Alternatives to Realist Consensus in Bio-Ontologies:
Taxonomic Classification as a Basis for Data Discovery and Integration

Beckett Sterner*
School of Life Sciences
Arizona State University
E-mail:beckett.sterner@asu.edu

Joeri Witteveen
Descartes Centre for the History and Philosophy of the Sciences and the Humanities
Utrecht University
&
Department of Science Education
Section for History and Philosophy of Science
University of Copenhagen

Nico Franz
School of Life Sciences
Arizona State University

*Corresponding author

Abstract:

Big data is opening new angles on old questions about scientific progress. Is scientific knowledge cumulative? If yes, how does it make progress? In the life sciences, what we call the Consensus Principle has dominated the design of data discovery and integration tools: the design of a formal classificatory system for expressing a body of data should be grounded in consensus. Based on current approaches in biomedicine and systematic biology, we formulate and compare three types of the Consensus Principle: realist, contextual-best, and coordinative. Contrasted with the realist program of the Open Biomedical Ontologies Foundry, we argue that historical practices in systematic biology provide an important and overlooked alternative based on coordinative consensus. Systematists have developed a robust system for referring to taxonomic entities that can deliver high quality data discovery and integration without invoking consensus about reality or "settled" science.

Keywords: social epistemology, data-centrism, ontology alignment

The possibility of a theory-neutral language for expressing empirical data has been an enduring question for philosophy of science (Kuhn 1996). Without a theoretical paradigm to guide the collection and interpretation of facts, science cannot get started — one simply acquires an ever-larger cabinet of curiosities. The very need for theory to make data meaningful, though, threatens the longevity of what we take to be scientific knowledge at any given time: new discoveries may always overthrow existing theory and render old investigations irrelevant. Big data is opening new perspectives on old questions like this one by deepening the theoretical foundations underpinning scientists' ability to find and use data as scientific evidence. Managing unprecedented magnitudes of data has forced scientists to articulate new conceptual assumptions and establish new social norms and organizations, leading to a period of increased reflective awareness of the central role that data play in science (Leonelli 2016). These changing circumstances provide an opportunity for integrative research centered on revitalized questions about the prerequisites for scientific progress: what epistemic principles should guide the design of computational systems to enable the accumulation and meaningful use of data over time?

New answers are emerging to this question, but their precise scope, content, and relative merits are still unclear (Leonelli 2016; Sterner and Franz 2017; Franz and Sterner 2018). In the life sciences, what we will call the Consensus Principle (CP) has dominated the design of data discovery and integration tools:

> (CP) The design of a formal classificatory system for expressing a body of data should be grounded in consensus.

We present three distinct versions of CP as a normative guide to designing data discovery and integration systems (Table 1). No single version presently seems adequate to the broader task of ensuring the cumulativeness of scientific knowledge over time, but considered jointly they provide the groundwork for a comparative study of the relative merits and trade-offs of each approach in concrete cases. We see this as a starting point for developing a robust information science perspective on formalized vocabularies for data discovery and integration. Data-centric biology thus represents an excellent opportunity for clarifying and operationalizing fundamental principles for the design of data infrastructure, especially classificatory theories, in order to provide empirical evidence for their performance.

The most visible formulation of CP in the life sciences has been Barry Smith and Werner Ceuster's "ontological realist" view, which has had an important influence on the adoption of computer ontologies into the biomedical sciences through the founding of the Open Biological and Biomedical Ontology (OBO) Foundry (Smith et al. 2007; Arp, Smith, and Spear 2015). Smith and Ceuster's realist conception of CP asserts that our classificatory theories of data should be grounded on consensus about nature in order to harnessing maximal efficiences of scale from computational processing. We suggest that while it presents a principled and well-ordered approach to ontology-building, scientific practice often calls for less restrictive and more flexible design principles.

Many contemporary ontologies that are endorsed by biologists depart from Smith and Ceuster's approach in adopting a weakened version of CP that only requires consensus on principles determining what is the "best" ontology for a body of data, but which permits the meaning of "best" to vary by context (Lord and Stevens 2010; Maojo et al. 2011; Sojic and Kutz 2012; Hoehndorf, Dumontier, and Gkoutos 2013; Franz and Goldstein 2013). While still consensus-based, these scientists' practices are not organized under an equivalently systematic and far-reaching theory of ontology design.

In contrast with the realist and contextual-best types of CP, we argue that historical practices in systematic biology provide an important and overlooked alternative based on coordinative consensus. This is a blind spot worth addressing, since systematists have honed an approach to data discovery and integration that departs in crucial ways from both realist and contextual design stances. Many areas of systematics continue to experience rapid change in, and dissent over, the meaning of classificatory terms. As a result, design principles that rely on consensus-building over the meaning of terms – whether grounded on realist or epistemic principles – have been of little service to the systematic research community. And yet, seemingly paradoxically, this hasn't precluded biologists from building classificatory systems based on the consensus principle. The key to their consensus-based solution has two parts. First, taxonomists use nomenclatural codes to govern the *application* of classificatory terms (i.e. taxonomic names) to taxonomic hypotheses, rather than specifying the *meanings* of those terms. Second, taxonomists follow the convention of including diagnostic criteria in the definitions of taxonomic names so as to enable the identification of new specimens. Together, these practices permit reliable translation of data across conflicting classifications.

Until recently, biologists have primarily used this system for coordinating the use of taxonomic names in human communication and reasoning, and they now face the challenge of formalizing and implementing this approach in a computational setting. We argue that shared principles for coordination rather than shared meaning does provide a sufficient foundation for a principled approach to designing computational data integration systems.

A comparative approach to big data trajectories

Every subfield of the life sciences appears to be leveraging big data in some fashion, from omics to biodiversity science. Sterner and Franz (2017) have described these instances of big data efforts as tracing out multiple "big data trajectories" over time that we can compare and analyze in order to arrive at deeper understanding of the role data plays in driving scientific change.[1] They define a big data trajectory as consisting of the activities of a group of researchers who (1) set out to expand the collective set of data available to address one or more shared problems of interest in such a way that (2) the researchers believe existing methods or resources available to the group are not adequate for the project and (3) they believe acquiring these methods or resources poses specific research problems separate from the original problems of shared interest" (Sterner and Franz 2017, 100). From an observer's perspective, one can now watch what happens as the big data movement plays out in different communities and evaluate how well it delivers on people's expectations. This provides us with a rich set of opportunities for comparative study as biologists begin to articulate lessons learned from their local experiences and apply them to new situations.

Looking ahead, a key issue will be how far the principles that scientists formulate for data discovery and integration in their local domains extend to other areas of science. A one-size-fits-all approach cannot be assumed and arguably should not be expected. Indeed, using our

---

[1] Our notion of a big data trajectory is distinct from Sabina Leonelli's concept of data journeys, although they are connected in important ways. Briefly, Leonelli uses journey to evoke how data travel across time and place from their original situations of production to new situations of use. In contrast, the idea of a big data trajectory is meant to describe the progress a scientific community makes as a function of increasing the amount of data available for a problem.

comparative perspective on big data trajectories will we show that different knowledge domains and scientific practices call for different solutions to the circulation of data between research sites.

