Ontologies for the Study of Neurological Disease

Alexander P. Cox[1], Mark Jensen[1], William Duncan[1], Bianca Weinstock-Guttman[3], Kinga Szigiti[3], Alan Ruttenberg[2], Barry Smith[1] and Alexander D. Diehl[3*]

[1] Department of Philosophy, University at Buffalo, Buffalo, NY, USA

[2] Department of Oral Diagnostic Sciences, University at Buffalo School of Dental Medicine, Buffalo, NY, USA

[3] Department of Neurology, University at Buffalo School of Medicine and Biomedical Sciences, Buffalo, NY, USA

ABSTRACT

We have begun work on two separate but related ontologies for the study of neurological diseases. The first, the Neurological Disease Ontology (ND), is intended to provide a set of controlled, logically connected classes to describe the range of neurological diseases and their associated signs and symptoms, assessments, diagnoses, and interventions that are encountered in the course of clinical practice. ND is built as an extension of the Ontology for General Medical Sciences — a high-level candidate OBO Foundry ontology that provides a set of general classes that can be used to describe general aspects of medical science. ND is being built with classes utilizing both textual and axiomatized definitions that describe and formalize the relations between instances of other classes within the ontology itself as well as to external ontologies such as the Gene Ontology, Cell Ontology, Protein Ontology, and Chemical Entities of Biological Interest. In addition, references to similar or associated terms in external ontologies, vocabularies and terminologies are included when possible. Initial work on ND is focused on the areas of Alzheimer's and other diseases associated with dementia, multiple sclerosis, and stroke and cerebrovascular disease. Extensions to additional groups of neurological diseases are planned.

The second ontology, the NeuroPsychological Testing Ontology (NPT), is intended to provide a set of classes for the annotation of neuropsychological testing data. The intention of this ontology is to allow for the integration of results from a variety of neuropsychological tests that assay similar measures of cognitive functioning. Neuropsychological testing is an important component in developing the clinical picture used in the diagnosis of patients with a range of neurological diseases, such as Alzheimer's disease and multiple sclerosis, and following stroke or traumatic brain injury. NPT is being developed as an extension to the Ontology for Biomedical Investigations.

1 INTRODUCTION

The field of neurology deals with a diverse domain of diseases related to the functioning of the nervous system in all its aspects, including diseases resulting from disorders of the central, peripheral, and autonomic nervous systems. Neurological diseases may exhibit both acute and chronic courses, affect a variety of cell types and anatomical regions of the body. They are manifested via a variety of mechanisms, including cell-autonomous disorders, unregulated protein aggregation, autoimmune conditions, and vascular pathology, which, depending on the disease, may occur alone or together in various combinations (Ropper *et al*., 2005; Merritt and Rowland, 2000). At a different level of granularity we see neurological diseases that affect cognitive as well as mental functioning. Following Ceusters and Smith (2010), we maintain that mental diseases are (at least primarily) special kinds of neurological diseases in the sense that the disorder, which serves as the material basis for the disease, is a part of an anatomical structure in the organism responsible for producing and maintaining cognitive representations and behavior. For example, a variety of neurological conditions result in dementia, such as Alzheimer's and Parkinson's disease, and many of the late-onset leukodystrophies.

We have recently begun building a new ontology for the domain of neurological diseases – the Neurological Disease Ontology (ND). ND is an ongoing project that aims to accurately represent every facet of neurological diseases in as much detail as possible. This includes their clinical presentation, diagnosis, treatment, physical manifestation, course of development, genetic and physical bases, and more. ND is still in the early stages of development, but is rapidly growing to include more of these facets. While our ultimate goal in developing ND is to provide a comprehensive account of all neurological diseases, it has three initial areas of focus: Alzheimer's disease (AD), multiple sclerosis (MS), and stroke and cerebrovascular events. At this time, the most progress has been made on AD and other diseases that result in dementia, but work is currently under way on representing MS and associated demyelinating diseases as well as on representing stroke and cerebrovascular disease.

