

# Preliminary Analysis of the OBO Foundry Ontologies and Their Evolution Using OQuaRE

Manuel QUESADA-MARTÍNEZ<sup>a</sup>, Astrid DUQUE-RAMOS<sup>b</sup>, Miguela INIESTA-MORENO<sup>a</sup>, Jesualdo Tomás FERNÁNDEZ-BREIS<sup>a,1</sup>

<sup>a</sup>Universidad de Murcia, IMIB-Arrixaca-UMU, CP 30100 Murcia, Spain

<sup>b</sup>Universidad de Antioquia, Calle 70 No. 52-21, Medellín, Colombia

**Abstract.** The biomedical community has now developed a significant number of ontologies. The curation of biomedical ontologies is a complex task as they evolve rapidly, being new versions regularly published. Therefore, methods to support ontology developers in analysing and tracking the evolution of their ontologies are needed. OQuaRE is an ontology evaluation framework based on quantitative metrics that permits to obtain normalised scores for different ontologies. In this work, OQuaRE has been applied to 408 versions of the eight OBO Foundry member ontologies. The OBO Foundry member ontologies are supposed to have been built by applying the OBO Foundry principles. Our results show that this set of ontologies is actually following principles such as the naming convention, and that the evolution of the OBO Foundry member ontologies is generating ontologies with higher OQuaRE quality scores.

**Keywords.** Ontology quality, Ontology evaluation, Ontology metrics, OQuaRE, OBO Foundry

## 1. Introduction

In recent years the biomedical community has increased its effort in the development of ontologies. As a consequence, the BioPortal repository<sup>2</sup> contains at the time of writing more than 500 biomedical ontologies. These ontologies change over time and the repositories contain all the versions. The frequency of release for new versions varies among ontologies. The availability of methods that support ontology developers in the analysis of the evolution of their ontologies would certainly contribute to improve the ontology development process, to make informed decisions about the effects of the changes made in the ontologies, and to detect whether certain modelling patterns or principles are applied. Some initiatives propose the analysis of ontologies using metrics. BioPortal calculates a set of metrics<sup>3</sup> for each ontology, so this provides some information about the quality of the ontology and helps ontology authors to make improvements. The analysis of ontologies based on metrics has also been used as a diagnostic task using structural, functional and usability profiling criteria [3]; using

---

<sup>1</sup> Corresponding Author: Jesualdo Tomás Fernández-Breis, email: [jfernand@um.es](mailto:jfernand@um.es)

<sup>2</sup> <http://bioportal.bioontology.org/>

<sup>3</sup> [https://www.bioontology.org/wiki/index.php/Ontology\\_Metrics](https://www.bioontology.org/wiki/index.php/Ontology_Metrics)

criteria such as philosophical rigor, ontological commitment, content correctness, and fitness for a purpose [5]; or presenting metrics for evaluating structural properties in the ontology [7]. All these methods are focused on the analysis of different aspects of one single ontology. In this paper we propose to apply a metrics-based approach for analysing the evolution of a set of ontologies. For this purpose, we are going to apply an adaptation of the OQuaRE framework [2] for supporting the analysis of ontology evolution processes within a common framework. The method will be applied to the OBO Foundry member ontologies. These ontologies are supposed to have been developed using the OBO Foundry principles. We are going to analyse not only the evolution of each ontology, but also the findings about the evolution of the ontologies as a group. We believe our method can contribute to generate new insights about the engineering of biomedical ontologies.

## 2. Materials and Methods

### 2.1. The OQuaRE framework

OQuaRE [1] is a framework for evaluating the quality of ontologies based on the standard ISO/IEC 25000:2005 for Software product Quality Requirements and Evaluation known as SQuaRE [4]. Briefly, OQuaRE proposes the evaluation of ontology quality using 3 levels: characteristics, subcharacteristics and metrics. The current OQuaRE version includes 8 characteristics, 29 subcharacteristics and 19 metrics. Each characteristic has a set of subcharacteristics associated that, in turn, have a set of metrics associated. The complete specification of OQuaRE can be found at<sup>4</sup>.

