# 5 EPISTEMIC POWERS

> Every proposition proposing a fact must, in its complete analysis, propose the general character of the universe required for that fact. There are no self-sustained facts, floating in nonentity.
>
> —ALFRED WHITEHEAD
> *Process and Reality* (1985, 11)

## TAXONOMIC STRUCTURE AND INTERNAL INTEGRITY

Power is a structural concept, making power both pervasive and elusive. In Miranda Fricker's (2009, 4) *Epistemic Injustice: Power and the Ethics of Knowing*, she defines social power as the "socially situated capacity to control other's actions." As human-made artifacts, all classifications are, at their core, expressions of social power. That power, as we have seen, can be epistemic or material, beneficial or detrimental. The primary role of biodiversity taxonomists is to posit empirically informed, hypothesis-based arrangements of living entities in natural world. They take expressed nomenclature and its associated taxon concept and propose unified taxonomic systems comprising species and their relationships. This is, at first glance, functionally different from the purpose of bibliographical or documentary systems, which are touted more as a means to facilitate retrieval than as a representation of the universe of knowledge *as it should be*. Indeed, bibliographical classifications, unlike those in biodiversity, are pragmatic, based more on the social organization of knowledge production (Bliss 1929). However, even if unintended, these practical bibliographical classifications have derivative impacts on how we value some forms of knowledge over others—and this has a great deal of epistemic influence. All classifications help us make sense of a complex world. Classification systems like Google's search engine, for example, take an incomprehensible volume of information and deliver it to us in ways that are digestible and, ideally, useful. By participating with this system, however, we begin to see and interpret our online world through the narrow and opaque lenses of algorithmic mediation.

In this way, classifications have an instrumental function, in line with Patrick Wilson's discussion of "bibliographical instruments" (Wilson 1968, chap. 4). Wilson believed that, to understand how much control we have within a bibliographical system, we must understand how it is logically arranged. Wilson states, "I cannot tell how much bibliographical control I have or could have simply by introspection, by memory of past success and failures, or by flexing my muscles. To discover what I can or might do if I would, I must discover what arrangements there are of which I can take advantage, what bibliographical instruments . . . are at my disposal" (Wilson 1968, 55–57). We must understand, for example, what it means that a subject inhabits one position in the instrument over some other position. And by extension, the better we understand an instrument's structure, the more we can understand the inferences that can be drawn from it. This chapter, then, focuses on these systemic aspects of classifications. I further argue that a classification's structure necessarily forms reduced and univocal interpretations of the world. I discuss classifications as exerting structural power in three senses: as mechanisms that (1) express which entities do and do not merit annunciation in the world; (2) obfuscate the inherent complexity present in nature; and (3) present nature in a univocal fashion, at the expense of alternative forms of knowledge.

Let us imagine a situation in which two taxonomists are given a "pile" of flash cards that represented a suite of taxon concepts and are asked to classify them in accordance with their own taxonomic commitments. Their first operation will be to take these taxon concepts, group them into species groups, and articulate relationships among them according to their preferred theoretical approaches to classification. For example, one might relate species by way of phylogenetics, whereas the other might choose to use phenetic approaches. Let us also further assume that we ask both these scientists to employ the same approach to constructing relationships—for example, doing so based on overall similarity using morphological approaches in accordance with phenetics. The chances that the two resulting taxonomies will be exactly the same are incredibly small, if nil, even if they both approach classification using the same theoretical foundations. One may lump more than split, thus creating a taxonomy with fewer overall species, while the other may mark small morphological differences as reason enough to merit a new species taxon, resulting in many species branches. This diversity of opinion is a natural part of the taxonomic landscape and the reason such taxonomies are said to be hypotheses. As stated by Clare Beghtol, "every classification system is a theoretical construct imposed on 'reality'" (2001, 99). And reality, as we know, is relative.

One rule of biological taxonomies (traditional ones, in any case) is that they must maintain a certain degree of internal integrity and consistency with regard to how they represent the external world. Taxonomies are only credible, after all, to the extent that their classifications maintain their commitments throughout their structure, in accordance with user's assumptions about how reality is composed. That the Catalogue of Life does not display this internal coherence is a prime criticism of its composition. As identified by Furner (2009b, 9), how we judge the success of a knowledge system has been a question of great concern to those seeking to classify knowledge, particularly in relation to how effectively these systems represent the external world. Taxonomies should be coherent, in that the rules of arrangement should be invoked uniformly within the classification. When a given taxonomy is consistent, a user can understand the logic behind how and why a given species is at a given location. Some believe that another hallmark of good classifications is that they should be simple and elegant to the extent that their taxonomic hierarchies are parsimonious and describe "the evolution of any particular set of characters using the smallest number of evolutionary changes" (Wiley and Lieberman 2011, chap. 6).

Yet we know that taxon concepts and taxonomic classifications change and are challenged over time. Taxonomic knowledge systems are products of contextually specific historical, cultural, and philosophical circumstances that evolve. Understanding that all classifications are unstable and subjective, even and especially within the natural sciences, helps me illustrate one of the general limits of universal classification systems: they are never truly correct or final. Universal systems, however constructed, are never truly universal. Further, the subjects we embed within them are not real, or definite, yet we continually treat them as if they are. And this truth applies to species: species taxa are not definite concepts, yet people outside of the biodiversity sciences continually speak about them as if they are statements of fact. As Richard Rorty wrote, "The desire for a theory of knowledge is a desire for constraint—a desire to find 'foundations' to which one might cling, frameworks beyond which one must not stray, objects which impose themselves, representations which cannot be gainsaid." (2009, 315). Giving structure to the abstract is one of the primary purposes of any classification—a theory for the presumed order of knowledge of natural biodiversity—biological or otherwise. Nature is, ultimately, abstract, and we necessarily create linguistic metaphors to make sense of it.