As a corollary to ND, we have begun development of the NeuroPsychological Testing Ontology (NPT) to represent neuropsychological assessments such as the Folstein Mini-Mental State Examination (MMSE), the Trail- Making Test, the Hopkins Verbal Learning Test, and the Wechsler Memory Scale. These standardized assessments are useful for identifying the presence and degree of cognitive impairment in patients (Lezak *et al*., 2004). An initial goal of the NPT project is to test hypotheses about the diagnosis of AD based on the results of neuropsychological assessments. Part of the development of NPT necessitates reference to aspects of cognitive functioning. For example, MMSE produces scores that are indicative of impairment in certain functional cognitive domains such as language, executive function, or memory. A challenge we have encountered is how to connect these commonly described cognitive domains to functioning on the side of the organism. We see this as an ex-

* To whom correspondence should be addressed: addiehl@buffalo.edu

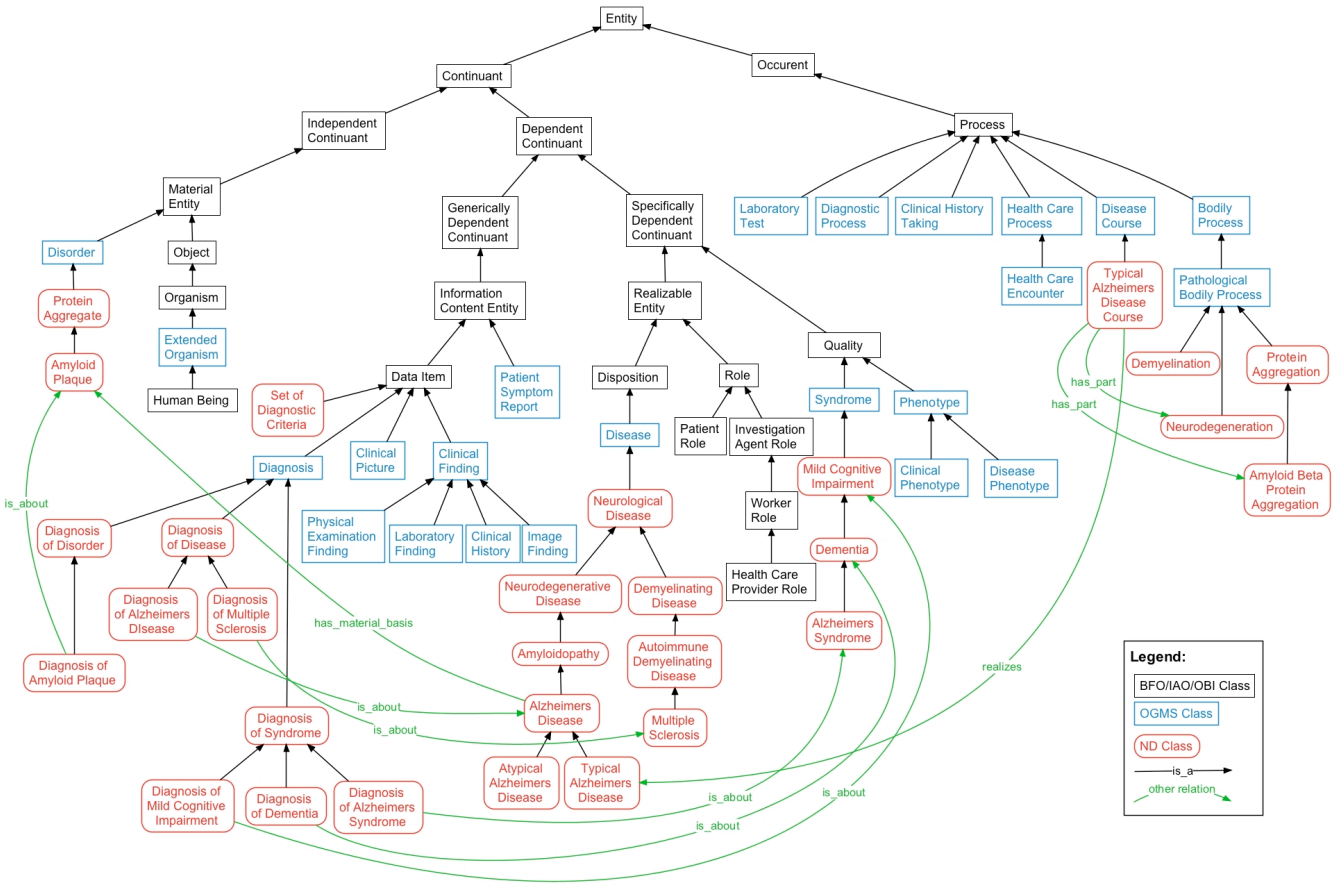

Figure 1: A subset of OGMS and ND and some connections to external ontologies.

cellent opportunity to connect ND and NPT with work in the Mental Functioning Ontology (MF) as well as with the Mental Disease Ontology (MD). Ideally we hope to drive development in both. For example, an extension of MD that represents dementia from the perspective of it being a mental disease or syndrome could then be linked via logically defined relations to classes in ND.

