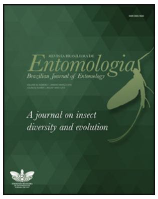
Revista Brasileira de Entomologia 61 (2017) 277–281

SOCIEDADE BRASILEIRA
DE ENTOMOLOGIA
FUNDADA EM 1937

REVISTA BRASILEIRA DE
Entomologia
A Journal on Insect Diversity and Evolution

www.rbentomologia.com



Review

Species descriptions and digital environments: alternatives for accessibility of morphological data

Thiago Sanches Ranzani da Silva

Universidade Federal do Paraná, Departamento de Zoologia, Curitiba, PR, Brazil

ARTICLE INFO

Article history:
Received 1 March 2017
Accepted 30 June 2017
Available online 14 July 2017
Associate Editor: Livia Pinheiro

Keywords:
Data accessibility
Interoperability
Annotations
Web-based taxonomy

ABSTRACT

Taxonomists' efforts throughout history provide significant amount of data that give support for establishing the specific identity of several groups of biological systems. In addition to identifying species, taxonomic research offers a wide range of biological information that can be used in other disciplines, *e.g.* evolution, ecology, integrated pest management. However, most of this information remains unappreciated due to certain aspects: (1) the advent of analytical tools have led to a shift in interest and investment in researches, focusing mainly in molecular studies; (2) the erroneous concept that the extensive data offered by taxonomic studies can be replaced by other datasets, separating it from its hypothesis-driven and investigative nature; (3) the final products found in taxonomic works are commonly restricted to a small group of researchers, due to its low accessibility and specific language. Considering this last aspect, web-based tools can be valuable to simplify the dissemination of the taxonomic product. Semantic annotation provide a condition in which species descriptions can be readily available and be far more extensive, enabling rapid exchange of countless data related to biological systems.

Introduction

Taxonomy is a fundamental science that provides support to all major fields in biology. Taxonomists produce much of the basic knowledge about species, providing crucial information to other disciplines that deal with any sorts of biological systems. However most taxonomic data remains unappreciated due to the fragmented condition and uneven accessibility of basic information on biodiversity. According to Miller et al. (2012), taxonomy allows species recognition through textual descriptions and images of specimens, and maps their distribution based on locality records, meaning that most of the vital knowledge used to preserve a particular species results from taxonomic works. With accelerated rates of extinction in many recent groups, summed to a loss of habitat in many regions of the world and the inexorable effects of climate change, taxonomy has never been more important (Isaac et al., 2004; Miller et al., 2012; Padial et al., 2010).

Biodiversity is one of the richest information provider fields for human knowledge (Peterson et al., 2010). However, only recent advances in cybertaxonomic infrastructure (*i.e.* the combination of digital and information technologies and informatics with taxonomy) made possible the creation of tools that unite biodiversity information in a way that they are available to a wide variety of users, including ecologists, wildlife and land managers, likewise making the existing information promptly accessible to the global taxonomic community (Godfray, 2002; Miller et al., 2012). Although there are plenty of websites devoted to specific taxa and projects, which are useful to those interested in questions inside that domain, the true potential of online tools lies in mechanisms that combine data from primary sources in a way that users can easily filter and recombine, making their use flexible toward any purpose (Peterson et al., 2010; Thessen and Patterson, 2011).

Taxonomic works are, generally, a summary of a set of elements, including descriptions in textual form, scientific names and nomenclature acts, literature references, images, records of species' occurrences and, recently, DNA sequences (Miller et al., 2012). The role of a taxonomist is to establish connections between specimens and nomenclature, express interpretations and hypothesis through textual elements and document observations using images. In a traditional taxonomic publication, all of those elements are combined in a single document. However, if they were to be released simultaneously and maintained as data sets linked to the original publication in an external environment, there would be direct effects on the taxonomic community, as well on other researcher groups (Deans et al., 2012; Godfray, 2002; Miller et al., 2012; Padial et al., 2010). In order to link data from various fields of knowledge,

---

E-mail: sanchesranzani@ufpr.br

http://dx.doi.org/10.1016/j.rbe.2017.06.005

disseminating phenotypic information in an open environment, the use of tools that enhance interoperability are vital. In this context, semantic annotations are the main choice to facilitate dissemination and retrieving of phenotypic data in a digital environment.

