Formal Ontology in Information Systems
C. Eschenbach and M. Grüninger (Eds.)
IOS Press, 2008

doi:10.3233/978-1-58603-923-3-21

Ontology (Science)

Barry SMITH[1]
Department of Philosophy, Center of Excellence in Bioinformatics and Life Sciences, and National Center for Biomedical Ontology, University at Buffalo, USA

Abstract. Increasingly, in data-intensive areas of the life sciences, experimental results are being described in algorithmically useful ways with the help of ontologies. Such ontologies are authored and maintained by scientists to support the retrieval, integration and analysis of their data. The proposition to be defended here is that ontologies of this type – the Gene Ontology (GO) being the most conspicuous example – are a *part of science*. Initial evidence for the truth of this proposition (which some will find self-evident) is the increasing recognition of the importance of empirically-based methods of evaluation to the ontology development work being undertaken in support of scientific research. Ontologies created by scientists must, of course, be associated with implementations satisfying the requirements of software engineering. But the ontologies are not themselves engineering artifacts, and to conceive them as such brings grievous consequences. Rather, ontologies such as the GO are in different respects comparable to scientific theories, to scientific databases, and to scientific journal publications. Such a view implies a new conception of what is involved in the authoring, maintenance and application of ontologies in scientific contexts, and therewith also a new approach to the evaluation of ontologies and to the training of ontologists.

Keywords: scientific method, expert peer review, ontology engineering, biomedical informatics, Gene Ontology, OBO Foundry

1 Introduction

For some time now the Gene Ontology (GO) [1] has enjoyed the status of a *de facto* standard vocabulary for the annotation of experimental data pertaining to the attributes of gene products. The GO has been widely applied to data drawn from experiments involving organisms and biological processes of many different types. It has also been subject to a series of logical reforms, which have enhanced the degree to which it can be exploited for algorithmic purposes. The GO is now routinely used in gene expression analyses of a wide range of biological phenomena, including phenomena relevant to our understanding of human health and disease.

The thesis to be defended here is that *the GO and its sister ontologies are a part of science*. This means (i) that these ontologies themselves are properly to be understood as results of scientific activity, analogous to journal publications (in some ways also to textbooks and databases), and (ii) that the processes involved in authoring, maintaining and evaluating them are a part and parcel of the activity of science.

In what follows I shall draw out some implications of this thesis, focusing my attentions on the GO and on the other biomedical ontologies participating in the OBO

---

[1] Corresponding author: Barry Smith, Department of Philosophy, 130 Park Hall, University at Buffalo, Buffalo, NY 14260, USA; E-mail: phismith@buffalo.edu.

Foundry initiative [2,3]. These provide the most conspicuous examples of ontology (science) in the sense here intended. The views expressed will appear to many to be self-evident; in their detail, however, they are still exploratory in nature (and thus they do not represent any settled policy of the Foundry initiative).

2 The OBO Foundry

The Open Biomedical Ontologies (OBO) repository was created in 2001 by Michael Ashburner and Suzanna Lewis as a means of providing convenient access to the GO and its sister ontologies at a time when resources such as the NCBO BioPortal [4,5] did not yet exist. The OBO Foundry was initiated by Ashburner, Lewis and Smith in 2005 as a collaborative experiment designed to enhance the quality and interoperability of life science ontologies from the point of view of both biological content and logical structure [3]. The Foundry initiative is based on the voluntary acceptance by its participants of an evolving set of principles designed to maximize the degree to which ontologies can support the needs of working scientists. The developers of nearly all of the ontologies within the OBO repository have committed themselves to participate in this initiative, which has spawned also the establishment of a number of new ontology projects within the Foundry framework.

2.1. The OBO Foundry Principles

The principles of the Foundry can be summarized, in their current version, as follows.

First, are syntactic principles to the effect that an ontology submitted to the Foundry must employ one or another common shared syntax, possess a unique identifier space, and have procedures for identifying distinct successive versions.

Second, are principles involving definitions: the Foundry requires that textual definitions (and, by degrees, equivalent formal definitions) be provided for all terms; that terms and definitions be composed using the methodology of cross-products (see below); and that ontologies use relations that are unambiguously defined according to the pattern set forth in the OBO Relation Ontology (RO) [6].

Third, ontologies are required to be open (available to be used by all without any constraint), to have a clearly specified and clearly delineated content, to have a plurality of independent users, and to be subject to a collaborative development process involving the developers of Foundry ontologies covering neighboring domains.

Finally, the Foundry embraces a principle of orthogonality. This asserts that for each domain there should be convergence upon a single ontology that is recommended for use by those who wish to become involved with the Foundry initiative. If an ontology is submitted which overlaps substantially with an existing Foundry ontology, then the two sets of developers will be invited to collaborate in the creation of a common, improved resource, through application of the sorts of evidence-based strategies applied in other parts of science to resolve the problems that arise where alternative theories of a single phenomenon are advanced by competing groups.

