



Ownership and Data Sharing

- Ownership
- Proprietary data
- HIPAA



Ownership and Data Sharing

- Ownership—PI, Institution, Funding source
- Proprietary data—IP Ownership / Redaction / Restrictions?
- HIPAA—Privacy / De-identification procedures?



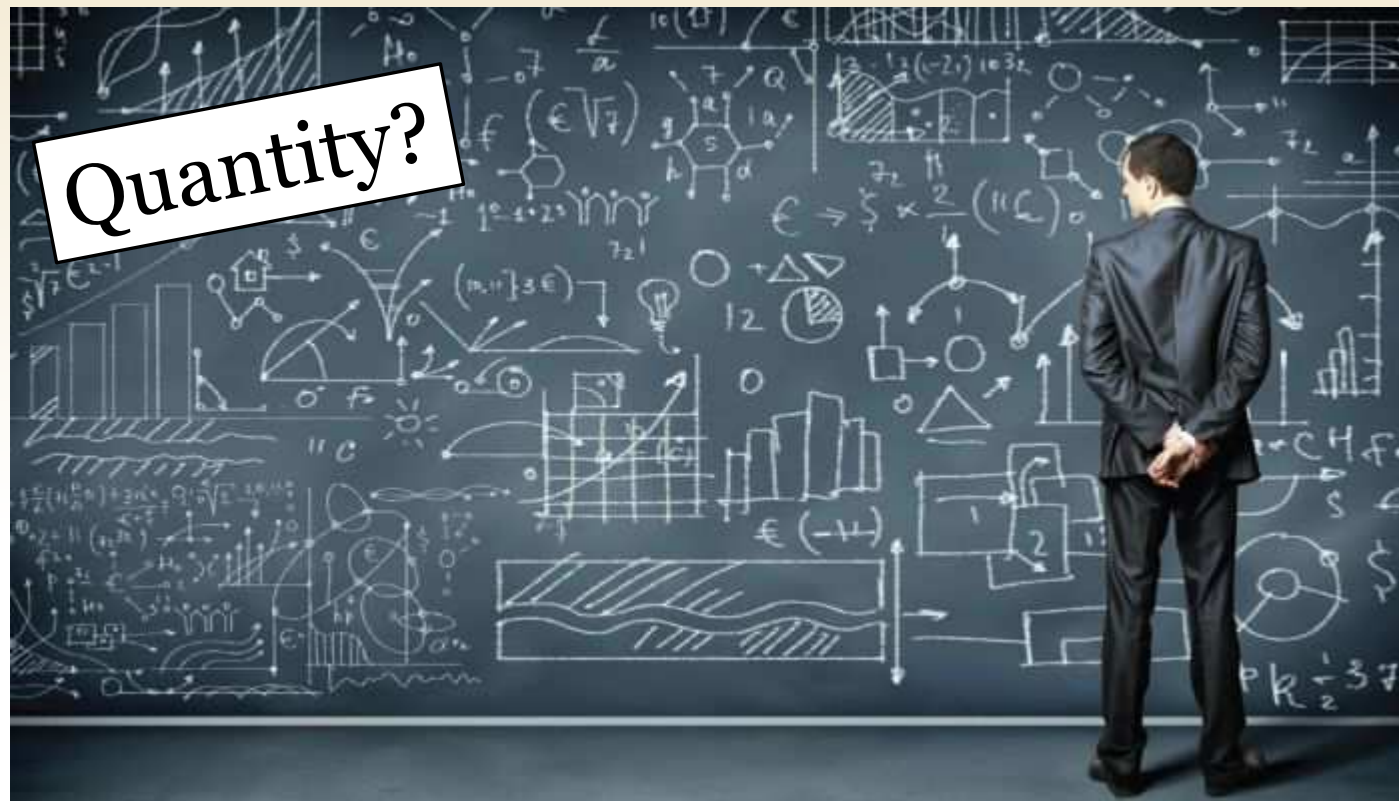
Data Types

Observational

Simulation

Derived

Compiled





Data Format

- Text
- Numeric
- Audiovisual
- Modeling
- Unique/ specific / proprietary
- Meta data / audit trails