In particular, we will focus on the underlying epistemological principles scientists use to warrant their choice of data classifications. Based on several case studies (Agar 2006), it appears that successful computerization of scientific work follows a general pattern: computers are first adopted to automate modular subtasks in a larger activity and then subsequently used to re-imagine the organization of work used to complete the activity as a whole. Agar's analysis highlights the special importance of analyzing the arguments people use to support the claim that adopting computer technology requires significant changes to historical practices. More broadly, this provides us with an analytical framework for computerization as an institutional process by which a group of individuals set out to make the use of computer technology indispensable for specific activities (Sterner and Lidgard 2014; Sterner and Lidgard 2018). The consensus principle in its different formulations serves as a normative, methodological resource for regulating scientists' practices for computerization projects. Clarifying its use and justification is therefore crucial to understanding broader landscape of what scientists mean by "data" (Leonelli 2016; Sterner 2018).

In this context, formalizing the languages that biologists use to describe their data has become increasingly important. Building a dictionary, glossary, or thesaurus for experts is not generally sufficient to make the leap from human to computer intelligibility: the definitions and semantic relations provided by the dictionary and thesaurus still fundamentally operate in a natural language setting where logical imprecisions and errors do not propagate via deductive inference into horribly wrong conclusions (Sterner and Franz 2017). "Dictionaries are prepared for human beings; their merely nominal definitions can employ the unregimented resources of natural language, can tolerate circularities and all manner of idiosyncrasy. In ontologies, however, definitions must be regimented in such a way that each reflects the position in the hierarchy to which the definiendum belongs" (Smith 2005, 15). Although colored by specific views about definition (see below), this quote captures the way that bringing computers into biologists' practices of representing and reasoning about biological data begins to transform the underlying language and use of data. These formalized systems for expressing data are classificatory theories that deserve philosophical analysis and criticism as much as other more traditional theories such as quantum mechanics or evolutionary theory (Leonelli 2016).

These classificatory theories serve at least two important functions for biology: data discovery and integration. Data discovery relies on searchable metadata to return datasets with high precision and recall (Remsen 2016). For example, the Global Biodiversity Information Facility's data portal enables one to query aggregated specimen and occurrence data using search terms such as the taxonomic name "*Hirundo rustica*." In contrast, data integration goes beyond discovery to combine data of different types or sources, often to provide input for a predictive model. For example, one could use the Gene Ontology to integrate data on the subcellular location of gene products with assay data on gene expression levels. Another example is to integrate data on molecular interactions across several species in order to annotate genes with predicted interactions or functions (Leonelli 2013).

Translating what data mean for human experts into forms that are intelligible to computers is therefore a fundamental challenge for data science. Observations and experimental results don't tell their own story: understanding their full significance requires contextualization in light of local aims, resources, analytical methods, and error repertoires (Leonelli 2016).

Human experts reliably integrate implicit knowledge and explicit standards to navigate subtle shifts in the meaning of data from one situation to the next. Computers, in contrast, are usually stuck playing the role of naïve reader to human discourse and thought. As more and more disciplines have embraced big data projects, scientists have encountered recurring obstacles in defining, preserving, and sharing the new information they are collecting, forcing a new level of reflective awareness about the nature and significance of data (Leonelli 2016; Millerand et al. 2013; Bowker 2000).

A realist view of the consensus principle

The general consensus principle we introduced above states that the design of a formal system for expressing a body of data should be grounded in consensus. The central question we examine in the remainder of the paper is the ideal nature of that common system. In this section, we look at pioneering work by Smith and Ceusters in the 2000s that introduced "ontological realism" as one solution (e.g. Smith 2004). Their approach endorses what we can call the *realist consensus principle* for ontologies:

> **Realist Consensus Principle (RCP)**: The design of a formal classificatory system for expressing a body of data should be grounded in a global consensus metaphysical interpretation of the reality those data describe.

According to Smith and Ceusters, that consensus metaphysical interpretation includes the reality of universals, although in later works they provide a more flexible view we will come back to below (Smith 2004; Smith 2005; Smith 2008a; Arp, Smith, and Spear 2015).

The need for an interpretation of the logical system highlights a fundamental point that extends to all computer-facilitated data integration approaches in use today. The basic logical structures that comprise any computer ontology are classes, individuals, and properties.[2] Relations can be between classes (e.g. class "X" is a class "Y"), between individuals and classes (individual "A" instantiates class "B"), and between individuals (individual "A" is the father of individual "B"). As the three examples show, ontologies can also specify classes of relations, such as "is the father of" or "is a part of."

Most frequently, computer ontologies instantiate some version of description logic which uses a restricted syntax compared to first-order logic in order to improve computational efficiency. The precise details aren't relevant for our purposes (see Baader et al. 2003; Keet 2018). For practical use, any type of description logic is further implemented in some standardized machine-readable syntax such as the OWL2 Web Ontology Language. The main point is that adhering to the abstract axioms of a description logic is not sufficient to guarantee that computer ontologies will be useful for data discovery and integration. Neither is following the standardized syntactic rules governing any implementation like OWL. None of these underlying resources determines a coherent semantics able to regulate the definition and interpretation of terms in the ontology such that they can be combined to form useful representations of biological data. This point is analogous to the distinction between syntactic and semantic views of scientific theories in philosophy of science (Winther 2015).

For simplicity, we treat Smith and Ceuster's ontological realism as containing two main elements: (1) an ontology called the Basic Formal Ontology (BFO), which is foundational for all

---

[2] We use terminology common to OWL here since it is the easiest to grasp intuitively, but researchers in the first-order logic and description logic communities use different terms for operationally equivalent ideas.

other bio-ontologies, and (2) a realist interpretation of the structures in an ontology. Their work also contains a number of supplementary guidelines that we do not discuss here.

The world of ontological realism has two types of entities: universals and particulars. Examples of universals include human being, enzyme, aspirin, and examples of particulars are you and me, the Planet Earth, this piece of cheese (Smith 2005). Smith and Ceusters give different general characterizations of universals and particulars in different papers, though see (Merrill 2010b). For example, in Smith (2005), universals are multiply located entities that exist in particulars, while particulars are entities with only one location in space at a time. In Smith and Ceusters (2010), they characterize universals as repeatable, i.e. as having potentially indefinitely many instances we can discover in principle. Smith also proposes that universals are distinguished from mere extensional sets by figuring in scientific laws of nature or having real definitions rather than nominal definitions (Smith 2005). Hence ontologies can include terms for what Smith calls "collections," which are sets formed from logical combinations of universals, but these collections are treated as artificial constructs rather than real entities in the world. In this manner, ontological realism aims to put a realist semantic interpretation on the basic syntactic structures of first-order logic. Things that can have predicates thus include universals as well as particulars. Smith also adds a further logical primitive, the instantiation relation, and stipulates that only particulars can instantiate universals (Smith 2005).

Two additional features of Smith and Ceuster's realist interpretation concern the ideal form of definitions for terms and the possibility of overlap in the domains of ontologies. As noted, ontological realism privileges real definitions over nominal ones, with the goal being to formulate necessary and sufficient conditions describing the essence of the universal. Ideally, this should take the form of Aristotelian differentia in a hierarchy of logical species and genera, i.e. species A is genus B with differentiating criterion C. In the Foundational Model for Anatomy, for instance, we can find Cell defined as "Anatomical structure, each instance of which has as its boundary the external surface of some maximally connected plasma membrane" (2018a). Smith and Ceusters also take an Aristotelian view of nature in the sense that they support the value of multiple ontological perspectives, so that a single domain can be described by ontologies that partition phenomena in different ways (Smith 2008a). As a result, ontological realism does not necessarily imply eliminative reductionism, e.g. about the reality of organisms in favor of molecules.