2.2. The Problem of Data Silos

The primary rationale for our insistence upon the principle of orthogonality is that it offers a potential solution to a pressing problem facing researchers in information-driven areas of biology and biomedicine, namely the problem of the siloing of data. In part in response to NIH mandates, many such researchers have recognized the need to find ways to present their data in a form that will make them more easily combinable and also accessible to wider communities of researchers. Ideally, this would be achieved by constraining terminologies and data schemes so that they converge on commonly accepted standards [7]. Unfortunately, however, there is normally still no clear answer as to what, in any given case, should serve as basis for such constraint. Many researchers therefore find that they have little choice but to create their own local schemes for description of their data.

The OBO Foundry proposes a solution to this problem that is incremental, modular, empirically based, and incorporates a strategy for motivating potential developers and users. Briefly, the ontologies in the Foundry are being built as orthogonal, interoperable modules which together form an incrementally evolving network. Scientists are motivated to commit themselves to developing ontologies falling within their domains of expertise because they themselves will need to use these ontologies in their own work in the future. Users are motivated by the assurance that the ontologies they adopt from the Foundry will be maintained in the future by scientists with the relevant sorts of expertise. Both arms of this strategy for motivation can be realized effectively only against a background in which the principle of orthogonality is accepted by all involved.

2.3. Benefits of Orthogonality

Further benefits brought by acceptance of the principle of orthogonality include:

First, it helps those new to ontology who need to know where to look in finding ontology resources relating to their subject-matter for which they can have reasonable assurance that they have been validated and will be used and maintained in consistent fashion by subject-matter experts; that they will work well with other established ontologies; and that the expertise acquired in adapting these resources to specific local needs will potentially be of general and lasting utility.

Second, it obviates the need for ‘mappings’ between ontologies, which have proved not only difficult (and expensive) to create and use, but also error-prone and hard to keep up-to-date when mapped ontologies change.

Third, it ensures the mutual consistency of ontologies, and thereby also the additivity of the annotations created with their aid by different groups of annotators describing common bodies of data. In this way, orthogonality contributes to the cumulativity of science and allows new forms of unmanaged collaboration.

Fourth, it rules out the sorts of simplification and partiality which may be acceptable under more pluralistic regimes, and thereby brings an obligation on the part of ontology developers to commit to strive for scientific accuracy and domain-completeness in their work.

Fifth, orthogonality provides support for the Foundry’s strategy of utilizing *cross-products in composing terms and definitions* [8,9]. This strategy is designed (i) to reduce the degree of arbitrariness typically involved in term composition in complex ontologies, and (ii) to ensure that Foundry ontologies are developed in tandem in such a way as to constitute a progressively more well-integrated modular network. The idea is that, where ontologies need to include complex representations (for example of: *effects of viral infection on cell function in shrimp*), these should be built up compositionally out of component representations (here: *virus*, *infection*, *cell*, *function, shrimp*) already defined within other, more basic feeder ontologies. By enforcing orthogonality (and the use of relations derived from the RO for term combination), we can go far towards ensuring a unique choice for such composition that serves at the same time to bind the more specialized ontologies to the benchmark feeder ontologies from which constituent terms are drawn.

Finally, orthogonality helps to eliminate redundancy and it serves the division of ontological labor in ontology development work. It allows communities from different biology disciplines to address the tasks of ontology building to different levels of detail and under different timetables. It makes possible the establishment of clear lines of authority, whereby experts in each domain are able to take responsibility for creating and maintaining a single, high-quality ontology that is tailored for that domain, adjustments to which are then passed on to those other ontologies which have used its resources in composing terms and definitions via cross-products, thereby bringing further benefits of cross-ontology synchronization.

2.4. What Orthogonality is Not

When ontologies are seen as analogous to scientific theories [10], then orthogonality is a principle which arises naturally. This is because it is a pillar of the scientific method that scientists should strive always to seek out and resolve conflicts between competing theories. This is why scientists have over centuries made a huge investment in intra- and interdisciplinary synchronization, illustrated for example in the use of common standard systems of measurement units. In order to allow detection of conflicts and testing of proposed resolutions, scientists must work as far as possible within a single universe of collaborators and scientific results must be (by definition) available to all.

Where, however, ontologies are conceived as engineering artifacts, then orthogonality is neither practically achievable nor, from the perspective of ontology creators, intrinsically desirable. Here there prevail quite different disciplinary mores, in some respects comparable to those obtaining in the world of commercial enterprise.

When the orthogonality principle has been subjected to criticism in engineering circles, this has sometimes been because the principle itself has been misinterpreted as resting on a view to the effect that there can be only one correct way to represent the entities in each domain of life science research. In fact, however, all of those involved in the Foundry initiative are aware that the Foundry ontologies represent nothing more than initial attempts to address difficult problems, and that they rest on hypotheses that are always subject to further revision and supplementation. We are also fully aware that, in this as in all other domains, scientific advance rests on the to-and-fro of criticism between the advocates of competing hypotheses. We thus see considerable benefit in the development of alternative sets of ontologies by other groups, even if at the same time we warn of a shared need for strategies to counter potential dangers of silo formation. We also envisage scenarios under which externally developed ontologies would be incorporated into the Foundry because they are of superior quality to those which they would then replace.

