Biodiversity Informatics, 20, 2026, pp. 310-313

BIODIVERSITY KNOWLEDGE SHORTFALLS: UNDERSTANDING, QUANTIFYING, AND BRIDGING GAPS

Angela P. Cuervo-Robayo[1,*], Luis E. Escobar[2] and Sara Gamboa[3,4]

[1] *Unidad de Síntesis en Sistemática y Evolución, Instituto de Biología, Universidad Nacional Autónoma de México, Ciudad de México, Mexico*

[2] *Department of Fish and Wildlife Conservation, Virginia Tech, Blacksburg, Virginia, 24061, United States of America*

[3] *Zentralmagazin Naturwissenschaftlicher Sammlungen (ZNS), Martin Luther University Halle-Wittenberg; Halle (Saale), 06108, Germany*

[4] *Centro de Investigación Mariña, Universidade de Vigo, Grupo de Ecoloxía Animal (GEA), MAPAS Lab; Vigo, 36310, Spain*

Abstract. Biodiversity knowledge shortfalls remain a central challenge in biodiversity science, limiting our ability to understand biodiversity patterns, assess environmental change, and inform conservation decisions. This editorial introduces the special issue "Biodiversity Knowledge Shortfalls: Understanding, Quantifying, and Bridging Gaps" published in the journal Biodiversity Informatics. The special issue brings together representative articles on how biodiversity informatics can help reduce knowledge gaps in species distributions, ecological interactions, community diversity, functional traits, and data accessibility and interoperability. Collectively, these articles demonstrate that bridging biodiversity knowledge shortfalls requires more than the accumulation of additional data. Progress in understanding biodiversity patterns depends on interoperable data infrastructures, robust analytical methods, innovative technologies, effective data mobilization strategies, and sustained collaboration among researchers, institutions, biological collections, and citizen scientists.

Editorial

The special issue Biodiversity Knowledge Shortfalls: Understanding, Quantifying, and Bridging Gaps was conceived to stimulate discussion on one of the most persistent challenges facing biodiversity science: the uneven distribution of knowledge about life on Earth. Although unprecedented volumes of biodiversity data have become available through global digitization efforts and data-sharing initiatives (Heberling et al., 2021; Steinke et al., 2025; Wilkinson et al., 2016), substantial gaps remain across the taxonomic, geographic, ecological, evolutionary, and functional dimensions of biodiversity knowledge (Hortal et al., 2015; Pertierra et al., 2025; Sayol et al., 2026). These shortfalls in biodiversity knowledge continue to limit our ability to understand biodiversity patterns and processes, assess the impacts of environmental change and the responses of species and ecosystems to those changes, and implement conservation strategies informed by the best available scientific evidence (Meyer et al., 2016; Nori et al., 2023). Although these knowledge shortfalls have traditionally been treated as distinct dimensions of biodiversity knowledge, many are closely interconnected, and progress in reducing one often contributes to narrowing others. Consequently, addressing biodiversity knowledge gaps requires complementary advances in biodiversity informatics, data mobilization, analytical methods, and biological collections.

The contributions assembled in this first installment illustrate the diversity of methodological approaches currently being developed to understand, quantify, and ultimately reduce biodiversity knowledge gaps. Collectively, articles in this collection highlight how advances in biodiversity informatics, data mobilization, analytical methodologies, and emerging technologies are

*** Corresponding author:** ancuervo@ib.unam.mx

helping transform fragmented information into actionable knowledge for research, conservation, and decision-making through increasingly interoperable biodiversity data infrastructures (Hardisty et al., 2013; Wilkinson et al., 2016).

One group of contributions focuses on digital infrastructures and analytical tools designed to improve the accessibility, interoperability, and usability of biodiversity information. These methodological articles include the Biodiversity Observatory System (BiOS; Roldán et al., 2026), ArctosR (Williams et al., 2026), and approaches connecting biological collections with digitized historical records and artificial intelligence-derived trait information (Bakış, 2026; Weeks et al., 2026). Together, these articles show how biodiversity informatics can transform increasingly diverse and scattered data into integrated and accessible knowledge systems.

For instance, BiOS (Roldán et al., 2026) addresses the lack of interoperability among biodiversity databases through an open-source FAIR-oriented framework that integrates taxonomic, distributional, genetic, ecological, conservation, and legislative information within a modular infrastructure linked by a shared taxonomic backbone. Similarly, ArctosR (Williams et al., 2026) expands the analytical potential of the classic Arctos platform (Cicero et al., 2024) by enabling reproducible programmatic access to occurrence records, specimen traits, tissue samples, and host-parasite and host-pathogen associations.. The novel data-integration tools presented in this special issue are expected to facilitate the integration and reuse of heterogeneous biodiversity information across ecological, evolutionary, conservation, and related research.