## CONTINUITY AND CONTINGENT EXISTENCE

In light of this abstraction, let us now switch our focus to the backdrop, so to speak, of biological classification. I want to expand on a rather large and amorphous concept: *nature*, and why I think it is important for me to assert, at least when discussing classification matters, that nature does not have natural joints, and that classifications have the often-negative popular effect of making it seem as if they do. If the experience of doing fieldwork with biodiversity scientists taught me anything, it was that saying as much to a room full of highly educated and accomplished individuals set the groundwork for some fiery (but lighthearted) repartee. And this is understandable; to say that there are no true, individual entities in nature and that, if we say there are, they are social and not natural, might sound like a slight against the informed, elegant work of taxonomists. Surely their work is incontrovertibly valid, and increasingly more essential, to the study of our natural ecosystems, especially given the challenges we face with a changing climate and rampant extinction. So, let me first establish what I do not mean. I do not mean that there are not, in fact, distinguishable natural entities in the world that are potentially categorizable as distinct. My argument is not an ontological one, questioning the existence of perceptively discrete concrete objects as possible entities that exist in the natural world. To my mind, there are organisms that scientists try to describe in this world—animals, plants, bacteria, and the like—and whether I am a materialist or an idealist makes no difference to my central argument, which is that nature does not produce simple classifications of organisms for us. Surely, when I was camping a few years back, when a pair of grizzly bear cubs rose to their hind limbs in front of me, the first question that came to mind was not, Are they (and their likely nearby parents) real? A grizzly is surely concrete enough that I'll back away slowly regardless of my ontological inclination. My assertion, of course, is much more nuanced than that. My argument is not that biological classification is wholly arbitrary, but rather quite the opposite: to be a taxonomist is to be carefully trained in method and to make informed decisions about what elements in nature *matter most* to distinguishing one living organism from another.

Species taxon concepts, however, *are* contested. As Kriti Sharma states, "No one expects a flower to be unitary—we know that some people will define the flower as a bloom atop a stem, and others will define it is as the whole plant down to its finest roots" (2015, 13). Nature, from the point of view of classification work, is most usefully defined as a process, and if there are any joints to be found, they are merely temporary. At the heart of this argument is, to a certain extent, an organicist approach to biological entities—a process-oriented philosophy that believes process, not things, to be, if not the most ontologically fundamental category, then certainly one of the primary characteristics of the natural world (Nicholson and Dupré 2018, 12–13). The view of process thinking is that life can be viewed as one process, or rather as a series of processes that mutually constitute one another at varying rates of transition. "A process-based metaphysics," Argyris Arnellos writes, "implies that living systems (such as cells, multicellular systems, ecosystems, organisms) are just temporary phenomena—the products of the dynamics of some processes" (2018, 200). Nature, as it were, is a set of processes, interconnections, and moments of transfer. Alfred Whitehead (1920, 13, 65) calls this "complex of nature" an undifferentiated fact that becomes a "entity for thought" that is defined, in part, by the innumerable interconnections between a string of "extensionless instants" perceived by a given viewer. Whitehead's philosophy presumes that objects, insofar as they are perceived, are *events* that are successively realized over time (1920, 126). Objects-as-events, then, are historically or genealogically connected perceptions that are effectively different each time they are experienced. The plant we see in our living room one day is temporally distinct, and thus a different event-object than the one previously perceived.

In *Interdependence: Biology and Beyond* (2015), Kriti Sharma argues for a systemic model of nature that highlights a dynamic, mutually constituted notion of the natural-world-as-entity within the biological sciences. Sharma describes two shifts that define this line of thought:

> The first is a shift from considering things in isolation to considering things in interaction. This is an important and nontrivial move. . . . To get to a thoroughgoing view of interdependence, I argue that a second shift is required: one from considering things in interaction to considering things as *mutually constituted*, that is, viewing things as existing at all only due to their dependence on other things. (2015, 2; emphasis original)

Sharma's approach is enticing, particularly because they make clear our perception itself is entangled in this mutual constitution and it is our perception that necessarily creates classes to make sense of a radically dependent space. This is not to say, as Sharma notes, that objects cease to exist in the world if we do not see them—after all, the grizzly bear that chases me away does not fail to be a grizzly bear (nor to be dangerous!) once I (or all people) flee to a safe distance. But, at the same time, without me, or anyone else as an observer, there is indeed nobody there to actively characterize that bear as being bear-like, brown, and in possession of skin-piercing, potentially murderous claws. "Only observers can perform the various actions necessary for experiencing phenomena as objects," states Sharma (2015, 22). We adhere attributes onto objects through naming and language. Objects require subjects to instantiate them within an ontology of other objects and relations.[1] To be the object, *tree*, for example, means that the object must fit with our notions about what we believe to be a tree and be defined in contradistinction to a host of other objects that do not express the tree-like properties of a tree (bushes, weeds, grasses, and the like). One result of this analytic approach is that, while objects of "nature" can exist independent of the subjects that encounter them, once we sense them, classify them, and name them, they become socialized and articulated within our instruments of knowledge.

