Draft version to be presented at FOIS 2008. Comments welcome.

Ontology (Science)

Barry SMITH
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. The ontologies created by scientists must, of course, be associated with implementations satisfying the requirements of software engineering. But these ontologies are not themselves engineering artifacts, and to conceive them as such brings grievous consequences. Rather, we shall argue, ontologies such as the GO are comparable to scientific theories, to scientific databases, or to scientific journal publications. Such a view implies a radically new conception of what is involved in the authoring, maintenance and application of ontologies in scientific contexts, and therewith also a radically 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 it 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 textbooks, databases, or journal publications, 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 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, including most notably the Ontology for Biomedical Investigations, which brings together some two dozen disciplinary communities from different domains of high-throughput biological and biomedical experimentation [[6]].

2.1.OBO Foundry Principles

The principles of the OBO 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) [[7]].

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 other 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 are invited to collaborate in the creation of a common, improved resource, through application of the sorts of strategies applied as a matter of course in other parts of science wherever 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 we believe that it offers a potential solution to a pressing problem facing researchers in information-driven areas of biology and biomedicine, namely the problem of data silos. Currently, the many groups involved in biomedical research find that they have little choice but to create their own local schemes for description of their data. Ideally, their efforts would be invested in the shared development of common schemes that would make their data interoperable with those of their colleagues, and there is now widespread recognition that it would be advantageous to constrain terminologies and data schemes so that they converge on commonly accepted standards [8]. Unfortunately, however, there is normally still no clear answer to the question as what, in any given case, should serve as basis for such constraint.

The OBO Foundry proposes a solution to this problem that is incremental, modular, empirically based, and such as to embody a coherent strategy for resolving the problem of motivating potential developers and users. Briefly, it is a solution which requires that ontologies be built as orthogonal, interoperable modules within an incrementally evolving network. Each constituent module within this network is authored and maintained by scientists who need to draw on its resources for their work in annotating (describing in algorithmically useful ways) both their own experimental results and also related results captured in scientific databases and in the journal literature. It is then the need for suitable resources to annotate such results which progressively dictates the ontologies' evolving content. Domain experts are hereby rewarded for their participation through the fact that they play a direct role in shaping the resources that they will need in their own work in the future.

The whole endeavor is entirely voluntary. Yet experience so far suggests that, in at least a significant portion of life science domains, adoption of the Foundry strategy will be considerable.

2.3.Benefits of Orthogonality

In addition to providing a strategy for resolving the silo problem, some further benefits of the principle of orthogonality include:

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

Second, the requirement of orthogonality obviates the need for 'mappings' between ontologies, which have proved not only difficult to create and use, but also error-prone and hard to keep up-to-date when mapped ontologies change.

Third, orthogonality 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.

Fourth, orthogonality provides support for the Foundry's strategy of utilizing cross-products in composing terms and definitions [[9],[10]]. This strategy is designed, both to reduce the degree of arbitrariness typically involved in term composition in complex ontologies, and to ensure that Foundry ontologies are developed in tandem in such a way as to constitute a progressively more well-integrated and hierarchically organized modular network. The idea is that, where ontologies need to include complex representations (for example of: *effects of viral infection on cell function*), these should be built up compositionally out of component representations (here: *virus*, *infection*, *cell*, *function*) 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 serves the division of ontological labor in ontology development work. It allows different disciplinary communities to address the task of ontology building at different speeds, at different levels of detail, and (initially at least) with different levels of axiomatic rigor. It makes it possible the establishment of clear lines of authority whereby experts in each single domain are able to take responsibility for creating and maintaining a single, high-quality ontology module tailored for that domain, adjustments to which are then passed on automatically to those other ontologies which use its resources in composing terms and definitions via cross-products, thereby bringing further benefits of cross-ontology synchronization.

2.4.Misinterpretations of Orthogonality

When ontologies are seen as analogous to scientific theories, then orthogonality is a principle which arises naturally. It is a pillar of the scientific method that scientists should strive always to resolve conflicts between competing theories.

When, however, ontologies are conceived as engineering artifacts, then orthogonality is neither practically achievable nor, from the perspective of ontology creators, intrinsically desirable. Ontology engineers gain benefits from a situation in which ontologies must normally be created anew for each new situation, and thus it is understandable that the orthogonality principle has been subject to criticism in engineering circles.