Semantic annotation is, basically, the process of attaching additional information to a concept, in a way that it is easily read and recovered by machines. It links multiple concepts across different domains, inside a digital platform. Hence, it enables the recovering of multiple knowledge, allowing the inclusion of particular concepts into broader perspectives. It can be used in ontology-based information retrieval queries for efficient data mining, spreading the current understanding of a specific concept to multiple domains. Therefore, its use may provide ways to extend the access to unappreciated data, such as those available in taxonomic works. Several studies applied semantic annotation to extend the application of phenotypic information to other domains of knowledge (*e.g.* Ćwiek-Kupczyńska et al., 2016; Hoehndorf et al., 2016; Serrano-Solano et al., 2017), suggesting that this process greatly enhances interoperability among systems that support this type of data.

Taking the aforementioned considerations into account, the intention of the present work is to discuss the importance of the taxonomic investigation, focusing on the means for retrieval of the taxonomic product, especially in the role played by semantic annotation in species descriptions. Discussions are based in a non-systematic literature review. Initially, I discuss the importance of the taxonomic work and the effectiveness of the taxonomist workflow to produce descriptions of biological systems. After that, I address the question of the accessibility of taxonomic data, focusing on the manner in which this type of information is presented, especially when data mining is concerned. In this sense, I discuss the construction of anatomical concepts and some conceptual disagreements concerning the application of homology assumptions in the elaboration of anatomic ontologies. Following this, I address the subject on how digital environments can be valuable to the dissemination of taxonomic products, emphasizing the role of semantic annotation in the extension of this kind of information.

Accessibility to the taxonomic information

Taxonomists provide much of the basic knowledge about biological systems, and are thus considered as information providers to the life sciences (Wilson, 2002; Gewin, 2002). They provide essential information to innumerous studies, making it possible to grasp the real dimension of biodiversity, as well as the causes and consequences of their diversification. It is such an integrative evidence that is a requisite to other disciplines in biology, such as conservation (Mooers, 2007; Forest et al., 2007), biomimetics (Badarnah and Fernández, 2015; Rajabi et al., 2015), and evolution of body plans (evo-devo) (Mallarino and Abzhanov, 2012), being also directly related applicable to fields such as commerce, biosecurity, environmental issues, fishery, medicine, mining prospection, public health, wildlife management, *etc.* (Kapoor, 1998). In particular, eco-morphological approaches benefit from taxonomic/morphologic products, which provides useful information for investigating the contribution of size and specific identity as related to function (revision in Sibbing and Nagelkerke, 2000). The relation between ecology and morphology has been clearly demonstrated in various animal taxa (Ricklefs and Travis, 1980; Miles and Ricklefs, 1984; Miles et al., 1987; Douglas and Matthews, 1992; Silva and Brandão, 2010). Researchers assume that species are more similar functionally in cases when they are morphologically more similar (Chalcraft and Resetarits, 2003).

Taxonomists examine specimens, classifying them in taxonomic concepts (*i.e.* hypothesis of species; Berendsohn, 1995, Berendsohn and Geoffroy, 2007 and Kennedy et al., 2006), organize their associated information (*e.g.* where and when they were collected and by whom) and describe their physical appearance through text annotations (Deans et al., 2012). These annotations compiled by taxonomists are transcribed in "descriptions", including diagnoses, which serve to formalize species that can be recovered by researchers in the future (Deans et al., 2012; Miller et al., 2012).

Taxonomic descriptions are a valuable source of knowledge concerning phenotypic diversity in the living world. These evidences, however, are generally restricted to taxonomic works, written by and consumed almost exclusively by taxonomists. Therefore, species' descriptions are not easily obtained and other scientists seldom reuse the phenotypic data resultant from it (Deans et al., 2012). Even when phenotypic data is obtained during the elaboration of non-taxonomic studies, it is rarely included in the finished study and, thus, not made publicly available (Zamir, 2013) or it is presented in an extremely heterogeneous and unordered manner, making its discovery problematic and interpretation impractical (Deans et al., 2015).