The problematics that such a processual view might have on the definition of species are quite obvious: if, indeed, everything is mutually contingent and event-based, how do we get to the point where concrete objects of any kind are distinct and locatable in time and space? In thinking about the limits of the body, Whitehead stated, "It is just as much part of nature as anything else there—a river, or a mountain, or a cloud. Also, if we are to be fussily exact, we cannot define where a body begins and where external nature ends" (1938, 21). A conundrum, indeed. John Dupré (Dupré 1993, 2014; Nicholson and Dupré 2018) has been a leading thinker in this area. "Many processes are bona fide individuals—they are concrete, countable, and persistent units. . . . In biology, processes are . . . dynamically stabilized at vastly different timescales: a matter of minutes for a messenger RNA molecule, a few months for a red blood cell, many decades for a human being, and up to several millennia for a giant sequoia tree" (Nicholson and Dupré 2018, 12). Temporal scales are an essential component in process theory, given that processes are not uniform in their velocity or rate of change. As humans, we are able to see the full systemic life cycle of a red blood cell or an RNA molecule. But witnessing a species evolve from one to another is an entirely different matter. Evolution, for one, is a property not of an individual member of a species but of populations and, as such, works slowly over the course of time. Watching evolution happen to a population is an impossible ask of any scientist. Thus, scientists look for other forms of evidence to help measure difference between genealogical lineages (phenetically, genetically, and so on) and infer evolutionary changes. An entity can appear concrete by way of an equilibrium of many intermingling dynamic systems functioning at a speed that exceeds the perceptual abilities of the viewing subject. A distinguishable eddy in a river, for example, is one concrete object-system that flows harmoniously as part of a larger river. As is an organism that lives within a larger ecosystem among many other seemingly distinct species.

The obvious reality is that we *do* sense and experience entities, and this is not antithetical to process-theoretical accounts. Dupré advocates for what he calls "promiscuous individualism," given that "there are various ways of drawing such boundaries, reflecting real biologically salient aspects of the multiply interconnected systems that make up the biological world" (2014, 241). The approach accepts that there can be many equally valid mechanisms to divide entities from one another. Depending on one's methodological approach, some systemic interactions appear in the fore- or background of an analysis. But, make no mistake, the dividing line between one cohesive entity-system and another is, indeed, blurry. An individual, then, in Arnellos's assessment, reflects the fact that "some [systems] work together so that they constitute *cohesive systems*, that is, systems that persist and manifest in a *form of stability* in the sense of a spatiotemporal integrity" (2018, 203; emphasis original). The individual can be described as a self-enclosed, constitutive system, constrained in such a way that matter and energy can be maintained, while still remaining open enough to engage with the external systemic environment (2018, 206).

If we were to adopt Kriti Sharma's (2015, 12) terminology, we'd say that things as they exist in our conception of the world—including objects such as organisms—exhibit themselves through "contingence existence." To exist contingently is not to exist individually and inherently outside of any given process, but rather to recursively depend on other equally contingent objects and subjects for being (2015, 15). So, like Whitehead and Dupré, Sharma sees the environment as mutually constituted by a series of ongoing object-object and object-subject relationships. Nothing exists on its own inherently, and as such, when we classify and categorize, it is imperative to understand the ecological conditions under which these entities came to be. This is an imperative concept to understand within the domain of information studies, primarily because it forces us to see that our classifications (of books, of documents, of species, of people, and so on) are entirely contingent not only on the internal, intellectual context of the classifier, but also on the emergent and inherently fuzzy boundaries between entities and concepts in the external world.

Such a view forces us to examine intersections of beings, identities, and qualities in light of their subjective and relational qualities. We *should* question the application of subjects—vis-à-vis Melissa Adler and Hope Olson's work—because any subject, or any combination of subjects, we apply to a document can never comprehensively describe it adequately. In addition, a much more ecologically and biologically focused driver is that we have to begin the long, complex process of reorienting our relationship to biological objects, the natural world, and the affective impact our classifications have on the entities that are increasingly vulnerable in the world.

## REDUCTION AND OUR MENTAL REALITY

Classifications extract entities from a natural world that is otherwise fluid, processual, and undifferentiated. Taxonomists stabilize the attributes of organisms and then use these attributes to distinguish one taxon from another. This activity reduces the noise of nature and distills it into individuated entities. This allows us to socialize the natural world by incorporating it into the linguistic and scientific discourses. Once the work of atomizing nature into discrete species taxa is complete, scientists must then follow up this task with the process of reconstructing those very entities back into an organic whole—into a classification.

Taxonomists make decisions about what goes into a classification system, where it is to be positioned, and what is to be left out, beyond the methodological boundaries of the system itself. Do we lump giraffes into one species, or do we break them up into four distinct species based on DNA testing (Fennessy et al. 2016)? The former conceals the possible existence of three others, while the latter makes it seem as though nature is somehow more wide-ranging in terms of its diversity. Reduction is an important concept in the process of building classifications, and the politics of this selection process is not to be undervalued. In *Counting Species: Biodiversity in Global and Environmental Politics* (2015b, 47), Rafi Youatt looks at the political history of biodiversity from the 1980s through the 2010s, focusing on how the concept of biodiversity has been part and parcel of, and perhaps primarily, a form of global political practice. "I take the position," Youatt notes in relation to bio-census activities, "that the global biodiversity census is as much about power and political life and the boundaries between nature and society as it is about scientific information gathering for conservation ends. . . . Here the focus is on considering the field of social power in which scientific efforts take place, and asking questions about the discourses, resources, and networks that make a biodiversity census possible" (2015b, 47–48). Yet, when people encounter biodiversity classification systems, they often experience them, in practice, as if they are complete.

Our typical understanding of biodiversity is often visual—diagrams and tree-relationships, for example, dominate the domain. Imagine being asked the question, How are dingoes related to domesticated dogs? Even if you know the answer, chances are, at some point, a tree diagram will have become one way you came to understand this relationship. And surely a schematic would need to be made to communicate your vision of the relationship with some other individual. We have been trained to use these representations over the course of hundreds of years. The Porphyrian tree is one of the oldest known tree diagrams, interpreted many times over back to the third century AD. The Porphyrian tree expresses the relationship between Aristotelean genus and species, along with branches diagramming his famous dichotomous divisions (Aristoteles 1995, pt. History of Animals, Parts of Animals, Categories; Lima 2014, 28).

