# Uncovering Definition Coverage in the OBO Foundry Ontologies

Daniel R. Schlegel
and Peter L. Elkin
Department of Biomedical Informatics
University at Buffalo, SUNY, Buffalo, NY, USA
Email: <drschleg, elkinp>@buffalo.edu

Selja Seppälä
Department of Health Outcomes and Policy
University of Florida, FL, USA
Email: sseppala@ufl.edu

***Abstract*—Definitions, both logical and textual, are an essential part of ontologies. Textual definitions help human users disambiguate and regularize their understanding and use of ontology terms to achieve intra- and inter-personal consistency and avoid errors, for example, when annotating scientific data, integrating databases with an ontology, or importing terms into other ontologies. Logical definitions are needed, among other things, for checking the consistency of the ontology and carrying out inferences, for example, over data that has been annotated with ontology terms. Despite the best efforts of ontology developers, it is not uncommon to see missing definitions. While the OBO Foundry explicitly states that its member ontologies should have a substantial fraction of their terms defined, these ontologies still often lack one or both kinds of definitions. Statistics on definition coverage in the OBO Foundry ontologies are scarce and it is difficult to tell what effectively constitutes a substantial fraction of terms in an ontology. In the present work, we examine the coverage of textual and logical definitions throughout the OBO Foundry ontologies in order to uncover the big picture and to give more detailed insight into logical definitions in these ontologies. We have found that textual definition coverage is reasonably good over the OBO Foundry ontologies (66%), but that the core ontologies exhibit a higher definition coverage (86%) than the non-core ones (64%). Logical definitions follow a similar trend, but with lower values — overall, the OBO Foundry has a 30% coverage, while core ontologies are better covered (53%) than non-core ones (28%).**

## I. Introduction

For an ontology to be of the highest quality, it must have both textual and logical definitions for its terms. Definitions serve many purposes. For example, good textual definitions allow for experts and non-experts alike to understand the content of an ontology and to use it in the manner the authors intended. Logical definitions are necessary for reasoners to verify that an ontology is consistent, and may make application of the ontology easier for users. Ideally, logical and textual definitions would convey the same information — it has been shown that there is an important correspondence between textual and logical definitions, and each can provide an accuracy check on the other [1], [2].

Producing definitions is difficult and time-consuming — often much more so than just creating hierarchical taxonomies. Thus, despite the best efforts of ontology developers and the existence of a number of tools and methods to populate ontologies with definitions (for example, [3]–[9]), it is not uncommon to see missing textual or logical definitions, if not both. This is also the case in the Open Biomedical Ontologies (OBO) Foundry [10] ontologies.

The OBO Foundry was created by the OBO consortium as a repository for biomedical ontologies. Currently the OBO Foundry contains 9 "core" ontologies and 128 non-core ontologies. The OBO Foundry ontologies are developed in a coordinated way according to a set of shared principles.[1] One of the OBO Foundry principles is about definitions. That principle states that member ontologies should have "textual definitions ... for a substantial and representative fraction [of terms], plus equivalent formal definitions (for at least a substantial number of terms)" [11]. The statement of this principle is rather vague and comes with no further specifications, eliciting an obvious question: How much is 'substantial'?

Some ontology browsers on the web, such as BioPortal,[2] provide statistics on the total number of classes, and the number of classes lacking textual definitions,[3] but they do not give us access to the big picture: definition coverage across the OBO Foundry ontologies. They also do not give us any insight into the existence of parts of or complete logical definitions.

In the present work, we examine the coverage of textual and logical definitions throughout the OBO Foundry ontologies. In particular, we aim to determine to what extent the principle of having textual and logical definitions for a substantial number of terms is upheld — even though what constitutes 'substantial' has not been formally defined. We also aim to determine if the prevalence of definitions is different between the core and non-core ontologies, if there are more textual than logical definitions, and if the size of ontologies has an effect on definitional coverage.

In Section II, we discuss what textual and logical definitions are within an ontology and, in Section III, we examine how we computed the number of these definitions for each ontology. We present results to the aims mentioned above in Section IV, which we discuss in Section V, before concluding with Section VI.

[1] OBO Foundry Principles, http://obofoundry.org/principles/fp-000-summary.html.