We plan to build ND over the long-term in a collaborative manner with other groups focused on representing particular neurological diseases as modules within ND. Our work is intended to be OBO-Foundry compliant and builds upon the paradigm established by Ontology for General Medical Sciences (OGMS) for the representation of entities in the domain of medicine and disease (Scheuermann et al. 2009).

2 METHODS

ND and NPT are being curated using both top-down and bottom-up approaches to the creation of classes within the ontology. A major aspect of the top-down approach for ND has involved analyzing what types of neurological diseases exist and how they ought to be represented within the ontology according to their relevant characteristics. Of some concern is how our strategy will fit with other disease ontologies. A key element includes deciding what other types of entities should be represented in ND in order to accurately represent the neurological diseases as well as how the relationships between these classes should be represented. For instance, the class 'neurological disease' currently includes 'neurodegenerative disease', 'infectious neurological disease', 'demyelinating disease', and 'vascular neurological disease' as four of its subclasses. The inclusion of these subclasses was driven by our decision to focus, as much as possible, on representing neurological diseases from the perspective of their etiology. For example, it is part of the logical definition for 'neurodegenerative disease' that all realizations of these diseases involve some process of neurodegeneration. This top-down approach provides ND with its primary structure.

Due to the complex nature of neurological diseases, as well as the diversity of perspectives from which they are studied and classified, we have also included additional immediate subclasses of 'neurological disease'. For example, 'central nervous system disease' and 'peripheral nervous system disease' are included as subclasses of 'neurological disease'. Currently we do not explicitly assert any disease

as a subclasses of these classes, however ND is being built using axioms that will allow an ontological reasoner to automatically create an inferred hierarchy of neurological disease types based on anatomical structure or genetic basis. This approach allows ND more versatility without committing it to a single perspective or creating confusion by switching between perspectives within the asserted hierarchy. Another example of this approach is creation of the defined class 'disease resulting in dementia', which has a limited number of asserted subclasses, and was created to provide a reference class from which to allow a reasoner to infer a hierarchy of all diseases that result in dementia.

While the top-down aspect of the project is essential to shaping the development of ND, it is the bottom-up aspect of the project that provides the bulk of the information. In particular, it is this approach that results in the creation and refinement of the definitions for terms in ND. We have consulted primary research articles, review articles, medical professionals, and other sources to inform the development of ND. This process has led to the inclusion of new terms in ND as well as more detailed classifications of particular neurological diseases. Both approaches are necessary for the completion of the project.

Development of NPT is based upon analyses of neuropsychological tests to drive the development of classes for the representation of neurological assays and their results. Many neuropsychological tests have multiple subtests, and these are being captured within the ontology as well. Neuropsychological tests assay domains such as verbal and visual-spatial memory, executive function, and linguistic functions. NPT is being developed to allow the integration of scores from different neuropsychological tests and subtests so that results for patients who have been tested using different protocols can be queried and grouped appropriately.

ND and NPT are built using Protégé 4.1 as OWL2 ontologies. The importation of classes from other ontologies according to the MIREOT standard has been achieved using OntoFox (Xiang, 2010).

Both ND and NPT are being developed according to OBO Foundry principles (Smith *et al.*, 2007) and is being done in cooperation with the related efforts to develop ontologies for representing Mental Disease (MD) and Mental Functioning (MF) (Hastings *et al.* 2012a and 2012b).

