

Spring 2026 STEAC Meeting Report

Introduction

The STEAC met on March 19th and 20, 2026 as a quorum of 9 members total, with six members in-person (Kelly Aho, Jennifer Edmonds, John Kochendorfer, Sydne Record, Daniel Rubenstein, Adrienne Sponberg) and three attending virtually (Kin Onn Chan, Samapriya Roy, Xi Yang). Fourteen NEON-Battelle staff attended various parts of the agenda (Hojun Song, Kelsey Yule, Tanya Maslak, Callie Puntenney, Kate Thibault, Darcy Gora, Paula Mabee, Chris McKay, Kim Nitschke, Claire Lunch, Jeff Coleman, Chau Tran, Steve Jacobs, Mike Kuhlman). The meeting was in-person and held at the NEON Headquarters in Boulder, Colorado. The following topics were discussed: 1) NEON Biorepository vision, sample use, and researcher derived data considerations, 2) Incentivizing community members to participate in data-driven working groups focused on operational advancements or targeted research questions, 3) Future resource investment strategies for community interactions with NEON staff. Two NSF officers (Dr. Charlotte Roehm and Dr. James Deshler) joined the conversation for a 1-hour Q&A session focused on these same three topics. STEAC recommendations are highlighted in italic text at the end of each summary.

Day One

Welcome and Stage Setting

Summary:

After introductions, NEON Director, Dr. Paula Mabee, updated the STEAC members on the NEON leadership's response to past recommendations from the advisory committee. These changes included moving forward with establishing a login process for all individuals accessing NEON data, adopting licensing for NEON use, and termination of select, infrequently used NEON data products to maintain budget integrity. After approval of minutes from February 2026, the meeting goals were reviewed and the agenda continued with a discussion of the Biorepository.

NEON Biorepository I: Vision

Summary:

Dr. Hojun Song (PI of NEON Biorepository and Professor at Arizona State University as of 2025) and Dr. Kelsey Yule (Senior Data Science Specialist) presented a vision for the NEON Biorepository located in Tempe, AZ. They set the stage by providing the original mandate for the Biorepository which was to maintain a long-term community resource optimized for monitoring and forecasting ecological responses to environmental change at a continental scale. The core responsibilities of the Biorepository include sample receipt and accession, curation and vouchering, long-term preservation, response to

external research communities, loans, specimen data availability, permitting/compliance, and quality control. Current staff include three full-time collections managers, one full-time sample preparator, and three full-time bioinformaticians. A brief history of the Biorepository, a virtual video tour of the facility, and an overview of the types of specimens maintained followed. The Biorepository is the most rapidly growing collection in the world and is unique in that all samples are completely digitized upon arrival into the Symbiota Data Portal, allowing for sample and specimen exploration and linkage to other NEON data products. The small mammal collections are the most extended as they include hair, DNA extracts, and blood, ear tissue, and fecal samples. In the coming year, NEON is making a huge investment in a capacity enhancement renovation of the Biorepository, so that the rapidly growing collections can be moved out of the space it is occupying in the ASU Natural History Museum, and into an immediately adjacent building location. The facility differs from a traditional natural history collection in terms of mission, sampling strategy, taxonomic focus, curation style, sample use policy, asset diversity, storage environment, data integration, and operational model.

The vision outlines how the Biorepository could move from a passive operator to an active community-driven research accelerator via three areas of unrealized potential for the Biorepository. One was the potential to reveal dark data in the bulk samples (e.g., invertebrate bycatch, macroalgae, microalgae, benthic macroinvertebrates, and zooplankton) by the use of emerging technologies (e.g., computer vision, next generation based, low-cost DNA sequencing, AI). The outcome would be to increase the user community and provide opportunities for innovation across disciplines. A second potential is to demonstrate the power of the hyper-extended specimen. Associated ecological data are powerful, but not always easy to integrate. Research could be accelerated by making trait data easier to extract at scale with computer vision and AI, leveraging DNA grade samples to enable genomic research, and investing in the Biorepository data portal to more seamlessly integrate these data layers. The outcome would be to increase synergy with the natural history collections community, extend value of data to ecological and evolutionary communities, and to accelerate continental scale research. A third unrealized potential is the capacity for the Biorepository to function as a physical hub for workforce development as a dynamic laboratory for training and a gateway to data literacy. Such an approach could accelerate research if the Biorepository were to host NSF postdocs, a Research Experience for Undergraduates site, and provide more outreach to K-12 groups. The outcome would be increased visibility and sample use by the community of next generation NEON users. Realizing this vision would require strategic planning, strategic investments, more personnel, and greater visibility.