[2] BioPortal, http://bioportal.bioontology.org/

[3] For a very narrow set of definition annotations.

## II. Textual and Logical Definitions in Ontologies

The intended meaning of an ontology term is specified, on the one hand, by a natural language definition — also called a *textual definition* — and other documentation; and on the other hand, by axioms that form its logical (or formal) definition.

### A. Textual Definitions

A textual definition is, ideally, a short sentence found as the object of an annotation property. In some cases, the annotation property is designated for that purpose and is introduced with an explicit annotation label. Examples of such labels are `iao:definition` (`IAO_0000115`) and `iao:elucidation` (`IAO_0000600`)[4] for annotation properties provided by the Information Artifact Ontology (IAO).

Not all ontologies use the dedicated IAO ontology annotation properties. We found that 14 out of the 119 OBO Foundry ontologies that we analyzed[5] use other types of annotations.[6] These ontologies use annotation properties with names related to 'definition', such as: `def`, `definition` (defined otherwise than in IAO), `external_definition`, `preferredDefinition`, and `hasDefinition`. Given the obviousness of the name, we as humans can deduce the definitional nature of the annotation.

Finally, textual definitions can appear in the more general `comment` annotation property without any specific indication. In this case, it is more difficult to identify them — at least, automatically. We excluded these cases from our inventory.

Textual definitions tell us about the properties of the instances of a class in an ontology. They state (i) the type of thing of which they are instances and (ii) one or more properties of these instances that differentiate them from instances of neighboring types. The general form of such a definition, namely 'An $A$ is a $B$ that has the property of $C$ing', is called an *Aristotelian definition*.[7] The first part of the definition ('$A$') is called the *definiendum* and corresponds to that which is defined. The second part ('$B$ that has the property of $C$ing') corresponds to the *definiens*, which is typically what we find in a definition annotation property. The definiens is further subdivided into a *genus* ('$B$') — *genus proximus* when it states the immediate superordinate type — and one or more *differentiae* ('$C$').

Consider the textual definition of the Ontology for Biomedical Investigations (OBI) term *sequencing facility contact person*:

*A person who is the contact representative at the sequencing facility.*

In this definition, *A person* is the genus that states that a sequencing facility contact person is a person. The differentia part serves to differentiate a sequencing facility contact person from other persons in virtue of the fact that this person has the role of being the contact representative of the sequencing facility of which they are a member.

Genus and differentia(e) combined provide ontology users with the necessary information to carry out the desired inferences when using an ontology term, while restricting their use of the term to the intended meaning within the ontology. Definitions thus have both a cognitive function and linguistic functions [14], such as term disambiguation and normalization. This is true for textual definitions, as well as logical definitions.

### B. Logical definitions

Ontologies are made up of a set of classes which are defined by *class expressions*. Class expressions represent sets of individuals which meet formally specified conditions for membership. A class expression might be an atomic class within the ontology, such as `bacteria`, or it might be an anonymous class defined by some combination of properties, such as `(has_part some flagellum)`, or it might be a union of atomic and anonymous classes to refer, for example, to the class containing instances of `bacteria` which have as part some `flagellum`.

Axioms define true things within the domain of interest. Some axioms, such as `SubClassOf` and `EquivalentClass`, define relationships between class expressions [15]. These two axiom types, specifically, constitute the logical definitions of the ontology terms. `SubClassOf` axioms correspond to definitions where each axiom is individually necessary for something to be an instance of the class that is being defined; `EquivalentClass` axioms correspond to definitions where all the axioms are individually necessary and jointly sufficient for something to be an instance of the class that is being defined. There are two other axioms which define relationships between classes — `DisjointClasses` and `DisjointUnion` — which we don't discuss here as they do not contribute to the logical definition as we have defined it.[8]

Like textual definitions, logical definitions ideally consist of at least one genus and one or more differentiae. Consider now the logical definition of the OBI term `sequencing facility contact person` seen above.

```
sequencing facility contact person
  equivalentClass
    Homo sapiens and
    (is_member_of organization some
      sequencing facility organization) and
    (has_role some contact representative role)
```

[4] Elucidations are technically clarifications, but are often used in lieu of definitions when they are difficult or impossible to write, as it is the case for upper-level terms in BFO, such as `bfo:entity` which has the elucidation "An entity is anything that exists or has existed or will exist."