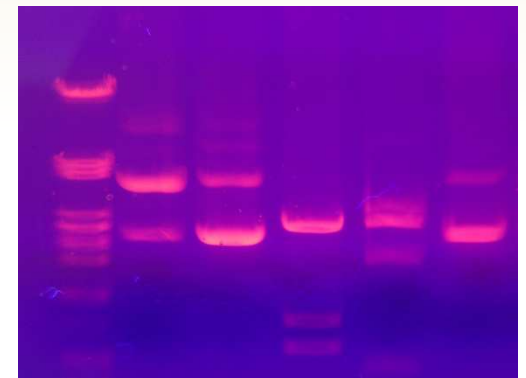
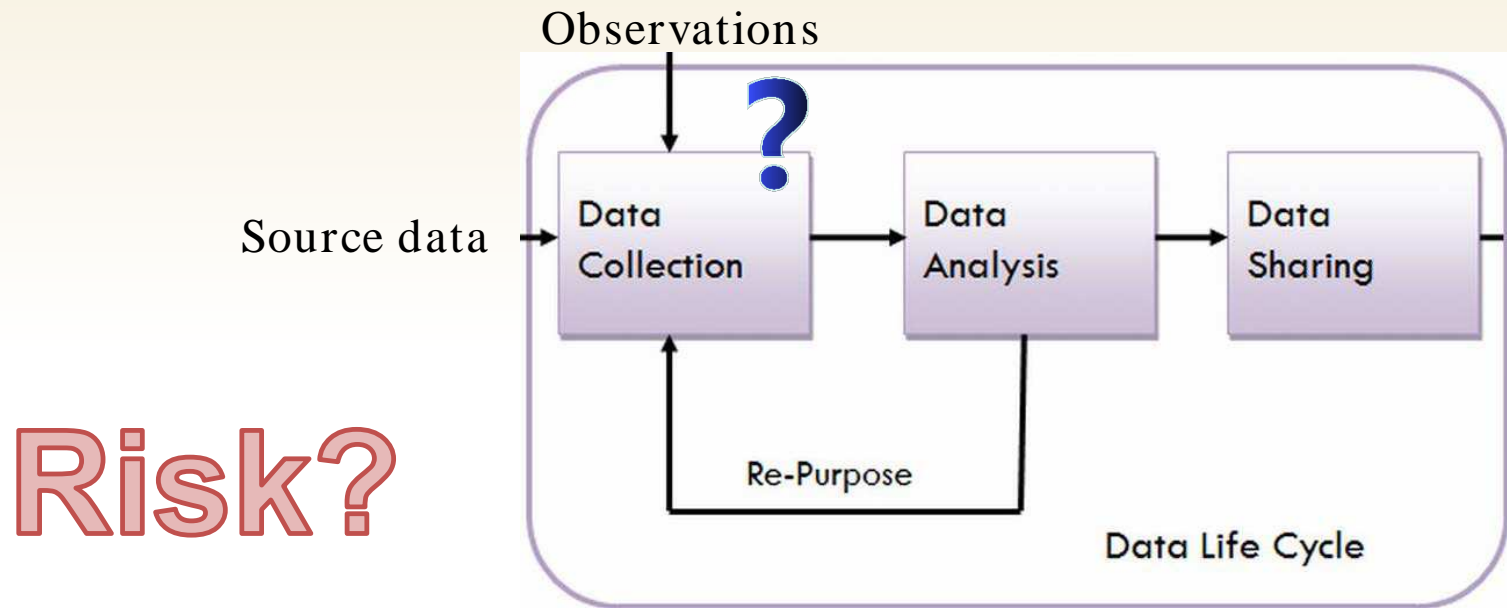