**Table 1.** OQuaRE metrics and a brief description of how we calculate them

OQuaRE metric	Description	OQuaRE metric	Description
ANOnto	mean number of annotation properties per class	NOCOnto	number of the direct subclasses divided by the number of classes minus the number of leaf classes
AROnto	number of restrictions of the ontology per classes	NOMOnto	mean number of object and data property usages per class
CBOnto	number of direct ancestor of classes divided by the number of classes minus subclasses of Thing	POnto	mean number of direct ancestor per class
CBOnto2	mean number of direct ancestor per class	PROnto	number of subclassof relationships divided by the number of subclassof relationships and properties
CROnto	mean number of individuals per class	RROnto	number of usages of object and data properties divided by the number of subclassof relations and properties
DITOnto	length of the longest path from Thing to a leaf class	RFCOnto	number of usages of object and data properties and superclasses divided by the number of classes
INROnto	mean number of subclasses per class	TMOnto	mean number of classes with more than 1 direct ancestor
LCOMOnto	mean length of all the paths from leaf classes to Thing	TMOnto2	mean number of direct ancestor of classes with more than 1 direct ancestor
NACOnto	mean number of superclasses per leaf class	WMCOnto	mean length of the path from Thing to a leaf class
		WMCOnto2	mean number of path from Thing to a leaf class per leaf class

Here, we will work at the level of quality metrics (see Table 1). In OQuaRE, the values of the metrics are transformed into *quality scores* by applying scaling functions. The current version of OQuaRE uses *quality scores* in the range [1, 5]: 1 – “Not Acceptable”, 2 – “Not Acceptable - Improvement Required”, 3 – “Minimally Acceptable”, 4 – “Acceptable” and 5 – “Exceeds Requirements”.

OQuaRE offers two scaling functions [2], which differ on how the metrics values are transformed into *quality scores*, and that provide complementary information:

<sup>4</sup> <http://miuras.inf.um.es/oquarewiki>

- *Static scaling function*: based on recommendations and best practices from the Software Engineering and Ontology Engineering communities. This method uses a predefined transformation function, so the value of a certain metric is always transformed into the same *quality score*.
- *Dynamic scaling function*: based on the observed values of the quality metrics of a corpus defined by a set of ontologies. The transformation function depends on the corpus of ontologies used, so the value of a metric is transformed into *quality scores* that depends on the corpus used.

We consider that the *static scale* is more appropriate for evaluating single versions of ontologies, whereas the *dynamic scale* can provide useful information about the evolution of an ontology. OQuaRE can indeed analyse consecutive versions of the same ontology providing information about the evolution of the ontology as it is further explained in [2]. The differences between consecutive versions are captured by the *mean change*, which is calculated using the *change in scale*. Every metric shown in Table 1 suffers a *change in scale* when the *quality score* for two consecutive versions is different. A *change in scale* can therefore be negative or positive. The magnitude of the *change in scale* is the absolute value of the difference between the scores. The *mean change* accounts for those metrics with *changes in scale* and produces one value by pair of versions. In addition to this, the *accumulative mean change* provides an overview of the changes produced between non-consecutive versions. For example, if the *mean change* of four consecutive versions is 0.2, -0.4 and 0 respectively, then, the *accumulative mean change* is -0.2. Finally, the OQuaRE framework is publicly available at<sup>5</sup> including web services to third-parties. OQuaRE uses R for data analysis.

## 2.2. OBO Foundry Member Ontologies

The Open Biomedical (OBO) Foundry initiative [6] proposes the creation of an orthogonal collection of ontologies by applying shared principles<sup>6</sup> for the coordinated evolution of ontologies. The ontologies of the OBO repository are either member or candidate ontologies. For an ontology to be a member, the OBO Foundry must have checked that they have been developed by following such criteria. In this work we analyse the eight member ontologies (June 2015). The number of versions for some ontologies was high, so we applied a sliding window filtering algorithm using the release date. Figure 2 (columns 1-3) describes the corpus using the BioPortal acronyms.