2.5. The Strategy of Reference Ontologies

Another criticism raised by engineers against the orthogonality principle is that it will cause problems for ontology users who require their own purpose-built ontologies to address specific needs. In fact, however, the Foundry offers a strategy to address such special purposes in ways that do not contribute to the formation of silos.

This strategy rests on a view of ontologies used in science as being divided into two kinds. On the one hand are the so-called *reference ontologies* [11], arranged orthogonally within the Foundry itself. On the other hand is a larger edifice of different types of *application ontologies* constructed on this foundation, the whole being connected together, prospectively, through application of the methodology of cross-products, and employing strategies for networking of the sort now being tested within the framework of the Semantic Web [12].

Reference ontologies are analogous, in different ways, to both scientific theories and textbooks. Each has its own subject-matter, which consists of the entities in reality addressed by the corresponding branch of biomedical science. Each seeks to maximize descriptive adequacy to this subject-matter by being built out of representations which are correct when viewed in light of our best current scientific understanding.

Application ontologies, in contrast, are comparable to engineering artifacts. They are constructed for specific practical purposes such as management of data in a multi-institution clinical trial [13]. The problems arise because such artifacts are still normally built afresh for each new trial or study. What results may then serve immediate needs perfectly well, but it creates snowballing obstacles as researchers need to reuse the associated data for other purposes – for example to share them with colleagues working on cognate phenomena, or to perform meta-analyses.

Our proposal is that application ontologies should as far as possible be developed from the start in alignment with a common set of reference ontologies such as are provided by the OBO Foundry [14]. This is achieved by employing terms residing in these reference ontologies (thus preserving their existing identifiers) and using them to build new terms via composition. Requests should be submitted to the relevant Foundry ontologies where needed terms are not available. Only in this way, we believe, can the tendency towards silo formation be counteracted and the associated obstacles to the retrieval, reuse and integration of data thereby be prospectively reduced.

A final criticism of the orthogonality principle, made by Musen, turns on the question of whether it is in fact possible to find, for each domain of reality, a single perspective that ‘is deserving of being canonized as a reference ontology’. From the standpoint of the ‘engineering faction’, as Musen sees it:

> [i]t would not make sense to talk about, say, an overarching ontology for biomedical investigations when, for example, the distinctions required to describe an experiment for publication are different from those required to describe an experiment to assist scientists in the execution of the experiment, which are different from those required to estimate the cost of the experiment to a sponsor, which are different again from those required to determine whether the experiment is ethical. (Personal communication)

This passage is interesting not least because there does in fact exist an Ontology for Biomedical Investigations (OBI) [15], which is one of the most successful new ontologies being created *ab initio* in accordance with OBO Foundry principles. Its goal is to provide controlled, structured representations for the design, protocols, instrumentation, materials, data and data analysis in biological and biomedical investigations of all types. This goal has been embraced by some two dozen communities representing different domains of high-throughput experimentation, ranging from flow cytometry to in situ hybridization and immunohistochemistry. We are learning from the experience of OBI development how difficult it is to do serious ontology work in a large and heterogeneous domain where such work necessarily involves the contributions of multiple groups of specialists with different sorts of domain expertise. To make such contributions work effectively within a single framework requires a complex process of a sort which mimics within a single ontology the process for cross-ontology coordination being applied by the Foundry as a whole.

OBI is, like every other ontology, a work in progress; it serves none of the context-specific purposes listed by Musen perfectly; but it serves all of them to some degree. OBI has been recognized as a reference ontology by the OBO Foundry because all of the communities involved recognize the advantages brought by creating a single ontology that can serve as a stable attractor for the many communities who need a framework that can already serve annotation of data describing how experimental results were obtained while being subjected to incremental improvement and expansion.

3 Science is Cumulative

Central to ontology (science) is the requirement that ontologies, like scientific theories, should be tested empirically. This requirement is realized in the context of the OBO Foundry through the work of the many biologists who contribute to the maintenance of its member ontologies from day to day [16]. This they do by aggressively using these ontologies in the annotation of new experimental results, in a development that has given rise to a new scientific profession of literature curator [17]. Because new results reported in the journal literature need to be annotated using corresponding reference ontologies, this generates new content for and corrections to these ontologies, thereby providing enhanced resources for literature curation in the future. This virtuous cycle is exemplified already in the work of a plurality of life science research communities, and the methodology has been thoroughly tested especially by the model organism research communities within the Gene Ontology Consortium [18].

Researchers in information-driven disciplines of contemporary biology are hereby realizing in a new form a pattern that has been characteristic of empirical science since its inception. Simplifying greatly, we can say that each branch of science is marked by the existence of a consensus core of established results surrounded by a changing penumbra of hypotheses that are to different degrees marked as problematic. This consensus core was earlier documented in textbooks. Increasingly, it will be documented also in ontological form.

Empirical science is *cumulative* in the sense that the consensus core of each discipline grows by absorbing hypotheses which began as problematic but have withstood attempts to refute them empirically. The process of cumulation is, of course, marked at every stage by setbacks and false starts and by the competition between theories referred to already above. Except in those rare periods in which sciences are undergoing revolutionary change, however – for example the change from Newtonian physics to special relativity – these factors will not be sufficient to dislodge the broad mass of propositions making up the consensus core.