Image courtesy of Bing

Organization and Storage

- Fixed (No changes)
- Updated with no changes
- Updated with changes—Change Control



Organization and Storage

- Location
- Accessibility / Security (including encryption)
- Back-up
- Retention / Archival
- Data Migration / Operating System Sustainability

Organization and Storage

Electronic Data

- Audit Trails
- Transposition Errors
- Software Compatibility
- Program Updates
 - Automatic
 - Impact to significant digits

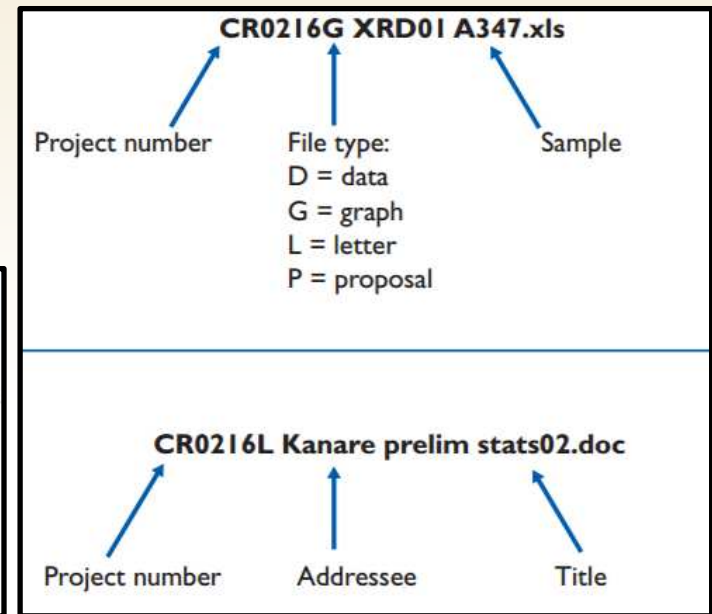
Application	Module	Changed Field	Old Value	New Value	Change Reason	User ID	Date
Documents	InfoCard Edit	Vault	ORNCES_LAB001...	delete	Changes made in...	System Administrator...	2-Aug-2016
Documents	InfoCard Edit	Document Number	SOP-LAB-001-...	SOP-LAB-001-...	Changes made in...	System Administrator...	2-Aug-2016
Documents	InfoCard	Owner	TRBRASEL	DGOODING	Changes made in...	System Administrator...	2-Aug-2016
Documents	InfoCard	Owner	TRBRASEL	DGOODING	Changes made in...	System Administrator...	2-Aug-2016
Documents	InfoCard	Owner	Matthew McGa...	DGOODING	Changes made in...	System Administrator...	2-Aug-2016
Documents	InfoCard Edit	Document Number	SOP-CAL-005-...	SOP-CAL-005-...	Deleted per Jac...	System Administrator...	2-Aug-2016
Documents	InfoCard Edit	Vault	ORNCES_CAL005...	delete	Deleted per Jac...	System Administrator...	2-Aug-2016
Documents	InfoCard	File Name	SOP-LAB-003-...	SOP-LAB-003-...	uploaded new fi...	Jacqu...	18-Jul-2016
Documents	Packets	packet_status	In Process	On Hold	wrong main file...	System...	18-Jul-2016
Documents	Packets	step_status	In Process	On Hold	wrong main file...	System...	18-Jul-2016
Documents	InfoCard	Title	HOBOWare	HOBOWare Lig...	Added File-JDA ...	Jacqu...	8-Jul-2016
Documents	InfoCard	File Name	SOP-LAB-001-...	SOP-LAB-001-...	changed file na...	Jacqu...	2-May-2016
Documents	InfoCard	Author	JDABENDR	CHMASSEY	Updated owner a...	Jacqu...	13-Apr-2016
Documents	InfoCard	Owner	JDABENDR	TRBRASEL	Updated owner a...	Jacqu...	13-Apr-2016
Documents	InfoCard	Creator	JDABENDR	CHMASSEY	Updated owner a...	Jacqu...	13-Apr-2016
Documents	InfoCard	Owner	TRBRASEL	DGOODING	Added SOP File...	Jacqu...	17-Mar-2016
Documents	InfoCard	Title	Spirit File...	Use and Main...	Added SOP File...	Jacqu...	17-Mar-2016
Documents	InfoCard	Document Number	SOP-MCS-000-...	SOP-MCS-000-...	Draft doc prove...	Prasar...	23-Feb-2016
Documents	InfoCard	Document Number	SOP-MCS-000-...	SOP-MCS-000-...	Draft doc using...	Prasar...	23-Feb-2016
Documents	InfoCard	Notes	No attachmen...	No attachmen...	Draft doc using...	Prasar...	23-Feb-2016
Documents	InfoCard	Document Number	SOP-MCS-000-...	SOP-MCS-000-...	Changed Draft b...	Prasar...	23-Feb-2016
Documents	InfoCard	Document Number	SOP-MCS-000-...	SOP-MCS-000-...	This is a draft...	Prasar...	23-Feb-2016
Documents	InfoCard	load_time	17 Feb 2016 ...	17 Feb 2016 ...	Replaced the ma...	Tasha...	17-Feb-2016
Documents	InfoCard	File Name	QAU-000-000...	QAU-000-000...	Replaced the ma...	Tasha...	17-Feb-2016

