought to the attention of the author of the ontology. These problems could have deleterious effects if the ontology is utilized beyond its original scope.

One significant complication is that, based on the metrics provided in BioPortal, hundreds of ontologies have not been updated in several years (if ever). Many of these ontologies are no longer maintained and reuse old versions of source ontologies that are long out of date. This leads to, for example, twelve versions of OBO REL/RO/BFO properties appearing throughout BioPortal's ontologies (as illustrated in Table I). This situation can impact ontology authors who decide to reuse the contents of these "dormant" ontologies (using, e.g., the BioPortal reuse plugin [26] for Protégé). In future work, we will investigate ways of warning ontology authors about potential issues when reusing an ontology's classes. We will also investigate semi-automated techniques for identifying and preventing issues when reusing content (which could, e.g., automatically align the different *part of* properties used in the SDO and other ontologies).

The errors reported on in this paper are from the year 2015. Checking on a sample of them in the current version of the BioPortal, we found that the errors mentioned here are still in existence because we did not alert the curators of the specific ontologies at the time. We can assume that many of the other errors still exist. In fact, a July 2018 scan of a sample of the ontologies reported on in the results revealed a number of ontologies whose latest BioPortal release predated 2015 (e.g., AERO, COGPO, DDI, CSEO, SYN, etc.). Moreover, all these ontologies had relatively significant numbers of visits at their BioPortal pages in the second quarter of 2018, indicating continued interest in them. Since many more ontologies have been added to the BioPortal in the interim, another review would probably uncover more errors, but we were not in a position to perform such a study. Although the examples are from 2015, they reflect the reality of some phenomena that curators and authors are liable to encounter when engaging in the practice of ontology reuse. The timeliness of the results is not critical since the purpose of the paper is to alert ontology designers and maintenance personnel, especially those new to the process of content reuse, to the kinds of problems and errors they are likely to face when creating an ontology with the aid of reuse.

Also in future work, we plan to offer a set of guidelines for ontology reuse in order to preempt some of the troubles described herein. Major aspects of those guidelines will deal with ontological commitment and the proper consideration of the hierarchical context of reused content. We will also review some of the software tools available to complement these guidelines.

VI. Conclusion

The reuse of content from existing ontologies is an important design principle that can facilitate the work of curators and authors when creating new ontologies. It can also help to ensure alignment of the new ontologies with previously modeled knowledge. However, the process of reuse is not a simple one, and there are potential pitfalls. In this paper, we studied a collection of BioPortal ontologies to determine what problems may have been introduced via reuse. We focused on three kinds of errors and presented heuristic methodologies to uncover these within a collection of ontologies. The results showed that significant errors could arise from reuse. This should encourage ontology maintenance personnel to be cautious and vigilant when adopting the reuse approach.

Acknowledgment

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under award number R01CA190779. The content is solely the responsibility of the authors and does not necessarily represent the views of the National Institutes of Health.

References

[1] "OWL Web Ontology Language Reference," available at https://www.w3.org/TR/owl-ref. Accessed May 11, 2018.

[2] "Protégé," available at http://protege.stanford.edu. Accessed May 13, 2018.

[3] P. Grenon, B. Smith, and L. Goldberg, "Biodynamic ontology: Applying BFO in the biomedical domain," *Ontologies in Medicine*, pp. 20–38, 2004.

[4] "OGMS – Ontology for General Medical Science," available at https://code.google.com/archive/p/ogms. Accessed May 13, 2018.

[5] E. Beisswanger, S. Schulz, H. Stenzhorn, and U. Hahn, "BioTop: An upper domain ontology for the life sciences – a description of its current structure, contents, and interfaces to OBO ontologies," *Applied Ontology*, vol. 3, no. 4, pp. 205–212, 2008.

[6] C. Ochs, Z. He, Y. Perl, S. Arabandi, M. Halper, and J. Geller, "Refining the granularity of abstraction networks for the Sleep Domain Ontology," in *Proc. Fourth Int'l Conference on Biomedical Ontology* (*ICBO 2013*), Montreal, Canada, Jul. 2013, pp. 84–89.

