

STEAC MEETING REPORT (06/06/2024)

The members of the STEAC met for an off cycle meeting on June 6, 2024, with a quorum of six members attending (Rich Fiorella, Shannon LaDeau, Steve Petruzza, Sydne Record, Daniel Rubenstein, and Meghan Avolio). Nine NEON-Battelle staff attended (Christine Laney, Chandra Earl [Biorepository], Chris McKay, Claire Lunch, Paula Mabee, Steve Jacobs, and Tanya Maslak, Steve Stone, Kate Thibault).

The meeting was virtual, and the following topic was discussed: I. Data portal Planning.

I. **Previous minutes for 05-15-2024.** Minutes were not yet ready for approval given the off cycle nature of this meeting, so they were not voted on.

II. **Data Portal Planning:**

Christine Laney from NEON continued a presentation on data portal planning that she began on 5/15/24. This off-cycle STEAC meeting was planned to go over the data portal planning including a request to NSF to use contingency funds from the risk register for the data portal that was time sensitive.

The presentation given to the STEAC was a practice run for a presentation to be given to NSF requesting use of contingency funds from the risk register to add three full time employees (FTEs) to the data portal development team over AY25-27. Current staffing is 1 FTE on this team. The request is motivated based on user feedback involving data access.

Christine began with some background on the NEON data portal, which includes infrastructure and code, public API, website, connections/transfers with NEON and hosting repositories, connections to 3rd party services. User feedback on the portal has been mixed. Balancing operations and developments is always a challenge with the data portal. There is a large backlog of user requested functionality and improvements.

The motivation behind the request for use of contingency funds from NSF is to accelerate data portal development without expanding operational scope. [Global Core Biodata Resource](#) declined NEON's application because the review team felt that data from NEON were not accessible. Researchers on the panel were used to being able to find sequence data and use it within a platform (e.g., NCBI). NEON data are contained in files that have to be downloaded and cannot be viewed directly. For context, here is a [list](#) of existing GCBR resources.

Development based on benefit for end users and mechanisms were presented that could be accomplished with one FTE and no contingency funding by AY28. These included: making relationships between data products, their revisions, and their releases easier to navigate and understand; improved data discovery, exploration of longer time series of data, improved accessibility of information prioritized by known issues with the data, and improved access for users with sight-related impairments.

Development based on benefit for end users and mechanisms were presented that could be accomplished with three FTEs and contingency funding by AY28. These included: more discoverable and downloadable AOP data and metadata, enabling omics researchers to discover, access, and use standardized genomic data, making it easier for researchers to query the completeness and validity of data, easier discovery of data (no 'dead end' searches), and making data available in pre-stacked data in downloaded packages.

The STEAC acknowledged that there is a risk of reputation if the data are not easy to use and discover. The STEAC noted that discussion at the end of the in-person meeting in April 2024 included concern that there were some data sets (e.g., AOP, microbial data) that have had data latency issues leading to lower use by the community. The STEAC also noted a need for more assessment on the use of the data portal to show whether or not the use of these contingency funds was successful. For instance, could survey responses on data use be increased by partnering with ESIIIL to add a survey to their participants about ease of using NEON data (or lack thereof). Aside from looking at the number of data downloads, could there be a pre/post contingency fund use survey reaching out to targeted communities? NEON also mentioned that there would be a NSF annual operations panel review this August that may involve users.