Although electronic availability of taxonomic treatments are rapidly growing, reflected in the change of publication requisites (ICZN, 2012) that foster digitalization endeavors (Balhoff et al., 2013), phenotypic descriptions are still made in natural language normally applying specialized anatomical terminology. This kind of description poses a challenge when researchers try to data mine it (however, *cf.* Cui (2012) and Thessen et al. (2012) for mechanisms to convert natural language descriptions in semantic annotations), probably due to the excessive amount of homonymies and synonymies that exist between anatomical concepts (Yoder et al., 2010).

Descriptions are constructed based on analogies and are seldom reused. According to Deans et al. (2012), descriptions are repeated when a taxon is revised, a mostly redundant process leading to the accretion of various text annotations of phenotypes for a small number of taxa, in which most of it turns out to be unused information. Owing to this inefficiency, many researchers consider descriptions as a nuisance (*cf.* Evenhuis, 2007) and, in some cases, have argued in favor of the delimitation of a taxon exclusively with molecular data (Hebert et al., 2003; Cook et al., 2010). Deans et al. (2012) mentions that the limited utility of descriptions may partially explain recurrent stagnation of funding and training of taxonomists (*cf.* Agnarsson and Kuntner, 2007; Carvalho et al., 2007) and the subsequent lack of morale that pervades this fundamental science.

There are three central issues, raised by Deans et al. (2012), that may be responsible for encumbering the elaboration of phenotypic semantic annotations: (i) the need for a physical printed version to validate a formal description (although this is mostly changed; see Cressey, 2011; Rinaldo and Norton, 2009); (ii) descriptions are normally composed in non-uniform natural language that are inefficient for data-mining; and (iii) descriptions do not commonly refer to logical definitions of concepts (*i.e.* homonymy, polysemy and synonymy are prevalent) (Yoder et al., 2010). Thus, the final product of taxonomic work remains largely unexploited and barely used by investigators in other disciplines.

However, according to Patterson et al. (2010), taxonomy possesses two singular features that makes its reuse possible in biodiversity informatics. The first is the universal use of the system of scientific names that enables treating names as metadata, which can be used to index information related to biodiversity. The second attribute refers to classification schemes. They transform lists of names into organizational structures (*i.e.* ontologies) responsible for grouping data, allowing the construction of generalized declarations and the inferences of taxa properties by users, permitting extensive or focused searches and facilitating browsing for information (Patterson et al., 2010). The value of names as metadata and classification as ontologies lead to the idea that names-based digital infrastructures can serve the biological sciences (Patterson et al., 2008). In this sense, taxonomists benefit from web-related initiatives and technology to improve the taxonomic workflow (Carvalho et al., 2007). Wheeler (2007) mentions that the combination of digital and information technologies and informatics with taxonomy is an efficient method to enhance the quality and amount of taxonomic products.

According to Will and Rubinoff's (2004) philosophical perspective regarding human discovery and knowledge of the natural world, to perceive the world through a restricted set of observable phenomena would result in an intellectually sterile landscape. Patterns of human perception derive from the understanding of all kinds of data, particularly from rich types as morphology. A holistic perception of organisms incorporating phylogenies, functional morphology, behavior, ecology, *etc.*, will help us make better-informed conservation decisions.

Still according to Will and Rubinoff (2004), it is expected that researchers in this field propose taxonomic hypothesis and develop identification tools essential for the rest of biological sciences. Taxonomic concepts are, or should be, constructed upon a multitude of data frequently obtained by non-systematists during their investigations. Reciprocity in the accessibility of data would enhance the value of the obtained final product.