The goal of the OBO Foundry can now be characterized as follows. *First*, and as it were on the object level, it is to provide a coherent and interoperable suite of controlled structured representations of the entities and relations described at any given stage in the consensus cores of each of the biological sciences. This framework is designed to be maximally stable, in order to provide a basis for the progressive cumulation of the scientific data described in its terms. At the same time the framework must be flexible enough to accommodate change as new experiments are performed, new results discovered, and old hypotheses refuted. *Second*, and on the meta-level, it is to establish ontology development itself as being, like statistics, a recognized part of the scientific enterprise. This brings the need to determine, incrementally and empirically, the consensus core of ontology (science), and to train a community of ontology experts who will be in a position to apply and to extend this core in their scientific work. The set of Foundry principles represents an initial glimpse of what this consensus core might contain. The overarching goal – whose significance we are only now beginning to understand – is to serve the ends of cumulativity (which means: preventing silos) in an era where the advance of scientific research is increasingly being mediated by computers, and thus increasingly subject to the influence of engineers whose incentives have sometimes been at odds with those of working scientists.

4 Ontology and Expert Peer Review

4.1. The Foundry Strategy

To become established as a properly scientific activity, ontology development must be subject to processes of evaluation of the same sort that are practiced in other parts of science. In this light, we believe that benefits can be gained from a view of ontologies as being, in crucial ways, analogous to scientific publications, and thus as subject to the discipline of *expert peer review*. The OBO Foundry has accordingly been experimenting with procedures designed to pave the way for the incorporation of the methodology of expert peer review into ontology development practice.

Progressively, each ontology submitted to the Foundry will be subject to review (1) by *Coordinating Editors* whose primary responsibility is that of harmonizing interactions (of content and of logic) between Foundry ontology development projects in neighboring domains;[2] and (2) by *Associate Editors* selected by those involved in the development and maintenance of the ontologies in the OBO repository, whose task is to provide input from these separate ontology developer communities.

The ontology peer review process will involve also ad hoc discipline-based reviewers, who will be selected on the basis of their specific scientific expertise, and who will evaluate ontologies not as computational artifacts but as representations of scientific domains. To this end it is important that there are ways to translate Foundry ontologies not only into multiple different computational formats [19] but also into something close to English [20]. In this way, ontologies such as the GO exist, and serve as objects for evaluation, in forms that are independent of specific computational implementations. In this respect, too, they are like scientific theories.

---

[2] Currently, the Foundry Coordinating Editors are, in addition to Ashburner, Lewis and Smith, also Christopher Mungall (a leader in the GO and model organism database communities), Alan Ruttenberg (principal scientist of Science Commons and Chair of the OWL Working Group), and Richard Scheuermann (principal investigator of the ImmPort Immunology Database and Analysis Portal and of the BioHealthBase Bioinformatics Resource Center projects).

4.2. Advantages of Expert Peer Review

As ontology engineers have criticized the principle of orthogonality, so also they have resisted the application to ontologies of the methodology of expert peer review [21]. It will thus be worth our while to summarize briefly some of the benefits that peer review has brought to the practice of science, benefits which have led to its adoption by scientific publishers, universities, and research and funding agencies in their quest for scientific quality.

Expert peer review provides an impetus to the improvement of scientific knowledge over time, as authors compete for scarce funding or for occupation of prestigious journal space [22,23]. It not only improves the quality of published papers through the *ex post* revisions fostered by reviewer comments, but also helps to discipline scientific communication as a result of the fact that authors are aware *ex ante* that their results need to be formulated in such a way that they will be intelligible to unknown, critical peers with powers of sanction.

Because peer review introduces an element of expert judgment independent of authors and editors, this lends it some of the functionality of an audit process. It serves as a filter to detect duplication, fraud or distorted information, and hence it is valued by regulatory agencies, which see it as providing a partial validation of scientific results.

These filters are of course not perfect. Thus far, however, no other vetting device has been offered that would do a better job. Moreover, some of the proposed alternatives have been shown to be marked by even more severe failings [24].

Filtering based on the judgment of experts brings benefits also to readers, since they need only absorb and collate vetted manuscripts, as opposed to all the manuscripts submitted to the relevant journals and to journal-like respositories. As Bug points out, such filtering promises to be especially useful in the field of biomedical ontology:

> Until there is a reliable vetting procedure, we cannot expect to re-use and extend existing ontologies effectively or with confidence for the purpose of bringing like data together in novel ways from across the biomedical data diaspora. Without vetting, we cannot expect to provide other developers with clear advice on what are the reliable ontological shoulders to build on. [25]

The fact that we currently often have multiple ontologies covering single domains at the same scope and level of granularity generates problems:

> how [can] a bioinformatics application developer determine which one to use? Even more importantly, if users pick at random from amongst the two or more ontologies covering the same domain, who will maintain the maps and software required to make deductions or inferences across the annotated data repositories which use these different ontologies to cover the same domain? [25]