One such criticism is that their adoption of this principle would somehow imply that Foundry members are asserting that all ontology development in biology and biomedicine should take place only within the confines of the Foundry itself. The principle is interpreted, in other words, as asserting not merely that there should ideally be one single ontology for each domain of life science research, but further that this single ontology must be one that has been approved for inclusion within the Foundry.

In fact, however, all of those involved in the Foundry initiative collaborate as a matter of course with other ontology developer and user groups. We are fully aware that 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.

2.5.The Strategy of Reference Ontologies

Another criticism is that the principle will cause problems for ontology users who require special-purpose ontologies in order to address their own specific practical needs even where corresponding legacy ontologies already exist. In fact, however, the Foundry provides what we believe is a coherent strategy to address such individual needs that is based, again, on the methodology of cross-products.

This strategy rests on a view of ontologies 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 *application ontologies* constructed on this foundation, the whole being connected together through application of the methodology of cross-products and employing strategies for networking of the sort currently being tested within the framework of the

Semantic Web. Just as clinical medicine relies on basic sciences such as anatomy or molecular biology to provide integration across medical specialisms, so, we believe, the biological and clinical information-processing applications of the future will need increasingly to rely on shared reference ontologies to integrate the data arising in their specialist domains of inquiry [11].

It is strictly speaking reference ontologies which are analogous to scientific theories. 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 constructed for specific practical purposes such as management of data in a multi-institution clinical trial [12]. The problems which need solving are thus not problems with ontology engineering artifacts as such. Rather, they are problems which arise where such artifacts are built afresh for each new trial or study. For while the latter may serve local needs perfectly well, they create snowballing obstacles as successive groups of researchers face the need to reuse data for other purposes – for example to share them with colleagues working on cognate phenomena, to perform meta-analyses, or to move to larger domains. 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. 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 prospectively reduced.

3 Science is Cumulative

Central to the OBO Foundry initiative is the requirement that ontologies, like scientific hypotheses, should be tested empirically. This requirement is realized not only through the fact that the Foundry's ontologies are authored by life scientists in response to needs arising within their work, but also through the role of biologist-curators of experimental literature in the maintenance of these ontologies from day to day [13]. We now have considerable experience in applying procedures to ensure that literature curation and ontology maintenance are able to work in tandem for the mutual benefit of each. Because new developments 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. Not only do existing ontologies expand in this way in step by step fashion, but also new ontologies come to be created in reflection of new avenues of research opened up for example by new technologies for experimentation.

This virtuous cycle is exemplified already in the work of a plurality of life science research communities in a development that has given rise to a new profession of scientific literature curator [14]. The methodology has been thoroughly tested by the model organism research communities within the Gene Ontology Consortium [15], who are realizing in a new form a pattern that has been characteristic of empirical science since its inception. Simplifying greatly, we can think of each branch of science as being 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. This 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 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 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 it needs to be flexible enough to accommodate change. Second, and on the meta-level, it is to establish ontology development itself as a recognized part of the scientific enterprise. This brings the need to determine, incrementally and empirically, the consensus core of ontology (science), and to nurture and 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 one first 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 *expert peer review*. The OBO Foundry has accordingly been experimenting with procedures designed to pave the way for the incorporation of a peer review methodology into ontology development practice.

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

We are experimenting also with procedures for involving ad hoc discipline-based reviewers, who will be included in the reviewing process in light of their specific scientific expertise, and who will evaluate ontologies not as computational artifacts but as representations of scientific domains. Already the reference ontologies made available within the Foundry exist in multiple different formats [[16]], and to serve such evaluation ways will need to be found to formulate ontology content using something close to a natural language such as English. In this way, ontologies such as the GO will exist, and serve as objects for evaluation, in forms which are independent of specific computational implementations. We are however developing also strategies for evaluation relating to implementation, addressing factors such as conformability to relevant logico-syntactic standards, support for interoperability, utilization of the cross-product methodology, sound naming policies, and so on.

