January 20, 2023

Office of Science and Technology Policy
Eisenhower Executive Office Building
1650 Pennsylvania Ave, NW
Washington, DC 20504

Re: RFI; National Biotechnology and Biomanufacturing Initiative

On behalf of the Center for Data Innovation (datainnovation.org), I am pleased to submit this response to the Office of Science and Technology Policy's (OSTP) request for information (RFI) on the National Biotechnology and Biomanufacturing Initiative (NBBI) and the steps OSTP can take to ensure a data ecosystem that supports a strong bioeconomy in the United States.[1]

The Center for Data Innovation studies the intersection of data, technology, and public policy. With staff in Washington, London, and Brussels, the Center formulates and promotes pragmatic public policies designed to maximize the benefits of data-driven innovation in the public and private sectors. It educates policymakers and the public about the opportunities and challenges associated with data, as well as technology trends such as open data, artificial intelligence, and the Internet of Things. The Center is part of the Information Technology and Innovation Foundation (ITIF), a nonprofit, nonpartisan think tank.

In response to the RFI, the Center offers four recommendations to bolster data for the bioeconomy: 1) OSTP should expedite and ensure proper implementation of open access policies for government-funded research publications and data; 2) OSTP should foster data sharing between government, industry, and academia using tools like cooperative research and development agreements (CRADAs) and data trusts; 3) OSTP should identify data divides in the bioeconomy to ensure that datasets are representative and that all Americans can participate; and 4) OSTP should treat data more like software or infrastructure and provide datasets with the routine maintenance afforded to other public goods.

Please find our response to the relevant question in the document below.

Sincerely,

[1] Office of Science and Technology Policy, "Request for Information; National Biotechnology and Biomanufacturing Initiative," Federal Register, December 20, 2022, https://www.federalregister.gov/documents/2022/12/20/2022-27600/request-for-information-national-biotechnology-and-biomanufacturing-initiative.

Gillian Diebold
Policy Analyst
Center for Data Innovation
gdiebold@datainnovation.org

(4) How can the Federal Government, in partnership with private, academic, and non-profit sectors, support a data ecosystem to drive breakthroughs for the U.S. bioeconomy? This may include technologies, software, and policies needed for data to remain high-quality, interoperable, accessible, secure, and understandable across multiple stakeholder groups.

There are four primary ways in which the U.S. government can support a data ecosystem that bolsters innovation in the bioeconomy. First, OSTP can implement its recently updated open access policy for federally funded research. OSTP's August 2022 memorandum, "Ensuring Free, Immediate, and Equitable Access to Federally Funded Research," recommends that federal agencies update their public access policies to make all peer-reviewed scholarly publications and related data supported by public funds publicly accessible by 2025.[2] Open access policies promotes innovation by ensuring all researchers, regardless of their financial resources, have access to the latest peer-reviewed government funded research. In particular, OSTP should work with federal agencies to develop guidance and best practices for sharing scientific data underlying federally funded research, such as by supporting the creation of data repositories and allocating a portion of grant funds for data storage and maintenance. Implementing these policies will improve the efficiency and integrity of federally-funded research for the bioeconomy, and deliver a number of social and economic benefits, like increased trust in scientific research, and easier data collection and reuse, and reduced need for duplication of data in the research process.[3] By simplifying data sharing and access, public access policies also facilitate the kind of cross-discipline work essential to the bioeconomy, a field that cross cuts multiple sectors like energy, chemicals, and agriculture.[4]

Second, OSTP should incentivize data sharing partnerships to solve some of the biggest issues facing the bioeconomy, like disease and other medical conditions, climate change, energy, food security, and supply chain resilience. Increased collaboration between government, industry, and academia played a critical role in positioning the United States as a leader in life sciences research, with the Bayh-Dole Act of 1980 incentivizing universities to take on the risks of life sciences innovation through formalized guidelines for discoveries stemming from federal funding.[5] The

---

[2] Alondra Nelson, "Ensuring Free, Immediate, and Equitable Access to Federally Funded Research," August 25, 2022, OSTP, https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf.

[3] Gillian Diebold, "Updated OSTP Guidance on Open Access Research a Win for Innovation," *Center for Data Innovation*, November 8, 2022, https://datainnovation.org/2022/11/updated-ostp-guidance-on-open-access-research-a-win-for-innovation/.