4.3. Creating Incentives for Investment of Effort in Ontology Development

The need for ontology resources on the part of scientific and clinical researchers is ever increasing. And while attempts are being made to meet this need through automatic generation of ontologies, there is an increasing recognition of the fact that successful ontology development will require a considerable contribution from human experts. Typically, however, ontology work – like its counterpart in the field of database development – brings rewards incommensurate with the effort that must be invested to yield seriously useful results. One set of incentives being brought into play within the

Foundry to address this problem rests on motivating factors relating to the exercise of *influence*. Briefly, individuals will be motivated to commit themselves to investment in ensuring the adequacy of a given resource if they know (a) that they themselves will be using that resource in the future, and (b) that they will gain benefits if that same resource is used also by watchful colleagues, some of whom will then embrace a similar commitment. The importance of this sort of motivation has been demonstrated already in open source endeavors in the field of software standards. Such endeavors are, as documented by Weber [26], subject to an ever-present danger of forking. This danger must be constantly counteracted if incentivizes for involvement are to be maintained. Weber shows that the open source process is most likely to achieve this end when addressing tasks that have the following characteristics:

1. Disaggregated contributions can be derived from knowledge that is … not proprietary.
2. The product is perceived as important and valuable to a critical mass of users.
3. The product benefits from widespread peer attention and review, and can improve through creative challenge and error correction.
4. There are strong positive network effects to use of the product.
5. An individual or a small group can take the lead and generate a substantive core that promises to evolve into something truly useful.
6. A voluntary community of iterated interaction can develop around the process of building the product.

The likelihood of success in realizing these characteristics seems to be highest where the community effort is organized on the basis of a pyramidal authority structure resting to a high degree on delegation. In each successive phase of the work, positions of authority are assigned by those already holding such positions to individuals who have demonstrated both commitment to the effort and relevant expertise. The Foundry is an attempt to realize a structure of this sort within the life science ontology domain.

4.4. The Strategy of Expert Peer Review of Ontologies

A second set of incentives is provided by bringing about a situation in which ontology developers would receive career-related credit by having their ontologies count as analogues of peer-reviewed scientific journal publications. This would allow citations of ontologies to be measured in the same way as are citations of other sorts. It would allow ontology reviewers to gain credit analogous to that currently awarded for membership in journal editorial boards. It would enable also crucial elements of a scientific career path for ontologists, given that career advance in academic institutions rests on the existence of peer review-based mechanisms for evaluation of scientific competence which, in the ontology domain, have hitherto been lacking.

As in the case of traditional journal submissions, so also in the case of ontologies, the peer review strategy which the OBO Foundry is pilot testing will be an iterative process, with recommendations for revision being addressed in successive versions of the ontology until a stage is reached where it is deemed suitable for publication.

One obvious problem for such a strategy turns on the fact that ontologies, in contrast to traditional journal publications, are subject to continuous update. This problem has however been addressed already by those publishers who have brought scientific databases within a peer review framework. The Nature Publishing Group (NPG), for example, has addressed the issue of data curation speed in relation to its Signaling Gateway[27] by employing wiki tools to allow responses submitted by users to supplement peer reviewed data. NPG is however careful to insist that, in experiments such as this, 'It must be made clear to the user … which information has been peer reviewed and which has not.' [28]

A further problem for ontology peer review turns on the special role of users. As Musen puts it, while the job of reviewing journal articles is performed 'rather well by scientists who are experts in the field and who can understand the work … described', the key question of whether an ontology makes the right distinctions about its domain

> can be answered only by application of the ontology to some set of real-world problems and discovering where things break down. The people best suited for making the kinds of assessment that are needed are not necessarily the best experts in the field, but the mid-level practitioners who actually do the work. Any effective system of peer review has got to capture the opinions of ontology users, and not just those of renowned subject-matter experts or of curators. [29]

These remarks are well taken. But we believe that they do not imply that there is some problem with the methodology of peer review as the Foundry conceives it. Expert users of ontologies are already included among the Foundry reviewers, and the OBO Consortium has established strategies for taking account of user input through a heavily utilized system of open access sourceforge trackers and email forums [30].

5 Ontology Evaluation via Democratic Ranking

5.1. A Strategy for Community Based Review of Ontologies

Scientific ontologies are often highly complex. They are subject to a high velocity of change, not only in virtue of scientific advance, but also because the associated computational technologies are themselves rapidly evolving. As new applications for ontology-based technology are identified, so new ontologies are being created, bringing problems of choice and validation to potential users. To address these problems the NCBO [31] and the Networked Ontology Consortium (NeOn) [32] are carrying out experimental tests of software-based strategies to support ontology assessment.

In essence, these strategies address goals addressed also by the Foundry editorial process. Both seek a particular kind of quality assurance of ontologies. Both rely on human reviews of ontologies. On the approaches advanced by NCBO and NeOn, however, the community of those involved in providing reviews is (potentially at least [33]) larger than on the more selective approach favored by the Foundry. This is because one key element of the NCBO and NeOn approaches is inspired by the 'open' Web-based systems for the rating of consumer goods developed by organizations such as amazon.com or eBay [34,35]. The resultant strategy for 'democratic ranking' is described by Lewen as one according to which 'everyone can write reviews about the ontologies', and 'some of the reviewers can (and should) be … experts'.