From my early days learning about natural history in elementary school, the "tree of life" image was a mainstay in understanding how species are interconnected and related. Dinosaurs were (and still, in part, continue to be) my jam. I'd labor for hours over different taxonomic charts. Relationship trees are aesthetically pleasing and easy to memorize. But, even before trees, encyclopedias were compiled to make the natural world portable. Pliny, for example, catalogued the known natural world in his *Naturalis Historia* (an original manuscript does not exist, but see 1472 for an example). Regardless of whether a catalogued entity was "real" in any traditional sense (the sciapod category, for example, was a mythical creature included in his original manuscript), Pliny's text allowed the general public the opportunity to peer into distant worlds—he made the distant intimate; the unknown, known, as an element into social consciousness. Linnaeus, similarly, created a nomenclatural pseudo-classification system that became part and parcel of what it means to broadly classify the natural world at all. To have the name *Canis lupus dingo* locates the dingo in the *Canis* genus, the *C. lupus* species, and *C. l. dingo* subspecies.

Originally, Linnaeus's naming system was intended as a technology to facilitate memory and, eventually, a collocating and retrieval mechanism for cogent and easy access to species identification and knowledge aggregation. The resultant nomenclatural system became part of the public imagination, as did the graphical structures that became part and parcel of the depiction of the natural world. These structural metaphors (Lakoff and Johnson 2003, 152) of the natural world are powerful and integrally shape our aesthetic and positional relationship to the natural world. Trees of life are sharply delineated and serve as a mechanism for wresting definition from a landscape defined by complexity. As Lakoff and Johnson note, "Metaphors have entailments through which they highlight and make coherent certain aspects of our experience" (2003, 156).

Classifications do not define reality, however, even if our particular understanding of the "real world" is defined by those same classifications. In fact, as many scholars have noted, the visual mechanics of classifications do influence and define how we negotiate and interpret our social world. All models, interfaces, visualizations, and diagrams—classifications and trees included—interpret underlying data. The resultant images make given conclusions seem intrinsic to the data and, thus, inherent to the world from which they were extracted. As Johanna Drucker notes, such visuals are performative in some respect; she calls them "acts of interpretation masquerading as presentation" (2014b, 10). The allure of the graphical representations of hierarchies has deep roots in the history of the human communication of knowledge (Lima 2014), and their power in the context of everyday settings should not and cannot be overlooked.

Even more insidious and pervasive are the classifications that go unseen amid the daily practices that structure some of our most important resources of general information and knowledge. Technological advances, especially those associated with internet and web environments, often function with the assistance of dozens of overlapping algorithmic classification systems. Classifications are now flattened, fragmented, and distributed throughout complex databases that shape the way we navigate our world. But because these classifications are not easily accessible and graspable through simplistic graphics, they are often not experienced or identified as proper classifications. Yet, they are, and the result is that, while these classifications still individuate and reduce a complex world, we often overlook their effects. Google is a prime example of a tool that uses algorithmic classification to rank, order, and provide access to the countless documents and webpages available on the web. Safiya Noble's (2018) work is paramount here, in that it lifts the veil on the seeming objectiveness of the platform. Noble shows how Google's search results have propagated racist ideology by "technologically redlining" and exacerbating the already present modes of racism prevalent in our society. Noble notes how Black identities, particularly young girls and women, are frequently sexualized and fetishized in online search environments, perpetuating the culture of harm and violence that these same communities experience in social spheres. Technology is nothing if not essentially human by design—design that is impressed by the makers' biases and inclinations.

Such misconceptions stem, in part, from the inevitable and necessary process of reduction. One danger of scientific classifications is that the discourses using them as evidence often assume that they represent nature—if not in full, at least in part. That they are partial and designed is rarely discussed, at least outside the biodiversity world where this reality is an obvious point. Henri Lefebvre writes,

> Reduction is a scientific procedure designed to deal with the complexity and chaos of brute observations. This kind of simplification is necessary at first, but it must be quickly followed by the gradual restoration of what has thus been temporarily set aside for the sake of analysis. Otherwise a methodological necessity may become a servitude, and the legitimate operation of reduction may be transformed into the abuse of *reductionism*. No method can obviate it, for it is latent in every method. Though indispensable, all reductive procedures are also traps. (2011, 105–106; emphasis original)

A trap, indeed, and an inevitable one. Lefebvre's point, in part, is that reduction has the potential to then be subsumed into our mental reality, and thus representations infiltrate the space of daily practice and the imaginative realm. A key to understanding a classification is to understand the methods, rules, and policies that dictate its ultimate form. The danger of the reductionism described by Lefebvre is of two kinds: on the one hand, the scientist can fall prey to the functionality and ease of their reductionist approach such that they become comfortable "curling-up happily" in the warmness of their niche (2011, 107). This, to me, is less dangerous in practice, given that the biodiversity specialists I encountered in my fieldwork were anything but uncritical. They were each well aware of their methodological approach, its pitfalls, and the dangers of artifacts that simplify complex and nuanced work.

This reflexive, critical attitude no doubt emerges in part from the longstanding debates between different taxonomic "factions." On the other hand, the more dangerous effect of reductionism, from Lefebvre's point of view, stems from the fact that the reflexivity exhibited by the biodiversity scientists is not typically practiced outside of the domain. So, whereas biodiversity scientists are fully aware of the limitations of classifications, nonspecialists do not have the expertise to negotiate these structures effectively and to critically understand their pitfalls and nuances. Taxonomic structures, such as they are graphically delivered, simplify and obfuscate the conflict between taxonomic approaches that is readily apparent to those within the discipline. In this way Lefebvre sees this kind of reductionism as a "tool in the service of the state and of power" (2011, 106). Individuals proceed as if these graphical statements are statements of fact and not as if they are hypothetical structures that merit careful scientific consideration.

