Symposium on the Interagency Strategic Plan for Big Data: Focus on R&D

NAS Board on Research Data and Information
October 23, 2014

Big Data Senior Steering Group (BDSSG)

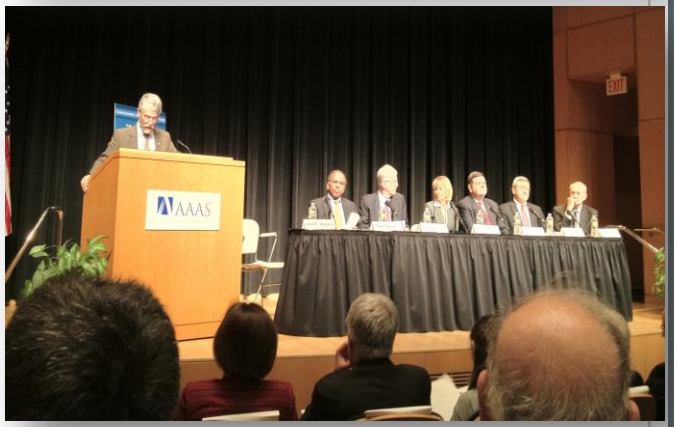
Allen Dearry, NIH, Co-Chair

Suzi Iacono, NSF, Co-Chair

George Strawn, NITRD, Co-Chair

National Big Data R&D Initiative

- BDSSG chartered in spring 2011 by the White House OSTP:
  - Co-chaired by NSF, NIH and NCO
  - Significant research community input
- Major Announcements in March 2012 at the launch: DARPA, DOE, DoD, NIH,NSF, USGS
- Shift to a “national” rather than federal initiative
- Data to Knowledge to Action Partnership event in November 2013
- Significant work over three years to get to where we are today



Member Agencies

- DOC
  - NIST
  - NOAA
- DoD
  - DARPA
  - NSA
  - OSD
  - Service Research Orgs
- DOE
  - NNSA
  - SC
- DHHS
  - AHRQ
  - NIH
- DoI - USGS
- Independent Agencies
  - DHS
  - EPA
  - NASA
  - NARA
  - NSF
  - Treasury

Vision

We envision a Big Data innovation ecosystem in which the ability to analyze, extract information from, and make decisions and discoveries based upon large, diverse, and real-time data sets enables new capabilities for federal agencies and the nation at large; accelerates the process of scientific discovery and innovation; leads to new fields of research and new areas of inquiry that would otherwise be impossible; educates the next generation of 21st century scientists and engineers; and promotes new economic growth.

Framework for Investments

FOUNDATIONAL RESEARCH to develop new techniques and technologies to derive knowledge from data

New CYBERINFRASTRUCTURE to manage, curate, and serve data to research communities

POLICY

New approaches for EDUCATION and WORKFORCE DEVELOPMENT

New types of interdisciplinary COLLABORATIONS, grand challenges, and competitions

BDSSG Selected Highlights
(Subcommittee Activities)

Core Technologies

- NSF-NIH Critical Techniques & Technologies for Big Data FY12-13 Joint Sollicitation
- Mid-Sized Project (five years, $150-$650K Direct/Year): 136 submissions, 11 awards
- Small Project (three years, $150,000 Direct/Year): 350 submissions, 34 awards

Domain Research/Cyberinfrastructure

- 2013 workshop on Data Sharing and Metadata Curation: Obstacles and Strategies to discuss
- Future strategies for managing scientific data and metadata for basic and applied research (https://www.nitrd.gov/nitrdgroups/images/b/ba/Summary_Report_BDSSG_Metadata_Workshop_May_29_2013.pdf)

Education & Workforce Development

- 2013 Joint Statistical Meeting Session on Recent Activities in Big Data: Curriculum Development and Funding Opportunities
- 2014 NSF workshop at NAS on training students to extract value from big data (http://sites.nationalacademies.org/DEPS/BMSA/DEPS_087192.htm)

Challenges & Competitions

- Hosted challenge on TopCoder with NASA Center of Excellence
- Focused on data fusion and moving from heterogeneous to homogeneous data
- Low response rate; may be result of platform

BDSSG Strategic Plan to Date

- April /May/July– Agency presentations
  - Game Changing Approaches and Success Factors
  - Five common themes synthesized for workshop
- June 26 – Internal Workshop to further develop strategies and themes document
  - Special guest presenters: Marjory Blumenthal (PCAST); Phil Bourne (NIH); Chaitan Baru (NSF)
  - Draft Strategies and Themes document created and released to members for further review

BDSSG Strategic Plan: Strategies and Themes (draft) (1/2)

- ***Strategy I***: Create **next generation capabilities** by leveraging emerging Big Data foundations, technologies, processes, and policies.
- ***Strategy II***: In addition to the generation of knowledge from data, also emphasize using trustworthy data and resulting knowledge to **make decisions and take confident action.**
- ***Strategy III***: Ensure the long term **sustainability**, access, and development of high value data sets and data resources.

Strategies and Themes (draft) (2/2)

- ***Strategy IV***: Improve the national landscape for Big Data **education and training** to fulfill increasing demand for both deep analytical talent and analytical capacity for the broader workforce.
- ***Strategy V***: Create **new gateways** that enable the interconnection and interplay of Big Data ideas and capabilities across agency mission.

BDSSG Coordination and Collaboration

- The purpose of the Big Data Strategic Plan is to collect and synthesize information that the BDSSG has gathered since it was established by the Administration in 2011.
- The goal is to provide a reference for Agencies who are planning or expanding Big Data programs.
- Many agencies have actively participated in the Strategic Planning Process, including DHS, DOE, NASA, NIH, NIST, NOAA, NSF, USGS.
- These and all other BDSSG Agencies, as well as the public, are participating in the drafting and review process over the next 8-10 months.

Request for Information

- Feedback from multiple big data stakeholders to inform development of a framework, priorities, and a strategic plan for the National Big Data R&D Initiative.
- Critical step in developing a cross-agency strategic plan that has broad community input and that can be referenced by all Federal agencies as they move forward over the next five to ten years with their own and collaborative big data R&D programs, policies, partnerships, and activities.
- Draft "vision and areas of interest" provided (https://www.nitrd.gov/nitrdgroups/images/0/09/Federal_BD_R&D_Thrusts_and_Priority_Themes.pdf)
- Input on gaps; high impact areas; research and infrastructure investments; education/training; new partnerships.
- Comments must be received by 11/14/2014.

BDSSG Strategic Plan Timeline

- August: Strategies and Themes Document review period
- September/October: Draft and release RFI
- January/February: Workshop to include industry and academia
- February/March– Draft strategic plan
- April/May –Review period
- June – Draft plan made available to Agencies