Although the relation between morphology and other disciplines is evident, with the obvious crucial role of taxonomy in the production of morphological data, this type of information is normally obtained through independent means of acquisition, *i.e.* they are not easily acquired through a database, which applies formal methods of presenting and retrieval of morphological information. They are typically acquired during the investigation process and are presented in a more-or-less formalized manner, making its reuse difficult, leading to a permanent independent acquisition of morphological data each time a trait would be addressed in broader perspectives. Acquisition of information through independent means usually relates to distinction of methods for obtaining, analyzing and presenting basic data.

In this perspective, presenting morphological information using semantic annotation tools would provide the means to acquire data in a formalized, analytic way, promoting extensive reuse of morphological products.

Annotations, ontologies and biological information

Annotations play an important role in the use and dissemination of information. Hence, its application is extremely vast, being useful in fields as linguistics, legal research, imaging, software engineering and computational biology. They represent commentaries or explanations attached to a document, such as texts, images or any other type of data. In an ontological context, annotations identify real-world entities alongside properties and relations that characterize the entities' attributes and role in their textual context, with respect to a reference ontology (Sanfilippo et al., 2006). Within a biological framework, more specifically in taxonomy, these commentaries or explanations can represent observations of a set of all phenotype expressed by a specimen or taxon (Deans et al., 2012).

Ontologies are formal representations of concepts and their logical relations inside a domain, supporting representations of knowledge with explicit semantics (Balhoff et al., 2013). In natural language processing and information retrieval, explicit semantics are vectorial representations of texts that uses structured sets of documents to store information that will be used by computer systems.

These ontologies are important because they are formal specifications of certain aspects of reality, and both human and computers can use it. They promote interoperability, which means communication between databases, like cross-searches (Mabee et al., 2007a). According to Washington et al. (2009) and Walls et al. (2012), databases can be combined, coherently computed and shared only by referencing standardized concepts. Biological ontologies have become standard tools to organize and access genomic and phenomic data from model species (Mungall et al., 2010). The application of these tools to represent and disseminate comparable descriptive data provides opportunities to perform connections between phenotypic and genomic information (Mabee et al., 2007a).

Deans et al. (2012), recently, proposed that the application of annotations in taxonomic descriptions, through semantic phenotypes, would allow efficient searches, wider integration, and automated reasoning techniques to data by a broader array of researchers. This would augment the value of taxonomic work and promote its reuse, whereas the discrete nature of semantic phenotypes makes them prone to extensive use. This means that individual observations could be made outside taxonomic collections and added to other systems, promoting the extension of this discipline.

According to Dahdul et al. (2010), taxonomists use the expressivity and richness of natural language to precisely describe the morphological variation observed among species. The construction of these descriptions is more or less formalized, like characters and character states in the systematic language (Sereno, 2007). By usually describing phenotypes binding each character to a specific state (implying homology hypothesis to each character), taxonomists can have difficulties to reconcile the relation of the homologies to the concept hierarchy represented by an ontology (Deans et al., 2012). That is, the hierarchical arrangement needed to structure an ontology differs significantly from the hierarchical arrangement needed to assess homology of characters in a phylogenetic context.

There are somewhat opposing views on how homology hypothesis can impend or not the construction of ontologies and, by extension, compromise the annotation of concepts. According to Deans et al. (2012), existent anatomic ontologies are primarily developed using definitions of structural concepts, a strategy that works well in a case in which these ontologies are elaborated based in one species. However, the same authors observe that the enhancement of the use of anatomic ontologies of multiple species and of high hierarchical levels force taxonomists to deal with evolutionary questions that are normally unfitting during the process of developing ontologies. According to Mungall et al. (2012), although the inference of homology is important from the perspective of connecting gene expression and phenotype across species, statements of homology can be controversial, subject to change and even contradictory, providing a limited set of potentially interesting results. In contrast, the same authors mention that, as one of many axes of classification, homology can be informative when grouping concepts, considering it is used in an integrative framework.