Beyond improving access to biodiversity repositories, other contributions exemplify how biodiversity informatics can unlock information that remains underused within biological collections. For example, Weeks et al. (2026) present a specimen-centered workflow to digitize field notebooks and connect their information on distributions, ecological interactions, phenology, taxonomy, and collecting history with natural history collections. Similarly, Bakış (2026) revisits the Raunkiæran shortfall from a biodiversity informatics perspective by proposing a framework that integrates biological collections, image-based phenomics, and artificial intelligence to generate quantitative trait data while preserving provenance, uncertainty, and intraspecific variation. In doing so, these contributions provide a roadmap for transforming digitized collections into dynamic sources of functional biodiversity knowledge. The digitalization articles in this special issue show how existing collections and their associated documentation can become dynamic sources of biodiversity knowledge.

A second group of articles focuses on the generation, curation, and mobilization of biodiversity data, showing that both data quality practices and emerging information sources strongly influence biodiversity inference and our ability to address existing knowledge shortfalls. For instance, Ronquillo et al. (2026) examine how occurrence data curation affects estimates of plant bioclimatic niches, emphasizing the importance of transparent and context-dependent curation protocols when addressing both the Wallacean shortfall, defined as the lack of knowledge of species distributions (Lomolino, 2004), and the Hutchinsonian shortfall, defined as the lack of knowledge about species' tolerances to abiotic conditions (Hortal et al., 2015). From a complementary perspective, Shepherd and Perrie (2026) illustrate how observation-based platforms such as iNaturalist function as sources of occurrence data and also as valuable tools for guiding specimen acquisition, linking community-generated observations with collection development to reduce Wallacean shortfalls while strengthening the long-term scientific value of biological collections.

A third group of articles examines gaps in species distributions, ecological interactions, and community diversity, providing approaches to identify missing information, quantify uncertainty, and guide future sampling. For instance, from an analytical perspective Barbosa (2026) proposes spatially adaptive calibration-area delimitation strategies to reduce uncertainty arising from uneven geographic sampling in species distribution modelling. At a broader ecological scale, Villa and Dáttilo (2026) examine the Eltonian shortfall in plant-hummingbird interaction networks across the Americas, revealing substantial geographic, environmental, and taxonomic biases and a large proportion of undocumented species interactions. Benites et al. (2026) focus on the Linnean shortfall by applying DNA metabarcoding to uncover the hidden diversity of leaf-litter arthropod communities in Mexican tropical dry forests. Their findings are a key example of the use of molecular approaches to reveal extensive undocumented biodiversity and generate essential baseline information for future ecological and conservation studies. In a final example, BioDyn (Herrador et al., 2026) provides a framework for exploring the spatiotemporal structure of species occurrences, enabling users to evaluate sampling effort, identify distributional biases, explore seasonal dynamics, including migratory systems and temporal variation, and detect environmental information gaps prior to formal modelling. Together, these contributions explain how complementary analytical and molecular approaches can reveal where biodiversity knowledge remains incomplete and help prioritize future research and data-collection efforts.

This collection of articles offer a broader conceptual reflection on the relationship between biodiversity knowledge and conservation practice. Drawing on examples from Mexico, Soberón (2026) argues that although advances in biodiversity informatics and the growing availability of digital biodiversity data have substantially reduced several knowledge shortfalls, information alone is insufficient to achieve effective conservation outcomes. Instead, successful conservation requires the integration of scientific understanding, stakeholder participation, governance processes, and negotiation among actors operating across multiple social and spatial scales.

Collectively, articles in this special issue show that bridging biodiversity knowledge shortfalls requires more than the accumulation of additional data. Progress in our knowledge of biodiversity will depend on the development of interoperable infrastructures, robust analytical methods, innovative technologies, effective data mobilization strategies, and stronger collaboration among researchers, institutions, biological collections, and citizen scientists. By bringing together these complementary perspectives, this special issue highlights both the complexity of biodiversity knowledge gaps and the growing array of tools available to address them. We hope that the contributions presented here are the contemporary advances in biodiversity informatics and could stimulate further discussion, foster new collaborations, and inspire continued efforts to make biodiversity knowledge more complete, accessible, and actionable for science, conservation, and policy.

Acknowledgments

LEE was supported by a research fellowship by the OECD Co-operative Research Programme: Sustainable Agricultural and Food Systems (CRP), and grants from the National Science Foundation STAR (2620646), CAREER (2235295), MCA (2322213), and HEGS (2116748) awards, and National Institutes of Health (K01AI168452). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders.

Literature Cited

Bakış, Y. 2026. Bridging the Raunkiæran Shortfall Through Specimen-Anchored, AI-Derived Trait Infrastructure. Biodivers. Inform, 20(2): 183–199. https://journals.ku.edu/jbi/article/view/25068

Barbosa, A. M. 2026. Spatially adaptive background delimitation strategies for species distribution models with unevenly surveyed occurrences. Biodivers. Inform, 20(2): 183-199. https://journals.ku.edu/jbi/article/view/25045

Benites, P., Astudillo-Pavón, K., Zárate-Hernández, F., Bautista-Briceño, N., Penteado-Dias, A. M., Zaldívar-Riverón, A., García-Bautista, A. 2026. Exploring the tip of the biodiversity iceberg: Spatio-temporal diversity and structure of arthropod leaf-litter communities in protected Mexican tropical deciduous forests revealed by metabarcoding. Biodivers. Inform, 20(2): 89–101. https://journals.ku.edu/jbi/article/view/24878

