BioEssays

EX LABORATORIO

Thoughts & Opinion

Why is there no service to support taxonomy?

Julia D. Sigwart[1] | **Chong Chen[2]** | **Ekin Tilic[1]** | **Miguel Vences[3]** | **Torben Riehl[1]**

[1]Senckenberg Research Institute and Museum, Marine Zoology Department, Frankfurt, Germany

[2]Japan Agency for Marine-Earth Science and Technology (JAMSTEC), X-STAR, Yokosuka, Japan

[3]Technische Universität Braunschweig, Zoological Institute, Braunschweig, Germany

Correspondence
Julia D. Sigwart, Senckenberg Research Institute and Museum, Marine Zoology Department, Frankfurt, Germany.
Email: julia.sigwart@senckenberg.de

Abstract

Increasing complexity and specialisation of modern sciences has led to increasingly collaborative publications, as well as the involvement of commercial services. Modern integrative taxonomy likewise depends on many lines of evidence and is increasingly complex, but the trend of collaboration lags and various attempts at 'turbo taxonomy' have been unsatisfactory. We are developing a taxonomic service in the Senckenberg Ocean Species Alliance to provide fundamental data for new species descriptions. This will also function as a hub to connect a global network of taxonomists, assembling an alliance of scientists working on potential new species to tackle both the extinction and inclusion crises we face today. The current rate of new species descriptions is simply too slow; the discipline is often dismissed as old fashioned, and there is a crisis level need for taxonomic descriptions to come to grips with the scale of Anthropocene biodiversity loss. Here, we envision how the process of describing and naming species would benefit from a service supporting the acquisition of descriptive data. Also see the video abstract here: https://youtu.be/E8q3KJor_F8

KEYWORDS
biodiversity, global science, marine inverterbates, species, taxonomy

INTRODUCTION

The early history of science typically focusses on solitary famous scientists exploring the multi-facetted 'natural sciences', in stark contrast to the increasingly baffling complexity and specialisation of modern knowledge. The advance of modern science benefits from large-scale scientific infrastructures maintained, long term, at national or global scale, ranging from particle accelerators to space stations to research vessels, and natural history museums. Our shared efforts are evidenced by upward trends in collaborative publications including hundreds-strong authorship consortia.[1] The vast majority of these authorships are linked to contributions of research technique, combining approaches contributed by multiple, complimentary specialist approaches.[2]

With our improving understanding of what species mean and their evolutionary patterns, modern integrative taxonomy depends on multiple, independent lines of evidence. The need for morphological diagnoses is increasing, not decreasing, with technological developments in biodiversity assessment. This has been the primary shortcoming of molecular barcode diagnoses as a tool for rapid species discovery. Recent advances in molecular techniques and AI-driven image recognition allow for astounding high-throughput automated monitoring of biodiversity.[3] But this leads to important, unanswered questions. Does a DNA barcode in a soil or water sample belong to a giant free-living organism, or a miniaturised parasite? Is the species in a photograph part of the endemic local fauna worthy of special protection, or is it perhaps easy to confuse with a widespread non-threatened species? Automated monitoring cannot be effectively integrated with conservation without names to assign to the organisms identified,[4] and DNA metabarcoding is blind without morphological information

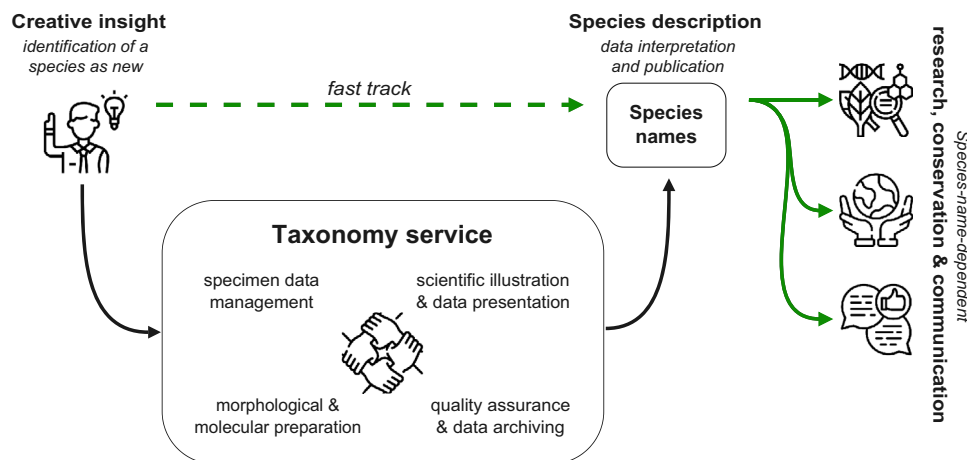Julia D. Sigwart and Chong Chen have contributed equally to this study.