[7] Z. He, C. Ochs, L. Soldatova, Y. Perl, S. Arabandi, and J. Geller, "Auditing redundant import in reuse of a top level ontology for the Drug Discovery Investigations ontology," in *Proc. Int'l Workshop on Vaccine and Drug Ontology Studies* (*VDOS-2013*), Montreal, Canada, Jul. 2013

[8] Z. He, C. Ochs, A. Agrawal, Y. Perl, D. Zeginis, K. Tarabanis, G. Elhanan, M. Halper, N. Noy, and J. Geller, "A family-based framework for supporting quality assurance of biomedical ontologies in BioPortal," in *Proc. 2013 AMIA Annual Symposium*, Washington, DC, Nov. 2013, pp. 581–590.

[9] S. Arabandi, C. Ogbuji, S. Redline, R. Chervin, J. Boero, R. Benca et al., "Developing a Sleep Domain Ontology," in *Proc. AMIA Clinical Research Informatics Summit*, 2010.

[10] D. Qi, R. D. King, A. L. Hopkins, G. R. Bickerton, and L. N. Soldatova, "An ontology for description of drug discovery investigations," *Journal of Integrative Bioinformatics*, vol. 7, no. 3, Mar. 2010.

[11] D. Zeginis, A. Hasnain, N. Loutas, H. F. Deus, R. Fox, and K. A. Tarabanis, "A collaborative methodology for developing a semantic model for interlinking Cancer Chemoprevention linked-data sources," *Semantic Web*, vol. 5, no. 2, pp. 127–142, 2014.

[12] "BioPortal," available at http://bioportal.bioontology.org/. Accessed May 7, 2018.

[13] C. Ochs, Y. Perl, J. Geller, M. Haendel, M. Brush, S. Arabandi, and S. Tu, "Summarizing and visualizing structural changes during the evolution of biomedical ontologies using a Diff abstraction network," *Journal of Biomedical Informatics*, vol. 56, pp. 127–144, 2015.

[14] M. R. Kamdar, T. Tudorache, and M. A. Musen, "A systematic analysis of term reuse and term overlap across biomedical ontologies," *Semantic Web Journal*, vol. 8, no. 6, pp. 853–871, 2017.

[15] A. Ghazvinian, N. F. Noy, and M. A. Musen, "How orthogonal are the OBO Foundry ontologies?" *Journal of Biomedical Semantics*, vol. 2(Suppl 2): S2, 2011.

[16] "The OBO Foundry," available at http://www.obofoundry.org/. Accessed April 30, 2018.

[17] B. Smith, M. Ashburner, C. Rosse, J. Bard, W. Bug, W. Ceusters et al., "The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration," *Nature Biotechnology*, vol. 25, pp. 1251–1255, 2007.

[18] C. Ochs, Y. Perl, J. Geller, S. Arabandi, T. Tudorache, and M. A. Musen, "An empirical analysis of ontology reuse in BioPortal," *Journal of Biomedical Informatics*, vol. 71, pp. 165–177, 2017.

[19] "Gene Ontology Consortium," available at http://www.geneontology.org. Accessed May 6, 2018.

[20] "Chemical Entities of Biological Interest (ChEBI)," available at [12]. Accessed April 2, 2018.

[21] M. Courtot, F. Gibson, A. L. Lister, J. Malone, D. Schober, R. R. Brinkman, and A. Ruttenberg, "MIREOT: The minimum information to reference an external ontology term," *Applied Ontology*, vol. 6, no. 1, pp. 23–33, 2011.

[22] "Computer-based Patient Record Ontology," available at https://code.google.com/archive/p/cprontology. Accessed May 13, 2018.

[23] B. Smith, W. Ceusters, B. Klagges, J. Köhler, A. Kumar, J. Lomax, C. Mungall, F. Neuhaus, A. L. Rector, and C. Rosse, "Relations in biomedical ontologies," *Genome Biology*, vol. 6:R46, 2005.

[24] C. Rosse and J. L. V. Mejino, "A reference ontology for biomedical informatics: The Foundational Model of Anatomy," *Journal of Biomedical Informatics*, vol. 36, no. 6, pp. 478–500, Dec. 2003.

[25] C. J. Mungall, C. Torniai, G. V. Gkoutos, S. E. Lewis, and M. A. Haendel, "Uberon, an integrative multi-species anatomy ontology," *Genome Biology*, vol. 13:R5, 2012.

[26] J. Nair, T. Tudorache, T. Whetzel et al., "The BioPortal Import Plugin for Protégé," in *Int'l Conference on Biomedical Ontology* (*ICBO 2011*), 2011, pp. 298–299.