However, the second element of ontological realism does enforce consistency with one fundamental ontology, BFO (Arp, Smith, and Spear 2015). In tension with the claim to realism, BFO provides a set of metaphysically fundamental terms that are roughly consistent with a common sense view of the world as "midsize dry goods" consistent with Newtonian mechanics (Lord and Stevens 2010). For example, BFO distinguishes between continuants, entities that only have spatial parts, and occurrents, entities which can have parts that extend in space as well as time. On this view, a cat is a continuant: its legs are parts of it at a given time, for instance, but its legs ten minutes ago are not part of it now. The cat walking, in contrast, is an occurrent because each step it takes forms a temporal part of the event as a whole. BFO also distinguishes between independent and dependent entities, which refers narrowly to whether a type of particular can only exist if a particular of another type also exists. The quality blue, for example, is dependent because it can only exist as a property of some actual object. To be consistent with ontological realism, any other ontology should at least define its universals in terms of the primitives provided by BFO using Aristotelian differentia. BFO therefore plays an important role for ontological realism in harmonizing the design of ontologies across diverse subject domains (Smith et al. 2007).

In general, the guiding slogan of ontological realism is to "describe what exists in reality, not what is known about what exists in reality" (Arp, Smith, and Spear 2015). In the early 2000s, Smith and Ceusters demonstrated a number of cases where existing terminologies failed to keep these separate, leading to basic reasoning flaws (Kumar and Smith 2003; Ceusters, Smith, and Goldberg 2005; Smith et al. 2005; Smith and Ceusters 2006). The underlying idea is straightforward enough: without a single, coherent semantics, ontologies can equivocate between referring to things in nature and things in our minds, such as beliefs and concepts. Consider a hypothetical scenario where "gene" is defined as "a concept used by biologists to denote coding regions of DNA" and "regulatory region" is defined as "a segment of DNA upstream of a coding region." Any instance of a gene is then a token concept, but any instance of a regulatory region is a token DNA molecule. If the ontology also includes the relation "regulatory region is part of gene," then observations associated with a particular regulatory region could be inferred to also be true of the associated gene concept. For ontologies with thousands of terms and many more individuals, this semantic slippage can easily lead to flawed search query results or datasets.

Historical interlude: deflating ontological realism

In 2010, Gary Merrill did for ontological realism more or less what Arthur Fine did for the realism debate in 1986. In *The Shaky Game*, Arthur Fine delivered a series of powerful critiques questioning the value of any interpretation, metaphysical or epistemic, that philosophers could add on top of the success of science (Fine 2009). Metaphysical interpretations based on inference to the best explanation do not provide a more secure foundation for science, for example, and in any case an anti-realist can substitute "reliable" for "true" in a metaphysical theory and get the same result. In contrast, Smith and Ceusters have consistently claimed that interpreting first-order logic in terms of universals leads to better results for ontology design (Smith 2004; Smith 2008a; Smith and Ceusters 2010), in contrast to a cluster of views they grouped under the label of "conceptualism." Merill's (2010b) critiques led to a broader discussion in the community, however, which illustrated a more expansive range of alternatives than Smith and Ceusters allowed. One effect of the discussion for scientists was to emphasize the need for clear evaluative criteria for ontologies that could connect fundamental design principles to empirical evidence.

In general, Smith and Ceusters have framed scientists' choice of approach as a stark opposition between ontological realism and conceptualism, which holds that words refer only to concepts, not the world. Smith and Ceusters moreover cast many pre-existing efforts at regulating scientific terminology as firmly in the conceptualist camp. Describing early versions of the SNOMED terminology and the Gene Ontology in 2008, Smith wrote: "Unfortunately, the new formalized biomedical terminologies were developed against the background of what are now coming to be recognized as a series of major and minor philosophical errors. Very roughly, the developers of terminologies made the assumption that we cannot have knowledge of the real world, but only of our thoughts. Therefore, they inferred, it is thoughts to which our terms (and our terminologies) necessarily refer" (Smith 2008b, 83).

In his 2010 critiques, Merrill sought to undermine the relevance of Smith and Ceuster's distinction between ontological realism and conceptualism by deflating the significance of believing in the reality of universals (or analogous ideas such as natural kinds, types, etc.) for scientific practice. In particular, he argued that Smith and Ceusters offered no way to operationalize the claim that ontologies should only include classes that were interpretable as universals.[3] On the other hand, Merrill also resisted being cast as a conceptualist — his main point was that these philosophical interpretations of the syntax of description logic were not contributing anything substantive to science. "The alternative to adopting a metaphysically based realist principle in science is not to adopt a metaphysical non-realist principle. It is rather not to adopt a metaphysical principle at all" (Miscellaneous authors 2010). Or, to borrow Fine's language from the realism debate, ontological realism was an "artificial additive" to the practice of building bio-ontologies.

Merrill's critique sparked a back-and-forth with Smith and Ceusters in print (Smith and Ceusters 2010; Merrill 2010a), as well as an extended discussion in the OBO Foundry's public e-mail discussion list (Miscellaneous authors 2010). This more informal discussion made especially clear the extent of the confusion and disagreement about ontological realism that existed among scientists building bio-ontologies. While the debate earned a lot of attention, it does not seem to have settled much except the need for more clarity on best practices. Michel Dumontier, for example, summed things up as follows:

> "There are two issues I'd like to raise to our communities: 1 - What metrics should we use to assess ontology quality and determine whether an ontology is 'good'. 2 - Can we re-factor existing, and well used terminologies… into 'good' ontologies for health care and the life sciences?" {Miscellaneousauthors:2010tj Jul 27, 2010; 9:16am.

Several participants in the OBO Foundry discussion also went on to publish more formal critiques of ontological realism as a practical guide to ontology design (Lord and Stevens 2010; Maojo et al. 2011; Sojic and Kutz 2012; Hoehndorf, Dumontier, and Gkoutos 2013).

There are thus a number of outstanding problems in this arena with considerable relevance for both philosophy and biology. A partial list would include: (1) How can ontological realism be operationalized in order to provide clear evidence of its impact on design choices? (2) Can ontological realism be extended to provide systematic guidance about instrumental aspects of scientific modeling, including the use of fictions, idealizations, and other practical expedients? (3) To what extent are uniform semantics really necessary for optimal data discovery and integration using ontologies? (4) What alternatives are there to ontological realism? It's to this last question that we will turn in the remainder of the paper.

A contextualist view of the consensus principle

While Smith and Cuesters have had important influence on bio-ontologies, especially in the 2000s, the broader community does not adhere strictly to the details of ontological realism. The well-established Drosophila anatomy ontology, for example, takes a pragmatic approach to inheritance hierarchies, permitting multiple inheritance in some cases, and also accommodates visual as well as textual definitions (Costa et al. 2013). Alternatively, while Arp, Smith, and Spear (2015) take species to be paradigm examples of universals, other biologists have strongly disputed this claim (Franz and Thau 2010). Moroever, the widely-used NCBI taxonomy ontology is a bare hierarchy of taxonomic names with no definitions that is automatically generated from

---

[3] This is related to the problem logical empiricists faced when they tried to define qualitative predicates (Goodman 1983).

other name databases (Federhen 2012; 2018b). Do biologists follow a coherent, systematic alternative to ontological realism in practice? We provide a preliminary answer here, though considerably more research is needed for a full accounting of what biologists actually do when building and revising their ontologies.