Evolutionary comparisons operate in higher levels in biological systems, namely in the level of continuity and modification of the phenotype throughout the tree of life (Mabee et al., 2007a). The study of phenotypic similarity due to the continuity of inherited information (*i.e.* homology) (Van Valen, 1982; Roth, 1984) is important to biologists in all levels and central to comparative biology (Bock and Cardew, 1999). This way, ultimately, conceptions between terms inside and between anatomic ontologies should be defined by declarations of relation of homology. However, the hierarchical arrangement of anatomic classes generally do not imply homologies (Deans et al., 2012). Developing a mechanism to define and use this relation is required for targeting queries of evolutionary biologists. Homology relations can be expressed as a mapping between ontologies in a database, allowing greater flexibility than codifying relations of homology inside an anatomic ontology. This means that homology hypothesis could be explicitly incorporated in ontologies through other mechanisms (non-hierarchical), such as accessory homology declarations (*i.e.* in another data structure) that unify anatomic concepts to taxa in a phylogenetic context (Deans et al., 2012). Homology attributions need to be based in the evidence of relation (*e.g.* position, developing or composition), and could be recovered from the literature or other sources (Mabee et al., 2007a).

The extended maintenance of anatomic terminologies, however, do not need the obligate link between anatomic concepts and homology hypothesis (see Edgecombe (2008) for a discussion on the use of the term "basipterigoid process"). According to Vogt et al. (2010), homology hypothesis transcend the observed by providing a possible explanation for the likeliness of traits and for the existence of kinds of traits through statements of a common evolutionary origin. This means that correspondence of structures in one trait to another only represents a necessary, but not sufficient, condition for homology. Hence, structural similarities can be expressible and communicable independent of explanatory hypotheses such as homology.

Extreme cases of morphological differences are found in anatomic systems under strong selection and/or possess great value as sources of specific diagnostic characters. Disparate terminologies result in a vast field of inconsistent declarations (Yoder et al., 2010). This circumstance inhibits future efforts to explore biological processes, including: (1) comparison of patterns of genic expression (Mabee et al., 2007a); (2) comparative morphology and phylogenetics, which is increasingly important when you try to incorporate the vast amount of information from the fossil record and rare taxa from which it is not possible to extract DNA; (3) phenotypic variability in an environmental context; (4) descriptive taxonomy cannot be efficiently consulted; (5) computerized reasoning, as well as logical reasoning, requires a well-defined semantic structure (Yoder et al., 2010).

A phenotype can be defined as a result of a given genotype in a given environment and can be described and annotated to facilitate comparisons (Washington et al., 2009). Currently, most of the morphological information is presented in natural language, following a more or less formalized approach, lacking a structured and ordered method that can be semantically annotated. In systematic characters, structures and their conditions can be found in the character description, in the description of the state of the character, or in both. At the same time, morphological descriptions of variation among species in the literature conform to general semantic formalisms (Washington et al., 2009). This means that morphological information is presented in an inordinate manner and hence is partially traceable. To circumvent this issue, descriptions of individual phenotypic characters could be recorded using formal representations, where they are presented as a set of structures followed by their corresponding conditions (*e.g.* shape, size, number) (Mabee et al., 2007a,b). Descriptions following strict methods of formalism are essential for semantic annotation, which, in turn, enables application of distinct axes of classification when structuring phenotypic concepts. This way, if phenotypic descriptions were based in an organized shared vocabulary they would be structured in a way that algorithms could be written to compare phenotypes using analytic tools (Washington et al., 2009).

The description of the biological diversity applying formalized methods of semantic annotation, along with computational tools consistent with taxonomic practice, provide a unique opportunity to standardize and consult phenotypic data in a more rigorous and enlightening way.

Conclusions

The importance of the taxonomic knowledge is unmatched for distinct fields that are concerned with the study and comprehension of biological systems. Identification, description, and assembling knowledge connected to species, as part of the taxonomic product, make it possible for researchers to recognize and establish limits among these units. Thus, they develop hypotheses and obtain available data from an extensive range of sources, be it in literature, collections or sampling events, presenting, by itself, a viable scientific effort. However, much of this information can be difficult to find or may be dispersed and fragmented, making data produced in taxonomic works difficult to access by researchers from other disciplines, hence impossible to reuse in other scientific communities.