Recommendations: The STEAC applauds the Biorepository's vision to transition from a passive operator to a community driven research accelerator. The STEAC recommends

that the Biorepository work on a strategic plan to help realize this vision that considers short, medium, and long-term goals and ways to achieve and assess them, including consideration of sustainability of the Biorepository beyond the 30-year horizon of NEON's funding.

NEON Biorepository II: Sample Use

Summary:

The NEON Biorepository facilitates the use of samples for a wide range of research purposes. Although traditionally most users have requested samples for making direct phenotypic measurements, more recently, requests for samples seek derived data associated with the samples. As a result, measurements have expanded to include things such as AI image analysis for species identification and trait measurements; phylogenomic reconstruction of mosquito host switching; detection of SARS-CoV-2 in wild rodent populations; characterization of gut and soil microbiomes; and much more.

The hope is that derived data products generated from the physical samples that are beyond the original scope of data to be provided by the observatory (e.g., images, trait measurements, blood panels) will increase sample value and use. Doing so, however, raises two central issues: 1) What features of the samples will increase their value for research use; and 2) What metrics characterizing use should the Biorepository staff create to increase utility and use, thus increasing repository sample attractiveness. It is becoming clear that features associated with the samples are what makes them attractive. Biorepository staff analysis of sample requests shows that invertebrate bycatch samples were requested the most, followed by small mammals, especially with respect to blood and fecal samples for disease and diet analyses, respectively. Metrics such as counting the number of enquiries, number of fulfilled requests, and number of requests for longitudinal samples over time were used to evaluate Biorepository use. All of these metrics have been increasing since 2019. Between 2021 and 2026 the total number of requests and fulfilled requests increased approximately 5 fold.

Sharing samples with a large research community raises several serious issues. The first is who pays for sample handling, preparation and transport? Another concern is the type of sharing: does sample use involve destructive, or non-destructive data collection? And finally, what criteria should be used in determining the extent to which consuming samples today limits their availability for use in the future? Currently, the repository follows a threshold rule that no more than a third of the samples can be borrowed if the intent will involve destructive analyses. Should such a hard and fast rule apply irrespective of context? One example presented to the STEAC where multiple users made a recent Biorepository request for the same samples illustrated that distributing

samples via a threshold rule puts NEON in the difficult position of determining which project is more valuable than another. Repository staff learned that there is no easy solution to this problem. During discussion with the potential users in this example presented, the staff learned that flexibility in sample use could transform competitive situations into mutualistic ones, potentially leading to sample sharing, thus helping maintain threshold criteria.

Recommendations: STEAC recommends that an outside advisory committee be formed to interact with potential sample users and develop case specific strategies much as IACUC committees do when making decisions on invasive animal use in proposed research. This advisory committee could be the Biorepository Technical Working Group, but if so the STEAC recommends that the membership of this group be refreshed to include more external members and a mix of expertise beyond collections researchers. The current Biorepository TWG is 50% NEON and Biorepository staff and external members are all curators of collections. STEAC also strongly supports the notion that once samples are used, the information gathered should be made available for future users since derived data increases sample value. Since NEON follows FAIR criteria--Findable, Accessible, Interoperable, and Reusable--for data sharing, STEAC recommends that sample users should be required to follow these principles and sign a contract indicating their responsibilities to the license, prior to receiving samples. Given that the NEON Biorepository is unique, in that it is a hyper-accessible 'big data' system, links to where the sample users publish or store derived data should be added to the NEON sample's data page. In this way, data providers having used samples can be credited for their contributions.

NEON Biorepository III: Researcher derived-data considerations

Summary:

Dr. Kelsey Yule and Dr. Hojun Song presented information on the status, current policies, and challenges associated with research products (e.g., data and media) derived by the research community from data published by the NEON Sample Portal. The presentation started with an overview of the NEON Sample Portal and highlighted the implementation of the Humboldt Extension to the Darwin Core Data Package, which allows for comprehensive metadata associated with NEON's Biorepository samples. Currently, the Sample Use Policy requires that external researchers provide all derived research products to NEON for potential publication on the NEON Sample Portal within two years of sample receipt or upon publication (whichever comes first). Compliance has been an issue, as many researchers have not been properly citing NEON in their publications, much less providing links to their derived data products.

