Highlights and Outcomes of the 2021 Global Community Consultation

Elizabeth R. Ellwood[‡], Andrew Bentley[§], Jutta Buschbom[|], Alex Hardisty[¶,#], Austin Mast[¤], Joe Miller[«], Anna Monfils[»], Gil Nelson[‡], Deborah L Paul[^,¤]

‡ iDigBio, Gainesville, United States of America
§ University of Kansas, Lawrence, KS, United States of America
| Statistical Genetics, Ahrensburg, Germany
¶ Cardiff University, Cardiff, United Kingdom
# School of Computer Science & Informatics, Cardiff University, Cardiff, United Kingdom
¤ Florida State University, Tallahassee, United States of America
« GBIF, Copenhagen, Denmark
» Central Michigan University, Mount Pleasant, United States of America
^ University of Illinois Urbana-Champaign, Urbana, United States of America

Abstract

International collaboration between collections, aggregators, and researchers within the biodiversity community and beyond is becoming increasingly important in our efforts to support biodiversity, conservation and the life of the planet. The social, technical, logistical and financial aspects of an equitable biodiversity data landscape – from workforce training and mobilization of linked specimen data, to data integration, use and publication – must be considered globally and within the context of a growing biodiversity crisis. In recent years, several initiatives have outlined paths forward that describe how digital versions of natural history specimens can be extended and linked with associated data. In the United States, Webster (2017) presented the "extended specimen", which was expanded upon by Lendemer et al. (2019) through the work of the Biodiversity Collections Network (BCoN). At the same time, a "digital specimen" concept was developed by DiSSCo in Europe (Hardisty 2020). Both the extended and digital specimen concepts depict a digital proxy of an analog natural history specimen, whose digital nature provides greater capabilities such as being machine-processable, linkages with associated data, globally accessible information-rich

biodiversity data, improved tracking, attribution and annotation, additional opportunities for data use and cross-disciplinary collaborations forming the basis for FAIR (Findable, Accessible, Interoperable, Reproducible) and equitable sharing of benefits worldwide, and innumerable other advantages, with slight variation in how an extended or digital specimen model would be executed.

Recognizing the need to align the two closely-related concepts, and to provide a place for open discussion around various topics of the Digital Extended Specimen (DES; the current working name for the joined concepts), we initiated a virtual consultation on the discourse platform hosted by the Alliance for Biodiversity Knowledge through GBIF. This platform provided a forum for threaded discussions around topics related and relevant to the DES. The goals of the consultation align with the goals of the Alliance for Biodiversity Knowledge: expand participation in the process, build support for further collaboration, identify use cases, identify significant challenges and obstacles, and develop a comprehensive roadmap towards achieving the vision for a global specification for data integration. In early 2021, Phase 1 launched with five topics: Making FAIR data for specimens accessible; Extending, enriching and integrating data; Annotating specimens and other data; Data attribution; and Analyzing/mining specimen data for novel applications. This round of full discussion was productive and engaged dozens of contributors, with hundreds of posts and thousands of views. During Phase 1, several deeper, more technical, or additional topics of relevance were identified and formed the foundation for Phase 2 which began in May 2021 with the following topics: Robust access points and data infrastructure alignment; Persistent identifier (PID) scheme(s); Meeting legal/regulatory, ethical and sensitive data obligations; Workforce capacity development and inclusivity; Transactional mechanisms and provenance; and Partnerships to collaborate more effectively. In Phase 2 fruitful progress was made towards solutions to some of these complex functional and technical long-term goals. Simultaneously, our commitment to open participation was reinforced, through increased efforts to involve new voices from allied and complementary fields. Among a wealth of ideas expressed, the community highlighted the need for unambiguous persistent identifiers and a dedicated agent to assign them, support for a fully linked system that includes robust publishing mechanisms, strong support for social structures that build trustworthiness of the system, appropriate attribution of legacy and new work, a system that is inclusive, removed from colonial practices, and supportive of creative use of biodiversity data, building a truly global data infrastructure, balancing open access with legal obligations and ethical responsibilities, and the partnerships necessary for success.

These two consultation periods, and the myriad activities surrounding the online discussion, produced a wide variety of perspectives, strategies, and approaches to converging the digital and extended specimen concepts, and progressing plans for the DES -- steps necessary to improve access to research-ready data to advance our understanding of the diversity and distribution of life. Discussions continue and we hope to include your contributions to the DES in future implementation plans.

Keywords

biodiversity data, GBIF, discourse, consultation, digital extended specimen, DES

Presenting author

Elizabeth R. Ellwood

Presented at

TDWG 2021

References

- Hardisty A (2020) What is a Digital Specimen? https://dissco.tech/2020/03/31/what-is-a-digital-specimen/
- Lendemer J, Thiers B, Monfils AK, Zaspel J, Ellwood ER, Bentley A, LeVan K, Bates J, Jennings D, Contreras D, Lagomarsino L, Mabee P, Ford LS, Guralnick R, Gropp RE, Revelez M, Cobb N, Seltmann K, Aime MC (2019) The Extended Specimen Network: A Strategy to Enhance US Biodiversity Collections, Promote Research and Education. BioScience 70 (1): 23-30. https://doi.org/10.1093/biosci/biz140
- Webster M (Ed.) (2017) The extended specimen: emerging frontiers in collections-based ornithological research. CRC Press [ISBN 978-1-4987-2915-4]