HPLS (2020) 42:8
https://doi.org/10.1007/s40656-020-0300-z

ORIGINAL PAPER

Coordinating dissent as an alternative to consensus classification: insights from systematics for bio-ontologies

Beckett Sterner[1] · Joeri Witteveen[2] · Nico Franz[1]

Received: 10 May 2019 / Accepted: 17 January 2020

Abstract The collection and classification of data into meaningful categories is a key step in the process of knowledge making. In the life sciences, the design of data discovery and integration tools has relied on the premise that a formal classificatory system for expressing a body of data should be grounded in consensus definitions for classifications. On this approach, exemplified by the realist program of the Open Biomedical Ontologies Foundry, progress is maximized by grounding the representation and aggregation of data on settled knowledge. We argue that historical practices in systematic biology provide an important and overlooked alternative approach to classifying and disseminating data, based on a principle of coordinative rather than definitional 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 Bio-ontologies · Big data · Data-centrism · Consensus principle · Coordination · Ontology alignment · Biodiversity informatics

The possibility of a theory-neutral language for expressing empirical data has been an enduring question for philosophy of science (Kuhn 1996). Without theoretical assumptions to guide the collection and interpretation of facts, science cannot get started—one simply acquires an ever-larger cabinet of curiosities. The fact that we rely on models and theories to make data meaningful, however, threatens

✉ Beckett Sterner
beckett.sterner@asu.edu

[1] School of Life Sciences, Arizona State University, Tempe, USA

[2] Department of Science Education, Section for History and Philosophy of Science, University of Copenhagen, Copenhagen, Denmark

the longevity of scientific knowledge, since new discoveries may always overthrow existing theory and render old investigations irrelevant. Big data is opening new perspectives on this classic problem 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 et al. 2011; Leonelli 2016). These circumstances provide an opportunity for novel integrative research on the prerequisites for scientific progress: what epistemic principles should guide the design of computational systems to enable the accumulation of useful data over time?

New answers to this question are emerging, but their precise scope, content, and relative merits are still unclear (Leonelli 2016; Sterner and Franz 2017; Franz and Sterner 2018). In this paper, we focus on analyzing what sort of consensus knowledge is sufficient to deliver reliable data discovery, aggregation, and reasoning. Data discovery concerns whether user queries to a database returns all the information it contains about the query subject and only information about that subject. Similarly, data aggregation and reasoning concern whether users can make inferences about data points and classes to derive logically valid or statistically accurate conclusions about groups.

In the biological and biomedical sciences, what we will call the *Definitional Consensus Principle* has dominated the design of data discovery and integration tools:

> **Definitional Consensus Principle (DCP)**: The design of a formal classificatory system for expressing a body of data should be grounded in a consensus about the definitions of the entities that are being classified.

The DCP has been motivated philosophically and operationalized in scientific practice in various ways. However, the variety of perspectives on (and debates over) ontology designs based on the DCP has concealed that it is not the only conceivable normative guide to designing data discovery and integration systems.

In this paper, we introduce an alternative *Coordinative Consensus Principle* and argue that designs based on this principle have distinctive advantages in certain contexts and for certain ends (Table 1). We thus claim that the broader task of ensuring the cumulativeness of scientific knowledge over time requires a comparative study of the relative merits and trade-offs of each approach to consensus formation 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.

We begin by discussing the most visible formulation of DCP in the life sciences, which has been spearheaded by Barry Smith and Werner Ceusters (Smith et al. 2007; Arp et al. 2015). What they call the "ontological realist" approach

Table 1 The three consensus principles compared and contrasted

	Definitional consensus principle (DCP)		Coordinative consensus principle (CCP)
	Realist interpretation (DCP/R)	Contextual interpretation (DCP/C)	
Grounds of consensus	Metaphysical truth	Epistemic agreement	Communicative expediency
Target of consensus	Meanings of the terms used for classified entities	Meanings of the terms used for classified entities	Application of the terms used for classified entities
Appropriate conditions	High agreement about meanings and uses of the terms used for classified entities	Moderate agreement about meanings and uses of the terms used for classified entities	Low agreement about meanings of classificatory terms, but high agreement about correct application of terms

towards ontology design 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). Smith and Ceusters's realist conception of DCP asserts that our classificatory theories of data should be grounded on consensus about nature in order to harnessing maximal efficiencies of scale from computational processing. We suggest that while it presents a principled approach to ontology-building, scientific practice often calls for less restrictive and more flexible design principles that can be contextualized to a field or topic.

Next, we show that many contemporary ontologies that are endorsed by biologists depart from Smith and Ceusters's approach in adopting a weakened version of DCP that only requires consensus on *some* principle determining what is the "best" ontology for a body of data, and which permits the meaning of "best" to vary by context (Lord and Stevens 2010; Maojo et al. 2011; Sojic and Kutz 2012; Hoehndorf et al. 2013; Franz and Goldstein 2013). While still consensus-based, these scientists' practices are not organized under an equivalently systematic and general approach to ontology design.

Having presented these two approaches toward ontology-building based on a consensus about the classificatory entities, we turn to a third alternative, based on coordination across conflicting views. Many areas of biological systematics continue to experience rapid change in, and dissent over, the meaning of taxon names. 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 (Epstein 2012; Remsen 2016; Vogt 2011; Witteveen 2015)?

We argue that the answer is yes, but that 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 relied on a consensus principle 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 methods and adopting practices 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 formation in systematics aims for consensus about the application of a name even in the face of disagreement about its precise meaning. We show that the underlying Coordinative Consensus Principle has been a cornerstone of systematic practice historically and argue that there are strong reasons for it to remain central to data discovery and integration in systematic biology in the age of data-centric biology.

1 A comparative approach to big data trajectories

Every field 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 efforts as tracing out multiple "big data trajectories" for research fields over time that we may then 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, p. 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 expectations. 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 (Bowker 2000; Millerand et al. 2013; Leonelli 2016). This provides rich opportunities for comparative study as biologists articulate and generalize lessons learned from local experiences.