The application of semantic annotation for describing biological diversity provides an addition to the utility of data produced in taxonomic works. Therefore, it serve as a tool for the extension of descriptive works focused in taxonomy, promoting the reuse of morphological information in other disciplines, such as evolution and ecology. Due to its unique structure, the construction of anatomic ontologies needs cooperation between computer scientists and taxonomists, so as to simplify computational platforms. This way it will be possible to continue taxonomic investigation as it is made today, promoting its expansion and encouraging more research in morphology yielding valuable data for the study of patterns, as well as helping elucidate processes in more generalized contexts.

Conflicts of interest

The author declares no conflicts of interest.

Acknowledgements

I am very grateful to Rodrigo Feitosa and John Lattke for reviewing preliminary versions of this manuscript, being also indebted to two anonymous reviewers, which gave valuable comments. The author is financially supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), grant number 40001016005P5.

References

Agnarsson, I., Kuntner, M., 2007. Taxonomy in a changing world: seeking solutions for a science in crisis. Syst. Biol. 56 (3), 531–539.

Balhoff, J.P., Mikó, I., Yoder, M.J., Mullins, P.L., Deans, A.R., 2013. A semantic model for species description applied to the ensign wasps (Hymenoptera: Evaniidae) of New Caledonia. Syst. Biol. 62 (5), 639–659.

Badarnah, L., Fernández, J.E., 2015. Morphological configurations inspired by nature for thermal insulation materials. In: International Association for Shell and Spatial Structures (IASS) Symposium.

Berendsohn, W.G., 1995. The concept of "potential taxa" in databases. Taxon, 207–212.

Berendsohn, W.G., Geoffroy, M., 2007. Networking taxonomic concepts – uniting without "unitary-ism". In: Biodiversity Databases – Techniques, Politics, and Applications., pp. 13–22.

Bock, G., Cardew, G. (Eds.), 1999. Homology. John Wiley & Sons.

Carvalho, M.R., Bockmann, F.A., Amorim, D.S., Brandão, C.R.F., de Vivo, M., Figueiredo, J.L., et al., 2007. Taxonomic impediment or impediment to taxonomy? A commentary on systematics and the cybertaxonomic-automation paradigm. Evol. Biol. 34 (3), 140–143.

Chalcraft, D.R., Resetarits Jr., W.J., 2003. Mapping functional similarity of predators on the basis of trait similarities. Am. Nat. 162 (4), 390–402.

Cook, L.G., Edwards, R.D., Crisp, M.D., Hardy, N.B., 2010. Need morphology always be required for new species descriptions? Invertebr. Syst. 24 (3), 322–326.

Cressey, D., 2011, July. Botanists shred paperwork in taxonomy reforms. Nature – News.

Cui, H., 2012. CharaParser for fine-grained semantic annotation of organism morphological descriptions. J. Am. Soc. Inf. Sci. Technol. 63 (4), 738–754.

Ćwiek-Kupczyńska, H., Altmann, T., Arend, D., Arnaud, E., Chen, D., Cornut, G., et al., 2016. Measures for interoperability of phenotypic data: minimum information requirements and formatting. Plant Methods 12 (1), 44.

Dahdul, W.M., Balhoff, J.P., Engeman, J., Grande, T., Hilton, E.J., Kothari, C., et al., 2010. Evolutionary characters, phenotypes and ontologies: curating data from the systematic biology literature. PLoS ONE 5 (5), e10708.

Deans, A.R., Yoder, M.J., Balhoff, J.P., 2012. Time to change how we describe biodiversity. Trends Ecol. Evol. 27 (2), 78–84.

Deans, A.R., Lewis, S.E., Huala, E., Anzaldo, S.S., Ashburner, M., Balhoff, J.P., et al., 2015. Finding our way through phenotypes. PLoS Biol. 13 (1), e1002033.

Douglas, M.E., Matthews, W.J., 1992. Does morphology predict ecology? Hypothesis testing within a freshwater stream fish assemblage. Oikos, 213–224.