[5] The remaining ontologies were either not accessible on the web, or could not be loaded into the OWL API.

[6] This applies to the following ontologies: BCGO, CEPH, CLO, DINTO, ERO, EXO, KISAO, MAMO, MFMO, MICRO, OVAE, SWO, VO, VTO.

[7] Significant work has been dedicated to what constitutes a good definition. See, for example [12] and ISO TC37's [13] work on ISO1087-1 and ISO1087-2.

[8] We acknowledge that `DisjointUnion` may be used by some ontology authors as a workaround for when the differentiae are not able to differentiate a class from its siblings. We will examine this phenomenon in future work.

This definition states that every instance of a `sequencing facility contact person` is an instance of `Homo sapiens`, is a member of some organization of type `sequencing facility organization`, and embodies the role of `contact representative role`. In this definition, the genus is `Homo sapiens` — the kind of thing a `sequencing facility contact person` is, and the differentiae are the restrictions upon the genus (the membership and role axioms).

In general, the genus of a logical definition for a class can be seen as the set of named classes with which it is in a SubClassOf or EquivalentClass relation. The differentiae are the axioms which restrict the extension of the class.[9]

## III. Methodology to Determine Definitional Coverage

Our study focuses on 119 ontologies out of the 128 present in the OBO Foundry. Indeed, 18 non-core ontologies were either currently unavailable on the Web due to broken links, or they failed to load using the OWL API [16].

For each of these ontologies, we automatically gathered the number of classes, textual definitions and elucidations, logical axioms on the classes, and non-anonymous subclass axioms. All extraction in this project was done using the Java OWL API version 4.2.3 [16].

### A. Identifying Textual Definitions

To identify textual definitions, we used the IAO annotation property `definition` used in 103 of the 119 ontologies in this study. We also examined the set of annotation properties used in the OBO Foundry ontologies which contained the string `def` but did not contain the strings `editor`, `source`, `citation`, `defines`, or `defined` to try to capture any non-standard annotation properties which might have been used to signal a definition. The string matching retrieved the following annotation properties:

- `def`
- `Definition`
- `skos:definition` (and other `definition` terms from other namespaces)
- `external_definition`
- `hasDefinition`
- `preferredDefinition`

We also included the IAO annotation property `elucidation` for ontologies that contain some primitive classes that cannot be, strictly speaking, defined. The use of elucidation is common in upper-level ontologies such as the Basic Formal Ontology (BFO), which is designed to represent the most general types of things in the world.

[9]It is possible for someone to abuse the OWL syntax by adding what they consider to be logical definition components in annotations, but those cannot be used for reasoning and we do not consider them here.

### B. Identifying Logical Definitions

For each ontology, we computed the number of classes that contain: (i) at least one genus; (ii) at least one differentia; and (iii) at least one of both. We consider a class that is specified by both a genus and one or more differentiae to have a "complete logical definition." To perform the computation, we created definitions for what it means for an axiom to contain a genus and one or more differentiae. In OWL parlance:

- An axiom contains a genus for the definition of class $c_1$ if the axiom contains some other class, $c_2$, where $c_2$ is not part of an object property restriction.
- An axiom contains one or more differentiae for the definition of a class if the axiom contains one or more object property restrictions.

This can be illustrated again with the logical definition of the OBI term `sequencing facility contact person` examined above. The axiom `Homo sapiens` is not part of an object property restriction, therefore, it was identified as a genus. Both `is_member_of organization some sequencing facility organization` and `has_role some contact representative role` are axioms that contain object property restrictions. These were identified as two distinct differentiae.

Throughout the rest of this paper we will talk about *coverage* of definitions or parts of definitions in ontologies, that is, the number of classes which have (i) at least one instance of the complete logical definition type, or (ii) at least one definition part under discussion. The reason for discussing coverage for logical definition parts, rather than discussing the number of genera or differentiae in classes on average, is that genera and differentiae can be expressed as several assertions, or as a single one using unions and intersections, so a count does not carry any particular meaning.