In this context, formalized languages have become increasingly important factors shaping how biologists understand and use data, and they constitute classificatory theories in their own right that deserve philosophical analysis as much as traditional theories such as quantum mechanics or evolutionary theory (Leonelli 2016; Sterner 2018). Classificatory theories for representing and reasoning about data serve at least two important functions for biology: data discovery and integration. Data discovery relies on searchable metadata so that user queries return appropriate 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 assay data measuring gene expression levels with data about where those gene products end up in cells. Another example would be to integrate data describing molecular interactions in order to predict new gene interactions or functions genes across several species (Leonelli 2013).

In this context, the DCP operates as a normative, methodological resource to regulate scientists' practices in the process of computerizing data representation and reasoning. We treat computerization here 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, 2018). 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

---

[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 data journeys 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.

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 critically analyzing the claims people make about the revolutionary consequences of adopting computer technology.

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 2003, p. 15). This quote nicely expresses the way that bringing computers into biologists' practices of representing and reasoning about data begins to transform the language and practice of science.

Looking ahead, a key issue will be how scientists should go about formalizing their classificatory terminologies in order to advance data discovery and reasoning in their local domains, and whether principles suited to one knowledge domain extend to other areas. A one-size-fits-all approach cannot be assumed and arguably should not be expected. Indeed, using our comparative perspective on big data trajectories, we will 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 articulating and contrasting three forms of epistemic consensus that scientists invoke to warrant their choice of data classifications. We start by illustrating two ways in which the Definitional Consensus Principle can be further articulated before turning to consider how both versions differ from a third alternative, the Coordinative Consensus Principle.

2 A realist view of the DCP

Precisely what sort of consensus does the Definitional Consensus Principle aspire to? In this section, we look at pioneering work by Smith and Ceusters in the 2000s that introduced "ontological realism" as one answer (e.g. Smith 2004). They have taken inspiration from the metaphysical realist literature in philosophy to formulate a methodology of ontology development in the biomedical sciences that posits universals and types. They thus endorse what we can call the *Realist Interpretation* of the Definitional Consensus Principle for ontologies:

> **Realist Interpretation of the Definitional Consensus Principle (DCP/R)**: 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.

The need for a further articulation of DCP reflects structural ambiguities inherent in the design of current technologies for data representation and reasoning. Computer ontologies possess well-defined syntaxes and reasoning rules, but their semantic interpretations can differ in philosophically fundamental ways.

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 these examples show, ontologies can also specify classes of relations, such as "is the father of" or "is part of." Computer ontologies commonly instantiate a type of description logic, which relies on a more restricted and efficient syntax compared to first-order logic and which is implemented in a standardized machine-readable syntax such as the OWL2 Web Ontology Language. As we'll see, simply 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, and neither is following the standardized syntactic rules governing any implementation like OWL, because these underlying standards don't specify a coherent semantics able to support reliable reasoning and description of biological data.

The Realist Interpretation of the DCP, then, aims to put a realist semantic interpretation on the basic syntactic structures of first-order logic. For simplicity, we treat Smith and Ceusters's ontological realism as containing two main elements: (1) an ontology called the Basic Formal Ontology (BFO), which is foundational for all other bio-ontologies, and (2) a realist interpretation of the structures in an ontology. Their work also contains a number of supplementary guidelines we do not discuss.

The world of ontological realism (and hence the BFO) has two types of entities: universals and particulars. Examples of universals include human being, enzyme, and aspirin. Examples of particulars are Queen Elizabeth II, the Planet Earth, and a piece of cheese (Smith 2003). In Smith and Ceusters (2010), they characterize universals as repeatable, i.e. as having potentially indefinitely many instances we can discover in principle.[3] Smith (2003) also proposes that universals are distinguished from mere extensional sets by figuring in scientific laws of nature or by having real definitions rather than nominal definitions. Hence, although ontologies can include terms for what Smith calls "collections," which are sets formed from logical combinations of universals, these collections are treated as artificial constructs rather than real entities in the world.

The guiding slogan of ontological realism is thus to "describe what exists in reality, not what is known about what exists in reality" (Arp et al. 2015). In the early 2000s, Smith and Ceusters demonstrated a number of cases where existing

---

[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.

[3] Smith and Ceusters give different general characterizations of universals and particulars in different places, though see Merrill (2010a). For example, in Smith (2003), universals are multiply located entities that exist in particulars, while particulars are entities with only one location in space at a time. 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 2003).

terminologies failed to keep these separate, leading to basic reasoning flaws (Kumar and Smith 2003; Ceusters et al. 2005; Smith et al. 2005; Smith and Ceusters 2006). Describing early versions of the SNOMED[4] 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 2008, p. 83). For more detail on how this approach has developed in the case of the Gene Ontology, see Leonelli (2010, 2012).

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. 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.

Two additional features of Smith and Ceusters'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" (Anonymous 2018). 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 2008). As a result, ontological realism does not necessarily imply eliminative reductionism, e.g. about the reality of organisms in favor of molecules.

However, ontological realism does enforce universal consistency of other bio-ontologies with the BFO (Arp et al. 2015). BFO provides a set of metaphysically fundamental terms, roughly consistent with a common sense view of the world as "midsize dry goods" consistent with Newtonian mechanics (Lord and Stevens

---

[4] Systematized Nomenclature of Medicine.

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).

3 A contextualist view of the DCP

Smith and Ceusters's realist view is not the only take on ontology design based on the DCP. In this section, we look at a contextualized, epistemic version of consensus that has been adopted in some recent bio-ontologies. To get a grip on the difference between these interpretations and operationalizations of the DCP, it will help to first consider an important critique of the realist view.