We propose that many practices of ontology-building in the biomedical domain are better characterized as employing design principles that are adapted to the particularities of scientific practice in their respective domains. This approach to ontology-building departs from the RCP by substituting local principles about what constitutes the best meanings for scientific concepts for a shared view of reality as the consensus criterion. We call this the contextual-best consensus principle for ontology-building:

> **Contextual-Best Consensus Principle (CBCP):** The design of a formal classificatory system for expressing a body of data should be grounded in a local consensus on principles for evaluating the quality of a definition.

This version of the consensus principle is contextual because each community designing an ontology must decide on which design principles are best. However, the CBCP still commits scientists to agreeing on substantive and general criteria that govern how they define terms in that ontology.

These criteria could rest on a particular metaphysical theory of their subject domain, as required by ontological realism, but can also be grounded in epistemic principles which, for example, characterize what counts as mature or settled science. As of September, 2018, the OBO Foundry principle for ontology maintenance states:

> "Tentatively, we consider scientific consensus to be reached if multiple publications by independent labs over a year come to the same conclusion, and there is no or limited (<10%) dissenting opinions published in the same time frame. In cases an area remains controversial, and no consensus is reached, then it is up to the ontology maintainer to either leave out the controversial term, or pick a viewpoint for practical considerations, and note the presence of controversy in an editor note" (The OBO Foundry 2018).

Apart from having an important say over what terms an ontology includes, ontology maintainers often play a crucial role in facilitating integration by creating lists of synonyms for terms that feature in an ontology. Where different research communities use different terms to refer to the same process or components, an ontology can bring these into contact by making new data available under familiar terms. However, this does call for expert judgment on part of the maintainer. Subtle differences in meaning or contextualized associations of a term might get lost in the translation process and can introduce unwarranted implications.

Another example is the choice among entomologists building the Hymenoptera Anatomy Ontology (HAO) not to use evolutionary homologies (Yoder et al. 2010; Seltmann, Austin, and Jennings 2012). In this case, the biologists argue that claims about evolutionary homologies are generally so uncertain and liable to change that they should be categorically excluded as classes in the ontology. We quote an extended excerpt here because it nicely illustrates how consensus views about best design can rest on context-dependent features of scientific knowledge:

> "Fundamentally, the HAO project rests on recognizing different instances of a topographically-defined concept as "the same"… The HAO employs the principle of "structural equivalence" to discuss topographical sameness. In biology, however, homology is often more explicit, referring to a more profound "sameness", because it expresses a theory about structures sharing a common

> evolutionary origin even if they appear structurally dissimilar… Homology in this evolutionary context is often dynamic, and may be controversial or involve conflicting hypotheses or quickly changing views. The dynamic nature of homology hypotheses conflicts with the HAO's goal of unambiguous circumscription of anatomical concepts, and, as such, overt reference to homology hypotheses are avoided in constructing HAO definitions" (Seltmann, Austin, and Jennings 2012, 79).

Although the authors recognize evolutionary homology as more fundamental, particular homology claims are often far from counting as "settled" science on any view. Rather than attempt to represent that uncertainty and conflict explicitly, the HAO designers move wholesale to structural equivalence as the basis for classifying data about anatomy.[4]

Ontological realism as well as the two contextualist examples use "settled" science as a basis for regulating the content of ontologies. Ontological realism characterizes settled science as approximately true to nature, while the OBO Foundry provides a more operational though *ad hoc* characterization in terms of number of confirming papers. As foundations for the continuity of scientific knowledge, then, ontologies under these approaches take current consensus as the best guide to future stability. As existing bio-ontologies accumulate more history, however, interesting complications are beginning to emerge as scientists introduce changes and align these to past versions (Bertone et al. 2013; Pesquita et al. 2013; Otero-Cerdeira, Rodríguez-Martínez, and Gómez-Rodríguez 2015). The history of science teaches us to be cautious about guaranteeing that any one aspect of scientific knowledge will stay constant over time, so it is worth exploring how much continuity is possible even in extreme cases of disagreement. This is one way in which systematic biology has important lessons for data science today.

Consensus in the face of lasting dissent

The diversity of life on Earth continues to challenge our understanding of the nature of taxonomic units and the best methods for studying them. The classification of biodiversity is one of the greatest challenges in biology in several ways: we have yet to discover much of the biodiversity on the planet, many known species have received only minimal attention so far, and biological species are some of the most complex phenomena we know. No consensus exists in systematics on the correct metaphysical nature of biological species (Wilkins 2009), nor is there one evidential standard that everyone can agree to follow when delimiting species boundaries (Conix 2018). Systematists are also often confronted with incomplete and conflicting data about taxa, and as a result many species classifications are best understood as ongoing research hypotheses rather than "settled" science. In these circumstances, ongoing instability and disagreements about the correct way to classify a taxonomic group are both expected and desirable, given that biologists continue to publish new empirical evidence and analyses (Franz and Thau 2010).

Nonetheless, taxonomic names and definitions are absolutely fundamental to understanding the biological significance of all data in the life sciences (Remsen 2016). In some taxonomic groups, biologists can rely on nomenclatural databases that assert a single coherent

---

[4] This empiricist concern about the theory-dependence of anatomical data has parallels to the earlier disputes between pheneticists, cladists, and evolutionary systematists about the best methodology for inferring classifications (Hull 1988; Sterner and Lidgard 2018), and would be a fruitful point of contact between philosophical analyses of homology and scientific practice.

classification system based on the judgment of one or more experts, e.g. the WoRMS database for fish (Costello et al. 2013). The present views of one or a few experts, however, are often unstable across the broader community of biologists as well as over time. In the largest analysis to date of 151 different taxonomic treatments of birds published from 1886 to 2014, only 11 of the 19,260 unique combinations (0.06%) of species names and taxonomic meanings were used consistently by all authorities (Lepage, Vaidya, and Guralnick 2014). In another analysis of two highly cited treatment versions that review primate taxonomy, published by the same author and separated by merely 12 years, one in three taxonomic names changed either their syntax, or their meaning, or both (Franz, Pier, et al. 2016). In either case, reconciling the remaining 33.0-99.9% of name usages requires some form of additional human processing, repeatedly, and in principle indefinitely

Assuming or imposing consensus views on the meanings of taxonomic names for the purpose of aggregating specimen and occurrence data also places a performance ceiling on data discovery and integration. Providing lists of synonyms – which we mentioned are often included in ontologies in the biomedical domain – does not address the complexity of the problem systematists face. They have not merely been looking for means to highlight many-to-one relationships between names and a known entity they refer to, but instead have to deal with many-to-many relationships between names and different (unsettled) hypotheses about the circumscription of the entity they refer to. The lack of a one-to-one correspondence between taxonomic names and biological entities leads to costly errors and missed opportunities for anyone that wants to find all and only published research and data about a particular taxonomic group (Remsen 2016). A database built on matching names rather than underlying meanings will have uncontrolled variation in aggregation errors across taxonomic groups depending on their specific scientific histories (Franz and Sterner 2018).

Our collective understanding of biodiversity is thus too fractious to be adequately grounded in a single substantive consensus view about the meaning of terms, whether it be a metaphysically realist or a contextual, epistemic one. Is the consensus principle still relevant, then, when circumstances dictate against selecting a single classificatory system for a body of data?

We argue that the answer is yes, but the nature of the consensus proves to be quite different. Instead of attempting to formulate consensus principles regarding classificatory content, systematists in the Linnaean tradition have agreed on consensus principles for the coordination and communication of their *dissensus*. They have laid down methods for coordinating their disputes about the proper circumscriptions of taxa by formulating principles that help them identify when they are disagreeing about the boundaries of the same taxon and when they are drawing boundaries of different taxa differently. In other words, consensus principles in systematics provide consensus about the application of a name (i.e. its extension) even in the face of disagreement about its precise meaning.