The presentation to the STEAC provided four example user-derived research products that have been produced by the research community from NEON data and discussed how those derived products are currently being incorporated into the Sample Portal. The examples comprised 1) species determinations, 2) sequence data, 3) species traits, and 4) digital images. Species determinations can easily be hosted via the Sample Portal with proper metadata and attribution to the deriving researchers. As of now, sequence data are published via other data portals; the Sample Portal links to the derived data on another portal and other associated publications. Species traits cannot be fully achieved on the Sample Portal at this time but could be in the future if software was developed. Some images can be hosted on the Sample Portal, however, there are limits in file types, attribution, and future proofing.

Feedback was requested on the following topics: 1) is it reasonable to require that researcher-derived products are published on the NEON Sample Portal, 2) strategies to ensure that researchers share their derived research products with NEON, 3) strategies to ensure that both the derived product and the original NEON data are cited by future users, and 4) how to prioritize software development required so that NEON can more fully host derived research products in the future.

The STEAC discussion centered around the role of the NEON Sample Portal, citation practices for derived research products, and incentives to encourage data sharing. The STEAC noted that archiving derived products could be a value-added to the NEON Sample Portal for the research community but also discussed potential issues around “mission creep” and duplicating the efforts/resources of other archives and data portals already in existence. Comparable derived data products from other types of NEON data are not hosted on the NEON Data Portal, but rather on other data portals (e.g., EDI), with proper attribution to NEON. Issues may arise when the NEON Sample Portal cannot fully host the derived products (e.g., the Sample Portal supports .jpeg files while researchers need .tif files), when derived products are hosted in multiple places (e.g., on an external portal/archive and on the Sample Portal), and when proper citation becomes confusing.

The STEAC also acknowledged that issues around derived research products are not unique to NEON and there is a larger community challenge around best practices for archiving, sharing, and citing derived research products. The STEAC noted that it is important for both the original NEON data and derived data product to be appropriately cited by future researchers. The STEAC also noted that there is a community-wide shift towards more open data sharing and FAIR data principles, which may result in more voluntary sharing of derived research products in the future and better citation practices by the community. Clearly highlighting the external researcher who derived the product

on the Sample Portal may be enough of an incentive to encourage data sharing with NEON as this exposure could amplify the efforts of the external researcher and lead to future research collaborations down the line.

Recommendations: Overall, the STEAC recommends that the Sample Use Policy require that derived research products be archived according to FAIR data principles somewhere (not necessarily on the NEON Sample Portal), and that a link to the derived research product and any associated publications be shared with NEON. However, the STEAC does not currently recommend that all derived products be published on the NEON Sample Portal given existing issues around the types of data that can be hosted, metadata requirements, potential flow-down citation issues, and potential for duplicate publication of derived products. The STEAC recommends that the Sample Use Policy “encourage” the sharing of some derived products for possible publication on the Sample Portal. In some cases, it may be appropriate for the NEON Sample Portal to host derived products that can be appropriately archived and clearly attributed to the external researcher (e.g., species determinations). In other cases, it may be appropriate for the NEON Sample Portal to link to the externally archived derived research product and associated publications.

Integrating AI workflows into NEON operations: community proposals and beetle pilot

Summary:

Dr. Kate Thibault presented a case study of community-driven research in NEON, which led to increased operational efficiency for NEON. Researchers tested multiple computer vision models for the efficacy of species ID of the NEON carabid (ground beetle) collection. The best model achieved 97% accuracy at the genus level and 94% at the species level. While further refinements (e.g., specific preparation requirements for specimens to facilitate imaging) and a pilot are needed to operationalize this approach, this study shows the potential for community-driven research that could help improve NEON operations. STEAC discussed how to incentivize and attract researchers to lead proposals to partner with NEON in similar research.