FIGURE 1 The Senckenberg model for a streamlined taxonomy workflow involving a commercial service that covers transferable technical aspects of species descriptions. This figure was designed using icons from Flaticon.com.

tied to the names. With growing anthropogenic impacts on ecosystems and their species, causing increasing rates of extinction, taxonomists are challenged with the task to describe the vast numbers of unknown species before they vanish.

ACCELERATING TAXONOMY: THE SENCKENBERG MODEL

The point has been made that large-scale species descriptions are achievable with concentrated effort, and that acceleration is critically needed.[5,6] Not only resources, but changes in working practice, are required. This includes actively embracing a more inclusive mentality among working taxonomists. Although taxonomy has become more collaborative in recent decades,[7] numbers of co-authors are lower and increasing more slowly compared to overall trends. The average number of authors on taxonomy papers was 1.9 in 2002 (vs. 4.1 for all life sciences) increasing to 3.2 (vs. 5.9) in 2018 (www.nlm.nih.gov/bsd/authors1.html accessed 12 February 2023). The specialisation of a taxonomist is naturally an organismal group, rather than a method. Taxonomists still produce superb multifaceted descriptions of new species that ensure the stability and usability of names. The perceived pressure to master and apply all possible techniques by a single researcher presents a Sysiphian burden to individual taxonomists.

In some fields, such as molecular biology, collaborative platforms have further developed into commercialised services. In the early 2000s, every major lab group invested in sequencing machines to ensure they could obtain timely and accurate data. High-throughput technologies later dramatically decreased sequencing costs, and commercial services have taken over.[8] Service units, responding to the varied needs of diverse clients, can drive the cutting edge of technological development, and implement standardised quality control measures.[9] With services, and to a certain extent scientific infrastructure, researchers can outsource expertise with an assurance of high quality results (Figure 1). This leads to the question, what could a taxonomic service unit provide?

The process of describing and naming a species could be usefully compared to the experience of building a house. There is a romantic allure to doing the entire project one's self, from design to final construction, and familiarity can give us the impression that it must be straightforward. This becomes less tractable when the eager builder encounters the wildly diverging demands for plumbing, electrical and design skills, and moreover the inscrutable demands of bureaucratic building codes and unappeasable safety inspectors. Very few people have the time, or the diverse talents, required to take on such a complex project. Some might attempt to do most of it, but hire a specialist to take on one particularly tricky aspect. Most of us, if we can afford it, would hire a building service that could arrange or subcontract all the different parts, expertly delivering the details but leaving us the creative satisfaction of seeing our vision fulfilled. What if such a service was available for taxonomic species descriptions?

The real intellectual, skilled and creative part of taxonomy is the process of recognising an organism as new to science, and formulating the hypothesis that it is an independent evolutionary lineage from all other described species. The additional laborious work that comes associated with publishing the description – producing images, drawings and sequence data as evidence for the new hypothesis – does not always require a researcher and was historically the remit of technicians. Research questions about morphology itself require intellectual involvement throughout the imaging process. However, in taxonomy, standard techniques of dissection, imaging and sequencing are usually applicable across a wide range of organisms. A more accessible approach to taxonomy should also embrace standardisation across phyla, without sacrificing details. Standardisation of data will allow for effective acceleration with emerging artificial intelligence and machine learning applications. Where techniques are transferable or repetitive, they can be done by a qualified technical service.

The race for research impact statistics pushes taxonomic publications to prioritise 'wider perspectives' (ecological contexts, phylogeography or evolution of morphological adaptations) at the cost of efficiency in naming species. Some modern descriptive work is

complex with lengthy details, such as isopod crustaceans where often every appendage is described exhaustively from each sex and the juvenile,[10] and bacteria where several lines of physiological characters from monocultures are expected in addition to morphological and molecular characterisation.[11] In many cases, it is clear that not all of these data are necessary to distinguish one species from another. This is interesting science, but not actually required for naming species, and doing everything at once slows taxonomy down.

The slow pace of taxonomic descriptions impacts all fields of sciences and beyond. There are numerous well-known cases of laboratory work on 'model' organism species that are misidentified (e.g. immortal jellyfish[12]) or even not yet named (e.g. fish[13]). Many species are easily confused with congeners; if researchers to some extents do not know what species they are working on, this fundamentally confounds replication in future research. Likewise, modern biodiversity monitoring is increasingly dependent on molecular data, which should ideally be connected to type material to ensure their reliability. Robust, integrative taxonomic descriptions using multiple lines of evidence from molecular and morphological diagnoses are the foundation for reliable identification.