The semantics of a contemporary Linnaean taxonomic name can be analyzed into three main components for our purposes here. The name itself, as a sequence of letters, carries no content relevant to its semantic meaning, even if it superficially appears to describe a property of the taxonomic referrent. (We skip over other syntactic information provided in names alone, such as Latin suffixes indicating rank and the binomial genus-species structure of species names.) It is the pair of name plus a fixed name-bearing element of a taxonomic group – its so-called "nomenclatural type" or "type" for short – that helps channel differences in hypotheses about its semantic content. The combination of name and type often fail to designate a unique referent, and in practice biologists are often uncertain about the correct referent of a name due to partial or contradictory data. Yet, by designating types for names, it becomes possible to "fix the reference" of a taxonomic name without specifying its full taxonomic meaning. In other words, nomenclatural types provide anchor points for the objective application of names against a background of subjective differences in where boundaries between groups need to be drawn. Using the device of types, taxonomists can establish whether they disagree about the circumscription of the same taxon or numerically different taxa.

For species names in particular, it is important for an author to make explicit what circumscription of organisms (or of lower-level taxonomic groups) it applies to, i.e. what kind of thing they take any individual species to be. Ideally, the specification of a circumscription should also make reference to the species concept it operationalizes, e.g. a biological (interbreeding) species concept or a phylogenetic species concept. We will refer to these circumscriptions as "taxonomic concepts," following established usage in the literature. In theory, the type specimen can only belong to one biological species, so the combination of the type plus a species circumscription should be sufficient to specify the application of a name.

Unlike the choice of a species concept, which given the current state of the species debate can be treated as somewhat arbitrary, taxonomic concepts have an empirical status as scientific hypotheses. To see this, consider the difference between identifying what a given name refers to with and without specifying a taxonomic concept. It is easy and uncontentious to say, "This taxonomic name refers to the biological entity that includes this type specimen." It is much harder epistemically to accurately identify and agree on which organisms other than the type specimen are also members of the designated species. In other words, even if we agree that there is a fact of the matter about the precise referent of the taxonomic name, there is still the epistemic challenge of correctly describing what that referent is. Taxonomic concepts can therefore be empirically accurate or inaccurate to varying degrees based on whether they include all and only the organisms that are actually part of the true referent. Even as the association between name and nomenclatural type remains fixed, then, it is legitimate and indeed desirable for biologists to revise or disagree about the correct taxonomic concept, especially as new traits and individuals are sampled.

Taxonomists' use of a method of ostensive reference fixing as a means of coordinating their comparisons of different descriptions has been a successful means to maintain a basic level of common knowledge about the semantic meaning of a term even as the precise nature of its referent remains disputed. As an early advocate of this method in botany put it succinctly, it has enabled taxonomists to "have a designation ready for the final entity, but also available for any number of approximating concepts which may follow each other with no unnecessary confusion" (Cook, 1898).

The value of taxonomists' approach to naming is therefore twofold. First, it enables systematists to designate a referent using ostension plus a supplementary description instead of an Aristotelian-style theoretical definition. Second, taxonomists can understand and share data classified under a name without having to agree on the name's referent (if any). The diagnostic description that supplements the ostensive component of the definition allows biologists to represent the name's meaning extensionally, i.e. as a set of organisms associated with the name. Along with supplementing the referential precision of the ostensive component, the descriptive component also provides operational criteria for determining which organisms belong under the name. A dissenting systematist can therefore represent a colleague's definition in a purely extensional way and side-step the question of what taxonomic entity the name actually denotes.

A coordinative view of the consensus principle

The principles and rules taxonomists rely on to share data without presupposing a consensus view of biodiversity have developed in circumstances very different from today's data-intensive research environments. Is it possible to translate these historical practices into an effective computational system for data discovery and integration? We argue that this is possible, at least in principle, by adapting the standards for the valid publication and application of names that we have described above. The relevant version of the consensus principle for systematics is then a matter of sharing a system of reliably coordinating the use of terms, not agreement about what exactly those terms refer to:

> **Coordinative Consensus Principle (CCP):** The design of a formal classificatory system for expressing a body of data should be grounded in a consensus standard for coordinating the use of classificatory terms, even if the meanings of those terms haven't been settled.

Before we address the CCP in more detail, we should first briefly clarify how taxonomic classifications relate to bio-ontologies. At the highest level of abstraction, taxonomic classifications also express logically nested hierarchies of terms that can be modeled using description logic (and its various implementations, e.g. in OWL). As noted earlier, the NCBI taxonomy, for example, is represented by the NCBITaxonomy ontology, though it provides no definitions. Indeed, biologists and philosophers continue to dispute the metaphysical issue of whether taxonomic names should be interpreted as referring to kinds or individuals, and hence whether a class in description logic is properly interpreted as a kind with organisms as instances or an enduring individual with organisms as spatio-temporal parts. In addition, ontologies can contain multiple logical hierarchies using different kinds of relationships. The Gene Ontology, for example, is a network composed of three different class hierarchies describing kinds of cellular location, molecular functions, and biological processes (Ashburner et al. 2000; Consortium 2017).

Moreover, we've seen that systematists use a combination of reference fixing devices (nomenclatural types) and hypotheses of the taxa that include them, at least for lower taxonomic ranks such as species, genera, and families. (A recent alternative to Linnaean classification, PhyloCode, uses ostensive, "node-pointing" definitions for all taxonomic names but in a different fashion that's outside our scope here.) The norm for bio-ontologies, especially ontological realism, has been the opposite (Klein and Smith 2010).

As with bio-ontologies, handling taxonomic names on a large scale without sacrificing performance requires researchers to explicitly state and formalize the logical relationships between names. Human experts are able to memorize the relevant classifications for a taxonomic group in order to disambiguate any polysemic name usages. They are similarly able to translate specimen identification back and forth between classifications based on an expert understanding of the logical and biological relationships between taxon definitions. Human non-experts and computers, however, typically lack the background knowledge to carry out this contextual disambiguation of taxonomic names. Automated ontology matching algorithms, for example, typically rely on similarity between the character strings of terms or their relationships to other terms in the class hierarchy. When we compare multiple taxonomic classifications for the same group, however, the issue is precisely that names are reused with different meanings and these meanings are revised to reflect alternative scientific hypotheses about the extensions of the taxa.

One approach is to align taxonomic concepts by representing their extensions spatially in a similar way to Venn diagrams (Franz and Peet 2009; Franz et al. 2015; Franz, Zhang, and Lee 2017). Region Connection Calculus (RCC) provides a suitable set-theoretic vocabulary for human experts to express the logical relationships between the extensions of taxonomic names in an intelligible form for computational search and reasoning. Figure 1 illustrates a set of results for taxonomic concepts of North American grasses in the "*Andropogon* complex." Franz et al. have used RCC-5, a version with five relationships: X is a subset of Y (X<Y), X contains the set Y (X>Y), X is congruent to Y (X==Y), X excludes Y (X ! Y), and X overlaps with Y (X><Y). In some cases these relations can be established deductively using the definitions of the taxonomic concepts. For example, if one name's definition uses the same type specimen as another, then we know the extensions of both names must at least overlap (i.e. X *not* ! Y). In other cases, the relationship between two names is not immediately apparent from their definitions and the trained judgment of a taxonomic expert is necessary to interpret the authors' intended semantics. Experts in each taxonomic group generally already know these semantic relationships implicitly, but expressing this knowledge is explicitly and in a machine-readable format requires substantive effort and scholarship.