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. Merrill similarly argued that Smith and Ceusters's neo-Aristotelian metaphysical realism does not provide a fundament for ontology design, but rather imposes a uniform philosophical interpretation on science where it is neither warranted nor required. The goal of ontologies in the empirical sciences is not to provide a foundational account of the processes and entities described by the relevant sciences, but rather to present a formal framework for facilitating (semi-)automated reasoning and computation (Merril 2010a, p. 104).

Merrill's (2010a, b) critique led to a broader discussion in the community and exposed a more expansive range of alternatives than Smith and Ceusters allowed. It made clear that the choice was not between an obviously flawed conceptualist design and a metaphysical realism based on first-order logic in terms of universals (Smith 2004, 2008; Smith and Ceusters 2010). Although, Smith and Ceusters were right to point to flaws in conceptualist-inspired ontology designs, these flaws were due to a misapplication of basic principles of formal logic, semantics and philosophy of language. They could be corrected without adopting a realist metaphysics of universals. Several other participants in the community 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 et al. 2013).

While Smith and Ceusters have had major influence on formulating design principles for bio-ontologies, especially in the early 2000s, their view is far from doctrinal in the broader biological community. 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 et al. (2015) take species to be paradigm examples of universals, other biologists have strongly disputed this claim (Hull 1976; Ghiselin 1974; Franz and Thau 2010). Moreover, the widely-used NCBI taxonomy ontology is a bare hierarchy of taxonomic names with no definitions that is automatically generated from other name databases (Federhen 2012).[5]

The current practices of ontology-building in the biomedical domain suggest that ontologies are better characterized as employing design principles that are adapted to the particularities of their respective domains. This approach to ontology-building departs from the DCP/R 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 Interpretation* of the Definitional Consensus Principle for ontology-design:

> **Contextual Interpretation of the Definitional Consensus Principle (DCP/C):** 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 DCP/C 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. For example, 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

---

[5] Note that we are not mentioning a design like this to endorse it. As Minelli (2017, this issue) has adroitly pointed out, there are major worries over the quality and durability of these and other name-based aggregators that include ‘grey’ non-Linnaean names. In the next section we will return to taxonomic names and consider different design solutions.

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 curators often play a crucial role in facilitating integration by creating lists of synonyms and homonyms 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 expressed choice among entomologists building the Hymenoptera Anatomy Ontology (HAO) not to use evolutionary homologies (Yoder et al. 2010; Seltmann et al. 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 (cf. Edgecombe 2008; Vogt 2017). 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 et al. 2012, p. 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.[6]

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,

---

[6] 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.

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 et al. 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.

4 A third way: consensus in the face of lasting dissent

The ability to coordinate meanings across conflicting views is critical for systematics because 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 and many known species have received only minimal attention so far. 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).

Given this state of dissensus about the accurate circumscription of significant taxonomic groups, taxon names are absolutely fundamental to reliably circulating and communicating biodiversity data (Remsen 2016). This has held true for the practice of biological taxonomy historically since Linnaean times (Müller-Wille 2017; Witteveen 2016, 2018), but it is exacerbated by both the technologies and demands of circulating data in the age of data-intensive biodiversity science (Patterson et al. 2010). In some taxonomic groups, biologists can rely on nomenclatural databases that assert a single coherent 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 et al. 2014). In another analysis of two highly-cited treatments of 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 et al. 2016c). 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. As a result, 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 (Guala 2016; Franz and Sterner 2018; Vaidya et al. 2018).

Facing persistent disagreement about what taxa exist in nature, or which parts of our classifications are settled knowledge, it might seem impossible for taxonomists to rely on consensus as a basis for communication and reasoning about their data. However, we argue that Linnaean nomenclature and classification uses a third form of CP based on agreement about procedural rules ensuring the translatability of data across classifications.

> **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 application of names of classified entities, even if the meanings and extensions of those names haven't been settled.

Simply put, the CCP can be interpreted as means of agreeing on how to reliably communicate and calibrate standing disagreements about the contents and boundaries of groups. It is a means of reaching a consensus about an actual, fluid state of dissensus. In the rest of this section, we show how taxonomists have implemented a system consistent with CCP. We also argue this system permits coherent reasoning about both referents and concepts, and thus presents a genuine alternative to the DCP.

To see how this is possible, we first need to look more carefully at three main components to taxon names today. The name itself, as a sequence of letters, carries no or minimal content relevant to its semantic meaning, even if it sometimes superficially describes a property of the taxonomic referent (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). What determines the application of a name is not its content, but the pair of the name plus a fixed name-bearing element of a taxonomic group—its so-called "nomenclatural type" or "type" for short. It is this practice of "anchoring" names to nature 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 also 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 (Dubois 2005; Witteveen 2015).

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 (Berendsohn and Geoffroy 2007; Franz and Peet 2009). The type specimen that falls within the circumscription that is covered by the taxonomic concept determines the application of a particular taxonomic name to that concept.[7]

Unlike the choice of a species concept, which given the current state of the species debate can perhaps 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.” While correct as a matter of principle, a statement of this kind fails to communicate anything about the present state of knowledge about the relevant species. 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.

The answer for how systematists can reach a consensus about the application of names in the absence of a consensus about the meanings of names therefore has two parts. First, taxonomists use nomenclatural principles and procedures to govern the *application* of taxon names to taxonomic hypotheses based on the designation of static nomenclatural types for each name. This uses ostensive reference to stabilize a name’s meaning by fixing its referent without requiring consensus about what that the true referent is. 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; see also Witteveen 2015).

Second, taxonomists follow the convention of including diagnostic criteria, drawings, pictures, and/or descriptions of (members of) new taxa when naming them, thus facilitating the recognition and potential revision of the circumscription by other taxonomists. As a result, even when biologists disagree about the group of organisms a name refers to in the world, they can still operationally assign new instances to the name according to any of its associated taxonomic hypotheses. In addition, experts can use these diagnostic criteria to make further inferences about how specimens assigned to a name under one taxonomic concept can be mapped onto other name-concept pairs.

