Building Secure Platforms for Research on Human Subjects: The Importance of Computer Scientists

Julia Ingrid Lane

New York University

Wagner Graduate School of Public Service

Julia.lane@nyu.edu

ABSTRACT

Businesses and government are using new approaches to decision-making. They are exploiting new streams of (mostly) digital personal data, such as daily transaction records, web-browsing data, cell phone location data, and social media activity; and they are applying new analytical models and tools.

Social science researchers, who are not trained in the stewardship of these new kinds of data, must now collect, manage and use them appropriately. There are many technical challenges: disparate datasets must be ingested, their provenance determined and metadata documented. Researchers must be able to query datasets to know what data are available and how they can be used. Datasets must be joined in a scientific manner, which means that workflows need to be traced and managed in such a way that the research can be replicated (Lane, 2017).

Computer scientists' expertise is of critical value in many of these areas, but of greatest interest to this group is the facilities in which data on human subjects are stored. The data must be securely housed, and privacy and confidentiality must be protected using the best approaches available. The access and use must be documented to meet the needs of data providers. Yet the technology currently used to provide access to sensitive data is largely artisanal and manual. The stewardship restrictions placed on the use of confidential administrative data prevent the use of best practices for research data management. As a result, links between data sources are rarely validated, results often are not replicated, and connected datasets, results, and methods are not accessible to subsequent researchers in the same field.

This is where computer scientists' expertise can come to play in building approaches that will enable sensitive data from different sources to be discovered, integrated, and analyzed in a carefully controlled manner, and that will, furthermore, allow researchers to share analysis methods, results, and expertise in ways not easily possible today.

Author Keywords

Privacy; confidentiality; secure data; social science

BIOGRAPHY

Julia Lane is a Professor at the NYU Wagner Graduate School of Public Service, at the NYU Center for Urban Science and Progress, and a NYU Provostial Fellow for

Innovation Analytics. Julia has led many initiatives, including co-founding the UMETRICS and STAR METRICS programs at the National Science Foundation. She conceptualized and established a data enclave at NORC/University of Chicago. This provides a confidential, protected environment within which authorized researchers can access sensitive microdata remotely and provides data producers with a secure dissemination platform. She also initiated and led the creation and permanent establishment of the Longitudinal Employer-Household Dynamics Program at the U.S. Census Bureau. She is co-leading the development of the Administrative Data Research Facility at NYU, joint with the University of Chicago and the US Census Bureau.

Julia is the recipient of the 2014 Julius Shiskin award and the 2014 Roger Herriot award. She is an elected fellow of the American Association for the Advancement of Science and a fellow of the American Statistical Association. She has been the recipient of over \$50 million in grants and has published many articles in leading journals, including *Nature* and *Science*.

REFERENCES

Lane, J.I., 2017. A call to action to build research data infrastructure. *Nat. Hum. Behav.* 1, 75.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author(s). Copyright is held by the owner/author(s).

HPDC'17, June 26–30, 2017, Washington, DC, USA.

ACM ISBN 978-1-4503-4699-3/17/06.

DOI: <http://dx.doi.org/10.1145/3078597.3091518>