While reference ontologies must be associated with implementations satisfying the requirements of software engineering, they are not themselves to be identified with such implementations. In this respect, too, they are like scientific theories.

4.2.Advantages of Expert Peer Review

As ontology engineers have criticized the principle of orthogonality, so also they have resisted the idea of expert peer review [17]. It will thus be worth our while to summarize briefly some of the benefits that peer review has brought to the practice of science, and which have led to its adoption by scientific publishers, universities, and funding agencies in their quest for scientific quality.

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 [18,19]. It thereby 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 must 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. Thus peer review serves as a filter to detect duplication, fraud or distorted information and hence is valued by regulatory agencies, who 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 which would do a better job. Moreover, some of the proposed alternatives have been shown to be marked by even more severe failings [20].

Peer review filtering also brings benefits to readers, since they need only read, 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 [21]:

> 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.

For as long as we have multiple ontologies covering a given domain at the same scope and level of granularity,

> how would 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?

4.3.Creating a System of Incentives for Investment of Effort in Ontology Development

Peer review by experts can not only help in this way to solve the silo problem; it can also yield a strategy to address the problems of providing incentives for ontology authorship and maintenance, where currently such work (like its counterpart in the field of database development) typically brings rewards incommensurate with the time and effort that must be invested to yield seriously useful results. The Foundry is thus exploring strategies by which a peer review seal of approval might contribute to the motivation of researchers to invest time and effort in advancing the quality and interoperability of ontologies. The goal is for ontology developers to receive career-related credit by having their ontologies count as analogues of peer-reviewed scientific journal publications on the basis of a strategy which would also allow the multiple developers typically involved in complex ontology endeavors to receive appropriate credit for their respective partial contributions. Ontology reviewers, similarly, would gain credit in the same way that membership in journal editorial boards is currently rewarded by academic institutions. We are aware that the number of peers with the competence to carry out such reviews is still quite small, and that we will face problems in training and marshalling the human resources needed to effect the sorts of serious review that will be required in relation to what is already a large and burgeoning body of target ontologies. On the other hand however adoption of a peer review system along the lines we have in mind will itself bring benefits in motivating experts to become involved in the processes of ontology development, training and evaluation, not least in that it will support the development of a coherent scientific career path for those who demonstrate the capacity for high quality work in both developing and reviewing ontologies.

It will also bring into play an additional set of motivating factors, relating to the exercise of *influence*, whose importance has been demonstrated already not only in science but also in open source endeavors on the field of software standards. As documented by Weber [[22]], the open source process is most likely to work effectively in tasks that have these characteristics:

1. Disaggregated contributions can be derived from knowledge that is accessible under clear, non-discriminatory conditions, not proprietary or locked up.
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 structure resting to a high degree on delegation, within which positions of authority are held by individuals whose expertise and commitment to the effort are acknowledged by all of those involved. The Foundry is an attempt to realize a structure of this sort within the ontology domain. Its principals are motivated to participate within the constrained environment provided by the Foundry because this gives them the opportunity to shape the ontology resources that will be available to them in their own work in the future.

4.4.Problems with the Strategy of Expert Peer Review of Ontologies

As in the case of traditional journal submissions, so also in the case of submitted ontologies, the peer review strategy which the OBO Foundry is currently 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 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, is addressing the issue of data curation speed in relation to its Signaling Gateway [23], by experimenting with the use of wiki tools in order 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.' [24]

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 the domain being modeled

> 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. [25]

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. This is because expert users of ontologies are already included among the Foundry reviewers, and because we have established strategies for taking account of user input through an elaborate and heavily utilized system of open access sourceforge trackers and email forums.

5 Ontology Evaluation via Democratic Ranking of Ontologies

5.1.A Strategy for Community Based Review

Software engineers, for understandable reasons, favor strategies in which software would substitute wherever possible for human experts [26]. Scientific ontologies are often highly complex artifacts. They manifest a high velocity of change, not only because of scientific advance, but also because the associated frameworks for computational reasoning are rapidly evolving. Moreover, as new applications for ontology-based technology are identified, this means that new ontologies are being developed, bringing problems of choice and validation to potential uses. To address these problems the NCBO [27] and the NeOn (Networked Ontology) Consortium [28] are carrying out experimental tests of software-based strategies to support ontology assessment.