## IV. Results

### A. Definition Covereage in Core and Non-Core Ontologies

Among textual definitions, we found that coverage within the 9 core ontologies was quite high, with 6 of the 9 having textual definitions for over 90% of their terms (see Table I). This includes very large ontologies like the ontology of Chemical Entities of Biological Interest (ChEBI) and the Protein Ontology (PR), both of which have over 100,000 terms. The outlier on the low end was the Human Disease Ontology (DO), which has definitions for just under 35% of its 9,247 terms. On average, core ontologies have textual definitions for 85.6% of their terms ($stdev = 21\%$). The coverage of textual definitions in non-core ontologies is lower, with an average of 63% ($stdev = 38\%$).

Among logical definitions, we again found reasonably good coverage among the core ontologies. Looking at only definition components, we found that an average of 91% of classes had one or more genera, and an average of 53% had at least one differentia ($stdev = 34\%$). Coverage for complete logical definitions was 53% ($stdev = 34\%$). We see such a large standard deviations for complete logical definitions and differentiae in core ontologies because coverage is very sporadic — DO and the Phenotype and Trait Ontology (PATO) have less than 10% coverage, while the Xenopus Anatomy and Development Ontology (XAO) and the Zebrafish Anatomy and Development Ontology (ZFA) have over 95%.

Once again, the coverage in non-core ontologies was lower, with an average of 86% of classes having genera ($stdev = 23\%$) and 34% having differentia on average ($stdev = 32\%$). Coverage of complete logical definitions was only 28% ($stdev = 29\%$) on average. The genera values are not much different from the core ontologies, since most ontologies have a taxonomic structure that requires at least one genera per term. There are exceptions to this. For example, the e-Mouse Atlas Project (EMAP) ontology is mostly flat, with only 525 genus axioms for its nearly 20,000 terms.[10]

### *B. Definition Coverage Accross OBO Foundry Ontologies*

Over the full set of analyzed OBO Foundry ontologies, we found textual definition coverage to be an average of 66% ($stdev = 37\%$) and complete logical definition covereage to be 30% ($stdev = 30\%$). Looking again at the logical definition components, there is an average genera coverage of 86% ($stdev = 23\%$), and average differentiae coverage of 36% ($stdev = 33\%$). The full set of coverage results is available in Table I.

| | Core | Non-Core | Total |
|---|---|---|---|
| Textual Definition Coverage | 86% | 64% | 66% |
| Logical Definition Coverage | 53% | 28% | 30% |
| Genera Covereage | 91% | 86% | 86% |
| Genera Only Coverage | 39% | 58% | 57% |
| Differentiae Covereage | 53% | 34% | 36% |
| Differentiae Only Covereage | 0% | 6% | 6% |

TABLE I
COVERAGE OF TEXTUAL DEFINITIONS, LOGICAL DEFINITIONS, AND PARTS OF LOGICAL DEFINITIONS ACROSS THE CORE, NON-CORE, AND SUM TOTAL OF THE ANALYZED ONTOLOGIES IN THE OBO FOUNDRY.

We can examine definitional coverage more closely by looking at the number of ontologies with a given range of coverage. Figure 1 shows the rate of coverage for both textual and logical definitions. Note that the logical definitions shown here are *complete* logical definitions — those with both a genus and differentia(e). Examining the graph, we can see that the trends are nearly opposite. Relatively few ontologies have poor textual definition coverage, while a large number have 90-100% coverage. On the other hand, a large number of ontologies have very poor logical definition coverage (0-10%), and few have good logical definition coverage. That said, it is not, in general, the case that ontologies which have a large number of textual definitions have a small number of logical definitions, or vise-versa. There are a few ontologies where this is the case, though. For example, the Cell Line Ontology (CLO) has 8% textual definition coverage and 93% logical definition coverage, while the Chemical Methods Ontology (CHMO) has 98% textual definition coverage and 7% logical definition coverage.

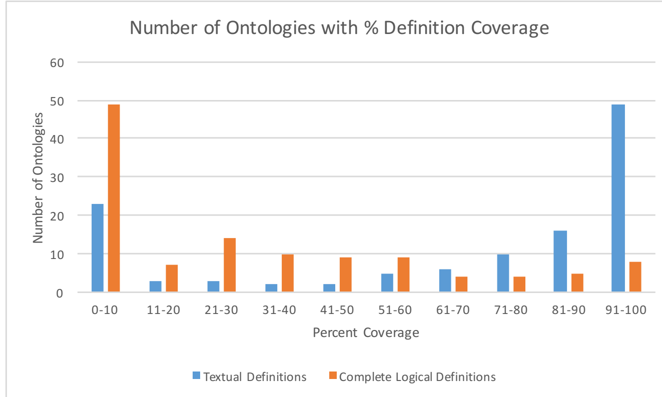


Fig. 1. The number of ontologies with percent coverage of textual and complete logical definitions.