## 3. Results

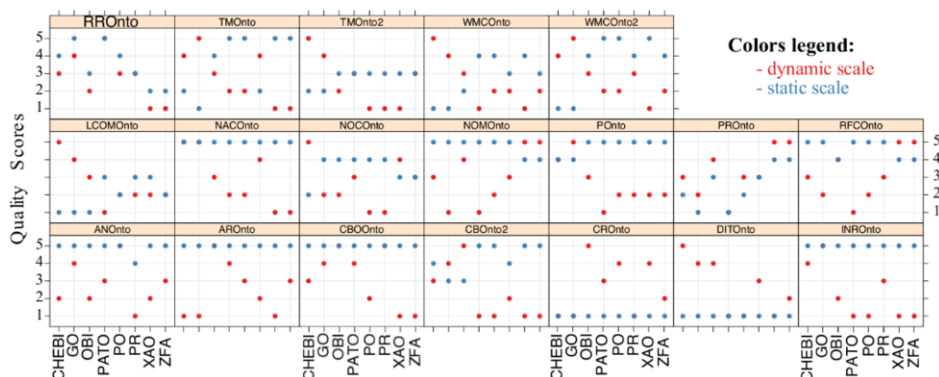
### 3.1. Comparative Analysis of the 8 OBO Foundry Member Ontologies

We describe the results of applying OQuaRE over the latest version of each ontology. The results are shown in Figure 1. There is one box by metric, the x-axis comprises the ontologies, and y-axis the *quality scores* for the *dynamic scale* (red) and for the static one (blue). Some metrics are analysed next. ANOnto is related to the “naming convention” OBO principle, which promotes the use of *rdfs:label* for the primary label

<sup>5</sup> <http://sele.inf.um.es/oquare>

<sup>6</sup> <http://www.obofoundry.org/principles>

and includes exactly one for every declared entity. Seven out of the eight ontologies get the highest score, “5”, using the *static scale*. This means that more than 80% of the entities have labels. However, PR does not reach this threshold, what could be reported to its developers. The *dynamic scale* complements this result since the scale is now obtained from the actual values of the metrics of the eight ontologies. According to this scale, GO gets the highest score, followed by PATO and ZFA. The analysis of the results with the *static scale* reveals that the highest scores are obtained in metrics that represent a proper use of annotations (ANOnto), the number of elements that can be related by properties (AROnto) and hierarchical relations (CBOnto, INROnto and NACOnto). The lowest scores are for CROnto (individuals per class) and DITOnto (depth of the hierarchies) although the CROnto score can be justified by the fact that those ontologies are not supposed to have individuals. Finally, more variability between ontologies is observed for the rest of metrics.



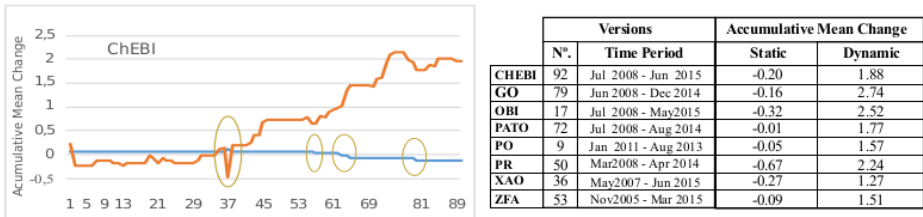
**Figure 1.** Quality scores for the latest version of each ontology in our experimental corpus.