Together, these two features of the contemporary practice of Linnaean naming constitute an alternative to ontology design based on the Definitional Consensus

---

[7] If the circumscription includes two or more type specimens, the name associated with the most senior type specimen is the valid/correct name for the species and the other names become (junior) synonyms.

Principle that we discussed earlier. Unlike approaches based on realist (DCP/R) and contextualist (DCP/C) principles of ontology design, the practice of Linnaean naming based on the designation of nomenclatural types and the articulation of taxonomic concepts establishes facts of the matter about the referents of taxonomic names in the absence of agreement about how to specify and individuate those referents. In accordance with the Coordinative Consensus Principle, the combination of names and types provides a means for communicating, testing, and revising taxonomic concept descriptions for empirical accuracy.

Although biologists regularly take the ostensive component of this shared convention to involve reference to real taxa, it is important to note that this realist interpretation is not mandated by the convention itself (and hence does not hark back to the DCP/R). Even if the "final entity" that Cook (1898) refers to in the quotation above proves to be illusive—or even illusory—the goal of coordinating and reliable communicating disagreements about candidates for accurate circumscriptions will be fulfilled by following the convention. Hence, including taxonomic concepts in the communication of names does not reflect an all-encompassing philosophy of language: systematists do not attempt to define "concept" in general, nor do they claim that meaning is always extensional rather than intensional, or that one form of definition is inferior in general. Instead, they simply stipulate a consensus convention within the local context of their professional community regulating what is required of taxonomic concept description for it to be admissible.

In this regard, the Linnaean system of naming based on types and taxonomic concepts resembles the contextualist interpretation of the DCP more than its realist interpretation. Yet it also departs from both interpretations in steering clear of imposing a consensus view on the definition of the individual classified entities. At most, it makes (minimal) requirements on the form of descriptions of taxonomic concepts in general.

5 Coordinative consensus-building and contemporary data science

From the outset, the Linnaean system and practice of naming has been attuned to the need for revision and facilitated the digestion of novel (cumulative and potentially conflicting) taxonomic data (Müller-Wille and Charmantier 2012; Dietz, this issue). Nevertheless, the historical circumstances that led to the original Linnaean system of naming and its contemporary version differ in important ways from the demands posed by 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 application of names that we have described above. We also suggest how the CCP generalizes beyond the hybrid ostensive reference system used by taxonomists.

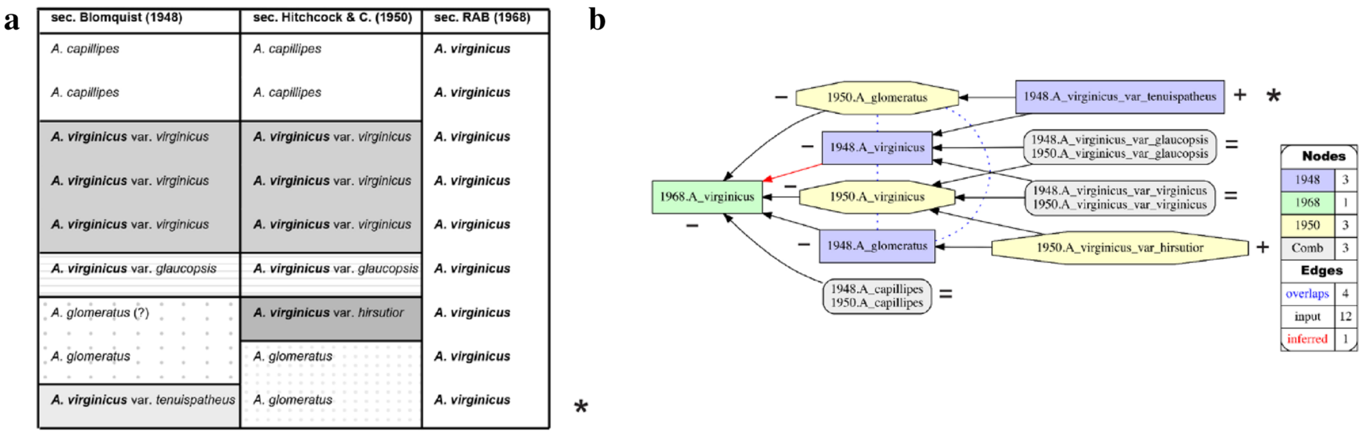
As with biomedical ontologies, handling taxonomic names on a large scale without sacrificing performance requires researchers to articulate the inferential relationships among names and meanings in a newly formalized way. Historically, human experts have been able to memorize the relevant classifications for a taxonomic group and translate specimen identifications across classification using substantial tacit background knowledge. Human non-experts and computers, however, generally lack this background knowledge and aren't able to out reliable contextual disambiguation of taxonomic names and their meanings (Sterner and Franz 2017). Automated ontology matching algorithms, for example, often rely on similarity between the spellings of class names and where they are located in the class hierarchy. Taxonomists regularly synonymize names with dissimilar character strings, of course, so name-based alignment alone has serious limitations.

Joint use of nomenclatural types and descriptive hypotheses, however, enables several fruitful inference strategies. For one thing, nomenclatural rules enforcing a strict association of a name and ostensive type ensure a minimal continuity and fixity of reference for the same name over time. In addition, name-usages associated with the same type generally have the same referent (i.e. point to the same taxon in the world), although they be associated with conflicting hypotheses about the extension of that taxon. This provides grounds for asserting some level of similarity in meanings of these names.