Edgecombe, G.D., 2008. Anatomical nomenclature: homology, standardization and datasets. Zootaxa 1950, 87–95.

Evenhuis, N.L., 2007. Helping solve the "other" taxonomic impediment: completing the eight steps to total enlightenment and taxonomic nirvana. Zootaxa 1407, 3–12.

Forest, F., Grenyer, R., Rouget, M., Davies, T.J., Cowling, R.M., Faith, D.P., et al., 2007. Preserving the evolutionary potential of floras in biodiversity hotspots. Nature 445 (7129), 757–760.

Gewin, V., 2002. Taxonomy: all living things, online. Nature 418 (6896), 362–363.

Godfray, H.C.J., 2002. Challenges for taxonomy: the discipline will have to reinvent itself if it is to survive and flourish. Nature 417, 17–19.

Hebert, P.D., Cywinska, A., Ball, S.L., 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B: Biol. Sci. 270 (1512), 313–321.

Hoehndorf, R., Alshahrani, M., Gkoutos, G.V., Gosline, G., Groom, Q., Hamann, T., et al., 2016. The flora phenotype ontology (FLOPO): tool for integrating morphological traits and phenotypes of vascular plants. J. Biomed. Semant. 7 (1), 65.

International Commission on Zoological Nomenclature, 2012. Amendment of Articles 8, 9, 10, 21 and 78 of the International Code of Zoological Nomenclature to expand and refine methods of publication. ZooKeys 219, 1–10.

Isaac, N.J., Mallet, J., Mace, G.M., 2004. Taxonomic inflation: its influence on macroecology and conservation. Trends Ecol. Evol. 19, 464–469.

Kapoor, V.C., 1998. Principles and Practices of Animal Taxonomy. Science Publishers.

Kennedy, J., Hyam, R., Kukla, R., Paterson, T., 2006. Standard data model representation for taxonomic information. OMICS: J. Integr. Biol. 10 (2), 220–230.

Mabee, P.M., Ashburner, M., Cronk, Q., Gkoutos, G.V., Haendel, M., Segerdell, E., Mungall, C., Westerfield, M., 2007a. Phenotype ontologies: the bridge between genomics and evolution. Trends Ecol. Evol. 22 (7), 345–350.

Mabee, P.M., Arratia, G., Coburn, M., Haendel, M., Hilton, E.J., Lundberg, J.G., Mayden, R.L., Rios, N., Westerfield, M., 2007b. Connecting evolutionary morphology to genomics using ontologies: a case study from Cypriniformes including zebrafish. J. Exp. Zool. B: Mol. Dev. Evol. 308 (5), 655.

Mallarino, R., Abzhanov, A., 2012. Paths less traveled: evo-devo approaches to investigating animal morphological evolution. Annu. Rev. Cell Dev. Biol. 28, 743–763.

Miles, D.B., Ricklefs, R.E., 1984. The correlation between ecology and morphology in deciduous forest passerine birds. Ecology 65 (5), 1629–1640.

Miles, D.B., Ricklefs, R.E., Travis, J., 1987. Concordance of ecomorphological relationships in three assemblages of passerine birds. Am. Nat., 347–364.

Miller, J., Dikow, T., Agosti, D., Sautter, G., Catapano, T., Penev, L., et al., 2012. From taxonomic literature to cybertaxonomic content. BMC Biol. 10 (1), 1.

Mooers, A.Ø., 2007. Conservation biology: the diversity of biodiversity. Nature 445 (7129), 717–718.

Mungall, C.J., Gkoutos, G.V., Smith, C.L., Haendel, M.A., Lewis, S.E., Ashburner, M., 2010. Integrating phenotype ontologies across multiple species. Genome Biol. 11 (1), R2.

Mungall, C.J., Torniai, C., Gkoutos, G.V., Lewis, S.E., Haendel, M.A., 2012. Uberon, an integrative multi-species anatomy ontology. Genome Biol. 13 (1), R5.