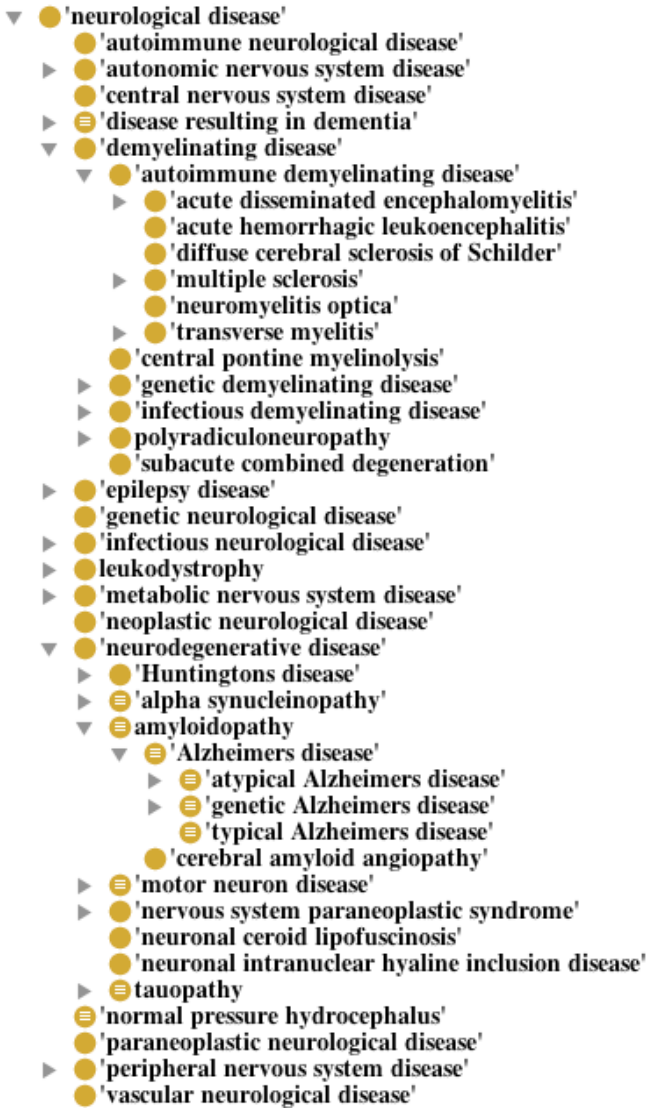

Figure 2: A portion of the ND disease hierarchy.

3 RESULTS

The Neurological Disease Ontology is being built according to OBO Foundry principles as an extension of OGMS, which provides a set of general reference classes related to diseases, their patients, and diagnoses (Scheuermann *et al.* 2009). OGMS follows the paradigm of the Basic Formal Ontology (BFO). Figure 1 illustrates the layers of granular-

Ontology Name	Use in ND
Basic Formal Ontology (BFO)	Top-level reference ontology
Ontology for General Medical Sciences (OGMS)	Mid-level reference ontology
NIF-Dysfunction *and* Disease Ontology (DO)	Externally referenced disease classes
Relation Ontology (RO)	Imported relation types
Protein Ontology (PR)	Select classes for proteins imported via MIREOT
Foundational Model of Anatomy (FMA)	Select classes for anatomical structures imported via MIREOT
IAO, PATO, ChEBI, GO, CL, and OBI	Select classes imported via MIREOT

Table 1. External ontologies used by the Neurological Disease Ontology.

ity captured by the relations between ND, OGMS, and BFO as well as IAO and OBI. Furthermore, we are ensuring that ND is compliant with the pre-release revised version of BFO – BFO 2.0, and the revised version of OGMS that is also compliant with BFO 2.0.

In building ND, we have relied upon a number of sources, including reference works, review articles, and other ontologies, such as NIF-Dysfunction and the Disease Ontology (DO) (Bug *et al.,* 2008; Larson & Martone, 2009). Based on these sources we have curated a high-level disease hierarchy that we believe presents a useful initial approach to categorizing neurological diseases, a section of which is shown in Figure 2. We go beyond earlier efforts at creating disease ontologies by providing textual definitions for every disease class and by incorporating logical definitions in order to relate classes for diseases and other entities in ND to other classes in ND and to separate ontologies (See Table 1 for a summary).

These high level disease classes provide a framework for the in depth curation of ND ontology modules intended to represent neurological diseases in extensive detail. At the University at Buffalo, our initial efforts are focused upon the areas of Alzheimer's disease and other diseases resulting in dementia, multiple sclerosis, and stroke and cerebrovascular disease. As an early stage ontology development project, ND currently contains approximately 400 classes; about 250 classes have textual definitions; more than 50 classes have logical definitions; more than 150 classes have external references; and there are nearly 200 children of the class 'disease'. In addition to disease classes, ND has a heavy focus on diagnosis, syndrome, disorder, and protein classes among others in order to fully represent all of the various aspects of neurological diseases.