More importantly, systematists publishing new classifications for a taxonomic group include extensive lists of synonyms as part of their name definitions. These synonym lists include cases of different names with the same types but also extend much further to include names that the authors judge based on their evidence and expertise to have the same meaning. The practice of making synonymy lists is largely a consequence of the nomenclatural rule of priority, which treats the first published name for a taxon as valid and all other names as subordinated synonyms. While the primary point of this rule is to establish a conventional fact of the matter about the correct name for a taxon relative to a particular classification system, it also ensures that systematists have to explicitly track the semantic relationships between names proposed by authors studying the same group over time.

Extensive databases of synonyms now exist that enable improved data aggregation beyond name-only alignments. The Integrated Taxonomic Information System (ITIS), for example, offers one of the most complete global lists of names and synonyms, with demonstrated value for improving recall and precision for information searches using taxonomic names (Guala 2016). Some databases also provide global taxonomic "backbone" hierarchies in addition to synonymy lists, but creating these backbones often involves stitching together local hierarchies in ways that deviate from any published view (Franz and Sterner 2018). Far from being a consensus view, then, parts of the backbone may assert claims that no expert endorses. Data aggregation based on name similarity plus synonymies is now recommended for current informatics workflows generating species distribution maps and ecological models (Kissling et al. 2017).

A more epistemically accurate but also resource-intensive 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, 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*



sec. Blomquist (1948)	sec. Hitchcock & C. (1950)	sec. RAB (1968)
A. capillipes	*A. capillipes*	***A. virginicus***
A. capillipes	*A. capillipes*	***A. virginicus***
A. virginicus var. *virginicus*	***A. virginicus*** var. *virginicus*	***A. virginicus***
A. virginicus var. *virginicus*	***A. virginicus*** var. *virginicus*	***A. virginicus***
A. virginicus var. *virginicus*	***A. virginicus*** var. *virginicus*	***A. virginicus***
A. virginicus var. *glaucopsis*	***A. virginicus*** var. *glaucopsis*	***A. virginicus***
A. glomeratus (?)	***A. virginicus*** var. *hirsutior*	***A. virginicus***
A. glomeratus	*A. glomeratus*	***A. virginicus***
A. virginicus var. *tenuispatheus*	*A. glomeratus*	***A. virginicus***

Fig. 1 Illustration of how logically aligning hypotheses about the extensions of taxon names 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 et al. (1968)—utilizing the name *A. virginicus* (adopted from Franz et al. 2016b). **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 et al. 2016b, c). 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

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.

Consider a simple example of how incorporating taxonomic concept alignments into the design of biodiversity data portals can impact the engagement and conclusions of users (for more details see Franz et al. 2016a; Franz and Sterner 2018). The data portal managed by the SouthEast Regional Network of Expertise and Collections (SERNEC) includes occurrence observations of some rare orchids that have been variously classified as forming one to three species and placed in four different genera. Currently, the SERNEC portal allows users to access and contribute its observational data only via a single taxonomic hierarchy defined and maintained locally by the portal, although scientists and wildlife managers more broadly continue to use four or more other classifications that feature names and concepts that conflict with SERNEC's view. As a result, aggregating data submissions from users based on the names they supply, plus synonym relations according to SERNEC's current view, can lead to false biological conclusions about the number and location of geographically restricted (i.e. endemic) orchid populations, with potentially important consequences for conservation decisions. Similarly, the portal's current search feature offers no way to visualize or download the same data according to different classificatory hypotheses. However, it is possible to articulate RCC-5 relationships between each name and associated concept under each taxonomic classification based on the type specimens and diagnostic descriptions provided by each classification's published source. With these relationships defined, it is then possible to reason from how names are applied to specimens under one classificatory view to how the same specimens would be correctly identified under another view despite conflicting concept hierarchies. The computer, on the back end of the portal, has sufficient information encoded about the logical relationships among these classifications so as to provide humans with the ability to engage with the portal as a reference standard from which they can deviate in a reproducible way rather than a consensus they do not believe in but against which they cannot directly express any conflicts.

Although formalized concept alignments are far from universally available for all taxonomic groups, they have been implemented in applied medium-scale platforms for decades. One of the earliest adopters was NatureServe, the Non-Governmental Organization (NGO) responsible for many species conservation status assessments in the U.S. (NatureServe 2019). Denis LePage has also implemented a version of RCC-5 alignments for all published bird classifications in AviBase (Lepage et al.

2014). The iNaturalist citizen science project has also recently implemented a partial version of RCC-5 alignments in order to help mediate community disagreements about preferred taxonomies and enable easier versioning for classifications over time (iNaturalist 2019).

While some aspects of CCP in taxonomy are specific to its rules on nomenclature and definition, the use of ostensive types is not necessary for implementing CCP in general. What matters is not following any particular approach to definition, per se, but reliably provisioning widely intelligible meta-level resources for translating data across views. In taxonomy these meta-resources include synonymy lists and operational concept definitions that are intelligible even to scientists holding alternative views.

In principle, similar techniques can also be applied to bio-ontologies. At a high level of abstraction, taxonomic classifications and bio-ontologies both 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 NCBI taxonomy ontology, though it provides no definitions. Computer ontologies can also contain multiple logical hierarchies such as in the Gene Ontology (GO), which forms a network based on three class hierarchies describing kinds of cellular location, molecular functions, and biological processes (Ashburner et al. 2000; Consortium 2017). Some level of synonym tracking is incorporated in the versioning process of the OBO-Edit platform (Day-Richter et al. 2007), but similarly strict expectations for compiling exhaustive synonymies are not enforced across the wide range of OBO-registered ontologies. Similarly, there are a variety of ways that OBO ontologies currently express mappings among terms and definitions from different ontologies, including for versioning over time (Seppälä et al. 2014), but no consensus practice like providing diagnostic concept descriptions has been implemented (Mungall 2019).

In sum, participants relying on CCP must be able to agree on what it takes to apply a term correctly according to its relevant definitions, but they do not have to agree on whether any of these definitions are even plausibly well-supported. Historically, humans have created and provided their own locally sufficient tacit knowledge or meta-languages for implementing CCP. In the future, though, the challenge will be lowering barriers to creating formalized meta-languages satisfying the needs of CCP and automating their use computationally.