The Senckenberg Ocean Species Alliance (SOSA; sosa.senckenberg.de) is now developing a service to support species descriptive data, that will be tested and made available through the next decade. The SOSA Discovery Unit aims to offer a service that covers a wide range of taxonomic groups, working closely with established experts to determine relevant standards. This could enable a service to eventually expand to non-taxonomists who discover potential new species – the network of supporting taxonomists can help vet whether they are indeed new species using the time freed by outsourcing. Most of the undescribed biodiversity resides in megadiverse countries which are mostly developing nations[14]; taxonomy must improve its inclusiveness. As taxonomic practice is closely interwoven with large museums and public collections, it is difficult for countries in the Global South to gain traction in taxonomic research. We envisage the SOSA discovery service unit to also function as a hub to connect its network of taxonomists with the Global South. This type of taxonomic description service could be replicated in future centres co-located with under-described faunas. Assembling a global alliance of scientists will help tackle both the extinction and inclusion crises we face today.

CONCLUSION

There has been no service for taxonomy because taxonomists did not demand one when their science became more complex – instead they just worked harder and harder. But this comes at a cost of time, time that we no longer have. A lack of a service for taxonomy impacts not only this discipline, but all others that are the users of species names: everyone in biosciences and beyond. Today, we need a service for taxonomy. A transformative approach to streamlining data provision and publication can revolutionise our task to understand life on Earth.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

ORCID

Julia D. Sigwart https://orcid.org/0000-0002-3005-6246
Torben Riehl https://orcid.org/0000-0002-7363-4421

REFERENCES

1. Fontanarosa, P., Bauchner, H., & Flanagin, A. (2017). Authorship and Team Science. *Journal of the American Medical Association*, *318*, 2433–2437.
2. Thelwall, M. (2020). Large publishing consortia produce higher citation impact research but coauthor contributions are hard to evaluate. *Quantitative Science Studies*, *1*, 290–302.
3. Miralles, A., Bruy, T., Wolcott, K., Scherz, M. D., Begerow, D., Beszteri, B., & Vences, M. (2020). Repositories for taxonomic data: Where we are and what is missing. *Systematic Biology*, *69*, 1231–1253.
4. IUCN. (2012). *IUCN Red List Categories and Criteria version 3.1* (2nd ed.). The International Union for Conservation of Nature (IUCN).
5. Orr, M. C., Ferrari, R. R., Hughes, A. C., Chen, J., Ascher, J. S., Yan, Y. H., Williams, P. H., Zhou, X., Bai, M., Rudoy, A., & Zhang, F. (2021). Taxonomy must engage with new technologies and evolve to face future challenges. *Nature Ecology & Evolution*, *5*, 3–4.
6. Wheeler, Q. D., Knapp, S., Stevenson, D. W., Stevenson, J., Blum, S. D., Boom, B. M., Borisy, G. G., Buizer, J. L., De Carvalho, M. R., Cibrian, A., & Donoghue, M J. (2012). Mapping the biosphere: Exploring species to understand the origin, organization and sustainability of biodiversity. *Systematics and Biodiversity*, *10*, 1–20.
7. Costello, M. J., Wilson, S., & Houlding, B. (2011). Predicting total global species richness using rates of species description and estimates of taxonomic effort. *Systematic Biology*, *61*, 871–871.
8. Pervez, M. T. M. J. u., Hasnain, S. H., Abbas, M. F., Moustafa, N. A., & Shah, S. S. M. (2022). A comprehensive review of performance of next-generation sequencing platforms. *BioMed Research International*, *2022*, 3457806.
9. Endrullat, C., Glökler, J., Franke, P., & Frohme, M. (2016). Standardization and quality management in next-generation sequencing. *Applied & Translational Genomics*, *10*, 2–9.
10. Riehl, T., & De Smet, B. (2020). Macrostylis metallicola spec. nov.— an isopod with geographically clustered genetic variability from a polymetallic-nodule area in the Clarion-Clipperton Fracture Zone. *PeerJ*, *8*, e8621.

11. Murray, A. E., Freudenstein, J., Gribaldo, S., Hatzenpichler, R., Hugenholtz, P., Kämpfer, P., & Reysenbach, A.-L. (2020). Roadmap for naming uncultivated Archaea and Bacteria. *Nature Microbiology*, *5*, 987–994.

12. Miglietta, M. P. (2023). On the perils of working on nonmodel organisms. *Proceedings of the National Academy of Sciences of the United States of America*, *120*, e2216683120.

13. Britz, R., Conway, K. W., & Rüber, L. (2021). The emerging vertebrate model species for neurophysiological studies is *Danionella cerebrum*, new species (Teleostei: Cyprinidae). *Scientific Reports*, *11*, 18942. https://doi.org/10.1038/s41598-021-97600-0

14. Canhos, D. A. L., Sousa-Baena, M. S., de Souza, S., Maia, L. C., Stehmann, J. R., Canhos, V. P., De Giovanni, R., Bonacelli, M. B. M., Losm, W., & Peterson, A. T. (2015). The importance of biodiversity E-infrastructures for megadiverse countries. *PLoS Biology*, *13*, e1002204.