> Not only does this approach scale (everybody can review), it is also very personalizable. It is up to the user to decide whether she values the opinion of a 'mere ontology user' more than the opinion of an 'ontology expert'. [36]

5.2. Problems with Democratic Ranking

On the democratic ranking approach ‘trust scores’ will be dynamically assigned to the authors of reviews by a larger community of users, on the basis of numerical responses to the question: *was this review helpful to you?* [37] It is then assumed, reasonably, that users will be drawn to the reviews of those who have received trust scores which are higher than the average. But will the latter also be those who have the necessary expertise, integrity, diligence, free time, and frankness to do their reviewing job properly? The empirical data that would enable us to answer this question are not as yet available. Already, however, there are reasons to question whether experts in scientific disciplines would devote their time to making contributions to an open ranking system of this sort:

1. The very idea that scientifically relevant decisions can be made on the basis of democratic vote will seem to them absurd. The evidence that this is so is easily acquired by talking to scientists. Their instinctive rejection of the idea turns on the fact that scientific decisions – as contrasted to decisions concerning, for example, choice of consumer goods – are tied logically to myriad further decisions made by other scientists, sometimes over long spans of time, on the basis of bodies of experimental evidence that are often too complex to be comprehended by any single person. It is for this reason that the processes of scientific decision-making are so involved, and why they have led to the evolution of formal and informal institutions which may seem cumbersome and antiquated to outsiders.

2. One such institution is the practice of reviewer confidentially, which brings the benefit of enhancing the ability of reviewers to express opinions frankly. On the strategies for ‘open’ democratic ranking advanced by NCBO and NeOn, this benefit will be lost. Certainly, the policy of open review, too, can bring benefits of its own: some may be motivated to write more thorough reviews, and perhaps thereby gain credit and acknowledgement. The prognosis for the success of such a policy is however poor, not least because the potential hazards for authors of negative reviews (ranging from career problems to lawsuits) will lead many potential reviewers to refrain from participating [33].

3. Career-related credit would seem not to accrue under the envisaged open ranking systems. Academic institutions do not promote on the basis of rankings assigned on the Web by non-experts.

4. A strategy centered on a user-based ranking of reviewers will it seems fall short of realizing one vital purpose inherent to the methodology of expert peer review, namely that of reducing search and decision costs on the part of those involved in research. For the reasons given by Bug in the passages quoted above, we should avoid the temptation to place these costs once more into the hands of researchers for the sake of an ‘openness’ whose benefits in the scientific context are as yet unproven.

5. Under the mix-and-match selection procedures envisaged by NCBO and NeOn, the views of experts will not only be subject to a filtering process involving non-experts, but also potentially diluted through admixture with non-expert views. This will at least diminish those sorts of motivation for serious investment of time in ontology development that rely on those who make such investments enjoying the opportunity to play a direct role in shaping resources which they themselves will need in the future.

We should thus view with caution arguments such as those advanced by Lewen to the effect that the proposed open ranking systems will solve a ‘problem with restricted reviewing systems’, namely that ‘they are very vulnerable to personal preferences, prejudices and reviewers' egos.' [38] For it is one important lesson of the enduring success of expert peer review in so many different fields over more than three centuries that some biases (roughly: the complex set of learned biases we call 'expertise') need to be imposed upon the mix in order to ensure even minimal coherence. The reliance on experts brings, to be sure, a certain tendency in favor of established (i.e. most commonly accepted) scientific paradigms. But it is not clear how the addition of more voices to the mix should help resolve this problem, particularly if so doing has the effect of driving away just those persons who are in the position of making contributions resting on scientific expertise.

As Sowa points out, the Web has brought about a situation in which

> [p]ublication is almost free, and we have the luxury of decoupling the reviewing process from the gatekeeping process. Metadata enables that decoupling … The metadata associated with each submission can indicate what tests were made, what the reviewers said, and what results the users, if any, obtained. Users can choose to see ontologies sorted by any criteria they want: in the order of best reviews, most thorough testing, greatest usage, greatest relevance to a particular domain, or any weighted combination. [38]

The problem, however, is that the obverse sign of these very same advances in the direction of publishing freedom is the potential for what we might call a *poisoning of the wells*. Sowa thus agrees with the Foundry on the importance of maintaining an expert peer review process having a level of rigor that is comparable to that of existing scientific journals. This view is supported also by the experience of open access journals such as *PLOS ONE* who have experimented with dual frameworks involving both expert peer review and community-based dialogue on published articles.

Certainly there is one sort of openness that is essential to the advance of science. Science progresses only if it is open to new hypotheses and to new criticisms of existing hypotheses. It is this which explains why there are multiple, independent publishers of scientific journals, and why new journals are constantly being established. It is this which explains, also, why Noy is right to warn against a situation in which reviewers would be restricted to 'the experts appointed by a closed board'.

