

Community action: Planning for specimen management in funding proposals

Andrew Bentley , Barbara Thiers , William E. Moser , Gregory J. Watkins-Colwell , Breda M. Zimkus , Anna K. Monfils , Nico M. Franz , John M. Bates , Kyria Boundy-Mills , Michael W. Lomas , Elizabeth R. Ellwood , Sinlan Poo , Dori L. Contreras , Michael S. Webster , Gil Nelson  and Jyotsna L. Pandey 

Andrew Bentley (abentley@ku.edu) is affiliated with the Biodiversity Institute at the University of Kansas, in Lawrence, Kansas, in the United States. Barbara Thiers is affiliated with the New York Botanical Garden, in the Bronx, New York, in the United States. William E. Moser is affiliated with the National Museum of Natural History, at the Smithsonian Institution, in Suitland, Maryland, in the United States. Gregory J. Watkins-Colwell is affiliated with the Yale Peabody Museum, in New Haven, Connecticut, in the United States. Breda M. Zimkus is affiliated with the Museum of Comparative Zoology at Harvard University, in Cambridge, Massachusetts, in the United States. Anna K. Monfils is affiliated with the Department of Biology at Central Michigan University, in Mount Pleasant, Michigan, in the United States. Nico M. Franz is affiliated with the School of Life Sciences at Arizona State University, in Tempe, Arizona, in the United States. John M. Bates is affiliated with the Negaunee Integrative Research Center, at the Field Museum, in Chicago, Illinois, in the United States. Kyria Boundy-Mills is affiliated with the Phaff Yeast Culture Collection, at the University of California Davis, in Davis, California, in the United States. Michael W. Lomas is affiliated with the National Center for Marine Algae and Microbiota, at the Bigelow Laboratory for Ocean Sciences, in East Boothbay, Maine, in the United States. Elizabeth R. Ellwood is affiliated with iDigBio, at the University of Florida, in Gainesville, Florida, in the United States. Sinlan Poo is affiliated with the Department of Conservation and Research at the Memphis Zoological Society, in Memphis, Tennessee, in the United States. Dori L. Contreras is affiliated with the Perot Museum of Nature and Science, in Dallas, Texas, in the United States. Michael S. Webster is affiliated with the Cornell Lab of Ornithology, at Cornell University, in Ithaca, New York, in the United States. Gil Nelson is affiliated with iDigBio, at the Florida Museum of Natural History, at the University of Florida, in Gainesville, Florida, in the United States. Jyotsna L. Pandey is affiliated with the American Institute of Biological Sciences, in Herndon, Virginia, in the United States.

Field- and lab-based research projects have the potential to amplify their impact through the generation, collection, and deposition of specimens into established repositories (Peterson 2010, Ball-Damerow et al. 2019, Boundy-Mills et al. 2019, NASEM 2020, Heberling et al. 2021, Hardisty et al. 2022). Specimens are broadly defined as organisms (including those that are genetically modified), parts of organisms, fossils and trace fossils, living organisms (e.g., microbes, algae), and nonbiological environmental samples (e.g., water, soil, rocks). Preservation of specimens is essential for two elements of modern science: validating past results and building on past discoveries. There have been calls for increased vouchering or deposition of biodiversity specimens into established repositories across the life sciences, including natural history disciplines and beyond (Funk et al. 2005, Peterson 2010, Schilthuizen et al. 2015, Buckner et al. 2021, Colella et al. 2021). Timely and properly archived physical specimens and their digitized data are accessible to an ever-increasingly diverse audience, support transparency and reproducibility in research, and provide critical information that can stimulate new avenues of research.

The biodiversity science community has also been actively engaging in interdisciplinary conversations about creating and sustaining open data resources that are FAIR (for *findable, accessible, interoperable, and reusable*; Wilkinson et al. 2016). The community is working toward the creation of an extended specimen network (Lendemer et al. 2020), which will rely on a dedicated workforce and infrastructure to maintain in-network existing specimens and data, engage with other data sources, and ensure that newly collected specimens and links seamlessly enter the data network. The societal value of managing extended specimens (Lendemer et al. 2020) is very closely linked to our present and projected capability to mobilize local to global specimen access and to build powerful tools to publish and share data under FAIR and CARE

principles. (for collective benefit, authority to control, responsibility, and ethics; Carroll, et al. 2020).