There is an obvious cost to following CCP, then, when consensus meanings are in fact available: translating across data classification schemes requires intensive labor to produce and maintain name-concept 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. As we saw, 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 collectively never been able to 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.

6 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, p. 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, p. 171). We have advanced this broader project here by formulating and articulating 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 (DCP/R), contextual standards for best definitions (DCP/C), and coordination conventions (CCP). We have shown that the first two principles differ from the last one with regard to the kind of consensus-building they facilitate. Whereas the DCP/R and the DCP/C provide normative guidelines for articulating a *substantive* consensus about definitions of individual classificatory items, consensus-building based on the CCP is merely *regulative*: it facilitates the coordination of disagreements about definitions. In a way, this renders the CCP a meta-level consensus principle: an approach to consensus formation concerning the principles and methods for tracking and monitoring of changing states of dissensus.

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. It has been beyond the scope of this paper to discuss which consensus principle is best suited to which area of data-intensive scientific practice. However, we have argued that limiting the content of scientific ontologies to areas of consensus, either ontological or epistemic, is not a universal solution. Our aim has been 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. At the very least, we have shown that this alternative approach to consensus-building cannot be rejected out of hand. We thus challenge the often implicit motivation for adopting a version of the DCP, that anarchy looms in the absence of a substantive consensus. As Barry Smith puts it: “We already face enormous challenges in assimilating the huge amounts of life science data being made available to researchers, and … [in the] need to ensure that these data work well together … the challenges would become even more intractable were different research groups addressing the same biological phenomena each encouraged to employ their own classifications in a spirit of tolerance and diversity" (Smith 2009). In contrast, we have argued that diversity and disagreement do not entail intractability in ontology-building. Rather than opting for an oxymoronic "dictated consensus" in response to being faced with a persistent dissensus, ontology-builders and database architects should consider alternative, coordinative varieties of consensus-based solutions.

We have argued that biological systematics provides an important example of the adoption of the such a coordinative consensus in practice. Both historically and today, systematists have relied on coordinative principles and practices for regulating the application of taxonomic names when neither metaphysical nor epistemic consensus is available. The coordination conventions and techniques employed in biological taxonomy generally and in pioneering domains of data-intensive biodiversity science in particular show that reliable data discovery and integration are possible even in situations that are far from consensus.

In closing, we note that we still 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 happened and do not reliably describe the discovery process. Understanding how scientists design ontologies in practice should therefore be 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? Barry 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.

Acknowledgments Our thanks to the editors for proposing and organizing the special issue on Taxonomy as an Information Science. Their comments and the referees' responses both helped substantially improve the final text. This work was also supported by the McDonnell Foundation via the Marine Biological Laboratory and ASU's Special Initiative Fund for Biodiversity Data Science.

Funding Funding was provided by Division of Social and Economic Sciences (Grant No. STS 1827993), Netherlands Organisation for Scientific Research (Grant No. 275-20-060) and National Science Foundation (Grant Nos. SES-1827993, DEB-1754731).

References

Agar, J. (2006). What difference did computers make? *Social Studies of Science, 36*(6), 869–907.

Anonymous. (2018). Foundational model of anatomy. The National Center for Biomedical Ontology. *Bioportal.bioontology.org*. Retrieved December 18 from https://bioportal.bioontology.org/ontologies/FMA.

Arp, R., Smith, B., & Spear, A. D. (2015). *Building ontologies with basic formal ontology*. Cambridge, MA: MIT Press.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: Tool for the unification of biology. *Nature Genetics, 25*(1), 25–29.

Berendsohn, W. G., & Geoffroy, M. (2007). Networking taxonomic concepts—Uniting without 'unitary-ism. In G. Curry & C. Humphries (Eds.), *Biodiversity databases—Techniques, politics, and applications. Systematics association special volume* (Vol. 73, pp. 13–22). Baton Rouge: CRC Taylor & Francis.

Bertone, M. A., Mikó, I., Yoder, M. J., Seltmann, K. C., Balhoff, J. P., & Deans, A. R. (2013). Matching arthropod anatomy ontologies to the Hymenoptera anatomy ontology: Results from a manual alignment. *Database, 2013,* bas057. https://doi.org/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), 643–683.

Ceusters, W., Smith, B., & Goldberg, L. J. (2005). A terminological and ontological analysis of the NCI thesaurus. *Methods of Information in Medicine, 44*(4), 498–507.

Conix, S. (2018). Integrative taxonomy and the operationalization of evolutionary independence. *European Journal for Philosophy of Science, 154*(2), 1–17.

Cook, O. F. (1898). The method of types. *Science*, *8*(198), 513–516.

Costa, M., Reeve, S., Grumbling, G., & Osumi-Sutherland, D. (2013). The Drosophila anatomy ontology. *Journal of Biomedical Semantics, 4*(1), 32.

Costello, M. J., Bouchet, P., Boxshall, G., Fauchald, K., Gordon, D., Hoeksema, B. W., 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. https://doi.org/10.1371/journal.pone.0051629.

Day-Richter, J., Harris, M. A., Haendel, M., The Gene Ontology OBO-Edit Working Group, & Lewis, S. (2007). OBO-Edit—An ontology editor for biologists. *Nucleic Acids Research, 23*(16), 2198–2200.

Dubois, A. (2005). Proposed rules for the incorporation of nomina of higher-ranked zoological taxa in the International Code of Zoological Nomenclature. *Zoosystema, 27*(2), 365–426.

Edgecombe, G. D. (2008). Anatomical nomenclature: Homology, standardization and datasets. *Zootaxa, 1950,* 87–95.

Epstein, M. (2012). Nomenclature, terminology and language. *Bionomina, 5,* 1–56.