In building NPT we have relied upon source tests, such as the Folstein Mini-Mental State Exam, as well as upon textbooks and articles about particular neuropsychological tests (Lezak *et al.*, 2004; Mitrushina *et al.*, 2005). NPT is built using the schema for representing assays that has been developed in OBI and consequently currently imports all of OBI. At a later point, we will rely upon a slimmed (MIREOTed) version of OBI. At the moment, there are more than 250 NPT specific classes, but we expect this to grow quickly as we add representations of additional neuropsychological tests. Figure 3 shows a portion of NPT for the representation of the MMSE.

4 DISCUSSION

Our use cases in building these ontologies include annotation of clinical studies in neurology as well as annotation of patient records. Particularly for the latter case we expect ND and NPT to complement each other, with ND providing terms for representing the diagnoses of patients based on their signs and symptoms, and associated phenotypes. NPT will provide a very detailed set of classes for annotation of neuropsychological measures that may be used in the formation of a patient's clinical picture, which is used to reach a diagnosis. These diagnostic conclusions are annotated as an

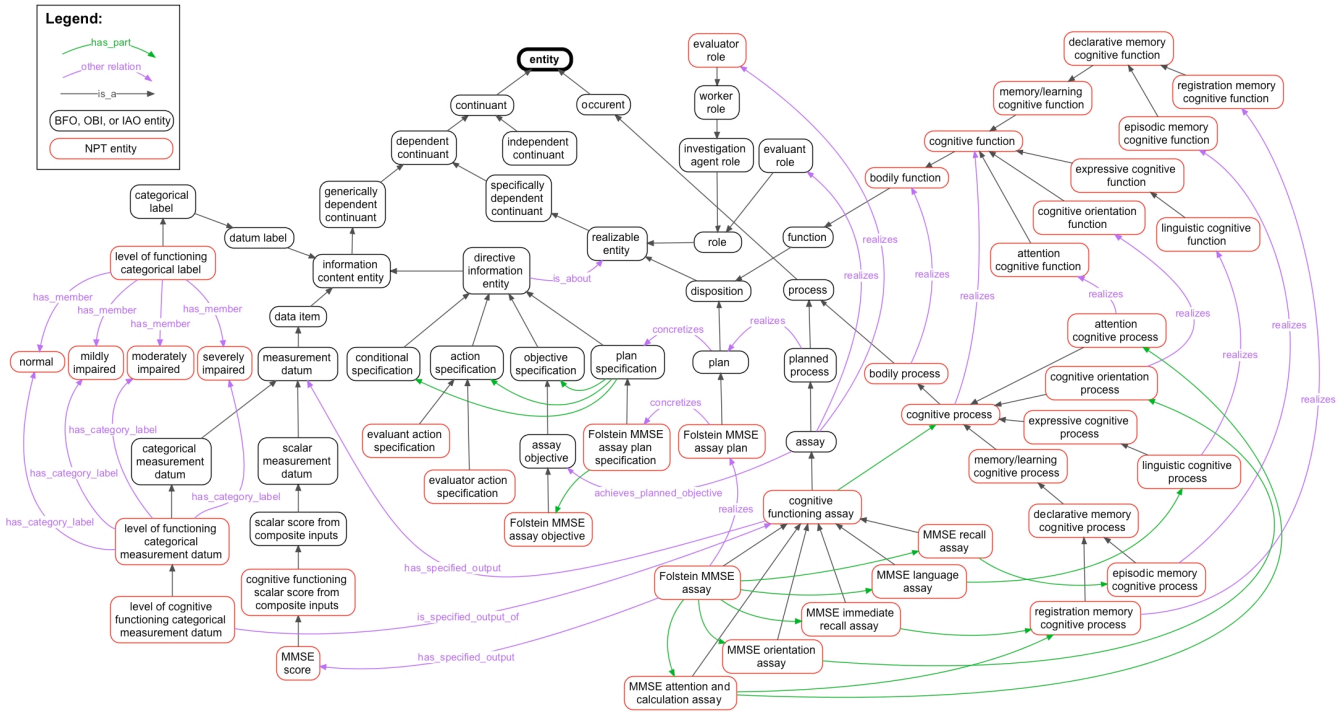

Figure 3: A portion of the representation of the MMSE assay in NPT.

instance of a diagnosis class in ND. The diagnosis classes are linked to the disease classes in ND, which themselves link via their logical definitions to other classes in ND such as the disorder which serves as the material basis of the disease, and then, in turn, to other ontologies such as PR.