Recommendations:

While STEAC sees the value in these types of research partnerships, we acknowledge that it could be difficult to find the perfect nexus of researcher interest and familiarity with NEON, funding, and NEON operational need for this type of research. STEAC sees great potential, though, in using NEON as a platform to develop new technologies (e.g., using cameras to track the progression of plant disease). STEAC encourages NEON to use its newsletter, workshops, and other communications with researchers to plant

seeds of interest highlighting particular data products or collections for this type of research as well as aligned funding sources.

NEON Strategic Interaction Plan (SIP)

Summary:

Tanya Masalak and Callie Puntenney led a presentation on NEON's community engagement, which has a dedicated budget that is < 5% of NEON's total budget. The history of engagement and outreach was summarized, and it was pointed out that engagement and outreach plans were described throughout the 2025 Five-Year Strategic Plan. NEON's engagement strategy is evolving as the network matures. There is an increased focus on impact metrics and shifting from outreach to deeper engagement.

An overview of the types of engagement and the metrics used to track them was provided. Four types of interactions were carefully defined by NEON leadership, incorporating guidance from interviews completed by an external evaluator group (Catalyst). In addition to defining the interaction types, metrics for determining success were defined. Types include (a) "Visibility", defined as activities prompting awareness outside of current NEON data users (e.g. exhibits, blogs, job fairs), (b) "Outreach", directed towards the operational and domain community, (c) "Education", events and activities used to broaden the use of NEON through activities provided by the Ambassador program and data skills webinars, and (d) "Engagement", with the goal of deepening and sustaining existing connections, potentially through advisory groups and NEON-hosted research events.

NEON leadership and staff then binned the 413 events completed in funding year AY25 into the newly defined categories to gauge past efforts and assess for the future how efforts will be distributed over the four categories as part of a new Strategic Interaction Plan (SIP).

Recommendations: The STEAC supports the shift in emphasis towards deeper engagement, and recommends performing an analysis of the return on investment of different types of engagement. The STEAC also recommends separating large, multifaceted activities into different types of engagement, dividing the effort into multiple categories. The STEAC recommends exploring community-led [stack exchange](#) type platforms or bulletin boards to promote community engagement and sharing of resources.

DAY TWO

Reflections on Day One

The group strategized on what questions to ask of NSF later in the day.

TWG evolution proposal

Summary:

Outside advisory groups consisting of scientists interested in and using the NEON data are a valuable part of the NEON operations, as they represent volunteers who are consulted for resolution of critical technical operational issues that arise as the NEON program matures. In this presentation, Kate Thibault and Tanya Maslak provided a summary of the historical use of Technical Working Groups (TWGs), the current TWG needs moving forward, and goals for a reimagined formatting. NEON leadership recommends dissolving 13 of the current TWGs whose scope of work has been completed or the TWG is infrequently consulted, reaching out to members of the larger scientific community with particular expertise on an ad hoc basis when specific technical issues need to be addressed. In addition, NEON staff would continue working with 9 currently active TWGs who are addressing NEON data collection strategies that are highly dynamic and involve rapidly evolving technology that requires frequent assessment.

NEON proposed that this reduction in the number of external advisory committees would be replaced with a new model involving scientists focused around research themes or specific research questions, co-led by a community member and a NEON scientist. The goal would be to drive the science forward with these groups, rather than address specific technical issues. The NEON leadership wanted to know what incentives could be provided to encourage members of the community to participate and lead these efforts.

A member of the STEAC discussed a TWG they participated in over the past several years that led to conference presentations and publications, an excellent example of the kind of new TWG model NEON is envisioning. In this TWG, members were motivated by the potential for future funding and by the goal of producing publications, which resulted in a deep examination of NEON data that was the “hook” that got people interested in pursuing scientific questions. The group felt using a few high-quality examples of these types of efforts to show the larger community, then piloting two or three of these newly designed, question-driven external groups would be a great way to proceed with the new model without overcommitting NEON staff.

Discussions also focused on how researchers external to NEON could initially connect with each other to potentially form new working groups. Suggestions included online

discussion boards where researchers could find common interests, the use of scientific conferences to convene informal groups to discuss potential themes or research ideas, and the use of the login information now required for use of NEON data to introduce researchers with common expertise to each other to start conversations.