Federhen, S. (2012). The NCBI taxonomy database. *Nucleic Acids Research, 40,* D136–D143.

Fine, A. (2009). *The shaky game: Einstein, realism, and the quantum theory*. Chicago: University of Chicago Press.

Franz, N., Gilbert, E., Ludäscher, B., & Weakley, A. (2016a). Controlling the taxonomic variable: Taxonomic concept resolution for a Southeastern United States herbarium portal. *Research Ideas and Outcomes, 2,* e10610.

Franz, N. M., Chen, M., Kianmajd, P., Yu, S., Bowers, S., Weakley, A. S., et al. (2016b). Names are not good enough: Reasoning over taxonomic change in the Andropogon complex. *Semantic Web, 7*(6), 645–667.

Franz, N. M., Chen, M., Yu, S., Kianmajd, P., Bowers, S., & Ludäscher, B. (2015). Reasoning over taxonomic change: Exploring alignments for the Perelleschus use case. *PLoS ONE, 10*(2), e0118247.

Franz, N. M., & Goldstein, A. M. (2013). Phenotype ontologies: Are homology relations central enough? A reply to Deans Et Al. *Trends in Ecology & Evolution, 28*(3), 131–132.

Franz, N. M., & Peet, R. K. (2009). Perspectives: Towards a language for mapping relationships among taxonomic concepts. *Systematics and Biodiversity, 7*(1), 5–20.

Franz, N. M., Pier, N. M., Reeder, D. M., Chen, M., Yu, S., Kianmajd, P., et al. (2016c). Two influential primate classifications logically aligned. *Systematic Biology, 65*(4), 561–582.

Franz, N. M., & Sterner, B. W. (2018). To increase trust, change the social design behind aggregated biodiversity data. *Database*. https://doi.org/10.1093/database/bax100.

Franz, N. M., & Thau, D. (2010). Biological taxonomy and ontology development: Scope and limitations. *Biodiversity Informatics, 7*(1), 45–66.

Franz, N. M., Zhang, C., & Lee, J. (2017). A logic approach to modelling nomenclatural change. *Cladistics, 34*(3), 336–357.

Ghiselin, M. (1974). A radical solution to the species problem. *Systematic Zoology, 23,* 536–544.

Guala, G. F. (2016). The importance of species name synonyms in literature searches. *PLoS ONE, 11*(9), e0162648.

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

Hoehndorf, R., Dumontier, M., & Gkoutos, G. V. (2013). Evaluation of research in biomedical ontologies. *Briefings in Bioinformatics, 14*(6), 696–712.

Hull, D. L. (1976). Are species really individuals. *Systematic Zoology, 25,* 174–191.

Hull, D. L. (1988). *Science as a process*. Chicago: University of Chicago Press.

iNaturalist. (2019). Taxon frameworks. *iNaturalist.org*. Retrieved March 29 from https://www.inaturalist.org/pages/taxon_frameworks.

Kissling, W. D., Ahumada, J. A., Bowser, A., Fernandez, M., Fernandez, N., García, E. A., et al. (2017). Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. *Biological Reviews, 93*(1), 600–625.

Kuhn, T. S. (1996). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Kumar, A., & Smith, B. (2003). The universal medical language system and the gene ontology: Some critical reflections. In *KI 2003: Advances in artificial intelligence. Lecture notes in computer science* (Vol. 2821, pp. 135–148). Berlin: Springer.

Leonelli, S. (2010). Documenting the emergence of bio-ontologies: Or, why researching bioinformatics requires HPSSB. *History and Philosophy of the Life Sciences, 32*(1), 105–125.

Leonelli, S. (2012). Classificatory theory in data-intensive science: The case of open biomedical ontologies. *International Studies in the Philosophy of Science, 26*(1), 47–65.

Leonelli, S. (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), 503–514.

Leonelli, S. (2016). *Data-centric biology: A philosophical study*. Chicago: University of Chicago Press.

Leonelli, S., Diehl, A. D., Christie, K. R., Harris, M. A., & Lomax, J. (2011). How the gene ontology evolves. *BMC Bioinformatics, 12*(1), 325.

Lepage, D., Vaidya, G., & Guralnick, R. (2014). Avibase—A database system for managing and organizing taxonomic concepts. *ZooKeys, 420*(420), 117–135.

Lord, P., & Stevens, R. (2010). Adding a little reality to building ontologies for biology. *PLoS ONE, 5*(9), e12258.

Maojo, V., Crespo, J., García-Remesal, M., de la Iglesia, D., Perez-Rey, D., & Kulikowski, C. (2011). Biomedical ontologies: Toward scientific debate. *Methods of Information in Medicine, 50*(3), 203–216.

Merrill, G. H. (2010a). Ontological realism: Methodology or misdirection? *Applied Ontology, 5*(2), 79–108.

Merrill, G. H. (2010b). Realism and reference ontologies: Considerations, reflections and problems. *Applied Ontology, 5,* 189–221.

Millerand, F., Ribes, D., Baker, K. S., & Bowker, G. C. (2013). Making an issue out of a standard: Storytelling practices in a scientific community. *Science, Technology and Human Values, 38*(1), 7–43.

Minelli, A. (2017). Grey nomenclature needs rules. *Ecologica Montenegrina, 7,* 654–666.

Müller-Wille, S. (2017). Names and numbers: 'Data' in classical natural history, 1758–1859. *Osiris, 32*(1), 109–128.

Müller-Wille, S., & Charmantier, I. (2012). Natural history and information overload: The case of Linnaeus. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 43*(1), 4–15.

Mungall, C. J. (2019). Never mind the logix: Taming the semantic anarchy of mappings in ontologies. *Douroucouli.wordpress.com*. Retrieved May 26 from https://douroucouli.wordpress.com/2019/05/27/never-mind-the-logix-taming-the-semantic-anarchy-of-mappings-in-ontologie/.