One of our aims mentioned in the introduction was to determine if there was a correlation between ontology size and definition coverage. To examine this, we broke the ontologies up into five groups: very small (0-99 terms, n=17); small (100-999, n=42); medium (1,000-9,999, n=44); large (10,000-99,999, n=11); and very large (100,000+, n=3). We found that the very small, small, medium, and very large groups had textual definitions for roughly 60-70% of their terms. The 11 ontologies in the large category formed the only outlier group, with a 33% coverage.

We looked at the coverage of logical definitions in three ways — the percent of classes with genera, the percent with differentiae, and the percent with both (a complete logical definition). Figure 2 shows a clustered bar chart with the coverage results for textual and logical definitions. The average number of differentiae rose in each group from very small to large, dropping off at very large. This drop-off can be explained: while the Protein Ontology (PRO) has very good differentiae coverage with 88% of its classes having at least one, the other two ontologies in this category the NCBI Taxonomy (NCBITAXON) Vertibrate Taxonomy Ontology (VTO) both contain no differentiae. The percent coverage of complete logical definitions rose slowly as ontology size grew.

## V. DISCUSSION

### *A. Explaining the Trends and Bridging the Gap*

Given the OBO Foundry principle emphasizing the inclusion of a substantial fraction of terms defined with a textual definition, we aimed to find out if the OBO Foundry ontologies would include more textual definitions than logical ones. This hypothesis was confirmed by our results, which surprisingly also showed that the coverage for textual definitions was especially high for larger ontologies such as ChEBI, GO, and PRO. These ontologies have a very large number of terms and, yet, a textual definition coverage of over 90%.

For some of the large ontologies, it is now fairly easy to submit term requests along with textual definitions and

[10]We don't count having `owl:Thing` as a direct superclass as a genera, since everything is by default a subclass of `owl:Thing`.

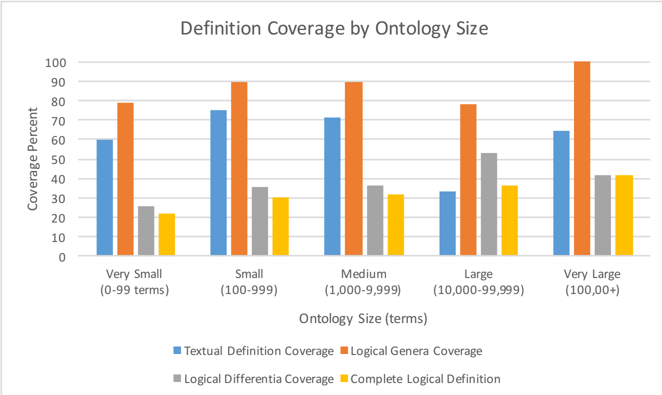


Fig. 2. The coverage of textual and logical definitions by ontology size. Both the genus and differentia components of the logical definitions are shown, along with coverage for the complete definitions containing both genera and differentiae

template-based logical definitions using tools like TermGenie [3], which may account for some of the higher coverage numbers among these larger ontologies. Other ontologies import a significant amount of data from other sources. ChEBI, for example, imports many of its classes (and perhaps other data) from the IntEnz and KEGG LIGAND databases.

For ontologies showing lower textual definition coverage, we can assume that at least the terms that have a textual definition constitute a 'represenatative' fraction of the ontology — recall that the OBO Foundry principle about definitions states that ontologies should have a definition for a 'substantial and representative' fraction of terms. The OBO foundry does not specify what is considered representative. However, if we posit that it concerns terms of the relevant ontology itself (as opposed to terms imported from other ontologies), then it might be the case that most of those terms within those small coverage results are indeed defined — this coverage would amount to a significant fraction of the ontology terms proper. Nevertheless, to test this hypothesis, we would need to add an extra test taking, for example, into account the domain space of the terms that have a textual definition. The resulsts could then be reported as textual definition coverage for the considered ontology terms and imported ontology terms.

