

Describing Scientific Results for Both People and Computers: The Dissemination of Science in a Connected World



Mark Musen

*Professor of Biomedical Informatics,
Stanford University*

*Director, Stanford Center for Biomedical
Informatics Research*

Wednesday, May 11, 2016
11:00-12:15 pm EDT

[Click here to join](#)

For more information contact:
E: cameron.mcrae2@mcgill.ca
T: +1.514.398.3299

Presentation Abstract

The scientific method requires nothing less than that experiments be reproducible and that the data be available for other scientists to examine and reinterpret. In an era when data are generated at rates and in quantities never before imaginable, there is an urgent need to understand the structure of datasets, the experimental conditions under which they were produced, and the information that other investigators may need to make sense of the data. Such information about scientific results needs to be disseminated online, and needs to be interpretable both by scientists and by computers. This goal requires the development of ontologies that can define scientific information in a standardized fashion, of frameworks for capturing the metadata that describe both how experiments are performed and what the resulting data values mean, and of software systems that support scientists in their work to report experimental results in the most complete and useful manner possible. The outcome is a paradigm for publishing scientific knowledge that one day may supplant the communication of experimental results in prose journal articles, and that will enable intelligent agents to comb online datasets, to bring interesting results to the attention of scientists, and to reanalyze and to integrate previous results in novel ways. Investigators at Stanford, Oxford, Yale, and Northrup Grumman are working together to study elements of this approach through the Center for Expanded Data Annotation and Retrieval (CEDAR), a center of excellence supported by the trans-NIH Big Data to Knowledge (BD2K) initiative.

About the Webinar Series

The **BRIDGE** webinar series is designed to prepare for the next generation of big data analytics, woven into transdisciplinary and intersectoral sciences, policy and innovation, and serving as catalyst for solutions at scale to better address the seemingly intractable problems that lie at the nexus of health and wealth production, distribution and consumption. A key to accelerate change lies in establishing bridges between sectoral big data, and between data and content. To foster real time learning, the **BRIDGE** webinar series brings together a new solution-oriented transdisciplinary translational paradigm for the four *Ms* of big data sciences used on both sides of the health and economic divide (*Machines, Methods, Models and Matter*).