Two recent federal initiatives, the Nelson Memo (OSTP 2022) and the White House report *Vision, Needs, and Proposed Actions for Data for the Bioeconomy Initiative* (OSTP 2023), prioritize FAIR data practices and open science concepts and thus align well with objectives of the broadly defined biodiversity community. The Nelson Memo, issued by the White House Office of Science and Technology Policy, calls on federal agencies to make publications and supporting data resulting from federally funded research freely available and publicly accessible immediately, without any embargo. The White House bioeconomy report describes the data and computational infrastructure needs required for a robust US bioeconomy. Importantly, the report also calls for enhanced specimen collection and digitization to leverage biodiversity to achieve bioeconomic goals.

Another significant development has been the enactment of the CHIPS and Science Act of 2022 (PL 117–167), which makes historic investments in science and technology, including providing a robust reauthorization for the National Science Foundation (NSF). Notably, the law supports a recommendation from the 2020 National Academies of Sciences, Engineering, and Medicine (NASEM) report *Biological Collections*, calling for funding agencies to “require a specimen management plan for all research proposals that include collecting or generating specimens that describe how the specimens and associated data will be accessioned into and permanently maintained in an established biological collection” (NASEM 2020). Furthermore, some biological collections are also defined as genetic resources; therefore, their stewardship and use are governed by international conventions such as the Nagoya Protocol (www.cbd.int/abs) and Marine Biodiversity of Areas Beyond National Jurisdiction (www.un.org/bbnj).

The Biodiversity Collections Network (BCoN, <https://bcon.aibs.org>), a national initiative that emerged from an NSF grant awarded to the American Institute of Biological Sciences (AIBS), has facilitated community discussions in partnership with the US Culture Collection Network (USCCN), an NSF funded research coordination network, about the need for a specimen management plan (SMP) requirement. This included a virtual discussion held in February 2023 where many voices in the community expressed support for such an endeavor (BCoN et al. 2023).

Successful long-term specimen care, curation, and data storage require expertise and a detailed and specific management plan. Biological collections are composed of an array of living and preserved organisms (including fossils), and their long-term maintenance requires curatorial expertise and knowledge of preservation techniques that are specific to the type of organism to maximize future research. Although specimens are closely linked with their associated data, physical objects differ widely in their requirements for physical care, method of preservation, and curatorial needs across disciplines. Specimens therefore have different and more diverse management practices when compared with data derivatives, especially concerning infrastructure and budget. Many federal agencies emphasize the need for and importance of specimen management. Box 1 lists some examples of sample management guidance currently in place at various federal agencies, but there is a need for standardized policies across the federal government that formalize the relationship between the researcher who generates specimens and the institution entrusted to conserve, curate, and provide access for research, innovation, intellectual property development, and advancing the bioeconomy. A clear and standardized SMP supports the data and integration required to achieve a truly extended specimen.

Recommendations for specimen management plan implementation

Implementation of a sound SMP has broad value for researchers, collections, their host institutions, publishers, and, in the case of publicly funded research, taxpayers as well (see box 2). Furthermore, additional support and guidance are often necessary to ensure that collections follow the CARE Principles for Indigenous Data Governance, which may limit or expand the availability of specimens and associated data. To this end, we recommend that funding agencies require SMPs for biological specimens collected or generated during government-funded research and that SMPs provide detailed information regarding the requirements for collection, digitization, curation, long-term maintenance, and funding of accessioned specimens as a part of all research projects in which specimens of any type are generated or collected.

Funding agencies would also be urged to provide specific guidance on specimen acquisition and deposition through further supplementation of current data management plan guidance. The increased scope of such a document would serve as an instrument to incentivize and promote the development, maintenance, coordination, and enhancement of services built around supporting physical and cyber infrastructures for specimens and their associated data.

