Department of Energy
Office of Science
Washington, DC 20585

Office of the Director

Dr. Bruce Hungate
Regents' Professor, Biological Sciences
Northern Arizona University
SLF Building 17, Room 300A
600 South Knoles Drive
Flagstaff, Arizona 86011

Dear Dr. Hungate:

On behalf of the Office of Science, I want to convey my sincerest appreciation for the outstanding work that the Biological and Environmental Research Advisory Committee (BERAC) and the Biological and Environmental Research (BER) Committee of Visitors completed on the review of the Biological Systems Science Division management processes. I also appreciate the ongoing efforts of the BERAC Subcommittee on International Benchmarking and look forward to the final report. Given recommendations in the 2017 Grand Challenges report and the 2018 Scientific User Research Facilities report from BERAC on developing more consistent, integrated, and distributable data across the BER programs, as well as the more recent focus on developing Artificial Intelligence and Machine Learning (AI/ML) technologies applicable to BER research, additional information and actionable items on these topic areas would be very useful.

I am therefore requesting BERAC to (1) review the existing and anticipated capabilities in data management and supporting infrastructures that are relevant to the breadth of BER science and (2) recommend a strategy for the next generation data management and analysis within a unified framework. This new framework should allow for the interoperability and compatibility of data, tools, and supporting information that span the biological, environmental, climate, and earth system sciences, and it should facilitate the analysis and synthesis of data for complex and multi-disciplinary research efforts across BER. For this assessment, data should include laboratory and field observations (e.g., ARM, AmeriFlux, ESS-DIVE, JGI, KBase, and NMDC), model-generated data (e.g., ESGF), simulated and stochastic data (e.g., ARM/LASSO), archives based on observations and models (e.g., Multisector Dynamics data and ILAMB), and relevant metadata, including uncertainty characterization, data provenance, and any tools used to generate the data.

In its analysis, BERAC should consider the need for models and data to interact and inform one another (e.g., the Model-Experimental [MODEX] approach or other approaches) and the following topics:

- Identify new science opportunities that could be possible within and across BER programs if a unified data framework were to be developed;
- Assess recommendations from recent AI/ML reports that could potentially be incorporated into a future data framework for BER (e.g., with a component that includes training data);
- Consider data management strategies and investments in other agencies that could be leveraged in developing the BER unifying framework;
- Provide a list and brief explanation of the components and specifications that would be needed in the development of a unified framework in service to BER science that is achievable in the next five years; and
- Examine the benefits of developing a unified data framework to the scientific research workforce, with particular attention to increased opportunities for enhancing career progression and which types of culture changes could help facilitate those benefits.

BERAC's review and recommendations generated should be informed by the Office of Science principles related to the management of digital research data, including making data available to the public to the greatest extent possible; the FAIR guiding principles for scientific data management; and the Office of Science Technology Policy report on Desirable Characteristics of Data Repositories for Federally Funded Research.

This review and subsequent report should provide sufficient information for BER to implement a major new data infrastructure for implementation as a phased approach over the course of five years. Results from this work, including a brief written report, should be presented at the Fall BERAC meeting in 2023.

Sincerely,

Asmeret Asefaw Berhe

Asmeret Asefaw Berhe
Director, Office of Science

cc: Gary Geernaert