Using the CCP, systematists are therefore able to deliver high quality data discovery and integration through the logical alignment of conflicting systems, yet it neither invokes consensus about "settled" science, nor does it endorse the pure conceptualism that Smith and Ceusters oppose. First of all, note that systematists do not attempt to define "concept" in general. They make no claim that meaning is always extensional rather than intensional, or that one form of definition is inferior in general. Instead, they stipulate a consensus convention within their professional community regulating what is required for an acceptable definition. Moreover, taxonomists regularly take the ostensive component of taxonomic name definitions to involve reference to real taxa (though this realist interpretation is not itself stipulated by any nomenclatural code). For example, type specimens in species definitions establish theoretical facts of the matter about the referents of taxonomic names, against which taxonomic concept descriptions can be tested and revised for empirical accuracy. Since systematists often face uncertainty or disagreement about the true referent of a name, they use taxonomic concepts to stand in for the true referent. Taxonomic concepts in this regard can provide reasonably reliable representations of what biologists using that classification system take the referent to be

There is an obvious cost to this approach when consensus is available: translating across data classification schemes requires intensive labor to produce and maintain the logical alignments. This effort is minimized when a single, universal classification for data is adequate for everyone involved. However, embracing the realist or best consensus approaches does not eliminate the need for alignment over time. As knowledge changes, the ability to translate historical data annotations ("legacy" data) into contemporary terms becomes valuable. In addition, bio-ontologies are not perfectly modular, so careful alignment is still often necessary to integrate data classified according to multiple overlapping ontologies. Finally, we have never been able to reliably predict just which aspects of existing knowledge will be overturned by future discoveries. Having well-developed approaches to translating data across fundamentally divergent terminologies may come in handy in the most unexpected places.

Conclusion

In *Data-Centric Biology*, Sabina Leonelli argues that "the real source of innovation in current biology is the attention paid to data handling and dissemination practices and the ways in which such practices mirror economic and political modes of interaction and decision making" (Leonelli 2016, 1). In its most productive forms, getting data to travel without losing their meaning "involves the use of computational tools to raise awareness of the conceptual, material, and institutional scaffolding required to package and interpret data, rather than hiding those aspects away" (Leonelli 2016, 171). We've advanced this broader project here by formulating three consensus-based principles that scientists have invoked to help settle matters of designing formal classificatory theories for scientific data: through metaphysical consensus about reality (RCP), contextual standards for best definitions (CBCP), and coordination conventions (CCP).

We've also argued for the value of the coordination conventions systematists have developed for regulating the application of taxonomic names when neither metaphysical nor epistemic consensus is available. It is by virtue of being epistemically and metaphysically lightweight that coordinative consensus-building does not lead to a "dictated consensus" being forced upon users of taxonomic databases by the database architects. The coordination conventions method show that reliable data discovery and integration are possible even in situations that are far from consensus.

The existence of different versions of the consensus principle strengthens the value of taking an information science perspective on data discovery and integration. Since there are multiple approaches, each community of scientists has an ongoing need for informed deliberation about which option is best for their situation. We've seen that limiting the content ontologies to areas of consensus, either ontological or epistemic, is not a universal solution, but this opens up many new questions. We hope that one outcome of this paper will be to illustrate alternative paths to investigating classificatory theories for biological data that don't center on consensus about the definitions of the phenomena. More generally, we suggest that the use of coordinating conventions can be a fruitful alternative means of ontology-building in domains of research for which there is a low degree of consensus about the metaphysical and epistemic status of the classificatory goods.

We also need a better understanding of how scientists operationalize their regulative principles for ontology design in different areas of science and how scientists provide empirical evidence on the effectiveness of these principles. It would be valuable to have a more comprehensive sense of how scientists employ the contextualist version of the consensus principle in different fields, for example. Designing an ontology is not unlike doing an experiment, in that published articles describing the result are usually post hoc rationalizations of what actually happpened and do not reliably describe the discovery process. Understanding how scientists design ontologies in practice is therefore an important locus for integrating historical, ethnographic, and philosophical methods. Important sources in this regard include public, online email lists, such as the OBO Foundry's discussion forum we cited above, as well as conversations at community meetings.

Another open question are the theoretical capabilities of different theories of reference to ensure the accumulation of scientific data over time. For example, what is the best way to combine ostensive and descriptive content in definitions to enable communication and reasoning despite fundamental disagreements or change of belief? Smith has suggested that definitions with ostensive content are generally less desirable compared to theoretical definitions using Aristotelian differentia, but this claim depends on a number of other presuppositions made by ontological realism that do not hold in general. Data discovery and integration offer a novel setting with great practical importance for exploring how theories of reference function in a dynamic context. Historical studies of continuity and discontinuity in scientific data and classificatory theories have a major role to play in that project.

References

Agar, Jon. 2006. “What Difference Did Computers Make?” *Social Studies of Science* 36 (6): 869–907. doi:10.1177/0306312706073450.

Arp, Robert, Barry Smith, and Andrew D Spear. 2015. *Building Ontologies with Basic Formal Ontology*. Cambridge, MA: MIT Press.

Ashburner, M, C A Ball, J A Blake, D Botstein, H Butler, J M Cherry, A P Davis, et al. 2000. “Gene Ontology: Tool for the Unification of Biology.” *Nature Genetics* 25 (1): 25–29. doi:10.1038/75556.

Baader, Franz, Deborah L McGuinness, Daniele Nardi, and Peter F Patel-Schneider, eds. 2003. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge: Cambridge University Press.

Bertone, Matthew A, István Mikó, Matthew J Yoder, Katja C Seltmann, James P Balhoff, and Andrew R Deans. 2013. “Matching Arthropod Anatomy Ontologies to the Hymenoptera Anatomy Ontology: Results From a Manual Alignment.” *Database* 2013 (0). Oxford University Press: bas057–bas057. doi:10.1093/database/bas057.

Blomquist, H L. 1948. *The Grasses of North Carolina*. Durham: Duke University Press.

Bowker, G C. 2000. “Biodiversity Datadiversity.” *Social Studies of Science* 30 (5): 643–83. doi:10.1177/030631200030005001.

Ceusters, Werner, Barry Smith, and Louis J Goldberg. 2005. “A Terminological and Ontological Analysis of the NCI Thesaurus.” *Methods of Information in Medicine* 44 (4): 498–507.

Conix, Stijn. 2018. “Integrative Taxonomy and the Operationalization of Evolutionary Independence.” *European Journal for Philosophy of Science* 154 (2). Springer Netherlands: 1–17. doi:10.1007/s13194-018-0202-z.

Consortium, The Gene Ontology. 2017. “Expansion of the Gene Ontology Knowledgebase and Resources.” *Nucleic Acids Research* 45 (D1): D331–38. doi:10.1093/nar/gkw1108.

Costa, Marta, Simon Reeve, Gary Grumbling, and David Osumi-Sutherland. 2013. “The Drosophila Anatomy Ontology.” *Journal of Biomedical Semantics* 4 (1). BioMed Central: 32. doi:10.1186/2041-1480-4-32.

Costello, Mark J, Philippe Bouchet, Geoff Boxshall, Kristian Fauchald, Dennis Gordon, Bert W Hoeksema, Gary C B Poore, et al. 2013. “Global Coordination and Standardisation in Marine Biodiversity Through the World Register of Marine Species (WoRMS) and Related Databases.” *PLoS ONE* 8 (1): e51629. doi:10.1371/journal.pone.0051629.