Finally, we recommend that grant-receiving institutions support their workforce and student researchers in adhering to best SMP practices, whereas specimen-receiving institutions

are funded through SMP mechanisms to provide information and infrastructure needed to support the deposition of specimens.

Suggested elements of a specimen management plan

Proposal submitters should engage and work collaboratively with a selected repository to produce an SMP. This communication should occur well in advance of the submitted proposal date and should include the following: (1) the name of the collections repository that has agreed to maintain the specimens along with a statement of collaboration from the repository; (2) the type and anticipated number of specimens, variety of preparations, and associated collection metadata and extended specimen data (including genotype, phenotype, images, sequences, etc.) that would be deposited following all national and international collecting, import, and export permitting agreements (Nagoya Protocol, etc.) that govern deposition of the collections; (3) plans for collecting specimens and preserving them within a repository that is in line with established best practices for the relevant organisms (including the expectations of specimen curation and care), to ensure long-term utility that complements recommendations for the digital extended specimen (Hardisty et al. 2022); (4) an estimated budget of the funding required by the repository to curate, digitize, and care for the material once received, and how those costs will be covered; (5) a plan for making the specimens available to the research community, such as through data publication in the collection's online catalog or other tools.

This minimum set of information could be used by reviewers to judge the soundness of the SMP for physical samples and their derivatives, which are equally important to preserve, if not more so, than the associated data addressed in the DMP.

A coordinated approach is needed to engage the living and preserved collections and their associated research communities to develop requirements and best practices for the collection, deposition, accession, and sustained curation of specimens and associated data that result from publicly funded research. In collaboration with NSF, the combined expertise of members of the Society for the Preservation of Natural History Collections (SPNHC), the AIBS, the Natural Science Collections Alliance, USCCN, BCoN, iDigBio (for *Integrated Digitized Biocollections*), the Global Biodiversity Information Facility, and others, could gather, create, and publish documents outlining the best practices for the collection, preservation or cultivation, and long-term maintenance of samples, which would be subjected to community review. To be of maximum utility, such documents should include cost projections for the curation and digitization of various categories of specimens and preparations. An online resource, such as the SPNHC website (<https://spnhc.org/resources>), could serve as a clearinghouse for publications, templates, and other resources. In addition, grant writers could use tools designed to help with these sections of grant proposals, such as DMP-Tool (<https://dmptool.org>) and ezDMP (<https://ezdmp.org>), to incorporate these guidance documents into their workflows for semiautomated creation of data management plans. Coordinated efforts at the national level, such as a biological collections action center, as was proposed in the 2020 NASEM Report and the CHIPS and Science Act, could provide the infrastructure necessary for the creation and dissemination of relevant resources.

Box 1. Examples of current sample management guidance at federal agencies.

At the **National Science Foundation**, the current NSF Proposal and Award Policies and Procedures Guide (NSF 2023a) allows proposers to request funds for “storage, preservation, documentation, indexing, etc., of physical specimens, collections, or fabricated items.” As it is currently written, the data management plan (DMP) guidance document associated with NSF **Biological Sciences (BIO)** directorate proposals (NSF 2020) mentions specimen curation but does not provide specific guidance or require sufficient information from a proposed project to ensure specimen deposition, digitization, care, and curation. However, following the passage of the CHIPS and Science Act, NSF BIO has implemented new specimen management requirements as part of select solicitations that provide the following guidelines: “PIs who propose to generate or collect specimens in the execution of their research that will be vouchered or deposited into some type of resource designed for enabling their reuse are required to include in the Data Management Plan a section titled “Specimen Management Plan,” which should include a description of which specimens and associated data will be permanently accessioned into and maintained in an established repository or other similar resources. PIs planning to deposit specimens, as well as those who do not, should indicate how their plans are consistent with community norms, best practices, or standards.”

