September 21, 2010

Dr. Myron Gutmann
Directorate for the Social, Behavioral and Economic Sciences
National Science Foundation

RE: SBE 2020: Future Research in the Social, Behavioral & Economic Sciences

Dear Dr. Myron Gutmann,

This report summarizes the results of two National Science Foundation (NSF) workshops on "***Knowledge Management and Visualization Tools in Support of Discovery***". It is a summary version of a much more comprehensive treatise on the workshops available as a PDF from this link: http://vw.slis.indiana.edu/cdi2008/whitepaper.html. The goal of the workshop series was to gather domain experts in science of science (SoS) fields to discuss the trajectory of current research and supporting infrastructures, and identify research areas and supporting technology needs over the next decade. This was done in the form of identifying areas of need and having participants propose concrete approaches to fulfill those research needs over the next decade.

The goals of the meeting series were to:
- Capture knowledge management and visualization needs with specific examples for SoS research.
- Present major challenges and opportunities for the design of more effective tools and Cyberinfrastructures in support of scholarly discovery, including a timeline of anticipated science and technology development that will impact tool development.
- Provide recommendations on how current lines of work in academia, government, and industry and promising avenues of research and development can be utilized to design more effective knowledge management, visualization tools and cyberinfrastructures that advance discovery and innovation in 21st century science.

Principal Findings

While the full workshop series report provides a comprehensive discussion of the most salient questions, contexts, strategies and implications for future research within and across these disciplines, there are several general, confirmative findings:
- **Science is interdisciplinary and global.** Researchers and practitioners need easy access to expertise, publications, software, and other resources across scientific and national boundaries.
- **Science is data driven.** Access to large amounts of high quality and high-coverage data is mandatory. The "long tail" of data producers/users is larger than the existing major databases and their users.
- **Science is computational.** The design of modular, standardized and easy to use cyberinfrastructures is key for addressing major challenges, such as global warming, or a deeper understanding of how science and technology evolves. Ideally, the "million minds" can share, combine, and improve expertise, data, and tools. It is advantageous for scientists to adapt industry standards, defacto or not, than to have to create their own tools.
- **Science uses many platforms.** Some sciences thrive on Web services and portals, others prefer desktop tools, while some require virtual reality environments, or mobile (handheld)

devices.

- **Science is collaborative.** A deeper understanding of how teams "form, storm, norm and perform" will improve our ability to compose (interdisciplinary/international) teams that collaborate effectively.

There were a number of findings specific to the workshop topic "Knowledge Management and Visualization Tools in Support of Discovery":

- **Formulas and visual imagery** help communicate results across scientific boundaries with different cultures and languages.
- **Users become contributors.** Researchers need more efficient means to share their datasets, software tools, and other resources with each other—just like they share images and videos via Flickr and YouTube today. Overly "topdown" approaches should be discouraged.
- **Science "Facebook".** Scholars need more effective ways to find collaborators, monitor research by colleagues and competitors, disseminate their results to a broader audience.
- **Advanced data analyses combined with visualizations** are used to identify patterns, trends, clusters, gaps, outliers and anomalies in massive amounts of complex data. Network science approaches seemed particularly useful in the selected SoS domains.

Principal Recommendations

The set of recommendations and proposed solutions for advancing SoS research specifically, and the scholarly research infrastructure in general, include:

- A decentralized, free "**Scholarly Database**" to keep track, interlink, understand and improve the quality and coverage of Science and Technology (S&T) relevant data. (see also page 76 and 77 in Appendix D)
- A "**Science Marketplace**" that supports the sharing of expertise and resources and is fueled by the currency of science: scholarly reputation. (see page 74 in Appendix D) This marketplace might also be used by educators and the learning community to help bring science to the public and out of the "ivory tower". (see page 89 in Appendix D)general
- A "**Science Observator**y" that analyzes different datasets in real-time to assess the current state of S&T and to provide an opportunity to develop predictive theories and models of the evolution of science and under several (actionable) scenarios. (see page 72 in Appendix D)
- "**Validate Science Maps**" to understand and utilize their value for communicating science studies and models across scientific boundaries, but also to study and communicate the longitudinal (1980-today) impact of funding on the science system. (see page 81 in Appendix D)
- An easy to use, yet versatile, "**Science Telescope**" to communicate the structure and evolution of science to researchers, educators, industry, policy makers, and the general public at large. (see page 87 in Appendix D) The effect of this (and other science portals) on education and science perception needs to be studied in carefully controlled experiments. (see page 88 in Appendix D)
- "**Science of (Team) Science**" studies are necessary to increase our understanding and support the formation of effective research and development teams. (see page 78 and 82 in Appendix D and http://stm.sciencemag.org/content/2/49/49cm24).
- "**Success Criteria**" need to be developed that support a scientific calculation of S&T benefits for society, such as economic and cultural impacts, beyond internal science metrics, such as citations. (see also page 88 in Appendix D)

- A "**Science Life**" (an analog to Second Life) should be created to put the scientist's face on their science. Portals to this parallel world would be installed in universities, libraries and science museums. (see page 80 in Appendix D)
- The portals would be "**fathered and mothered**" by domain experts, as well as learning experts. Their effect on education and science perception should be rigorously evaluated in carefully controlled experiments and improved from a learning science standpoint. (see page 91 in Appendix D)

Free printed copies of the full report can also be requested via email from Beth Works (bworks@indiana.edu).

Sincerely, the workshop organizers

Katy Börner, School of Library and Information Science, Indiana University, 10th Street & Jordan Avenue, Wells Library 021, Bloomington, IN 47405, USA. E-mail: katy@indiana.edu
Luís M. Bettencourt, Santa Fe Institute, 1399 Hyde Park Rd, Santa Fe NM 87501; T-5 MS B284, Theoretical Division, Los Alamos National Laboratory, Los Alamos NM 87545, USA. E-mail: LMBETT@LANL.GOV
Mark Gerstein, Molecular Biophysics and Bioinformatics & Computer Science Departments Bass 432, 266 Whitney Ave., Yale University, New Haven, CT 06520, USA. E-mail: Mark.Gerstein@yale.edu
Stephen Uzzo, New York Hall of Science, 47-01 111th Street, Flushing Meadows Corona Park, NY 11368, USA.E-mail: suzzo@nysci.org