From this, however, it does not follow that our only alternative is a situation in which users would be allowed to draw only on the expertise of those reviewers willing to have their opinions made subject to review under an 'open' democratic voting process. A middle way would draw precisely on the lessons learned from scientific publishing by allowing users of ontology selection software to draw also on the contributions of experts employing traditional methods of ('closed') peer review. The NCBO BioPortal [5] might, for example, provide its users with the option to bypass the democratically supplied rankings of reviewers and to move directly to a prevalidated list of ontologies such as is provided by the Foundry (and we note that, to the degree that users of ontologies are encouraged to take advantage of such an option, silo problems will be averted in the future). The Foundry might itself then also be able to benefit from comments of Bioportal users, for example in supporting vetting of Foundry ontologies for errors, or in assessing the degree to which the terms used in ontologies might gain consensus approval on the part of significant numbers of users.

6 Conclusion: Ontology (Science) vs. Ontology (Engineering)

We can summarize our arguments by pointing to certain special features which are possessed by reference ontologies created to serve scientific purposes. Such ontologies are developed to be (1) common resources (thus they cannot be bought or sold), (2) for representation of well-demarcated scientific domains; they are (3) subject to constant maintenance by domain experts, (4) designed to be used in tandem with other, complementary ontologies, and (5) independent of format and implementation.

Sadly, the view still predominating in engineering circles is that ontologies need possess none of these features because ontologies are of their nature engineering artifacts [39]. It is as if all ontologies, both inside and outside science, are assigned by default the status of application ontologies. This leaves no room for any foundation of application ontologies in reference ontologies, and thus undermines what we believe to be the only promising strategy for addressing the problem of silo creation. Indeed it reinforces those very expectations on the part of many ontology engineers which have done so much to cause this problem in the first place.

We believe, in this light, that if we are to have a chance of resolving the silo problem, then recognition of this fact must bring in its wake a new approach to the training of ontologists working in support of scientific research, based on a new set of expectations to the effect that the authoring and maintenance and evaluation of scientific ontologies is an incremental, empirical, cumulative, and collaborative (i.e., precisely, scientific) activity that must be carried out by experts in the relevant scientific domains. Practitioners of ontology (science) will need to learn to see ontologies in contexts in which they are required to work well not only from a logical and a technological point of view, but also from the point of view of supporting the advance of science. To bring about these changes it may be necessary also to address the degree to which educational opportunities for ontologists are still largely confined to departments of computer science, and thereby also to address the degree to which tenure decisions for ontologists are made on a basis which awards too little weight to the contributions made by ontology to scientific research.

These recommendations receive support from Akkermans and Gordijn [40], who point out that computer scientists and knowledge engineers still standardly conceive ontologies as computer science artifacts, which means that they still see an ontology developed to serve (for example) biology as 'just another application' of their own computational expertise, and thus as something that is of lesser scientific importance than core computer science issues for example in logic or in systems for ontology mapping [40]. This 'self-limiting approach' will in the end 'not be able to exploit the full potential of the ontology idea', and Akkermans and Gordijn accordingly insist that the ontologies developed for scientific purposes need to be taken much more seriously as first-class citizens by computer scientists and knowledge engineers.

Empirical evidence of the benefits to be gained from such a move has been accumulating for some time, and we are gratified that this recognition is now beginning to make itself manifest also within the framework of the Semantic Web. There, too, we can see how mature ontologies, often resting on input from the OBO Foundry, are finally beginning to be put to serious scientific use [41].

Acknowledgements. This work was funded through the NIH Roadmap for Medical Research, Grant 1 U 54 HG004028. With thanks to Robert Arp, Holger Lewen, Suzanna Lewis, Mark Musen, Natasha Noy, Mitsu Okada, Richard Scheuermann, John Sowa, and Holger Stenzhorn for helpful comments.

References

[1] http://www.geneontology.org/. Last accessed June 24, 2008.

[2] http://obofoundry.org/. Last accessed June 24, 2008.

[3] B. Smith *et al.*, The OBO Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration, *Nature Biotechnology* **25** (2007), 1251–1255.

[4] D. Rubin *et al.*, The National Center for Biomedical Ontology: Advancing Biomedicine through Structured Organization of Scientific Knowledge, *OMICS*, **10** (2006), 185–198.

[5] http://www.bioontology.org/ncbo/faces/index.xhtml. Last accessed June 24, 2008.

[6] B. Smith *et al.,* Relations in Biomedical Ontologies, *Genome Biology*, **6** (2005), R46.

[7] J. J. Cimino and X. Y. Zhu, The Practical Impact of Ontologies on Biomedical Informatics, *Yearbook of Medical Informatics*, 2006, 124-35.

[8] D. P. Hill, J. A. Blake, J. E. Richardson, M. Ringwald, Extension and integration of the Gene Ontology: combining GO vocabularies with external vocabularies, *Genome Research*, **12** (2002),1982–1991.

[9] C. J. Mungall, Obol: Integrating Language and Meaning in Bio-Ontologies, *Comparative and Functional Genomics* **5** (2004), 509–520.

[10] C. Rosse and J. L. V. Mejino, Jr. “The Foundational Model of Anatomy Ontology”, in A. Burger, *et al.* (eds.), *Anatomy Ontologies for Bioinformatics: Principles and Practice*, New York: Springer, 2007.

