

August 10, 2022

Ji Lee
NITRD National Coordination Office
2415 Eisenhower Avenue
Alexandria, VA 22314

RE: RFI Response – Federal Big Data Research and Development Strategic Plan Update

Dear NITRDE Interagency Working Group Members:

On behalf of the Council of Graduate Schools (CGS), I am submitting comments to the Request for Information (RFI) on the Federal Big Data Research and Development Strategic Plan Update. The CGS membership welcomes the opportunity to provide suggestions on the ways in which the National Science Foundation (NSF) and the Networking and Information Technology Research and Development (NITRD) National Coordination Office (NCO) can update the ***Federal Big Data Research and Development Strategic Plan***.

For over 60 years, CGS has served as the national organization dedicated to advancing graduate education and research. Our membership includes 500 institutions of higher education in the United States, Canada, and abroad representing over 1.8 million graduate students. Collectively, our members grant 87 percent of all U.S. doctorates and majority of U.S. master's degrees. As an association, we are proud to represent a diverse and dynamic group of higher education institutions ranging from public and private research-intensive institutions to regional comprehensive institutions, minority-serving institutions, and historically black colleges and universities.

In 2021, CGS and the Pervasive Data Ethics for Computational Systems (PERVADE) hosted a two-day workshop in which CGS graduate deans, experts in Big Data, and practitioners in research integrity education met to discuss the ethical challenges and opportunities for preparing graduate students to use and analyze big data sets. Following this workshop, CGS and PERVADE issued the report, [“Preparing Graduate Students for the Ethical Challenges of Big Data.”](#) The CGS membership commends this report to your attention and encourages the NITRD Big Data Interagency Working Group (NITRD BG IWG) to consider the report as you update the Federal Big Data Research and Development Strategic Plan.

The five major conclusions and recommendations from this workshop are as follows:

1. The ethical challenges of research involving big data are relevant to a large population of master's and doctoral students and should be broadly integrated into graduate research training and not be relegated to specialized training programs or courses.
2. The research and graduate education communities should evaluate current Responsible Conduct of Research (RCR) curricula and ensure that they address the challenges of big data. Research with large datasets is changing the way universities need to teach several categories

of RCR training, including collaborative science; data acquisition, management, sharing and ownership; and human research protections.

3. Graduate deans have the most responsibility for the quality of graduate student research training and play a leading role in supporting and facilitating institution-wide collaborations. The graduate dean community has a strong track record of supporting communication and collaboration across campus with the goal of improving and expanding student learning and professional development.
4. Plans to expand graduate research training to address ethical issues in big data research should include the participation of a broad range of stakeholders, such as faculty, students, IRB review boards, vice provosts, etc.
5. Universities, organizations that support graduate education, and funders should increase their efforts to develop centrally available resources that prepare graduate students for the ethical challenges of using large datasets.

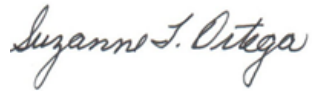
During the two-day workshop, several noteworthy points were made by participants, discussion group leaders, and keynote speakers including:

- **RCR Training:** All workshop participants agreed that data ethics should be expanded within RCR courses. It was also suggested that federal funding agencies could play a role in strengthening training requirements by issuing new RCR guidance and requirements. Participants also suggested that ethics training for values and ideas are not well covered by current RCR and IRB training resources, such as the risks of de-identification of participants and the potential for harm to those individuals.
- **Disciplinary Research and Ethics Training:** On day two of the workshop, Dr. Casey Fiesler, an assistant professor at the University of Colorado-Boulder, made the point that it is important to embed ethics training within disciplinary research and training. “Dr. Fiesler suggested a set of priorities for faculty and campus leaders as they identify opportunities within the graduate training curricula to engage students in ethical concerns. These priorities include: (1) integrating ethics content into the curriculum so that they are gaining exposure to these issues at the same time they are learning the methods, (2) creating shared resources between disciplines, (3) facilitating conversations among stakeholders, and (4) making experts in the field available during these conversations.”
- **Micro-credentials and Badging:** Participants agreed that providing micro-credentials/badging in the ethical use of big data methods is important to the professional development of graduate students. Moreover, continuing education courses that offer credits for researchers using big data research methodologies as these approaches change over time could help

students/researchers document their skills and knowledge for future employers and government funding agencies.

Thank you for the opportunity to submit comments to the RFI on updating the Federal Big Data Research and Development Strategic Plan. Please do not hesitate to contact me or the CGS staff if you have questions about the CGS/PERVADE report cited in this comment letter.

Sincerely,

A handwritten signature in cursive script that reads "Suzanne T. Ortega".

Suzanne T. Ortega

President