Federal Definitions of “Data” and Related Terms (& See below for “John’s Definitions”)

Data

• Open Data Policy Memo, May 9, 2013 -- All structured information, unless otherwise noted

• The National Institutes of Health (NIH) Grants Policy Statement defines “data” as “recorded information, regardless of the form or medium on which it may be recorded, and includes writings, films, sound recordings, pictorial reproductions, drawings, designs, or other graphic representations, procedural manuals, forms, diagrams, work flow charts, equipment descriptions, data files, data processing or computer programs (software), statistical records, and other research data.”

NIH definition of “final research data” vs “data”

1. **What do you mean by final research data?**
   By "final research data", we mean recorded factual material commonly accepted in the scientific community as necessary to validate research findings. Final research data do not include laboratory notebooks, partial datasets, preliminary analyses, drafts of scientific papers, plans for future research, peer review reports, communications with colleagues, or physical objects, such as gels or laboratory specimens.

   [LINK]

• The Federal Acquisition Regulations refer to “recorded information, regardless of form or the media on which it may be recorded,” and includes technical data and computer software.

• The Department of Defense Acquisition Regulations (DFARS) defines “technical data,” as “recorded information, regardless of the form or method of the recording, of a scientific or technical nature (including computer software documentation but not software programs, source code, etc.).”

• The Environmental Protection Agency (EPA) defines raw data as “any laboratory worksheets, memoranda, notes or exact copies thereof that are the result(s) of original observations and activities of a study and are necessary for the reconstruction and evaluation of the report of that study.”

• NASA defines “data” as “recorded information, regardless of form, the media on which it may be recorded, or the method of recording, created under the grant. The term includes…data of a scientific or technical nature, and any copyrightable work in which the recipient asserts copyright, or for which ownership was purchased, under the grant.”

• OMB Circular A-110/2CFR 215 defines research data as “the recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not any of the
following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues.

Open Data

- Open Data Policy Memo, May 9, 2013 -- Publicly available data structured in a way that enables the data to be fully discoverable and usable by end users, generally encompassing the following principles: Public, Accessible, Described, Reusable, Complete, Timely, and Managed Post Release.

Dataset

Open Data Policy Memo, May 9, 2013 -- A collection of data presented in tabular or non-tabular form.

Information

OMB Circular A-130 -- Any communication or representation of knowledge such as facts, data, or opinions in any medium or form, including textual, numerical, graphic, cartographic, narrative, or audiovisual forms.

Information Life Cycle

OMB Circular A-130 -- The stages through which information passes, typically characterized as creation or collection, processing, dissemination, use, storage, and disposition.

John’s Definitions -- Same words having different meanings (intended or unintended) leads to confusion regarding the general discussion, web site organization, and institutional responses to “Data Sharing” issues:

“Data Sharing”

1. In a multiple institution agreement, the participating institutions agree to share all accumulated data with each other, during and after completion of the project. (Contract terms and conditions.)

2. An individual researcher agrees to share data with research peers via:
   
   A. Peer-reviewed publication
   B. Publication that is not peer-reviewed (personal web site; institutional data warehouse, etc.)
   C. Providing raw data upon request
   D. Providing raw data after a reasonable amount of time (to permit publication)

3. A researcher (& the researcher’s institution) agree to “share” (i.e., specifically to publish) research data and findings, to make the results available for the general public.

4. A research institution agrees to make research-related information (including data?) available to public requests for information, at various stages (Pending; Funded: After the project has ended) in the research grant cycle.

5. A university creates policies and procedure for “sharing data” (course information; grades; schedules; etc.) across campus. This can include data STORAGE policies and procedures (quantity permitted on various platforms; free vs. restricted access for read/edit rights; no charge/charging for storage, etc.)