Recommendations: The STEAC supports dissolution of the 13 current TWGs that have completed their scope of work or are infrequently consulted, as well as maintenance of the 9 remaining TWGs for a limited time frame. The STEAC recognizes that technical issues will likely arise several times for all data products over the lifespan of the project, but that NEON staff members have a broad network of experts in the field to whom they can reach out to for advice as needed. For the new model envisioned for TWG 3.0, the STEAC recommends renaming the working groups DaWGs (Data Working Groups), piloting 2-3 teams with a community member as the lead and a focus on moving scientific research forward.

The STEAC felt incentivizing researchers by supporting efforts to secure funding of workshops would be a viable manner to support initial work by DaWGs, and that these community led approaches are ultimately a component of the newly designed SIP that was presented earlier in the agenda, and therefore careful planning for synergy between the two efforts (DaWGs and SIP) will allow for more efficient and impactful use of NEON data in driving the science forward. Finally, the STEAC recommends that NEON leadership continue to explore how best to navigate the transition from NEON-led to community-led working groups, potentially by examining successful models used by other NSF funded centers that made a similar switch.

Discussion with NSF NEON Program Director, Dr. Charlotte Roehm, and Dr. James Deshler

This portion of the agenda began with a brief summary of the STEAC meeting for Dr. Roehm and Dr. Deshler. Jim Deshler then opened by remarking that all NSF awards are in the midst of a reorganization and that NSF divisions no longer exist. NSF Directorate for Biological Sciences (BIO) solicitations are being consolidated within three thematic areas. NSF facilities continue to be reorganized, with the goal to make reorganization as invisible and smooth as possible to researchers.

The STEAC asked how the thematic areas within NSF relate to NEON. In addition, the STEAC asked if NSF can help scaffold broader use of NEON data through more specific funding opportunities.

Dr. Deshler said that community members can seek NSF funding to do NEON-based work. For example, researchers using AI can apply to the active solicitation on AI and

continental scale ecosystem research. NEON was built to support all areas of BIO, so within the NSF BIO theme, NEON proposals will be welcome. PIs will no longer have to worry about which solicitation to choose. PIs no longer write to programs, but the innovation and capacity building calls are still open. He said that program officers like to see projects that leverage NEON data. In addition, NSF has funding for workshops, and NEON can reach out to NSF for workshop support.

The STEAC explained that the Biorepository is lean, and engagement with the scientific community is limited due to the lack of resources. As the Biorepository expands and continues to provide unique data that are unlike a typical natural history museum collection, the STEAC asked for clarification regarding the NSF officers' metrics of success. As the collection grows it will become more valuable - it is currently in only its 7th of 30 years. The STEAC also asked if it would be feasible to reconsider the Biorepository as a research accelerator, as a more active body, rather than a passive receiver of requests.

The NSF officers encouraged people to find opportunities for submission of NSF proposals focused on innovation. They suggested that NEON cannot oversell the value outreach. If researchers don't know how good the NEON data are, it won't get used. The one person at NSF who thinks the most about repositories is Katarina Dhittmar, and they suggested that NEON contact her.

The STEAC explained that regarding TWGs, NEON is considering piloting new data-focused groups. As an example, a stream-focused TWG organically transformed into a working group producing papers and proposals after solving the initial technical problem it was focused on addressing. The STEAC asked about potential approaches for community members seeking to answer research questions using new techniques or methods to enhance NEON data collection, including funding opportunities. NSF responded that while the number of funding calls are being reduced, the innovation and capacity calls are still open, and program officers like to see projects that leverage NEON data. NSF still supports collections like the Biorepository, for example, there was a special track for postdocs supporting collections, and it was highly successful. NSF is holistically interested in leveraging NSF investments such as NEON, and new programs focused on AI and innovation are future opportunities.

The STEAC also emphasized the value of NEON to early career faculty who could potentially leverage these consistent, baseline, environmental monitoring data as an investment to build their own research question upon, freeing up funds for more specific measurements that are not collected within the scope of NEON. Messaging this benefit

to pre-tenure faculty may result in more proposals that formally incorporate NEON collaboration.

STEAC Business and report planning

The agenda concluded with STEAC members gathering in a closed-door session to further discuss the recommendations they would like to propose to Battelle in response to the issues raised throughout the meeting by NEON leadership, as well as a timeline for writing the annual report to be publicly posted on NEON's website.