Padial, J.M., Miralles, A., De la Riva, I., Vences, M., 2010. The integrative future of taxonomy. Front. Zool. 7 (1), 1.

Patterson, D.J., Faulwetter, S., Shipunov, A., 2008. Principles for a names-based cyberinfrastructure to serve all of biology. Zootaxa 1950, 153–163.

Patterson, D.J., Cooper, J., Kirk, P.M., Pyle, R.L., Remsen, D.P., 2010. Names are key to the big new biology. Trends Ecol. Evol. 25 (12), 686–691.

Peterson, A.T., Knapp, S., Guralnick, R., Soberón, J., Holder, M.T., 2010. The big questions in biodiversity informatics. Syst. Biodivers. 8, 159–168.

Rajabi, H., Ghoroubi, N., Darvizeh, A., Dirks, J.H., Appel, E., Gorb, S.N., 2015. A comparative study of the effects of vein-joints on the mechanical behaviour of insect wings: I. Single joints. Bioinspiration Biomim. 10 (5), 056003.

Ricklefs, R.E., Travis, J., 1980. A morphological approach to the study of avian community organization. Auk, 321–338.

Rinaldo, C., Norton, C., 2009. BHL, The Biodiversity Heritage Library: An Expanding International Collaboration.

Roth, V.L., 1984. On homology. Biol. J. Linn. Soc. 22 (1), 13–29.

Sanfilippo, A., Tratz, S., Gregory, M., Chappell, A., Whitney, P., Posse, C., et al., 2006. Ontological annotation with wordnet. In: Proc. Int. WordNet Conference GWC.

Serrano-Solano, B., Ramos, A.D., Hériché, J.K., Ranea, J.A., 2017. How can functional annotations be derived from profiles of phenotypic annotations? BMC Bioinform. 18 (1), 96.

Sereno, P.C., 2007. Logical basis for morphological characters in phylogenetics. Cladistics 23 (6), 565–587.

Sibbing, F.A., Nagelkerke, L.A., 2000. Resource partitioning by Lake Tana barbs predicted from fish morphometrics and prey characteristics. Rev. Fish Biol. Fish. 10 (4), 393–437.

Silva, R.R., Brandão, C.R.F., 2010. Morphological patterns and community organization in leaf-litter ant assemblages. Ecol. Monogr. 80 (1), 107–124.

Thessen, A.E., Cui, H., Mozzherin, D., 2012. Applications of natural language processing in biodiversity science. Adv. Bioinform., 2012.

Thessen, A.E., Patterson, D.J., 2011. Data issues in the life sciences. ZooKeys 105, 15–51.

Van Valen, L.M., 1982. Homology and causes. J. Morphol. 173 (3), 305–312.

Vogt, L., Bartolomaeus, T., Giribet, G., 2010. The linguistic problem of morphology: structure versus homology and the standardization of morphological data. Cladistics 26 (3), 301–325.

Walls, R.L., Athreya, B., Cooper, L., Elser, J., Gandolfo, M.A., Jaiswal, P., et al., 2012. Ontologies as integrative tools for plant science. Am. J. Bot. 99 (8), 1263–1275.

Washington, N.L., Haendel, M.A., Mungall, C.J., Ashburner, M., Westerfield, M., Lewis, S.E., 2009. Linking human diseases to animal models using ontology-based phenotype annotation. PLoS Biol. 7 (11), e1000247.

Wheeler, Q.D., 2007. Invertebrate systematics or spineless taxonomy. Zootaxa 1668, 11–18.

Will, K.W., Rubinoff, D., 2004. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. Cladistics 20 (1), 47–55.

Wilson, E.O., 2002. What is nature worth? Wilson Q. (1976) 26 (1), 20–39.

Yoder, M.J., Miko, I., Seltmann, K.C., Bertone, M.A., Deans, A.R., 2010. A gross anatomy ontology for Hymenoptera. PLoS ONE 5 (12), e15991.

Zamir, D., 2013. Where have all the crop phenotypes gone? PLoS Biol. 11 (6), e1001595.