### 3.2. Evolution of the Ontologies

The evolution of the ontologies has been studied by ontology. We have created eight independent corpus, including the versions of the same ontology. The trajectories of the *quality scores* for each ontology and scale can be inspected in our webpage<sup>7</sup>. The results of the accumulative mean change for ChEBI are shown in Figure 2. Blue and orange lines represent respectively the scores for the static and dynamic scales. ChEBI is relatively stable for the *static scale*, so this ontology does not suffer many changes in terms of *quality scores*. The four versions remarked by ovals are the ones with higher changes in the *quality scores*. If we analyse the results for the *dynamic scale*, the *accumulative mean change* is negative until version 37, turning into positive since then.

The *accumulative mean change* for the eight ontologies is shown in Figure 2 right. There, we can observe that it is negative for the *static scale*, and positive for the *dynamic scale*. Our experience reveals that the *dynamic scale* is more appropriate for analysing the evolution of ontologies, since it is able to capture smaller differences between values of the metrics. The *dynamic scale* shows that the evolution of the OBO members has produced ontologies with higher *quality scores* according to OQuaRE.

<sup>7</sup> <http://miuras.inf.um.es/oquare/mic2017>



**Figure 2.** Left) Profile of the *accumulative mean change* for ChEBI ontology. Right) *Accumulative mean change* score for the different member of our corpus.

## 4. Conclusions

The developers of biomedical ontologies need support to analyse the quality of their ontologies and how the evolution of the ontologies is affecting to such quality. We have proposed a method based on the OQuARE framework, which has been applied to study the evolution of the OBO Foundry member ontologies. Our results show that the ontologies follow the “naming convention” principle and that the evolution process followed is generating ontologies with higher quality scores, which is made explicit by the *accumulative mean change* metric. As future work, a deeper analysis of the relation of the OBO Foundry principles and OQuARE metrics will be performed. We believe that this kind of method may contribute to gain insights on the engineering of biomedical ontologies and support ontology developers in generating better resources.

## Acknowledgements

This work has been partially funded by to the Spanish Ministry of Economy and Competitiveness and by the Fundación Séneca through grants TIN2014-53749-C2-2-R and 19371/PI/14 respectively. Also partially funded by the FEDER Program.

## References

- [1] A. Duque-Ramos, J. T. Fernández-Breis, R. Stevens, and N. Aussenac-Gilles. OQuARE: A SQuAREbased approach for evaluating the quality of ontologies. *Journal of Research and Practice in Information Technology*, 43(2):159–176, 2011.
- [2] A. Duque-Ramos, M. Quesada-Martínez, M. Iniesta-Moreno, J. T. Fernández-Breis, and R. Stevens. Supporting the analysis of ontology evolution processes through the combination of static and dynamic scaling functions in OQuARE. *Journal of Biomedical Semantics*, 7(1):63, 2016.
- [3] A. Gangemi, C. Catenacci, M. Ciarmita, and J. Lehmann. Modelling ontology evaluation and validation. In *European Semantic Web Conference*, pages 140–154. Springer, 2006.
- [4] ISO (International Organization for Standardization). ISO/IEC 25000:2005, Software Engineering - Software Product Quality Requirements and Evaluation (SQuARE) - Guide to SQuARE, 2005.
- [5] J. Rogers. Quality assurance of medical ontologies. *Methods Inf. Med.*, 45(3):267–274, 2006.
- [6] B. Smith, M. Ashburner, C. Rosse, J. Bard, W. Bug, W. Ceusters, L. J. Goldberg, K. Eilbeck, A. Ireland, C. J. Mungall, N. Leontis, P. Rocca-Serra, A. Rutenberg, S.-A. Sansone, R. H. Scheuermann, N. Shah, P. L. Whetzel, and S. Lewis. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat Biotech*, 25(11):1251–1255, 11 2007.
- [7] S. Tartir and I. B. Arpinar. Ontology evaluation and ranking using OntoQA. In *ICSC '07: Proceedings of the International Conference on Semantic Computing*, pages 185–192, Washington, DC, USA, 2007. IEEE Computer Society.