The Social, Behavioral and Economic Perspective of Big Data Research

Designed data

Data originated from designed sources, e.g.:

- **scientific instruments,**
- **administrative data,**
- **large-scale surveys,**
- **brain research,**
- **large-scale simulations, and etc.**

Organic data

Data produced without explicit data collection design, e.g.:

- **data generated by mobile apps,**
- **ubiquitous sensing apps,**
- **social interactional data from social networking sites, twitter feeds, click streams, and etc.**

SBE + Big Data

FY 2014:
- **computer science**
- **info. science**
- **psychology**
- **geo sciences**
- **statistics**

FY 2015:
- **computer science**
- **info. science**
- **psychology**
- **geo sciences**
- **statistics**
- **economics**
- **business**
- **sociology**
- **political science**

Resource Implementations for Data Intensive Research in the Social Behavioral and Economic Sciences (RIDIR)

We encourage and support communities to develop capabilities dedicated to the creation of new, large-scale, next-generation data resources and relevant analytic techniques to advance fundamental SBE research.

http://www.nsf.gov/pubs/2015/nsf15523/nsf15523.htm

FY 15 SBE BIGDATA Award

✧ **1546259: Population** ... **at Birth from Surveys** ... **Registrations**

❖ **PI: Michael Rendall (U** ... **(UCLA)**

❖ **Amount: $1,201,219**

| ACS: American Community Survey |
| CPS: Cross-sectional Current Population Survey |
| NLSY: National Longitudinal Survey |
| PSID: Panel Survey of Income Dynamics |
| SIPP: Survey of Income and Program Participation |

❖ **To aggregate birth information from the decennial census, ACS, CPS, NLSY, PSID, and SIPP, then to produce large times-series life-course data about poverty and social mobility.**

FY 14 SBE BIGDATA Award

✧ **1447634: Iterative Crowdsourced Hypothesis Generation**

❖ PI: James Bagrow, University of Vermont

❖ To build a crowdsourcing web platform to use the knowledge of interested non-experts (Hunch) and the algorithmic power of computers (Crunch) to discover and test causal relationships.

   o Algorithms identify potential relationships and users are asked to validate them.

   o Users are able to propose their own hypotheses that can subsequently be validated, creating an accelerating feedback loop of scientific discovery.

FY 15 RIDIR Awards

1539228: Longitudinal Intergenerational Family Electronic Micro-Database (LIFE-M)
PI: Bailey, Martha (University of Michigan-Ann Arbor)
Amount: $ 2,129,972

1539129: Enabling Access to and Analysis of Shared Daylong Child and Family Audio Data
PI: Warlaumont, Anne (University of California – Merced)
Amount: $ 985,439

1539302: Modernizing Political Event Data for Big Data Social Science Research
PI: Brandt, Patrick T. (University of Texas at Dallas)
Amount: $ 1,497,358