For the most part, logical definitions have not received the same attention — ChEBI and GO have complete logical definitions for around 40% of their terms, while it is much higher in PRO, with an 88% coverage. Across all of the OBO Foundry, ontologies only had complete logical definitions for only 30% of classes on average. The fact that logical definitions are very difficult to construct may explain this. In order to build logical definitions, it is often necessary to use terms outside of the ontology's primary domain, requiring the ontology developers to build or import more ontological structure. For example, `Homo sapiens` in our example from OBI was imported from NCBITAXON. This need for importing outside terms and structures was experienced by the Gene Ontology (GO) group when they put forth a concerted effort to have formal definitions: many terms had to be imported from other OBO Foundry ontologies [5]. This process can be helped along using tools such as OntoFox [17] which allow users to easily find suitable terms and axioms.

Solutions for bridging the gap in textual and logical definition coverage already exist. For example, some data sources, such as Wikipedia and the Unified Medical Language System (UMLS), may provide existing textual definitions [4], or logical relationships which may be transformed into the kind used in ontologies for logical definitions. The advantage of this kind of solution is that it can be partially automated, leaving the manual task to post-editing and validation work.

Given the opposite trends in textual vs. logical definition coverage (Figure 1), there may also be a significant impact of tools that build textual definitions from logical ones, and vise-versa. Some work has already been done on translating logical definitions to textual ones in ontologies (*e.g.,* [6]) and in general knowledge representation systems (*e.g.,* in the SNePS system [18], [19]). Significant work has also been done on transforming natural language text into formal logics (*e.g.,* [20]–[23]). The challenge would reside in making the methods and tools more easily available and usable for the OBO Foundry ontology developers.

### *B. Quantifying 'Substantial'*

Now that we have concrete results, we should be able to answer two questions:

- What is the range of definition coverage that constitutes a 'substantial' coverage of the ontology terms?
- Do the ontologies that are currently members of the OBO Foundry meet its principle on definitions?

However, the answers to these questions are not trivial. If we consider that 'substantial' equates with the average definition coverage measured over the core ontologies, then an adequate coverage to be included in the OBO Foundry would be to have 86% of the ontology terms specified with a textual definition, and 53% of the terms with a complete logical definition.

To expect that all ontologies have coverage which is as complete as the core ontologies is probably unrealistic. Therefore let us consider that 'substantial' equates with the average coverage measured over all of the (analyzed) ontologies in the OBO Foundry, then an adequate coverage to be part of the OBO Foundry would be much lower, namely at 66% of the ontology terms specified with a textual definition and 30% with a complete logical definition.

Next, we must consider whether we should have different levels of 'substantial' coverage for textual and logical definitions.[11] We would argue "No," the roughly 65% value for textual definitions is a reasonable value for significance, and logical definitions should be held to this standard as well.

[11] This is related to the idea of significance, which obviously varies with the topic under consideration — a significant percentage of parking spaces with land mines hidden in them could be rather small, say 10%, but both of the things we are considering are definitions and therefore, we would argue, should be held to the same significance standard.

### C. Meeting the OBO Foundry Principles on Definition Coverage

When looking at the set of ontologies as a whole, the OBO Foundry principle seems to be met, at least for textual definitions: textual definition coverage can be qualified as 'substantial' overall. That said, many ontologies in the OBO Foundry have no or very few textual definitions,[12] and a concerted effort should be made to improve them. Coverage of logical definitions is much lower and, in most cases, does not meet the bar of "a substantial number of terms." If we use the value of 65% which we came to earlier, we find that only 19 out of the 119 analyzed ontologies meet the bar, and only 15 have both substantial percentages of textual and logical definitions.

### D. Addressing Definition Quality

The OBO Foundry principles do not say anything explicitly about definition quality — nor does our study. The original formulation of the principle stated that "terms should be defined so that their precise meaning within the context of a particular ontology is clear to a human reader." For a definition to be 'clear', it needs to meet certain quality standards. In this respect, the OBO Foundry seems to be committed to quality standards as well.

Even though the OBO Foundry ontologies have a substantial fraction of their terms defined, this doesn't mean that they always meet expected definition quality standards, such as clarity. Furthermore, a large definition coverage with definitions of poor quality amounts to a small coverage in definitions.