## UNIVERSALITY AND THE MYTH OF TOTALITY

One cannot speak of the reductionary qualities of classifications without also broaching the concept of universality. When we speak of universality in relation to classification, we often mean "this classification organizes, or can organize, everything." Yet, what actually defines universal structures, in practice, is that they simplify complex sets of ideas and express them in a univocal fashion. In his *Quiddities* (1989), Quine, reflecting on the notion of the "Universal Library," highlights the inherent contradiction and absurdity of this "melancholy fantasy." As illustrated in Borges's "Library of Babel" (1999, 112–117), the infinite space of the library is untenable and maddening, where rationality "is an almost miraculous exception." With no comprehensive library index, despondent librarians are useless and unable to unlock sense from this senseless infinity of letters, words, and text. In a move fitting of Ockham, Quine whittles down the infinite to a similarly confounding sensibility, that all this complexity can emerge from the finiteness of Morse code: "The ultimate absurdity is now staring us in the face: a universal library of two volumes, one containing a single dot and the other a dash. . . . The miracle of the finite but universal library is a mere inflation of the miracle of binary notation: everything worth saying, and everything else as well, can be said with two characters" (1989, 42–43). And therein lies the ever-problematic notion of claiming universality: to be truly universal, one needs to represent all knowledge, and yet, in practice, this is an impractical possibility. The only solution is to fall back to the (seemingly) infinite possibilities of language.

For one, if classifications, by definition, reduce the natural world to a series of perceptible and documented tokens (linguistic or otherwise), then how can we *also* purport that classifications are also universal? Using the term universal and speaking of its discursive use in both information studies and biodiversity could cause some confusion, so some clarification is in order. Within the realm of information studies, universal classification systems are those general systems that are able to expand and adapt to the entirety of extent knowledge—that is, systems that are able to classify documents in any and all disciplines. For example, systems such as the Dewey Decimal Classification and the Library of Congress are created in a way that all disciplinary literature can fit within their established scheme, and as such, they are universally capable of expanding and absorbing concepts within its scheme. These types of classifications are often associated with the modernist turn in the information sciences (Mai 2011).

On the face of it, this is a virtuous goal. As recently discussed by scholars in the subdiscipline of knowledge organization (Dahlberg 2017; Gnoli and Szostak 2014), to be universal is to enhance cross-disciplinary conversations and support system interoperability. Some of these same theorists have posited that a universal method based on nature's fundamental properties can more accurately organize knowledge from multiple domains. "The theory of integrative levels claims that the natural world is organized in a series of levels of increasing complexity: from physical particles and molecules, through biological structures, to the most sophisticated products of human thought" (ISKO Italia 2004) (see also, Gnoli and Ridi 2014; Gnoli and Poli 2004; Szostak 2008). This adherence to a static understanding of external phenomena (consistent enough, in practice, that the organization of *all* knowledge can perpetually be conformed to this schematic) is of little use, however, in domains as specific as biodiversity taxonomy, where the "reality" in question is a shifting ground of concepts and taxonomic arrangements. These reductionist approaches not only presuppose a permanence and consistency to our general knowledge that just does not exist, they also overlook the socially situated, culturally defined unfolding of our knowledge production practices.

As we look to biodiversity systems such as the Catalogue of Life, these consensus structures are implemented to facilitate cross-group data sharing and data aggregation from disparate sources. Yet, as demonstrated, a classification not only *organizes* information, it also *creates* very distinct epistemic boundaries that proliferate throughout society in intended and unintended ways. As such, discussions regarding what universality means must attend to not only what it facilitates, in terms of discovery, but also what it inhibits and epistemically distributes via the epistemic commitments that underlie its universal modes of aggregation.

In the biodiversity sciences, universality is used similarly, but also differently, in certain respects. Of course, unlike general documentary classifications systems, the practical function of most biodiversity taxonomies (say, for example, a taxonomy covering all weevils) is not to represent all knowledge, but rather to represent an organism-relationship model that attends to the particular theoretical and epistemic orientations of one scientist or group as it applies to a finite set of entities or concepts. In theory, a taxonomic approach *can* be used to organize all species. But taxonomists specialize and function on small scales. This is partly why consensus classifications have grown so contentious—it is a different matter to organize *everything*, since no one person can accomplish this to any effective degree. This falls in line with Hjørland and Albrechtsen's (1995) articulation of the domain classifications, an approach that Hjørland has vehemently supported for years (2009). Within the boundaries of these species-specific classifications, what is universal is not, necessarily, that the classification itself can include all species. What is emphasized as universal is the particular methodological approach to articulating species and their relationships. If a scientist uses a cladistic approach, for example, to organize weevils, that same approach can be used to examine any segment of the animal kingdom. But, even more than that, when you listen to the discourse of biodiversity scientists, to be universal actually means to have a classification method that can be of use to all scientists as a mechanism of construction. In some spaces, the assumption behind a *truly* universal system is that it represents a "natural" organizational principle that can be used as a baseline against other methodological approaches (Doolittle 1999b).

As Mai has noted, the impulse of "modern" classification systems to standardize and universalize knowledge organization systems around notions of "exclusiveness" and "exhaustivity" overlooks the application of standards in individual domains (1999, 548–551). In recent years, a more historically informed notion of classification has taken shape, one in which "unificationism" has given way to a "generation of theories, principles and methods that emphasize both the cultural and historical specificity of classification practices and their emancipatory function" (Furner 2013, 32)—a topic we attend later in this text.