Divisions within the **Geological Sciences Directorate (GEO)** at NSF have provided more informative resources and guidance on sample management. For example, the Division of Ocean Sciences provides detailed sample archiving requirements for physical geological samples but only encourages the archiving of voucher biological specimens (www.nsf.gov/pubs/2017/nsf17037/nsf17037.jsp). The Division of Earth Sciences (EAR) recently announced updates to its Data and Sample Policy (www.nsf.gov/pubs/2023/nsf23131/nsf23131.jsp), which articulates expectations for the content of DMPs, including requirements for when and how data and samples must be made publicly accessible after collection (NSF 2023b). EAR has also made available a list of resources to facilitate adherence to the new requirements (NSF 2023c).

The US **Long Term Ecological Research Network (LTER)**—a cross-cutting program administered through BIO, GEO, and the Directorate for Social, Behavioral, and Economic Sciences—requires that for proposals that include specimen collection (e.g., organisms, parts of organisms, fossils including trace fossils, microbial isolates) “the Data Management Plan must include a description of how the specimens and associated data will be accessioned into and maintained in an established biological collection” (www.nsf.gov/pubs/2024/nsf24520/nsf24520.htm?org=NSF).

At the **Department of the Interior (DOI)**, the **National Park Service (NPS)** has outlined curatorial requirements for specimen collection at national parks that apply to both NPS and non-NPS researchers—for example, at Denali National Park and Preserve (www.nps.gov/dena/learn/nature/research-collecting.htm) and Acadia National Park (www.nps.gov/acad/learn/nature/collecting-specimens.htm).

The US **Geological Survey (USGS)** provides resources for collection management (www.usgs.gov/programs/national-geological-and-geophysical-data-preservation-program/resources-collection).

The US **Fish and Wildlife Service (USFWS)** provides guidance on the collection and donation of fish, wildlife, and plant specimens (www.fws.gov/policy-library/701fw5, www.fws.gov/policy-library/e1126fw2).

The US **Department of Agriculture (USDA)** provides guidance on germplasm deposition (www.ars.usda.gov/plains-area/fort-collins-co/center-for-agricultural-resources-research/paagrpru/docs/plants/pages/deposit-germplasm). The **National Institute of Food and Agriculture (NIFA)** provides management and access policies for institutional scientific collections that are financially supported by NIFA (USDA 2016). The Agricultural Research Service provides guidance on scientific collections management and access (USDA 2014).

Evaluation of specimen management plans

Reviewers of funding requests for projects that include the generation of specimens may be unfamiliar with sampling and curation protocols. For that reason, we suggest a potential rubric for the evaluation of an SMP that includes responses to the following questions: (1) Is an appropriate repository named in the proposal, and does the proposal contain evidence that the repository has agreed to house the specimens? (2) If the agreement is for a particular term, are there clear plans in place for the preservation, care, curation, and distribution of specimens beyond that term? (3) Does the proposal demonstrate awareness of standards for discipline-specific preparation and storage of the specimens, and is the work needed to accomplish this accounted for in the task analysis and the budget of the proposal? (4) Is there a provision for digitizing and sharing specimen data through a generally

accessible portal? (5) Does the proposal indicate that the grant recipient will provide all the necessary legal documentation (e.g., associated permits or accession documents) to the repository outlining ownership or stewardship conditions? (6) For collections made outside of US national borders, does the proposal outline the specific actions the researchers will take to comply with import or export requirements, the Nagoya Protocol, biosafety, biosecurity, and other relevant regulations? (7) Does the proposal contain an explicit accounting of costs for acquisition, processing, and care, with a consistent, transparent calculation?

Conclusions

In summary, for collections willing to partner with new research projects, SMPs will maximize the return on long-term investments in collections infrastructure and ensure that they are

Box 2. Value-added benefits of a specimen management plan.

Robust SMPs increase the reproducibility, utility, and long-term value of federally funded, specimen-based research that benefits numerous stakeholders in the following ways:

Researchers: Communication between researchers and repositories prior to lab and fieldwork and during the grant writing phase would ensure institution-specific protocols and best practices for the generation, collection, and preservation of specimens, data, and products are followed. It would lead to more timely curation of specimens and aid in meeting reproducibility and data availability requirements of publishers and grant agencies when disseminating project results. It would also facilitate the exchange of information on relevant state, national, and international collecting and permitting laws and regulations, institutional animal care and use committee or institutional review board protocols, USDA-APHIS (Animal and Plant Health Inspection Service), biosafety, biosecurity, and ethical guidelines. Finally, it would enable the repository to provide inventories of existing specimens and data in their or other institutions, as well as information about complementary research projects, thereby preventing research duplication, ensuring ethical collecting and providing collaboration opportunities.