[4] Marcy Gallo, "The Bioeconomy: A Primer," (Washington, DC: Congressional Research Service, September 2022), https://crsreports.congress.gov/product/pdf/R/R46881.

[5] Stephen Ezell, "The Bayh-Dole Act's Vital Importance to the U.S. Life-Sciences Innovation System," (ITIF, March 2019), https://itif.org/publications/2019/03/04/bayh-dole-acts-vital-importance-us-life-sciences-innovation-system/.

federal government can promote collaboration in biotechnology and biomanufacturing research with data sharing partnerships. For example, cooperative research and development agreements (CRADAs) enable no-cost data sharing between federal agencies and those in industry, academia, nonprofits, and even other federal agencies.[6] For example, the National Oceanic and Atmospheric Administration (NOAA) has established the Big Data Project through a series of CRADAs to make the organization's vast stores of environmental data more accessible to the private sector and the public through cloud-based platforms.[7] In the UK, the government has promoted the use of "data trusts"—legal frameworks that allow multiple parties to collect and share data responsibly—to make data more accessible and usable, including for medical research.[8] OSTP should evaluate opportunities to leverage existing data sharing models, such as CRADAs and data trusts, to facilitate data sharing for biotechnology research.

Some biomedical research could benefit from access to more health data.[9] Unfortunately, there are few options in the United States for patients to contribute voluntarily their health data for use in medical research. OSTP should explore how to make it as easy for patients to donate their medical data as it is to register to be an organ donor. In addition, OSTP should examine how privacy laws, including both those at the state and federal level, impact health information data collection and sharing for biomedical research. Health data sharing systems are a quasi-public good and as such, private actors have much less incentive in organizing them than the government. Data-driven medicine, including data-driven drug development, holds great promise, but it will not reach its full potential without more widespread access to health data for research.

Third, OSTP should identify and address data divides in the bioeconomy and ensure that datasets are representative. The data divide refers to the social and economic inequalities that arise from a lack of data collection or use of data about certain individuals or communities.[10] Dataset representation concerns the extent to which a given dataset sufficiently describes the characteristics

---

[6] "Cooperative Research and Development Agreement (CRADA)," National Institute of Standards and Technology, last modified September 2019, https://www.nist.gov/tpo/partnerships/cooperative-research-and-development-agreement-crada.
[7] Alexander Kostura and Daniel Castro, "Three Types of Public-Private Partnerships That Enable Data Innovation," August 1, 2006, Center for Data Innovation, https://datainnovation.org/2016/08/three-types-of-public-private-partnerships-that-enable-data-innovation/.
[8] "How can legal mechanisms associated with data trusts enhance participation in healthcare research?," Data Trust Initiative, n.d. https://datatrusts.uk/trusts-in-health-research.
[9] Joshua New, "The Promise of Data-Driven Drug Development," (Center for Data Innovation, September 2019), https://www2.datainnovation.org/2019-data-driven-drug-development.pdf.
[10] Gillian Diebold, "Closing the Data Divide for a More Equitable U.S. Digital Economy," (Center for Data Innovation, August 2022), https://datainnovation.org/2022/08/closing-the-data-divide-for-a-more-equitable-u-s-digital-economy/.

of a larger population.[11] If a dataset excludes a certain population outright or excludes certain details about that group, new technologies may either lack the intended impact or even be outright harmful. Creating more inclusive datasets can help ensure that all Americans benefit from an innovative bioeconomy.

Lastly, OSTP and other federal agencies should treat data more like software or infrastructure and provide datasets with the maintenance afforded to other public goods. For example, just as software developers maintain a list of known bugs, data creators should maintain a list of known limitations or errors in datasets. Most public goods are treated as an ongoing process that require routine upkeep to ensure their safety and quality. Managing a data ecosystem more like software or infrastructure requires that same level of attention and will increase the quality and usefulness of data. To that end, OSTP should encourage agencies to fund not only the creation, but also the maintenance, of datasets to ensure they are routinely updated, cleaned, and secured.[12]

[11] Ibid.

[12] National Academies of Sciences, Engineering, and Medicine, *Toward a 21st Century National Data Infrastructure: Mobilizing Information for the Common Good*, (Washington, DC: The National Academies Press, 2022), https://www.nationalacademies.org/our-work/toward-a-vision-for-a-new-data-infrastructure-for-federal-statistics-and-social-and-economic-research-in-the-21st-century.