NIH Data Science

Allen Dearry, PhD

Office of Data Science, NIH OD

Office of Scientific Information Management, NIEHS
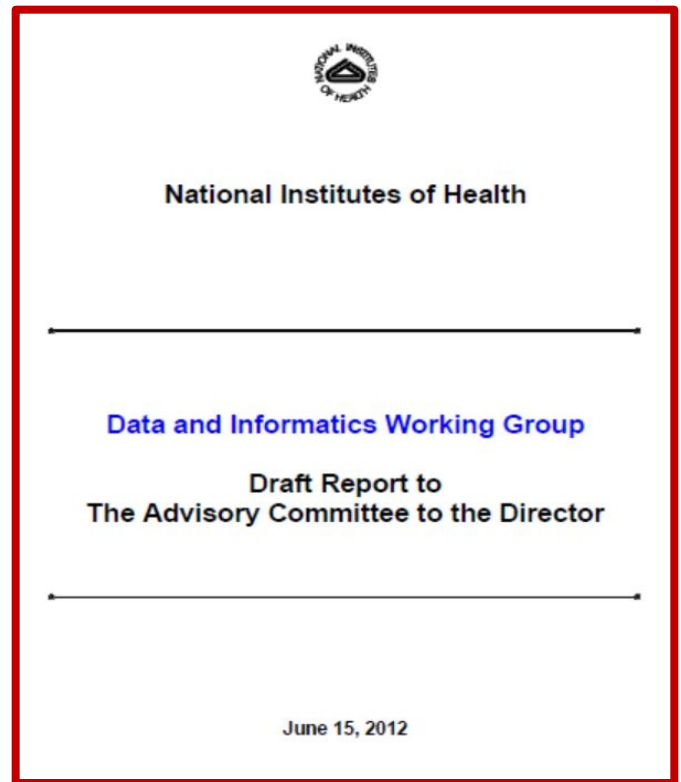
NSF BDPI Meeting ▶ Arlington, VA ▶ April 21, 2016

National Institutes of Health
Office of the Director
Data Science at NIH

What Are the Big Problems to Solve?

1. Locating the data

2. Getting access to the data

3. Extending policies and practices for data sharing

4. Organizing, managing, and processing biomedical Big Data

5. Developing new methods for analyzing biomedical Big Data

6. Training researchers who can use biomedical Big Data effectively

National Institutes of Health

Data and Informatics Working Group

Draft Report to
The Advisory Committee to the Director

June 15, 2012

http://acd.od.nih.gov/diwg.htm

NIH-Supported Data Sharing Repositories

- Model organisms: Drosophila, zebrafish, C elegans, rat, mouse

- Universal Protein Resource, Protein Data Bank

- NCBI: dbGAP, dbSNP, GenBank, Gene Expression Omnibus

- Neuroscience Information Framework, NDAR, fMRI, Traumatic Brain Injury, Parkinson's

- https://www.nlm.nih.gov/NIHbmic/nih_data_sharing_repositories.html

BD2K Information and Funding Opportunities

- RFIs
  - Current: Supporting team science in biomedical research
  - Upcoming: Data sharing; data citation; metrics/value of data repositories; data management in core facilities
  - http://grants.nih.gov/grants/oer.htm
- FOAs
  - Open Science Prize: Innovative, Ground-Breaking Technology with Open Data
  - Development of Technologies in Biomedical Computing, Informatics, Big Data Science
  - https://datascience.nih.gov/bd2k/announcements

Data Science at NIH

▶ https://datascience.nih.gov/adds

▶ bd2k@nih.gov

▶ @NIH_BD2K

▶ #BD2K, #BigData