Risk?

Organization and Storage

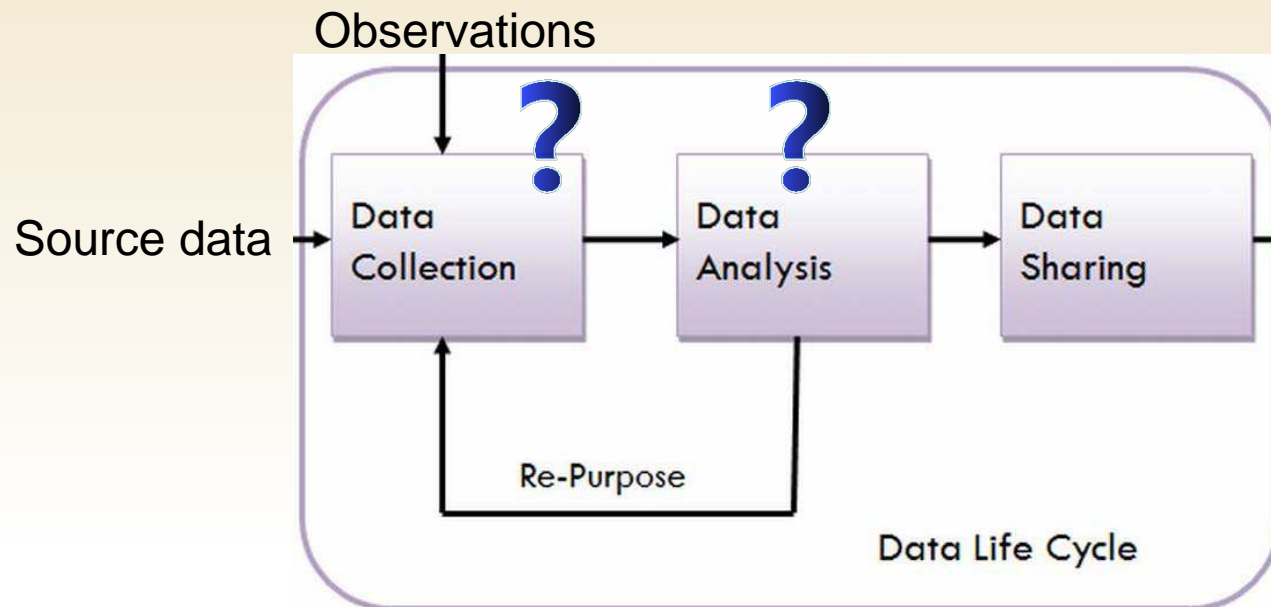
Electronic Data

- Standard File Naming System
 - Brief, descriptive, consistent, dated
 - Plans for edits and changes