In developing ND and NPT we recognize the need to coordinate with other ontology development efforts in related domains. In Ceusters and Smith (2010), for instance, the framework for what are now named the Mental Functioning Ontology (MF) and the Mental Disease Ontology (MD) was presented. Neurological diseases by their very nature often affect cognitive and mental functioning, for instance in any disease that results in dementia, such as Alzheimer's disease, and often lead to mental diseases, such as depression in MS or epilepsy patients. In developing ND we will need to ensure representation of conditions such as dementia or depression are coordinated with MF and MD, such that a class representing a clinical phenotype of "depression in conjunction with multiple sclerosis" may have a parent class of "depression" in MD. Moreover we feel that our work can aid in a bottom-up approach to developing MD and MF.

Furthermore, we believe our work on NPT will be valuable for the annotation of neuropsychological data not just for patients with neurological disease, but also for studies of general mental functioning and in testing in patients with mental diseases. Thus, our work on NPT will hopefully prove of value for a number of related domains in addition to that of neurological diseases, and will eventually be complemented by ontologies for other types of assessments of nervous system function and anatomy, such as an MRI imaging ontology.

ACKNOWLEDGEMENTS

We would like to thank Ralph Benedict, Ph.D., of the Department of Neurology, University at Buffalo, for guidance in understanding neuropsychological testing, and Naveed Chaudhry, Marcus Ng, and Donat Sule for assistance with term development in ND.

REFERENCES

Bug, W. J., Ascoli, G. A., Grethe, J. S., Gupta, A., Fennema-Notestine, C., Laird, A. R., et al. (2008). The NIFSTD and BIRNLex vocabularies: Building comprehensive ontologies for neuroscience. Neuroinformatics, 6(3), 175-194.

Ceusters, W. and Smith, B. (2010). Foundations for a realist ontology of mental disease. *Journal of Biomedical Semantics*, **1**(1), 10.

Hastings, J., Smith, B., Ceusters, W., Jensen, M., and Mulligan, K. (2012a). The mental functioning ontology. Available at http://code.google.com/p/mental-functioningontology/

Hastings, J., Smith, B., Ceusters, W., Jensen, M., and Mulligan, K. (2012b). Representing mental functioning: Ontologies for mental health and disease. See proceedings for this workshop: Towards an Ontology of Mental Functioning, ICBO 2012, Graz, Austria.

Larson, S. D., and Martone, M. E. (2009). Ontologies for Neuroscience: What are they and What are they Good for? *Frontiers in Neuroscience*, 3(1), 60-67.

Lezak, M. D., Howieson, D. B., and Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). Oxford: Oxford University Press.

Merritt, H. H. and Rowland, L. P. (2000). *Merritt's Neurology* (10th ed.). Philadelphia: Lippincott Williams & Wilkins.

Mitrushina, M., Boone, K. B., Razani, J. and D'Elia, L. (2005). Handbook of Normative Data for Neuropsychological Assessment (2nd ed.). Oxford: Oxford University Press.

Ropper, A. H., Adams, R. D., Victor, M., Brown, R. H., and Victor, M. (2005). *Adams and Victor's Principles of Neurology* (8th ed.). New York: McGraw-Hill Medical Pub. Division.

Scheuermann, R., Ceusters, W., and Smith, B. (2009). Toward an ontological treatment of disease and diagnosis. In AMIA Summit on Translational Bioinformatics, San Francisco, California, March 15-17, 2009, pages 116–120. Omnipress.

Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L. J., Eilbeck, K., Ireland, A., Mungall, C. J., The OBI Consortium, Leontis, N., Rocca-Serra, P., Ruttenberg, A., Sansone, S.-A., Scheuermann, R. H., Shah, N., Whetzel, P. L., and Lewis, S. (2007). The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat Biotechnol*., **25**(11), 1251–1255.

Xiang Z, Courtot M, Brinkman RR, Ruttenberg A, and He Y. (2010). OntoFox: web-based support for ontology reuse. *BMC Res Notes*, **3**, 175.