Federhen, Scott. 2012. “The NCBI Taxonomy Database.” *Nucleic Acids Research* 40: D136–43. doi:10.1093/nar/gkr1178.

Fine, Arthur. 2009. *The Shaky Game: Einstein, Realism, and the Quantum Theory*. Chicago: University of Chicago Press.

Franz, N M, and R K Peet. 2009. “Perspectives: Towards a Language for Mapping Relationships Among Taxonomic Concepts.” *Systematics and Biodiversity* 7 (1): 5–20. doi:10.1017/S147720000800282X.

Franz, Nico M, and Beckett W Sterner. 2018. “To Increase Trust, Change the Social Design Behind Aggregated Biodiversity Data.” *Database* 2018 (January). Oxford University Press. doi:10.1093/database/bax100.

Franz, Nico M, and David Thau. 2010. “Biological Taxonomy and Ontology Development: Scope and Limitations.” *Biodiversity Informatics* 7 (1): 45–66. doi:10.17161/bi.v7i1.3927.

Franz, Nico M, Chao Zhang, and Joohyung Lee. 2017. “A Logic Approach to Modelling Nomenclatural Change.” *Cladistics* 34 (3): 336–57. doi:10.1111/cla.12201.

Franz, Nico M, Mingmin Chen, Parisa Kianmajd, Shizhuo Yu, Shawn Bowers, Alan S Weakley, and Bertram Ludäscher. 2016. “Names Are Not Good Enough: Reasoning Over Taxonomic Change in the Andropogon Complex.” *Semantic Web* 7 (6). IOS Press: 645–67. doi:10.3233/SW-160220.

Franz, Nico M, Mingmin Chen, Shizhuo Yu, Parisa Kianmajd, Shawn Bowers, and Bertram Ludäscher. 2015. “Reasoning Over Taxonomic Change: Exploring Alignments for the Perelleschus Use Case.” *PLoS ONE* 10 (2): e0118247. doi:10.1371/journal.pone.0118247.

Franz, Nico M, Naomi M Pier, Deeann M Reeder, Mingmin Chen, Shizhuo Yu, Parisa Kianmajd, Shawn Bowers, and Bertram Ludäscher. 2016. “Two Influential Primate Classifications Logically Aligned.” *Systematic Biology* 65 (4): 561–82. doi:10.1093/sysbio/syw023.

Goodman, Nelson. 1983. *Fact, Fiction, and Forecast*. 4 ed. Cambridge, MA: Harvard University Press.

Hitchcock, A S, and A Chase. 1950. *Manual of the Grasses of the United States*. 2nd ed. Vol. 200. Washington, DC: US Department of Agriculture.

Hoehndorf, Robert, Michel Dumontier, and Georgios V Gkoutos. 2013. “Evaluation of Research in Biomedical Ontologies.” *Briefings in Bioinformatics* 14 (6): 696–712. doi:10.1093/bib/bbs053.

Hull, David L. 1988. *Science as a Process*. Chicago: University of Chicago Press.

Keet, C M. 2018. *An Introduction to Ontology Engineering*. Maria Keet.

Klein, Gunnar O, and Barry Smith. 2010. “Concept Systems and Ontologies: Recommendations for Basic Terminology.” *Information and Media Technologies* 5 (2): 720–28. doi:10.11185/imt.5.720.

Kuhn, Thomas S. 1996. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Kumar, Anand, and Barry Smith. 2003. “The Universal Medical Language System and the Gene Ontology: Some Critical Reflections.” In *KI 2003: Advances in Artificial Intelligence*, 2821:135–48. Lecture Notes in Computer Science. Berlin: Springer. doi:10.1007/978-3-540-39451-8_11.

Leonelli, Sabina. 2013. “Integrating Data to Acquire New Knowledge: Three Modes of Integration in Plant Science.” *Studies in the History and Philosophy of Biological and Biomedical Sciences* 44 (4): 503–14. doi:10.1016/j.shpsc.2013.03.020.

Leonelli, Sabina. 2016. *Data-Centric Biology: A Philosophical Study*. Chicago: University of Chicago Press.

Lepage, Denis, Gaurav Vaidya, and Robert Guralnick. 2014. “Avibase – a Database System for Managing and Organizing Taxonomic Concepts.” *ZooKeys* 420 (420): 117–35. doi:10.3897/zookeys.420.7089.

Lord, Phillip, and Robert Stevens. 2010. “Adding a Little Reality to Building Ontologies for Biology.” *PLoS ONE* 5 (9): e12258. doi:10.1371/journal.pone.0012258.

Maojo, V, J Crespo, M García-Remesal, D de la Iglesia, D Perez-Rey, and C Kulikowski. 2011. "Biomedical Ontologies: Toward Scientific Debate." *Methods of Information in Medicine* 50 (3): 203–16. doi:10.3414/ME10-05-0004.

Merrill, G H. 2010a. "Realism and Reference Ontologies: Considerations, Reflections and Problems." *Applied Ontology*. doi:10.3233/AO-2010-0080.

Merrill, Gary H. 2010b. "Ontological Realism: Methodology or Misdirection?" *Applied Ontology* 5 (2). IOS Press: 79–108. doi:10.3233/AO-2010-0076.

Millerand, F, D Ribes, K S Baker, and G C Bowker. 2013. "Making an Issue Out of a Standard: Storytelling Practices in a Scientific Community." *Science, Technology & Human Values* 38 (1): 7–43. doi:10.1177/0162243912437221.

Miscellaneous authors. 2010. "Ontological Realism and OBO Foundry Criteria." *OBO Discuss*. http://obo-discuss.2851485.n2.nabble.com/Ontological-Realism-and-OBO-Foundry-Criteria-tp5293729.html.

Otero-Cerdeira, Lorena, Francisco J Rodríguez-Martínez, and Alma Gómez-Rodríguez. 2015. "Ontology Matching: A Literature Review." *Expert Systems with Applications* 42 (2): 949–71.

Pesquita, Catia, Daniel Faria, Cosmin Stroe, Emanuel Santos, Isabel F Cruz, and Francisco M Couto. 2013. "What's in a 'Nym'? Synonyms in Biomedical Ontology Matching." In *The Semantic Web – ISWC 2013*, 8218:526–41. Lecture Notes in Computer Science. Berlin: Springer. doi:10.1007/978-3-642-41335-3_33.

Radford, A E, H E Ahles, and C R Bell. 1968. *Manual of the Vascular Flora of the Carolinas*. Chapel Hill, NC: University of North Carolina Press.

Remsen, David. 2016. "The Use and Limits of Scientific Names in Biological Informatics." *ZooKeys* 550 (4): 207–23. doi:10.3897/zookeys.550.9546.

Seltmann, K, A Austin, and J Jennings. 2012. "A Hymenopterists' Guide to the Hymenoptera Anatomy Ontology: Utility, Clarification, and Future Directions." *Journal of Hymenoptera Research* 27 (2): 67–88. doi:10.3897/JHR.27.2961.

Smith, Barry. 2004. "Beyond Concepts: Ontology as Reality Representation." In *Proceedings of FOIS*, edited by Varzi Achille and Laure Vieu.

Smith, Barry. 2005. "The Logic of Biological Classification and the Foundations of Biomedical Ontology." In *Logic, Methodology and Philosophy of Science*, edited by P Hájek, 505–20. Invited Papers From the 10th International Conference in Logic Methodology and Philosophy of Science. London: Elsevier.