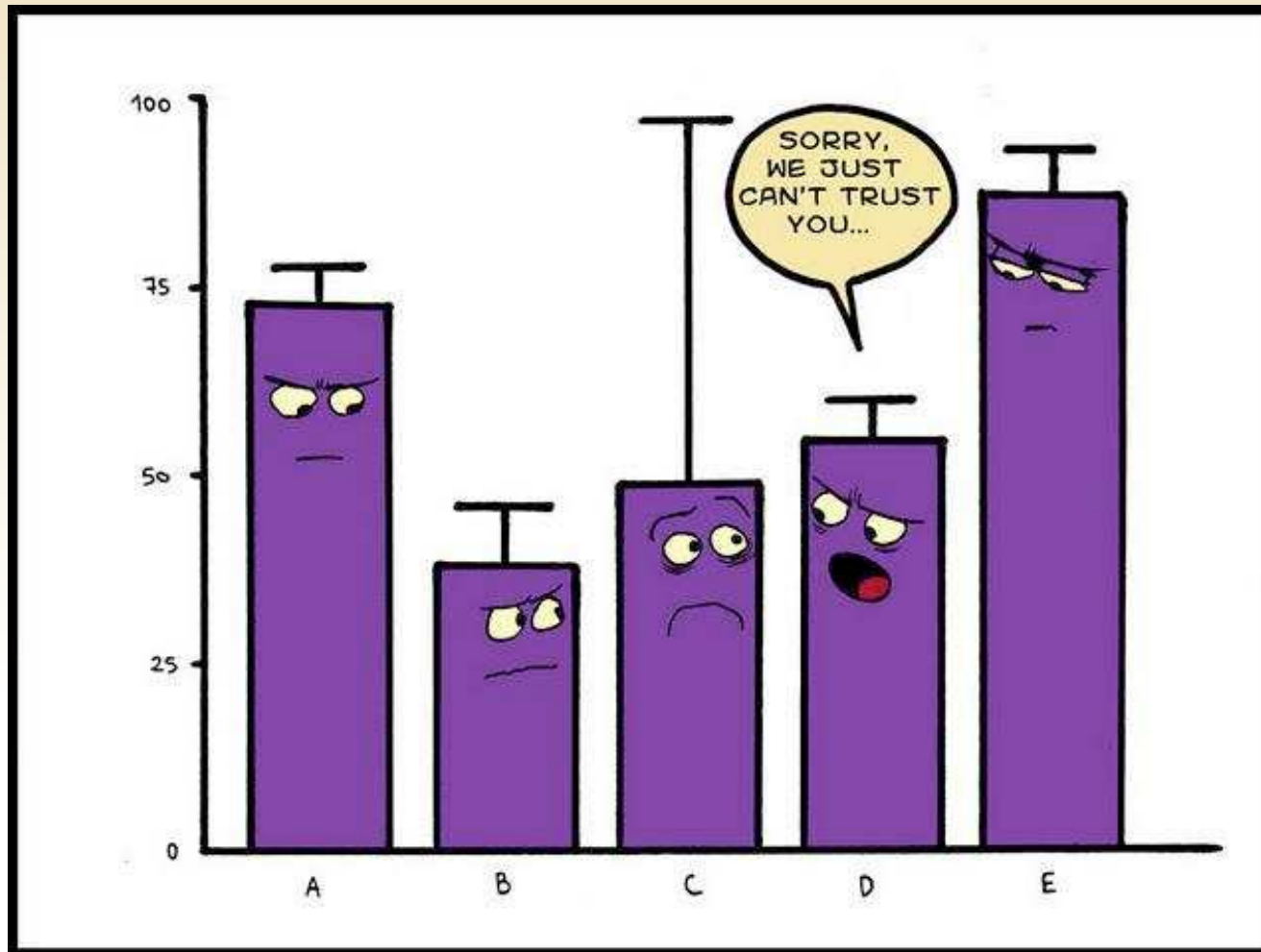




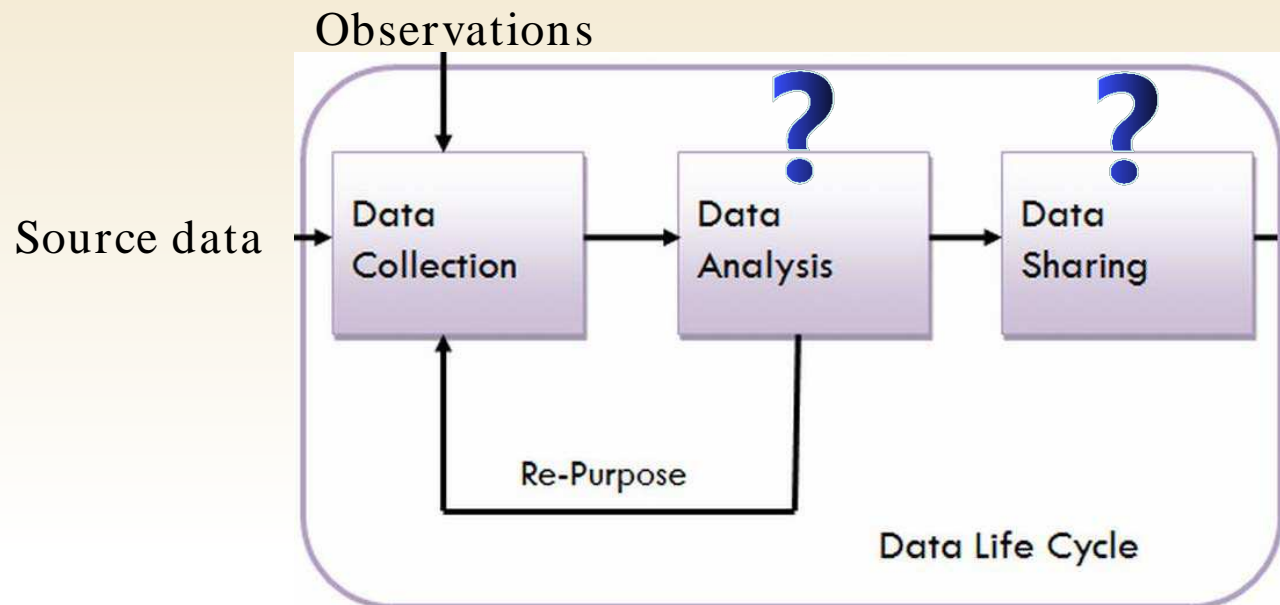
Data Manipulation



Data Manipulation



Mechanisms for Data Sharing



Mechanisms for Data Sharing

- Email
- Online repositories
- Supplemental to publication
- CD / USB
- Sharing agreements
- Mixed
- Conditions / Exclusions

Mechanisms for Data Sharing

Data Sharing / Citation

- Creator(s)
- Title (version)
- Publication year (or release date)
- Publisher or Data Repository
- Identifier



Summary of the eight standards and three levels of the TOP guidelines

Levels 1 to 3 are increasingly stringent for each standard. Level 0 offers a comparison that does not meet the standard.