Collections and Institutions: Collections would be funded at a negotiated rate to support the workforce and equipment needed to catalog, digitize, maintain, and curate specimens for an agreed-on period, following NSF's mandate to fund the entire research endeavor. The funds would be directed to collections that are in active use and capable of growth. Likewise, SMPs may incentivize institutional leaders and communities to build more robust, sustainable, and powerful infrastructures and services around their managed specimens, thereby maximizing the specimens' future scientific and societal impact. By disseminating funds at the local and regional level, agencies would create an equitable allocation of resources in support of a more sustainable infrastructure for the digital extended specimen. Collections would also receive high-quality specimens that were legally collected and preserved in a format compatible with their storage and data that adhere to standards and best practices compatible with their collection management systems.

New investments in recruitment and training, whether directly through grant funds or through complementary institutional funds made possible by these additional funds, would sustain and diversify the collections workforce, per the NSF's stated objectives of increasing equity for excellence in STEM (science, technology, engineering, and math) while simultaneously promoting diversity, equity and inclusion improvements in the collections community.

Funding sources: Public funding agencies would receive a larger and earlier return on their investment in collections to leverage future research. They would have better metrics to understand the landscape of repositories actively involved in the curation of specimens generated through research. And small and underresourced collections that receive accessions from taxpayer-funded research would receive federal funding through SMP mechanisms in amounts that reflect the collection's value to federally funded research.

Additional stakeholders, such as publishers, data users, the general public, and early career scientists, also benefit from SMPs in ways similar to the groups above. Collectively, this increases the reproducibility and utility of funded, specimen-based research.

adequately safeguarded and ethically curated for the support of critical research, education, and innovation into the future. Both researchers and collections professionals have expressed the concern that an SMP requirement would create additional burdens for their work, most recently in the February 2023 virtual discussion led by AIBS, BCoN, USCCN, and the Natural Science Collections Alliance (BCoN et al. 2023). However, we assert that a detailed and mutually agreed-on plan would have the opposite effect. With knowledge of accepted best practices for specimen preparation and documentation, researchers can incorporate realistic time and budget commitments for this work. Receiving collections institutions can feel confident of the value resulting from accessions and can incorporate these efficiently. Developing a consensus on the requirements for an SMP and how to evaluate such plans is key for the efficiency of this process and the accrual of value-added benefits described above.

References cited

Ball-Damerow JE, Brenskelle L, Barve N, Soltis PS, Sierwald P, Bieler R, LaFrance R, Ariño AH, Guralnick RP. 2019. Research applications of primary biodiversity databases in the digital age. *PLOS ONE* 14: e0215794. <https://doi.org/10.1371/journal.pone.0215794>.

[BCON] Biodiversity Collections Network, American Institute of Biological Sciences, US Culture Collection Network, and Natural Science Collections Alliance et al. Kick-off webinar and discussion on the need for a specimen management plan requirement. BCON (7 February 2023). <https://bcon.aibs.org/event/kick-off-webinar-discussion-on-the-need-for-a-specimen-management-plan-requirement>.

Boundy-Mills K, et al. 2020. Preserving US microbe collections sparks future discoveries. *Journal of Applied Microbiology* 129: 162–174. doi:10.1111/jam.14525.

Buckner JC, Sanders RC, Faircloth BC, Chakrabarty P. 2021. Science Forum: The critical importance of vouchers in genomics. *eLife* 10: E68264. <https://doi.org/10.7554/eLife.68264>.

Carroll SR, et al. 2020. The CARE Principles for indigenous data governance. *Data Science Journal* 19: 43. <https://doi.org/10.5334/dsj-2020-043>.