NatureServe. (2019). Methodology and guidelines. *Help.natureserve.org*. Retrieved August 20 from http://help.natureserve.org/biotics/#Methodology/MethodologyGuidelines.htm.

Otero-Cerdeira, L., Rodríguez-Martínez, F. J., & Gómez-Rodríguez, A. (2015). Ontology matching: A literature review. *Expert Systems with Applications, 42*(2), 949–971.

Patterson, D., Cooper, J., Kirk, P., Pyle, R., & Remsen, D. (2010). Names are key to the big new biology. *Trends in Ecology & Evolution, 25*(12), 686–691.

Pesquita, C., Faria, D., Stroe, C., Santos, E., Cruz, I. F., & Couto, F. M. (2013). What's in a 'Nym'? Synonyms in biomedical ontology matching. In *The semantic web—ISWC 2013. Lecture Notes in Computer Science* (Vol. 8218, pp. 526–541). Berlin: Springer.

Radford, A. E., Ahles, H. E., & Bell, C. R. (1968). *Manual of the vascular flora of the Carolinas*. Chapel Hill, NC: University of North Carolina Press.

Remsen, D. (2016). The use and limits of scientific names in biological informatics. *ZooKeys, 550*(4), 207–223.

Seltmann, K., Austin, A., & Jennings, J. (2012). A hymenopterists' guide to the Hymenoptera anatomy ontology: Utility, clarification, and future directions. *Journal of Hymenoptera Research, 27*(2), 67–88.

Seppälä, S., Smith, B., & Ceusters, W. (2014). Applying the realism-based ontology-versioning method for tracking changes in the basic formal ontology. In P. Garbacz & O. Kutz (Eds.), *Formal ontology in information systems* (pp. 227–240). Amsterdam: IOS Press.

Smith, B. (2003). The logic of biological classification and the foundations of biomedical ontology. In *Invited papers from the 10th international conference in logic methodology and philosophy of science, Oviedo, Spain* (pp. 19–25).

Smith, B. (2004). Beyond concepts: Ontology as reality representation. In V. Achille & L. Vieu (Eds.), *Proceedings of FOIS 2004 international conference on formal ontology and information systems* (pp. 73–84). Fairfax, VA: IOS Press.

Smith, B. (2008). New desiderata for biomedical terminologies. In K. Munn & B. Smith (Eds.), *Applied ontology: An introduction* (pp. 83–107). Boston: De Gruyter.

Smith, B. (2009). Biometaphysics. In R. Le Poidevin, P. Simons, A. McGonigal, & R. Cameron (Eds.), *The Routledge companion to metaphysics* (pp. 537–544). London: Routledge.

Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., et al. (2007). The OBO foundry: Coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology, 25*(11), 1251–1255.

Smith, B., & Ceusters, W. (2006). HL7 RIM: An incoherent standard. *Studies in Health Technology and Informatics, 124,* 133–138.

Smith, B., & Ceusters, W. (2010). Ontological realism: A methodology for coordinated evolution of scientific ontologies. *Applied Ontology, 5*(3–4), 139–188.

Smith, B., Ceusters, W., Klagges, B., Köhler, J., Kumar, A., Lomax, J., et al. (2005). Relations in biomedical ontologies. *Genome Biology, 6*(5), R46.

Sojic, A., & Kutz, O. (2012). Open biomedical pluralism: Formalising knowledge about breast cancer phenotypes. *Journal of Biomedical Semantics, 3*(Suppl 2), S3.

Sterner, B. W. (2018). Review of Sabina Leonelli's data-centric biology: A philosophical study. *Philosophy of Science, 85*(3), 540–550.

Sterner, B. W., & Franz, N. M. (2017). Taxonomy for humans or computers? Cognitive pragmatics for big data. *Biological Theory, 12*(2), 99–111.

Sterner, B. W., & Lidgard, S. (2014). The normative structure of mathematization in systematic biology. *Studies in the History and Philosophy of Biological and Biomedical Sciences, 46,* 44–54.

Sterner, B. W., & Lidgard, S. (2018). Moving past the systematics wars. *Journal of the History of Biology, 51,* 31–67.

The Gene Ontology Consortium. (2017). Expansion of the gene ontology knowledgebase and resources. *Nucleic Acids Research, 45*(D1), D331–D338.

The OBO Foundry. (2018). Principle: Maintenance. *Obofoundry.org*. Retrieved September 22 from http://www.obofoundry.org/principles/fp-016-maintenance.html.

Vaidya, G., Lepage, D., & Guralnick, R. (2018). The tempo and mode of the taxonomic correction process: How taxonomists have corrected and recorrected North American bird species over the last 127 years. *PLoS ONE, 13*(4), e0195736. https://doi.org/10.1371/journal.pone.0195736.

Vogt, L. (2011). Signs and terminology: Science caught between language and perception. *Bionomina, 4,* 1–41.

Vogt, L. (2017). Assessing similarity: On homology, characters, and the need for a semantic approach to non-evolutionary comparative homology. *Cladistics, 33,* 513–539.

Wilkins, J. S. (2009). *Species: A history of the idea*. Berkeley, CA: University of California Press.

Witteveen, J. (2015). Naming and contingency: The type method of biological taxonomy. *Biology and Philosophy, 30,* 569–586.

Witteveen, J. (2016). Suppressing synonymy with a homonym: The emergence of the nomenclatural type concept in nineteenth century natural history. *Journal of the History of Biology, 49*(1), 135–189.

Witteveen, J. (2018). Objectivity, historicity, taxonomy. *Erkenntnis, 83*(3), 445–463.

Yoder, M. J., Mikó, I., Seltmann, K. C., Bertone, M. A., & Deans, A. R. (2010). A gross anatomy ontology for Hymenoptera. *PLoS ONE, 5*(12), e15991. https://doi.org/10.1371/journal.pone.0015991.