	LEVEL 0	LEVEL 1	LEVEL 2	LEVEL 3
Citation standards	Journal encourages citation of data, code, and materials—or says nothing.	Journal describes citation of data in guidelines to authors with clear rules and examples.	Article provides appropriate citation for data and materials used, consistent with journal's author guidelines.	Article is not published until appropriate citation for data and materials is provided that follows journal's author guidelines.
Data transparency	Journal encourages data sharing—or says nothing.	Article states whether data are available and, if so, where to access them.	Data must be posted to a trusted repository. Exceptions must be identified at article submission.	Data must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Analytic methods (code) transparency	Journal encourages code sharing—or says nothing.	Article states whether code is available and, if so, where to access them.	Code must be posted to a trusted repository. Exceptions must be identified at article submission.	Code must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Research materials transparency	Journal encourages materials sharing—or says nothing.	Article states whether materials are available and, if so, where to access them.	Materials must be posted to a trusted repository. Exceptions must be identified at article submission.	Materials must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Design and analysis transparency	Journal encourages design and analysis transparency or says nothing.	Journal articulates design transparency standards.	Journal requires adherence to design transparency standards for review and publication.	Journal requires and enforces adherence to design transparency standards for review and publication.
Preregistration of studies	Journal says nothing.	Journal encourages preregistration of studies and provides link in article to preregistration if it exists.	Journal encourages preregistration of studies and provides link in article and certification of meeting preregistration badge requirements.	Journal requires preregistration of studies and provides link and badge in article to meeting requirements.
Preregistration of analysis plans	Journal says nothing.	Journal encourages preanalysis plans and provides link in article to registered analysis plan if it exists.	Journal encourages preanalysis plans and provides link in article and certification of meeting registered analysis plan badge requirements.	Journal requires preregistration of studies with analysis plans and provides link and badge in article to meeting requirements.
Replication	Journal discourages submission of replication studies—or says nothing.	Journal encourages submission of replication studies.	Journal encourages submission of replication studies and conducts blind review of results.	Journal uses Registered Reports as a submission option for replication studies with peer review before observing the study outcomes.

SCIENTIFIC STANDARDS

Promoting an open research culture

Author guidelines for journals could help to promote transparency, openness, and reproducibility

By B. A. Nosek,* G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, M. Contestabile, A. Dafoe, E. Eich, J. Freese, R. Glennerster, D. Goroff, D. P. Green, B. Hesse, M. Humphreys, J. Ishiyama, D. Karlan, A. Kraut, A. Lupia, P. Mabry, T. Madon, N. Malhotra, E. Mayo-Wilson, M. McNutt, E. Miguel, E. Levy Paluck, U. Simonsohn, C. Soderberg, B. A. Spellman, J. Turrillo, G. VandenBos, S. Vazire, E. J. Wagenmakers, R. Wilson, T. Yarkoni

Transparency, openness, and reproducibility are readily recognized as vital features of science (1, 2). When asked, most scientists embrace these features as disciplinary norms and values (3). Therefore, one might expect that these valued features would be routine in daily practice. Yet, a growing body of evidence suggests that this is not the case (4–6).

A likely culprit for this disconnect is an academic reward system that does not sufficiently incentivize open practices (7). In the present reward system, emphasis on innovation may undermine practices that support verification. Too often, publication requirements (whether actual or perceived) fail to encourage transparent, open, and reproducible science (2, 4, 8, 9). For example, in a transparent science, both null results and statistically significant results are made available and help others more accurately assess the evidence base for a phenomenon. In the present culture, however, null results are published less frequently than statistically significant results (10) and are, therefore, more likely inaccessible and lost in the “file drawer” (11).