In essence, these strategies address the same goals as those addressed by the Foundry editorial process. Both seek a particular kind of quality assurance when it comes to ontology selection. Both rely on human reviews of ontologies. On the kinds of approach advanced by NCBO and NeOn, however, the community of those involved in providing such reviews is (potentially, at least [29]) much larger than on the more selective approach that is favored by the Foundry. One key element of these approaches is inspired by the systems for the rating of consumer goods developed by organizations such as amazon.com or eBay [30,31]. The resultant strategy for 'democratic ranking', as described by Holger Lewen, is 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'. [32]

5.2.Problems with Democratic Ranking

On the democratic ranking approach, in contrast, experts would be those whose reviews received above-average 'trust scores' dynamically assigned by the larger community of users on the basis of numerical responses to the question: *was this review helpful to you?*

Ideally, users will be drawn to the reviews of those who have been positively evaluated by other users, and the latter will be those reviewers who have the necessary free time, expertise, integrity, diligence, and frankness to do their job properly.

Unfortunately, however, it is highly questionable whether experts in scientific disciplines would in fact devote their time to making contributions to an open ranking system of this sort [33].

First, 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 with decisions concerning, for example, choice of consumer goods – are tied logically to myriad further decisions made by other scientists 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 come to draw on institutions which, to outsiders, will seem cumbersome and antiquated.

One such institution is the practice of reviewer confidentially. This means however that there is a second problem faced by the strategies for democratic ranking advanced by NCBO and NeOn, namely that the benefits of confidentiality – effectively, the ability to express opinions frankly – will be lost. Certainly, the policy of open review, too, brings benefits: some will 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 [29], not least because of the potential hazards (potentially including lawsuits) which arise to authors of negative reviews.

Third, under the mix and match selection procedures described 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 the ontology resources of the future.

Fourth, career-related credit would seem not to accrue under such a system (academic institutions do not promote on the basis of rankings assigned on the Web by non-experts). More generally, there is a danger that a strategy centered on a user-based ranking of reviewers will fall short of realizing the vital purposes, inherent to the methodology of expert peer review, 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.

Certainly there is one sort of openness that is essential to the advance of science. Science progresses only if it is open to new ideas 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, too, why Noy is right to warn against a situation in which reviewers would be restricted to 'the experts appointed by a closed board' [34]. As already mentioned, however, the Foundry initiative is fully aware that scientific advance requires a constant interchange of criticism between competing groups of experts and why it welcomes the development of competing federations of ontologies resting on competing sets of principles, in order that the empirical benefits of different approaches to ontology development and evaluation can be measured and compared.

Interestingly, John Sowa has advanced an even more ambitious version of the line of thought underlying open ranking of ontologies. As he 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. [35]

At the same time, however, there is the potential for what we might call a poisoning of the wells that is the obverse sign of these very same advances in the direction of publishing freedom. Sowa thus agrees with the Foundry on the importance of maintaining a 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* in experimenting with dual frameworks involving a combination of expert peer review coupled with community-based dialogue on articles published, and where it is still the component of peer review that plays the dominant role.

Lewen sees the newly proposed open ranking systems as solving a 'problem with restricted reviewing systems', namely that 'they are very vulnerable to personal preferences, prejudices and reviewer's egos.' [35] We believe, however, that it is one important lesson of the enduring success of the peer review methodology in so many different fields over more than three centuries that some biases (roughly: the complex set of learned biases we call 'expertise') may 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 those who are in the position of making contributions resting on their expertise.