[11] C. Rosse and J. L. V. Mejino Jr. A Reference Ontology for Bioinformatics: The Foundational Model of Anatomy, *Journal of Biomedical Informatics*, **36** (2003), 478–500.

[12] H. Stenzhorn, *et al.* “Simplifying Access to Large-Scale Health Care and Life Sciences Datasets “, The Semantic Web: Research and Applications (Lecture Notes in Computer Science 5021 (2008), 864-868.

[13] R. D. Shankar, *et al.*, Epoch: An Ontological Framework to Support Clinical Trials Management. *International Workshop on Healthcare Information and Knowledge Management*, 2006, 25–32.

[14] J. F. Brinkley, *et al.* A Framework for Using Reference Ontologies as a Foundation for the *Semantic Web*, *Proceedings of the AMIA Fall Symposium*, 2006, 96–100.

[15] http://obi.sourceforge.net/. Last accessed June 24, 2008.

[16] D. P. Hill, *et al.*, Gene Ontology annotations: What they mean and where they come from, *BMC Bioinformatics*, **9** (Suppl. 5) (2008), S2.

[17] http://tesuque.stanford.edu/biocurator.org/. Last accessed June 24, 2008.

[18] Gene Ontology Consortium, The Gene Ontology (GO) project in 2006, *Nucleic Acids Research*, **34** (Database Issue) (2006), D322-D326.

[19] http://www.berkeleybop.org/ontologies/. Last accessed June 24, 2008.

[20] Rassinoux A. M., *et al*., Coupling Ontology Driven Semantic Representation with Multilingual Natural Language Generation for Tuning International Terminologies, *Medinfo*, **12** (2007), 555–9.

[21] http://ontolog.cim3.net/forum/ontology-summit/2008-03/msg00076.html. Last accessed June 24, 2008.

[22] T. Jefferson, *et al*. Effects of editorial peer review – A systematic review. *Journal of the American Medical Association*, **287** (2002), 2784–2786.

[23] M. Ware, *Peer Review: Benefits, Perceptions and Alternatives* (PRC Summary Papers 4), 2008.

[24] Editorial: Peer Review and Fraud. *Nature* **444** (2006), 971–972.

[25] http://ontolog.cim3.net/forum/ontology-summit/2008-03/msg00070.html. Last accessed June 24, 2008.

[26] S. Weber, *The Success of Open Source*, Cambridge, MA: Harvard University Press, 2004.

[27] *Nature Molecule Pages*: http://www.signaling-gateway.org/molecule/. Last accessed June 24, 2008.

[28] B. Riley, Database Publication Presents Unique Challenges for the Peer Reviewer, *Nature* (2006), doi:10. 1038/nature04993.

[29] http://ontolog.cim3.net/forum/ontology-summit/2008-03/threads.html/. Last accessed June 24, 2008.

[30] http://www.obofoundry.org/cgi-bin/discussion.cgi.

[31] M. Musen, If We Build It, Will They Come? Available for download at: http://ontolog.cim3.net/cgi-bin/wiki.pl?ConferenceCall_2007_12_06/. Last accessed June 24, 2008.

[32] http://www.neon-project.org/. Last accessed June 24, 2008.

[33] Despite enthusiasm for the concept, open peer review was not widely popular, either among authors or by scientists invited to comment. *Nature* (December, 2006). DOI:10.1038/nature05535.

[34] N. Noy, R. Guha, and M. Musen, User Ratings of Ontologies: Who Will Rate the Raters? *Proceedings of the AAAI 2005 Spring Symposium on Knowledge Collection*. Available for download at: http://smi.stanford.edu/smi-web/research/details.jsp?PubId=1112. Last accessed June 24, 2008.

[35] B. Smith, The Evaluation of Ontologies: Editorial Review vs. Democratic Ranking, *Proceedings of InterOntology 2008* (Tokyo, Japan, 26-27 February 2008), 127–138.

[36] http://ontolog.cim3.net/forum/ontology-summit/2008-03/msg00074.html. Last accessed June 24, 2008.

[37] H. Lewen, K. S. Supekar, N. F. Noy, M. A. Musen, Topic-Specific Trust and Open Rating Systems: An Approach for Ontology Evaluation. *Workshop on Evaluation of Ontologies for the Web* (EON2006). Available at: http://smi.stanford.edu/smi-web/research/details.jsp?PubId=1142.

[38] http://ontolog.cim3.net/forum/ontology-summit/2008-03/msg00076.html. Last accessed June 24, 2008.

[39] http://ontolog.cim3.net/forum/ontology-summit/2008-03/msg00067.html/. Last accessed June 24, 2008.

[40] H. Akkermans and J. Gordijn, Ontology Engineering, Scientific Method and the Research Agenda, in S. Staab and V. Svatek (Eds.), *EKAW 2006* (Lecture Notes in Artificial Intelligence, 4248), 2006, 112-125.

[41] A. Ruttenberg, *et al*. Advancing Translational Research with the Semantic Web, *BMC Bioinformatics*, **8** (Suppl 3), (2007), S2.