## INTERNAL EPISTEMIC PRACTICES

Expanding this, what does it mean to "uniformly implement a method" in a taxonomic space? In short, most taxonomies are built along a set of uniform rules. By following a rule throughout the taxonomic space, a taxonomist can validly argue for one particular arrangement over any other. Understanding these rules is precisely what Patrick Wilson was referring to when speaking of bibliographical systems as instruments. And, this instrumentation is important to understand, especially, with respect to the Catalogue, how we can critique its shortcomings. Along with every subsidiary taxonomy the Catalogue collects comes a particular taxonomic structure that represents the opinion of the experts that created it. "The source databases are diverse in their origin, their purpose and therefore their structure. A key challenge for the Catalogue of Life has been the integration of this disparate data, and a standard dataset has been established for that purpose" (Species 2000 2015c). The distinctions between one condition and another represent deep-seated theoretical and philosophical divides about how relationships can and should be built within classification systems, and how species are lumped into nested and hierarchical taxa.

An in-depth explanation of biological taxonomy is outside the scope of this book, but I can broadly describe some examples to show how different schools of thought go about creating taxa and connecting them in networked relationships. Marc Ereshefsky identifies evolutionary taxonomy, cladism, and pheneticism as three possible ways to construct a biological taxonomy (2007, 50–51). Evolutionary taxonomists believe that the emergence of new species taxa can occur through two distinct processes: cladogenesis and anagenesis. Cladogenesis is the splitting (branching) of a "single genealogical lineage" (2007, 52), through, for example, the process of occupying new adaptive zones (a population of a species becomes geographically isolated from the rest of the population and adapts with new genetic characteristics) (SæTher 1979, 308–309; Ereshefsky 2007). Anagenesis, on the other hand, is the gradual change (divergence) (Ereshefsky 2007, 52) over time of a species until it becomes distinguishable as a new species. The mechanisms and qualities used by taxonomists to assess "significant" (Ereshefsky 2007, 52) enough changes to warrant a new species for anagenic change is a subjective process (Vaux, Trewick, and Morgan-Richards 2016) and a source of much debate in the taxonomic arena. Cladogenesis produces monophyletic taxa: taxa that include an ancestral species and all of its descendants (Grant 2003). Anagenesis, however, produces paraphyletic taxon groups: taxa that contain some but not all of the descendants of a particular taxon (see figure 5.1).

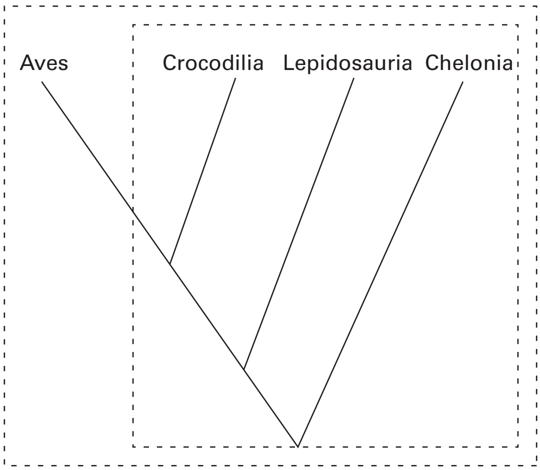


**Figure 5.1**
Example of monophyletic (outer box) and paraphyletic groups (inner box). Evolutionary taxonomists would define the class Reptilia as containing only lizards, snakes, and crocodiles, excluding Aves as part of this schematic; thus, it is paraphyletic. Cladists, on the other hand, would be unwilling to exclude Aves, because groups should include all the descents of a taxon, and thus a Reptilia group that includes Aves is monophyletic. Figure adapted from Marc Ereshefsky's *Poverty of the Linnaean Hierarchy: A Philosophical Study of Biological Taxonomy* (2007, 54) and "The Reptile Database" (Uetz 2016).

One might look at figure 5.1 and ask why this seemingly minute distinction matters. After all, the tree is essentially the same, regardless of whether you understand the class Reptilia to contain birds or not. What differs is the way each school *interprets* this diagram to represent some argument about how animal groups form taxa. However, while the differences seem subtle in figure 5.1, stating that these taxa occupy different class *positions* means they occupy vastly different spaces in the nested taxonomic hierarchy of the Catalogue database. In figure 5.2, an evolutionary point of view (according to Ereshefsky's model) situates birds (Aves) as a different class. Theoretical and interpretive distinctions amplify themselves in online taxonomic structures.

Cladistics, in contrast to evolutionary taxonomy, does not accept paraphyly as part of their construction of taxa.[2] Cladists would balk at separating Aves from the Reptilia group since cladists base the construction of their taxa on the concept of shared traits and common ancestry (Ereshefsky

- kingdom: Animalia
  - phylum: Acanthocephala Rudolphi, 1802
  - phylum: Annelida
  - phylum: Arthropoda
  - phylum: Brachiopoda
  - phylum: Bryozoa
  - phylum: Cephalorhyncha
  - phylum: Chaetognatha
  - phylum: Chordata
    - class: Actinopterygii
    - class: Amphibia
    - class: Appendicularia
    - class: Ascidiacea
    - class: Aves
    - class: Cephalaspidomophi
    - class: Elasmobranchii
    - class: Holocephali
    - class: Leptocardii
    - class: Mammalia Linnaeus, 1758
    - class: Myxini
    - class: Reptilia
    - class: Sarcopterygii
    - class: Thaliacea
  - phylum: Cnidaria
  - phylum: Ctenophora
  - phylum: Dicyemida
  - phylum: Echinodermata

**Figure 5.2**

Catalogue of Life 2016 Annual Checklist taxonomic tree depicting the separate placement of the class Aves from class Reptilia in the tree structure (Species 2000 2020). CC-BY 4.0, Catalog of Life, used by permission.