Smith, Barry. 2008a. "Chapter 4: New Desiderata for Biomedical Terminologies." In *Applied Ontology*. An Introduction. Boston: De Gruyter. doi:10.1515/9783110324860.83.

Smith, Barry. 2008b. "New Desiderata for Biomedical Terminologies." In *Chapter 4: New Desiderata for Biomedical Terminologies*, edited by Katherine Munn and Barry Smith, 83–107. Boston: De Gruyter. doi:10.1515/9783110324860.83.

Smith, Barry, and Werner Ceusters. 2006. "HL7 RIM: an Incoherent Standard." *Studies in Health Technology and Informatics*. 124: 133-138.

Smith, Barry, and Werner Ceusters. 2010. "Ontological Realism: a Methodology for Coordinated Evolution of Scientific Ontologies.." *Applied Ontology* 5 (3-4): 139–88. doi:10.3233/AO-2010-0079.

Smith, Barry, Michael Ashburner, Cornelius Rosse, Jonathan Bard, William Bug, Werner Ceusters, Louis J Goldberg, et al. 2007. "The OBO Foundry: Coordinated Evolution of

Ontologies to Support Biomedical Data Integration." *Nature Biotechnology* 25 (11): 1251–55. doi:10.1038/nbt1346.

Smith, Barry, Werner Ceusters, Bert Klagges, Jacob Köhler, Anand Kumar, Jane Lomax, Chris Mungall, Fabian Neuhaus, Alan L Rector, and Cornelius Rosse. 2005. "Relations in Biomedical Ontologies." *Genome Biology* 6 (5): R46. doi:10.1186/gb-2005-6-5-r46.

Sojic, Aleksandra, and Oliver Kutz. 2012. "Open Biomedical Pluralism: Formalising Knowledge About Breast Cancer Phenotypes." *Journal of Biomedical Semantics* 3 (Suppl 2): S3. doi:10.1186/2041-1480-3-S2-S3.

Sterner, Beckett W. 2018. "Review of Sabina Leonelli's Data-Centric Biology: a Philosophical Study." *Philosophy of Science* 85 (3): 540–50. doi:10.1086/697732.

Sterner, Beckett W, and Nico M Franz. 2017. "Taxonomy for Humans or Computers? Cognitive Pragmatics for Big Data." *Biological Theory* 12 (2). Springer Netherlands: 99–111. doi:10.1007/s13752-017-0259-5.

Sterner, Beckett W, and Scott Lidgard. 2014. "The Normative Structure of Mathematization in Systematic Biology." *Studies in the History and Philosophy of Biological and Biomedical Sciences* 46: 44–54. doi:10.1016/j.shpsc.2014.03.001.

Sterner, Beckett W, and Scott Lidgard. 2018. "Moving Past the Systematics Wars." *Journal of the History of Biology* 51: 31–67. doi:10.1007/s10739-017-9471-1.

The OBO Foundry. 2018. "Principle: Maintenance." *Obofoundry.org*. Accessed September 22. http://www.obofoundry.org/principles/fp-016-maintenance.html.

Wilkins, John S. 2009. *Species: A History of the Idea*. Berkeley, CA: University of California Press.

Winther, Rasmus Gronfeldt. 2015. "The Structure of Scientific Theories." Edited by Edward N Zalta. *The Stanford Encyclopedia of Philosophy*. doi:10.3389/fpls.2012.00102/full.

Yoder, Matthew J, István Mikó, Katja C Seltmann, Matthew A Bertone, and Andrew R Deans. 2010. "A Gross Anatomy Ontology for Hymenoptera." Edited by Corrie S Moreau. *PLoS ONE* 5 (12): e15991. doi:10.1371/journal.pone.0015991.

2018. "NCBITaxon Ontology." *Obofoundry.org*. Accessed September 22. http://purl.obolibrary.org/obo/ncbitaxon.owl.

Figures

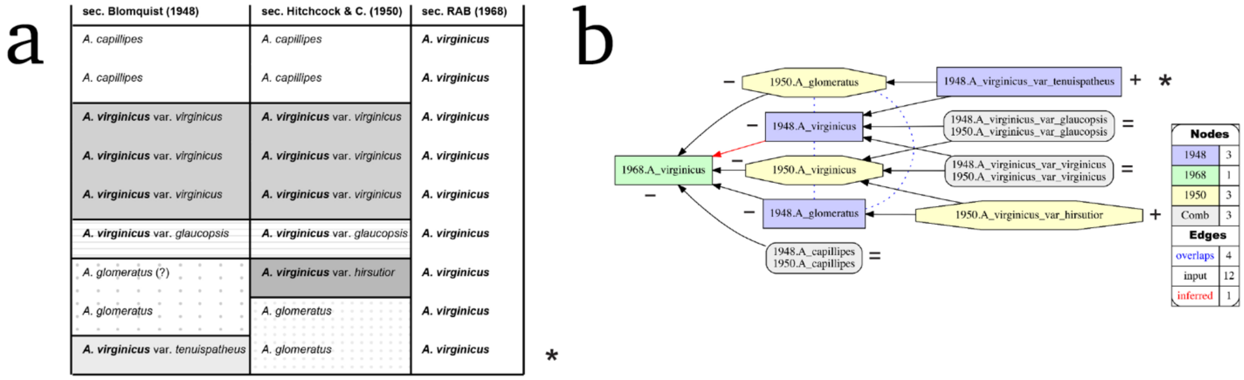

Fig. 1 Illustration of how logically aligning the extensions of classificatory terms allows search and reasoning services for data associated with specific names. Representations of the semantic relations between three classifications—Blomquist (1948), Hitchcock and Chase (1950), and Radford, Ahles, and Bell (1968)—utilizing the name *A. virginicus* (adopted from Franz, Chen, et al. 2016). **a** Tabular alignment of lowest-level taxonomic concepts; each horizontal row corresponds to a congruent lineage of taxonomic concepts labeled with different but (locally) valid names. **b** Directed acyclical graph alignment visualization ("meaning transition map") of the three classifications, as logically inferred with the Euler/X multi-taxonomy alignment toolkit (Franz et al. 2015; Franz, Chen, et al. 2016; Franz, Pier, et al. 2016). Taxonomic concept labels are abbreviated. *Arrows* symbolize inverse proper inclusion (<), whereas *dashed lines* indicate overlap (><). Regions with *rounded corners* symbolize congruent concept regions; whereas other regions with various (*gray* to *white*) shadings and either *rectangular or octagonal* shapes represent concepts unique to the respective input taxonomy (i.e., either 1948, 1950, or 1968). Additional annotations: = congruent concepts, identical names; +: unique concept, unique name; − unique concept, non-unique name; * (**a** and **b**): low-level concept lineage congruent with *A. virginicus* var. *tenuispatheus* sec. Blomquist (1948), as singled out in the main text

Table 1: The three consensus principles compared and contrasted.

	Realist	**Contextual-Best**	**Coordinative**
Grounds of consensus	*Metaphysical truth*	*Epistemic agreement*	*Communicative expediency*
Target of consensus	*Meanings of classificatory terms*	*Meanings of classificatory terms*	*Standard for defining classificatory terms*
Appropriate conditions	*High agreement about meanings and uses of classificatory terms*	*Moderate agreement about meanings and uses of classificatory terms, respectively*	*Low agreement about meanings of classificatory terms; high agreement about extensions of definitions*