Definitions fulfill cognitive as well as linguistic functions which justify both (i) their systematic inclusion in ontologies and (ii) the inclusion of relevant content in an appropriate form adequate for the uses that are made of ontologies [2], [14]. Consequently, their form and content should meet certain quality standards that allow definitions to fulfill these functions. In future work, we will examine definition quality more closely, including methodologies for automating the process of evaluating quality as much as possible.

## VI. Conclusion

Definitions, both logical and textual, are essential components of an ontology. The OBO Foundry has the noble goal of creating a repository for ontologies developed using a shared set of principles, including some (vague) requirements for including definitions. Until now, there has been no analysis of the "big picture" — how well, overall, ontologies are conforming to this principle.

In this study, we have shown that:

- if we are to take the OBO Foundry principles as being true for the ontologies in the Foundry collectively, then a substantial, thus acceptable, definition coverage rate is likely around 65%;
- there is substantial coverage of textual definitions in the OBO Foundry ontologies;
- coverage for logical definitions is not substantial;
- there is still work to be done to achieve full coverage.

Logical definitions cover less terms than we would expect for high quality ontologies and require more attention. We found that logical definition coverage improved as ontology size grew, a finding that might be counter-intuitive. The highlights of the OBO Foundry are the core ontologies, which have better coverage of textual and logical definitions than their non-core counterparts.

In future work, we will focus on definition quality across the OBO Foundry ontologies. This will require developing a quality metric for both textual and logical definitions (as well as how well textual and logical definitions align, where both are present). We hope to automate the evaluation task to the greatest degree possible so that ontology curators (and ontology repository maintainers) can obtain information about the coverage and quality of their definitions in real time.

[12] Including: CDAO, CTENO, EMAP, EMAPA, FIX, FLOPPO, KISAO, MF, MRO, NCBITAXON, TAXRANK, VTO, and ZFS.

## References


[1] S. Seppälä, Y. Schreiber, and A. Ruttenberg, "Textual and logical definitions in ontologies," in *Proceedings of The First International Workshop on Drug Interaction Knowledge Management (DIKR 2014), The Second International Workshop on Definitions in Ontologies (IWOOD 2014), and The Starting an OBI-based Biobank Ontology Workshop (OBIB 2014)*, R. D. Boyce, M. Brochhausen, P. E. Empey, M. Haendel, W. R. Hogan, D. C. Malone, P. Ray, A. Ruttenberg, S. Seppälä, C. J. Stoecker, and J. Zheng, Eds., vol. Vol-1309. Houston, TX, USA: CEUR Workshop Proceedings (CEUR-WS.org), October 6-7 2014, pp. 35–41. [Online]. Available: http://ceur-ws.org/Vol-1309/

[2] S. Seppälä, Y. Schreiber, A. Ruttenberg, and B. Smith, "Definitions in ontologies," *Cahiers de lexicologie*, vol. 4, no. Numéro thématique "Au coeur de la définition", forthcoming.

[3] H. Dietze, T. Berardini, R. Foulger, D. Hill, J. Lomax, D. Osumi-Sutherland, P. Roncaglia, and C. Mungall, "Termgenie - a web-application for pattern-based ontology class generation," *Journal of Biomedical Semantics*, vol. 5, no. 1, p. 48, 2014. [Online]. Available: http://www.jbiomedsem.com/content/5/1/48

[4] G. Jiang, H. Solbrig, and C. Chute, "Using semantic web technology to support icd-11 textual definitions authoring," *Journal of Biomedical Semantics*, vol. 4, no. 1, p. 11, 2013. [Online]. Available: http://www.jbiomedsem.com/content/4/1/11

[5] C. J. Mungall, M. Bada, T. Z. Berardini, J. Deegan, A. Ireland, M. A. Harris, D. P. Hill, and J. Lomax, "Cross-product extensions of the gene ontology," *Journal of Biomedical Informatics*, vol. 44, no. 1, pp. 80–86, 2011, ¡ce:title¿Ontologies for Clinical and Translational Research¡/ce:title¿. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1532046410000171

[6] R. Stevens, J. Malone, S. Williams, R. Power, and A. Third, "Automating generation of textual class definitions from owl to english," *Journal of Biomedical Semantics*, vol. 2, no. Suppl 2, p. S5, 2011.