Cicero, C., Koo, M. S., Braker, E., Abbott, J., Bloom, D., Campbell, M., Cook, J. A., Demboski, J. R., Doll, A. C., Frederick, L. M., Linn, A. J., Mayfield-Meyer, T. J., McDonald, D. L., Nachman, M. W., Olson, L. E., Roberts, D., Sikes, D. S., Witt, C. C., Wommack, E. A. 2024. Arctos: Community-driven innovations for managing natural and cultural history collections. PLOS ONE, 19(5): e0296478. https://doi.org/10.1371/journal.pone.0296478

Hardisty, A., Roberts, D., The Biodiversity Informatics Community. 2013. A decadal view of biodiversity informatics: Challenges and priorities. BMC Ecology, 13(1): 16. https://doi.org/10.1186/1472-6785-13-16

Heberling, J.M., Miller, J.T., Noesgaard, D., Weingart, S.B. and Schigel, D., 2021. Data integration enables global biodiversity synthesis. Proceedings of the National Academy of Sciences, 118(6), p.e2018093118. https://doi.org/10.1073/pnas.2018093118

Herrador, T. A., Reati, C. A., Huais, P. Y., Osorio-Olvera, L., Contreras-Díaz, R. G., Peluc, S. I., Nori, J. 2026. BioDyn: A tool for the spatiotemporal analysis of intra-specific geographic data. Biodivers. Inform, 20(2): 275-285. https://journals.ku.edu/jbi/article/view/25258.

Hortal, J., De Bello, F., Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., Ladle, R. J. 2015. Seven shortfalls that beset large-scale knowledge of biodiversity. Annu. Rev. Ecol. Evol. Syst., 46(1): 523–549. https://doi.org/10.1146/annurev-ecolsys-112414-054400

Meyer, C., Weigelt, P., Kreft, H. 2016. Multidimensional biases, gaps and uncertainties in global plant occurrence information. Ecology Letters, 19(8): 992–1006. https://doi.org/10.1111/ele.12624

Nori, J., Maas, B., Brum, F., Villalobos, F. 2023. Addressing Knowledge Shortfalls in Conservation Science: A long way to go, as quickly as possible. Biol. Conserv, 287, 110314. https://doi.org/10.1016/j.biocon.2023.110314

Lomolino, M.V. 2004. Conservation biogeography. Frontiers of Biogeography: new directions in the geography of nature (ed. by M.V. Lomolino and L.R. Heaney). Sinauer Associates, Sunder- land, Massachusetts.

Roldán, A., Golomb Durán, T., Far, A. J., Capa, M., Arboleda, E., Cancellario, T. 2026. BiOS: An Open-Source Framework for the Integration of Heterogeneous Biodiversity Data. Biodivers. Inform, 20(2): 275-285. https://journals.ku.edu/jbi/article/view/25254.

Ronquillo, C., Stropp, J., Hortal, J. 2026. Data curation on species occurrence records placed in unrecognized distributions shapes our understanding of plant bioclimatic niches. Biodivers. Inform, 20(2): 89–101. https://journals.ku.edu/jbi/article/view/24925.

Shepherd, L., Perrie, L. (2026). Leveraging the Citizen Science Platform iNaturalist for the Targeted Collection of Natural History Specimens. Biodivers. Inform, 20(2): 102-110. https://journals.ku.edu/jbi/article/view/24994

Soberón, J. 2026. Conservation science and the knowledge shortfalls. Biodivers. Inform, 20(2): 111-117. https://journals.ku.edu/jbi/article/view/24814

Steinke, D., Gemeinholzer, B., Martínez-Meyer, E., Noesgaard, D., Young, A., Schigel, D. (2025). Globally aggregated biodiversity data impact predictive and descriptive research. Proceedings of the National Academy of Sciences, 122(50): e2519119122. https://doi.org/10.1073/pnas.2519119122

Villa, A. R., Dáttilo, W. 2026. Finding the missing links: Knowledge shortfalls in plant–hummingbird interaction networks across the Americas. Biodivers. Inform, 20(2): 89–167-182. https://journals.ku.edu/jbi/article/view/25141.

Weeks, A., Best, J. H., Mann, W., Nazaire, M., Niñi, A., Sheik, M., Shenoy, K. B., Fahringer, A. T. 2026. Adapting digital humanities tools for biodiversity science: A case study in field notebook digitization and the Extended Specimen Network. Biodivers. Inform, 20(2): 118-131. https://journals.ku.edu/jbi/article/view/25170.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P. A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velerop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B. 2016. The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3(1): 160018. https://doi.org/10.1038/sdata.2016.18

Williams, H. R., Colella, J. P., Barve, V., Koo, M. S., Cobos, M. E. 2026. ArctosR: An R package to interface with the Arctos Museum Database API. Biodivers. Inform, 20(2): 200-218. https://journals.ku.edu/jbi/article/view/25022