March 2, 2020

Comments submitted online to: OpenScience@ostp.eop.gov

Sean C. Bonyun,
Chief of Staff
Office of Science and Technology Policy
Executive Office of the President
Eisenhower Executive Office Building
1650 Pennsylvania Avenue
Washington, DC 20504

RE: Request for Public Comment on Draft Desirable Characteristics of Repositories for Managing and Sharing Data Resulting from Federally Funded Research (85 *Federal Register* 3085)

Dear Mr. Bonyun:

Public Responsibility in Medicine and Research (PRIM&R) appreciates the opportunity to comment on the Office of Science and Technology Policy's Request for Public Comment on Draft Desirable Characteristics of Repositories for Managing and Sharing Data Resulting from Federally Funded Research, published January 17, 2020.

PRIM&R is a nonprofit organization dedicated to advancing the highest ethical standards in the conduct of research. Since 1974, PRIM&R has served as a professional home and trusted thought leader for the research protections community, including members and staff of human research protection programs and institutional review boards (IRBs), investigators, and their institutions. Through educational programming, professional development opportunities, and public policy initiatives, PRIM&R seeks to ensure that all stakeholders in the research enterprise understand the central importance of ethics to the advancement of science.

PRIM&R endorses the White House Office of Science and Technology Policy's (OSTP) efforts to improve the consistency of guidelines that federal R&D-funding agencies provide to their grantees and other stakeholders about best practices in long-term storage of data from federally funded research. We especially appreciate the current step of

developing a proposed, common set of desirable characteristics of data repositories that agencies can use to support their current Public Access and data sharing efforts. As the request for public comment notes, this kind of forward thinking has the potential not only to improve government-operated repositories, but also to lead to better and more consistent practices across repositories run by non-governmental entities.

PRIM&R has long believed that harmonization of federal policies around research can be an important and effective means of supporting the conduct of responsible research, as long as it does not negatively affect the interests and welfare of research subjects. Harmonization can reduce policy redundancies that do little to add to research oversight and drain limited research resources, and can foster the consistent adoption of best practices. Harmonization of policies is clearly desirable in the data sharing and management space.

In 2018, we submitted comments in response to the National Institutes of Health (NIH)'s RFI on Proposed Provisions for a Draft Data Management and Sharing Policy for NIH Funded or Supported Research,[1] in which we expressed concerns about the proliferation of data repositories that follow a variety of rules and procedures. We pointed out that this has the potential to weaken the overall value of the data sharing enterprise. More recently, in January 2020, we submitted comments to the NIH on their Draft Policy for Data Management and Sharing,[2] in which we requested the NIH itself play a role in vetting grantees' proposed data repositories and sharing platforms to ensure they support the secure and ethical sharing of data.

The OSTP's proposed recommendations on repository governance issues are a welcome step in the right direction in terms of promoting harmonization of policies that both reduce burden and enhance responsible research. To that end, we hope the final document will include a strong recommendation that the Subcommittee on Open Science member agencies put language in their grants and contracts *explicitly* requesting adherence to this common set of desirable characteristics in data repositories. Such a move will amplify the benefits of harmonization, and, likely, the utility of the data sharing enterprise.

PRIM&R also appreciates that the draft acknowledges that there are important additional human subject protections considerations when the data repository involves human data, and that these considerations are relevant even if that data is deidentified. To that end, we support the draft's general language on privacy, but urge that as the OSTP further develops its common set of characteristics and considerations, or provides further guidance in this area, it include language about the need for repositories themselves to have in place mechanisms for preventing or discouraging reidentification of deidentified data, in addition to enforcing submitters' data use restrictions. PRIM&R has publicly commented on

---

[1] Response to the National Institutes of Health's RFI on Proposed Provisions for a Draft Data Management and Sharing Policy for NIH Funded or Supported Research. (2018). Public Responsibility in Medicine and Research (PRIM&R).
[2] Response to the National Institutes of Health's Draft NIH Policy for Data Management and Sharing and Supplement Draft Guidance. (2020). Public Responsibility in Medicine and Research (PRIM&R).

reidentification issues extensively and we would happy to serve as a resource on this important topic if that is of interest.

PRIM&R for the most part endorses the OSTP's current list of "Desirable Characteristics of Repositories for Managing and Sharing Data Resulting from Federally Funded or Supported Research" as appropriately comprehensive and flexible. Below we offer a few additional considerations we think might improve the two sets of desirable characteristics outlined in the draft:

Desirable Characteristics for All Data Repositories:

- We strongly urge the OSTP to add to the list of desirable characteristics that data repositories have a mechanism for ensuring credit for data generators. Giving those who generate data credit for their contributions to the scientific enterprise will incentivize researchers to share their data in the spirit of open science. We direct the OSTP to recently released expert recommendations on how data repositories can play a role in ensuring data generators receive credit for making their data available for future reuse.[3]

Additional Considerations for Repositories Storing Human Data (Even if De-Identified):

- We agree there should be "plans for addressing violations of terms-of-use by users and data mismanagement by the repository." These plans should construe "terms-of-use" as broadly as possible and explicitly include research service agreements. We would like to also note the government as a whole needs to reconsider what penalties should be levied if research subjects' rights are violated during the course of data sharing. It also needs to assess how to determine who should be held responsible for such violations. We believe limiting penalties to just a rescission of funding is likely to be insufficient and an inadequate deterrent to future bad actors.

- We believe the "Fidelity to Consent" consideration as written is likely to be inadequate as a guide for repository developers or those who are evaluating data management plans. We agree that researchers have an obligation to use data in a manner consistent with original consent, as a matter of respect for persons, and data repositories, as gatekeepers for such uses, should do their part to limit dataset access to uses consistent with consent. To that end, we urge OSTP to make clear that repositories that store human data have a responsibility to establish mechanisms for attaching permissions granted in the original consent, as machine-readable metadata, to the data itself.

  Furthermore, we note that ensuring that future uses of data are consistent with consent may not always be straightforward. It is not clear what it means to be faithful to consent when, for example, (1) the original consent was silent regarding

---

[3] Credit Data Generators for Data Reuse, Pierce, H., Dev, A., Statham, E., & Bierer, B. (2019). *Nature.*

whether and in what ways data would be shared or used in the future, or (2) the original consent promised that the data that is stored and shared would remain deidentified, when today's technologies and methodologies, including the aggregation of data sets, make permanent deidentification impossible. Given these complexities, we suggest future policies on this important topic provide additional guidance, perhaps including examples, about what fidelity to consent means or entails in these sorts of circumstances.

Thank you again for the opportunity to comment and for the OSTP's work on this important issue. We hope our comments on the current draft will be useful in your next stage of policymaking in this area. PRIM&R stands ready to provide any further assistance or input that might be useful. Please feel free to contact me at 617.303.1872 or ehurley@primr.org.

Respectfully submitted,

Elisa A. Hurley, PhD
Executive Director

cc: PRIM&R Public Policy Committee, PRIM&R Board of Directors