[7] T. Wächter and M. Schroeder, "Semi-automated ontology generation within obo-edit," *Bioinformatics*, vol. 26, no. 12, pp. i88–i96, 2010. [Online]. Available: http://bioinformatics.oxfordjournals.org/content/26/12/i88.abstract

[8] T. Wächter, G. Fabian, and M. Schroeder, "Dog4dag: semi-automated ontology generation in obo-edit and protégé," in *Proceedings of the 4th International Workshop on Semantic Web Applications and Tools for the Life Sciences*. ACM, 2011, pp. 119–120.

[9] Z. Xiang, J. Zheng, Y. Lin, and Y. He, "Ontorat: automatic generation of new ontology terms, annotations, and axioms based on ontology design patterns," *Journal of Biomedical Semantics*, vol. 6, no. 1, p. 4, 2015. [Online]. Available: http://www.jbiomedsem.com/content/6/1/4

[10] B. Smith, M. Ashburner, C. Rosse, J. Bard, W. Bug, W. Ceusters, L. J. Goldberg, K. Eilbeck, A. Ireland, C. J. Mungall *et al.*, "The obo foundry: coordinated evolution of ontologies to support biomedical data integration," *Nature biotechnology*, vol. 25, no. 11, pp. 1251–1255, 2007.

[11] OBO Techniucal WG, "Principle: textual definitions," 2016, http://www.obofoundry.org/principles/fp-006-textual-definitions.html.

[12] R. Arp, B. Smith, and A. D. Spear, *Building Ontologies with Basic Formal Ontology*. Cambridge, MA: MIT Press, 2015, forthcoming.

[13] International Standards Organization, "ISO/TC 37 terminology and other language and content resources," http://www.iso.org/iso/iso_technical_committee.html%3Fcommid%3D48104, 2016.

[14] S. Seppälä, A. Ruttenberg, and B. Smith, "The functions of definitions in ontologies," in *9th International Conference on Formal Ontology in Information Systems (FOIS 2016)*, Annecy, France, July 6-9 forthcoming.

[15] B. Motik, P. F. Patel-Schneider, B. Parsia, C. Bock, A. Fokoue, P. Haase, R. Hoekstra, I. Horrocks, A. Ruttenberg, U. Sattler *et al.*, "Owl 2 web ontology language: Structural specification and functional-style syntax (second edition)," *W3C recommendation*, 2012.

[16] M. Horridge and S. Bechhofer, "The owl api: A java api for owl ontologies," *Semantic Web*, vol. 2, no. 1, pp. 11–21, 2011.

[17] Z. Xiang, M. Courtot, R. R. Brinkman, A. Ruttenberg, and Y. He, "Ontofox: web-based support for ontology reuse," *BMC research notes*, vol. 3, no. 1, p. 175, 2010.

[18] S. C. Shapiro and W. J. Rapaport, "The SNePS family," *Computers & Mathematics with Applications*, vol. 23, no. 2–5, pp. 243–275, January–March 1992.

[19] D. R. Schlegel and S. C. Shapiro, "Visually interacting with a knowledge base using frames, logic, and propositional graphs," in *Graph Structures for Knowledge Representation and Reasoning, Lecture Notes in Artificial Intelligence 7205*, M. Croitoru, S. Rudolph, N. Wilson, J. Howse, and O. Corby, Eds. Berlin: Springer-Verlag, 2012, pp. 188–207.

[20] M. Craven, J. Kumlien *et al.*, "Constructing biological knowledge bases by extracting information from text sources." in *ISMB*, vol. 1999, 1999, pp. 77–86.

[21] M. Shamsfard and A. A. Barforoush, "Learning ontologies from natural language texts," *International journal of human-computer studies*, vol. 60, no. 1, pp. 17–63, 2004.

[22] D. Sánchez, A. Moreno, and L. Del Vasto-Terrientes, "Learning relation axioms from text: An automatic web-based approach," *Expert Systems with Applications*, vol. 39, no. 5, pp. 5792–5805, 2012.

[23] S. C. Shapiro and D. R. Schlegel, "Natural language understanding for soft information fusion," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*. IFIP, July 2013, 9 pages, unpaginated.