As both Noy and Musen have pointed out, however, the democratic ranking approach proposed by the NCBO can in fact incorporate also some of the advantages of approaches based on expert peer review. It would do this by providing users with the option to bypass the user-supplied rankings of reviewers and to pass directly to the lists of ontologies belonging to federations, like the OBO Foundry, which have been pre-validated via some process of expert peer review, ontologies which work well together and in relation to which the selection problems have been already solved. Others, however, will ignore this option, because the Foundry's adoption of the principle of orthogonality will be for them too restrictive. It is then conceivable that the Foundry itself could benefit from the subjection of application ontologies and of ontologies covering new domains to the sort of open review which the NCBO strategy allows. A successful realization of the democratic ranking based approach might also bring useful supplementation to the realization of some of the needs addressed by expert peer review – for example in vetting ontologies for errors or fraud, 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 now distinguish certain special features which are possessed by reference ontologies, like the GO, which have been developed in such a way as to be analogous to scientific theories. Such ontologies are: (1) developed to be common resources (thus they cannot be bought or sold), (2) developed and validated by domain experts, (3) recognized as being always subject to further development, and (4) independent of format and implementation. Sadly, the view still predominating in ontology engineering circles is that *ontologies are of their nature engineering artifacts* [25], so that it is as if all ontologies, both inside and outside science, would be assigned by default the status of application ontologies. This, however, 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 ontology engineers – to the effect that ontologies in general are the sorts of things that can be created at will to address each new set of needs on the part of each new set of users – 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 the needed changes it may be necessary to readdress the degree to which educational opportunities for ontologists are confined to departments of computer science.

In this we receive support from Akkermans and Gordijn [36], who point out that computer scientists and knowledge engineers still standardly conceive ontologies as computer science artifacts, which means that they still see ontology development to serve 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. [36]. They thereby adopt 'a self-limiting approach that in the end will not be able to exploit the full potential of the ontology idea'. Akkermans and Gordijn thus insist that the ontologies developed for scientific purposes need to be taken much more seriously as *first-class citizens* in computer science and knowledge engineering.

Empirical evidence of the benefits to be gained from the recognition of the distinctive character of ontologies created in the service of science has been accumulating for some time within the context of the GO and OBO Foundry endeavors. We are gratified that this recognition is now beginning to make itself manifest also within the framework of the Semantic Web, where scientifically serious ontologies, often involving input from the OBO Foundry, are finally beginning to distinguish themselves from the surrounding wine and pizza landscape [37].

Acknowledgements. This work was funded by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant 1 U 54 HG004028. With thanks to Robert Arp, Holger Lewen, Suzanne Lewis, Mark Musen, Natasha Noy, Mitsu Okada, 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 [] http://obi.sourceforge.net/. Last accessed June 24, 2008.

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

8 [] J. J. Cimino and X. Y. Zhu, The practical impact of ontologies on biomedical informatics, *Yearbook of Medical Informatics* 2006:124-35.

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

10 [] C. J. Mungall, Obol: Integrating Language and Meaning in Bio-Ontologies. *Comparative and Functional Genomics* 2004(5(7)):509–520.

11 [] C. Rosse and J. L. V. Mejino Jr. A reference ontology for bioinformatics: The Foundational Model of Anatomy. *Journal of Biomedical Informatics*. 2003; 36: 478–500.

12 [] R. D. Shankar, S. B. Martins, M. J. O'Connor, D. B. Parrish, and A. K. Das, Epoch: an ontological framework to support clinical trials management. In: *Proceedings of the international Workshop on Healthcare information and Knowledge Management*, 2006, 25-32.

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

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

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

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

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

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

19 [] M. Ware, *Peer review: benefits, perceptions and alternatives* (PRC Summary Papers 4), London: Publishing Research Consortium, 2008.

20 [] Editorial: Peer review and fraud. *Nature* 444 (2006), 971–972.

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

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

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

24 [] B. Riley, Database publication presents unique challenges for the peer reviewer, *Nature* (2006), doi:10. 1038/nature04993.

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

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

27 [] 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.

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

29 [] 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.

30 [] 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.

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

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

33 [] 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) at the 15th International World Wide Web Conference* (WWW 2006), Edinburgh, UK. Available at: http://smi.stanford.edu/smi-web/research/details.jsp?PubId=1142.

34 [] Personal communication of Natasha Noy.

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

36 [] H. Akkermans and J. Gordijn, Ontology Engineering, Scientific Method and the Research Agenda, in S. Staab and V. Svatek (Eds.): *EKAW 2006* (LectureNotes in Artificial Intelligence, 4248)¸ 2006, 112–125.

37 [] A. Ruttenberg, *et al*. Advancing translational research with the Semantic Web. *BMC Bioinformatics*, 8 (Suppl 3), (2007), S2.