2007, 55). Cladists reject the decision to separate birds from the Reptilia group because there is no hard and fast way to delineate when one species split off into another.

Phenetics, our final approach under discussion, produces a distinctly different representational hierarchical arrangement. Numerical taxonomy, or phenetics, is a probabilistic method that groups organisms together based on the general premise that those with the most phenotypic overlap are necessarily more closely related. As articulated by Robert Sokal,

> Numerical taxonomists contend that evolutionary importance is undefinable and generally unknown and that no consistent scheme for weighting characters before undertaking a classification has yet been proposed. To weight characters on the basis of their ability to distinguish groups in a classification . . . is a logical fallacy. Since the purpose of employing the characters is to establish a classification, one cannot first assume what these classes are and then use them to measure the diagnostic weight of a character. (1966, 109)

Numerical taxonomy was touted as the apex of empirical science when it burst onto the scene, particularly because such analysis, in theory, would always produce the same result from laboratory to laboratory over time (Sneath and Sokal 1973, 11). Codes were created for each particular object trait: "'hairiness of a leaf' might be coded as follows: hairless: 0, regularly haired: 2, densely haired: 3" (Sokal 1966, 114). Of course, as is the case with any method deemed empirical, one must critically assess the choice of variables or characteristics that undergo analysis for clustering. A notable weakness of numerical taxonomy is the fact that a relationship between organisms is defined solely by "similarity . . . operating on the assumption that the total phenotype accurately effects genotype. [Numerical systematists] believe that an unweighted measure of overall similarity provides an accurate determination of relationship. In so doing, pheneticists ignore the possibility of evolutionary convergence" (a circumstance where unrelated organisms evolve similar traits because of environmental influence) (Pietsch 2015).

Every classification is but one way to represent a phylogeny that is altogether too complex to be represented in any one graphical structure (Hull 2001, 227)—there are far too many variables (known and unknown) involved in the process of evolution. As D. L. Hull states, "Any one phylogeny can be classified legitimately in many different ways. . . . Only the most generic and impressionistic inferences about phylogeny can be drawn from an evolutionary classification" (2001, 227). Regardless of which method of taxonomic arrangement prevails as dominant in our contemporary climate, elements of these schools demarcate sharp differences in the way classifications are constructed and interpreted as modes of argumentation. What matters most is that each classification produced by any of these methods, or any other, must adhere to ontologically and epistemically informed rules; these metaphysical and epistemic commitments then permeate the entire taxonomic structure.

The Catalogue, however, unlike these description-based approaches, is not concerned with any one approach to producing taxonomic knowledge. Each section of the Catalogue's management classification might adhere to different commitments depending on the source of the taxa in question. For example, the editorial group for the management classification of fungi might choose one "master" taxonomic approach, whereas another authoritative approach may be implemented for birds. In other words, the Catalogue is not uniformly internally consistent, which is a major issue for many in the taxonomic field.

## RETRIEVAL OR DESCRIPTION?

A pertinent question here is, What are biodiversity taxonomies attempting to accomplish, particularly in light of the uniform application of method to any given instrument? That is, what is their function from the standpoint of *use*? Jonathan Furner proposes one way that we can potentially assess classification systems:

> I think it is possible to distinguish two conceptions of the goal of the practice of KO [knowledge organization], and this distinction corresponds roughly to the one Raya Fidel draws between two conceptions of the goal of indexing. On the one hand, we have the document-centered view that indexers should aim to assign index terms to documents (or documents to index terms) in whichever way it is that produces the most accurate representation of that content. On the other hand, there is the user-centered view that indexers should aim to associate documents with those terms that are most likely to be used by searchers looking for those documents. Similarly, I think that, on the one hand, we have a description-oriented conception of the goal of KO, being to build systems that do well at helping people produce accurate descriptions and representations of documents. And on the other hand, we have a retrieval-oriented conception of the goal of KO, being to build systems that do well at helping people find the documents they think they want to find. (2009b, 9)

This distinction between description- and retrieval-oriented approaches to classification seems to me a possible way to categorize the two basic kinds of taxonomies that currently flourish within the biodiversity world. As Tony Rees, manager of the Divisional Data Centre, CSIRO Marine and Atmospheric Research, in Tasmania, articulated this on the Taxacom biodiversity listserv: "I look upon biological classifications as serving two purposes—first, to illustrate our current best guess/es as to the relationships between organisms, and second, to provide a recognisable navigation structure so that persons entering the classification can (hopefully) find their way to their particular organisms of interest" (2009). This quote brings to light the two distinct aspects of the biodiversity instrument. For Furner, these two approaches are rooted in the larger question about how one is to *evaluate* classification systems as systems that "represent relationships of identity between classes of documents," and "help people find the right labels for classes of documents that about those identities, and help people find those documents" (2009b, 4).

Although there exists no standard by which the true "goodness" of a classification system can be quantitatively assessed, a number of variables can be identified to critique the effectiveness of classification schemes. Furner indicates that the basic role of a classification system is to adequately represent the identity of some external reality: "The main claim that I would like to make about the importance of identity for KO is not that an understanding of identity is helpful in analyzing the structure of aboutness and relevance. It is that there is a sense in which identity is actually the goal of KO" (2009b, 12). The rubric for assessing whether classification systems "work well" can be distilled to a series of factors, from within either a "description-oriented" or "retrieval-oriented" notion of classification (2009b, 9–10).