Colella JP, Stephens RB, Campbell ML, Kohli BA, Parsons DJ, Mclean BS. 2021. The open-specimen movement. *BioScience* 71: 405–414. <https://doi.org/10.1093/biosci/biaa146>.

Funk VA, Hoch PC, Prather AL, Wagner WL. 2005. The importance of vouchers. *Taxon* 54: 127–129. <https://doi.org/10.2307/25065309>.

Hardisty AR, et al. 2022. Digital extended specimens: Enabling an extensible network of biodiversity data records as integrated digital

objects on the Internet. *BioScience* 72: 978–987. <https://doi.org/10.1093/biosci/biac060>.

Heberling JM, Miller JT, Noesgaard D, Weingart SB, Schigel D. 2021. Data integration enables global biodiversity synthesis. *Proceedings of the National Academy of Sciences* 118: e2018093118. <https://doi.org/10.1073/pnas.2018093118>.

Lendemer J, et al. 2020. The extended specimen network: a strategy to enhance US biodiversity collections, promote research and education. *BioScience* 70: 23–30. <https://doi.org/10.1093/biosci/biz140>

[NASEM] National Academies of Sciences, Engineering, and Medicine. 2020. *Biological Collections: Ensuring Critical Research and Education for the 21st Century*. National Academies Press. <https://doi.org/10.17226/25592>.

[NSF] National Science Foundation. 2020. Directorate for Biological Sciences: Updated Information about the Data Management Plan Required for All Proposals. NSF. www.nsf.gov/bio/pubs/BIODMP_Guidance.pdf.

[NSF] National Science Foundation. 2023a. *Proposal and Award Policies and Procedures Guide*. NSF. Report no. (PAPPG) NSF 23-1. <https://beta.nsf.gov/policies/pappg/23-1-ch-2-proposal-preparation#2D2fvib>.

[NSF] National Science Foundation. 2023b. EAR Data and Sample Policy, Division of Earth Sciences. NSF. www.nsf.gov/geo/geo-data-policies/ear/ear-data-policy-jul2023.pdf.

[NSF] National Science Foundation. 2023c. Resources for EAR Data and Sample Policy, Division of Earth Sciences. NSF. www.nsf.gov/geo/geo-data-policies/ear/ear-data-policy-resources-jul2023.pdf.

O'Brien KM, et al. 2022. The time is right for an Antarctic biorepository network. *Proceedings of the National Academy of Sciences* 119: e2212800119. doi:10.1073/pnas.2212800119

[OSTP] Office of Science and Technology Policy. 2022. Memorandum for the Heads of Executive Department and Agencies: Ensuring Free, Immediate, and Equitable Access to Federally Funded Research. Executive Office of the President. www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf.

[OSTP] Office of Science and Technology Policy. 2023. Vision, Needs, and Proposed Actions for Data for the Bioeconomy Initiative. Executive Office of the President. www.whitehouse.gov/wp-content/uploads/2023/12/FINAL-Data-for-the-Bioeconomy-Initiative-Report.pdf.

Peterson AT. 2010. The critical role of permanent voucher specimens of hosts and vectors in public health and epidemiology. *Emerging Infectious Diseases* 16: 341–342. <https://doi.org/10.3201/eid1602.091241>.

Schilthuizen M, Vairappan C, Slade E, Mann D, Miller J. 2015. Specimens as primary data: Museums and "open science." *Trends in Ecology and Evolution* 30: 237–238. <https://doi.org/10.1016/j.tree.2015.03.002>.

[USDA] US Department of Agriculture. 2014. Agricultural Research Service Scientific Collections Management and Access Policy September 2014. USDA. www.ars.usda.gov/ARSUserFiles/00000000/NPS/ARS%20Policy%20for%20Scientific%20Collections.pdf

[USDA] US Department of Agriculture. 2016. National Institute of Food and Agriculture Permanent Scientific Collections Management and Access Policy. USDA. https://iwgsc.nal.usda.gov/sites/default/files/NIFA_Policy_for_Scientific_Collections_Draft_to OSTP_10-7-2016.pdf.

Wilkinson M, et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3: 160018. <https://doi.org/10.1038/sdata.2016.18>.