POLICY The situation is a classic collective action problem. Many individual researchers lack

Downloaded from <http://science.sciencemag.org/> on June 9, 2016

sciencemag.org SCIENCE

Data Management and Sharing Plan

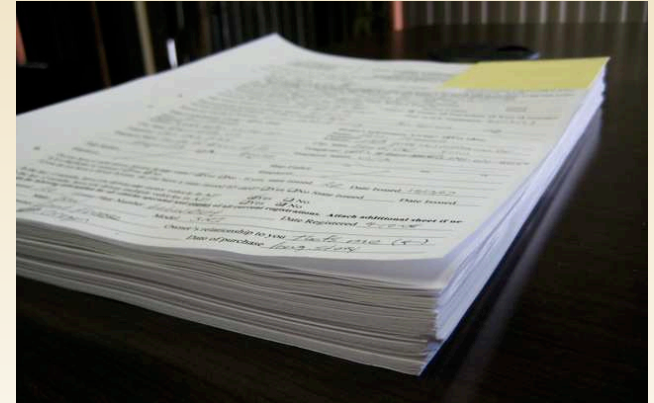
- Data Description / Types
- Data Standards for Format and Content
- Mechanisms for Access and Sharing (Provisions, Privacy Protection, confidentiality, Security, Intellectual Property, etc.)
- Provisions for Data Reuse and Redistribution
- Archiving / Long-term Preservation and access
- Other...



Data Quality: GDP Principles

BONUS!

- Atributable
- Legible
- Contemporaneous
- Original
- Accurate



ALCOA+

Complete, Consistent, Enduring, Readily Available

Part 4: Case Studies

Case Study 1—Data Sharing

Your research study will include data from approximately 500 subjects being screened for three bacterial sexually transmitted diseases (STDs) at an inner city STD clinic. The final dataset will include self-reported demographic and behavioral data from interviews with the subjects and laboratory data from urine specimens provided. Because the STDs being studied are reportable diseases, you will be collecting identifying information.

Even though the final dataset will be stripped of identifiers prior to release for sharing, there remains the possibility of deductive disclosure of subjects with unusual characteristics. Identify options (i.e., conditions) for sharing the data.

Case Study 2—Data Management

You will be conducting a single-dose drug study in a mouse model to determine if Drug X will have an impact on cholesterol values. You plan to collect weekly blood samples over a 28-day period to statistically analyze LDL and HDL levels.

Describe processes that might be in place for the storage, protection and retention of the laboratory data during and after study completion.

Case Study 3—Data Organization

You are conducting a study that will generate the following electronic files. Develop a plan to organize the electronic folders and develop a standard naming system for the files that will allow for tracking of any changes. (Note: Consider if it is more useful to order the files by date, by author, or by subject, for example).

Project Number: JJ2A2016

Study conducted: January - September, 2016

- PCR data, analyzed in batches performed weekly by 3 different trained research associates (.scn files)
- Gel images, collected weekly by 2 different trained research associates (.tif files)
- Compiled blood chemistry data updated 2-3 times per week (.xlsx files)
- Compiled hematology data updated 2-3 times per week (.xlsx files)
- Statistical analysis in GraphPad Prism (.pzfx files)

Summary

- **Who** owns the data (resource)?
- **What** are your data?
- **When** will the data be available?
- **Where** will the data be retained?
- **How** will the data be organized, secured and shared?



References – Data Management

https://dmptool.org/dm_guidance (General guidance)

<http://www.icpsr.umich.edu/files/datamanagement/DataManagementPlans-All.pdf>
(Guidelines)

<http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/dmp/plan.html>
(Sample plan)

<http://lgdata.s3-website-us-east-1.amazonaws.com/docs/2784/1033892/DM101Checklist101314.pdf> (Checklist)

<https://obamawhitehouse.archives.gov/blog/2016/02/22/increasing-access-results-federally-funded-science> (White House Memorandum Archives)

<http://libguides.northwestern.edu/datamanagement/federalagency#s-lg-box-wrapper-7012284> (Other Federal funding agencies)

http://www.hhmi.org/sites/default/files/Educational%20Materials/Lab%20Management/Making%20the%20Right%20Moves/moves2_ch8.pdf (Data Management and Lab Notebooks)

<https://grants.nih.gov/grants/NIH-Public-Access-Plan.pdf> (NIH Public Access Plan)

<http://libraries.ucsd.edu/services/data-curation/data-management/dmp-samples.html>
(Example DMPs)