According to Furner, description-oriented classifications ask two basic questions of a designed system: (1) How correct, just, and fair, is any given ontology in relation to the natural, real world? and (2) How internally coherent is the infrastructure itself in exemplifying a unified ontological system with an internal logic? (2009b, 9). In the biodiversity realm, this means creating classifications that provide a consistent model that includes a fair and accurate representation of biological organisms, as well as one that provides a classificatory system that depicts things the "the way things really are, or the way somebody thinks things are" (2019, 9). Traditional, internally consistent taxonomies can be defined as functioning in this way; they are systems that invoke a unified methodology throughout, and thus, provide a consistent model of the natural world. Because descriptive systems are consistent, a species location in a system tells you about how it operates in relation to all other entities within that system. In terms of the Catalogue, we might call most of the contributed taxonomies descriptive. This was certainly the case for our previously discussed WeevilBase, for example, as it was brought into various global infrastructure. A standing question that remains in this case, however, is how we define "reality" as a goal of a descriptive system when nature is characterized by change. Entities evolve, and so our classifications will always be catching up to this moving target.

Retrieval-oriented classifications, on the other hand, are evaluated by way of a classification's ability to facilitate the locating of documents or required information by a user (2009b, 10). These systems, in theory, aim to model classifications in ways that are expected, or understood, by users. Terms Furner associates with this level of assessment include effectiveness, efficiency, and usefulness to the user. In the biodiversity world, this would mean creating a system that has the ability to adjust to meet the expectation of many kinds of users and points of view: "different people see reality in different ways," says Furner (2009b, 9). Such a view prioritizes variables other than description as the core role of a system. The Catalogue, for example, is attempting to use consensus to find an epistemic middle ground that meets the needs of many user expectations and applications. It is thus more pragmatic. Evaluating from a retrieval standpoint, some might say that the Catalogue is more effective because it is more comprehensive and global, even if the internal taxonomies might contradict one another—that it is more efficient because it is a one-stop shop, and more useful because it can also serve as a backbone taxonomy.

My intention in making this somewhat fluid distinction between description- and retrieval-oriented classifications is to bring attention to two broad approaches to classification in biodiversity sciences: one is the product of individualized scientific work and hermeneutic development, and the other is a space that attempts to prioritize consensus to unify information access and facilitate communication. This boundary tends to mark the divide between those that support and those that do not support a generalized taxonomic model such as the Catalogue of Life. To be sure, Furner also notes that the division between description- and retrieval-oriented approaches is artificial, for these approaches often commingle in practice. Historically, Linnaean nomenclature—terms that with their genus-species designation gesture toward a general sense of rank—were created for memory retrieval, after all. The trick, as with all retrieval systems, is to balance correctness (however that is defined within a certain context) with facility of use (for expert and nonexpert audiences alike).

The biodiversity taxonomic instrument is a complex machine. As we have seen, systems necessarily reduce nature to distinct entities and, in the process of creating classifications, must consistently reassemble these entities in ways that adhere to a certain set of epistemic commitments. From a process-oriented perspective, this is problematic, as it erases the system-oriented context from which these entities are derived. Nonetheless, classifications are practically necessary if we are to communicate effectively about nature and manage our relationship to it. The problem is that there are many ways to communicate this information. Taxonomic commitments vary widely in practice, and they are often fundamentally incommensurable. One problem with rigid, description-oriented taxonomic systems is that they are univocal and do not represent, or pretend to attend to, multiple points of view. This is, in part, because these descriptive taxonomies serve hypothesizing functions—they are arguments. The benefit of them, of course, is that they maintain the intellectual integrity of those who built them.

But even while these descriptive taxonomies may be incommensurable, structures like the Catalogue of Life are nonetheless trying to commingle these taxonomic opinions in spaces that represent consensus. The assumption is that, by aggregating multiple sources of data, retrieval-oriented concerns can be met. However, given the traditional, epistemic functions of classifications, these systems run counter to taxonomic expectations. In chapter 6, we switch our focus to the space of the Catalogue in order to better understand its intended benefits, even while its ultimate function is up for debate.

# Power of Position

## Classification and the Biodiversity Sciences

## By: Robert D. Montoya

## Citation:

*Power of Position: Classification and the Biodiversity Sciences*
**By:** Robert D. Montoya
**DOI:** 10.7551/mitpress/12245.001.0001
**ISBN (electronic):** 9780262369961
**Publisher:** The MIT Press
**Published:** 2022

This book is freely available in an open access edition thanks to TOME (Toward an Open Monograph Ecosystem) – a collaboration of the Association of American Universities, the Association of University Presses, and the Association of Research Libraries – and the generous support of Arcadia, a charitable fund of Lisbet Rausing and Peter Baldwin, and the UCLA Library. Learn more at the TOME website, available at: openmonographs.org

The MIT Press

This book is freely available in an open access edition thanks to TOME (Toward an Open Monograph Ecosystem)—a collaboration of the Association of American Universities, the Association of University Presses, and the Association of Research Libraries—and the generous support of Arcadia, a charitable fund of Lisbet Rausing and Peter Baldwin, and the UCLA Library. Learn more at the TOME website, available at: openmonographs.org.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Adobe Garamond Pro by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Montoya, Robert D., author.
Title: Power of position : classification and the biodiversity sciences / Robert D. Montoya.
Description: Cambridge, Massachusetts : The MIT Press, [2022] | Series: History and foundations of information science | Includes bibliographical references and index.
Identifiers: LCCN 2021033972 | ISBN 9780262045278 (paperback)
Subjects: LCSH: Biology—Classification. | Life sciences—Classification. | Cladistic analysis.
Classification: LCC QH83 .M68 2022 | DDC 570.1/2—dc23/eng/20211221
LC record available at https://lccn.loc.gov/2021033972