

Data Management Tools

Lizzy Rolando, Georgia Tech

Aaron Trehub, Auburn University

August 6, 2013

A brief history of how we got here



The march of data, 3000 BC – 2010 AD



2011-2013



National Institutes
of Health

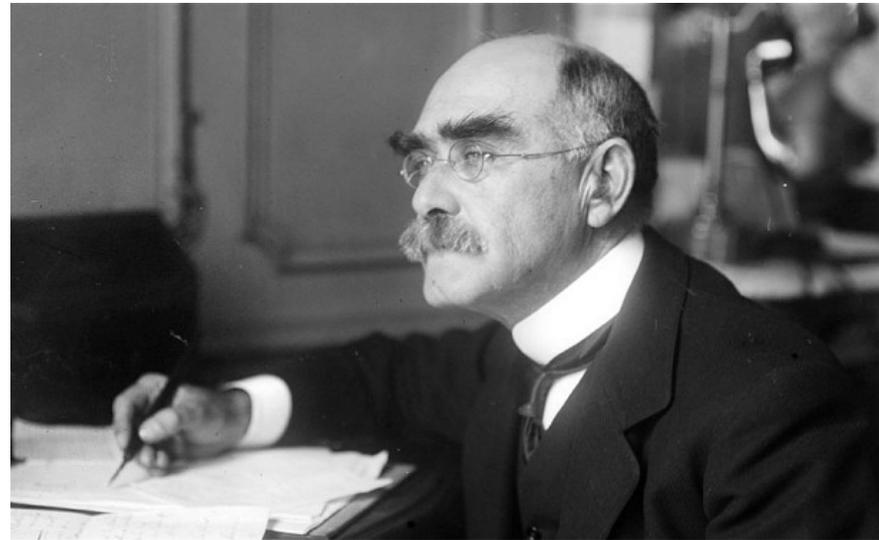


NATIONAL
ENDOWMENT
FOR THE
HUMANITIES

Etc.

Kipling on data management

“I KEEP six honest serving-men
(They taught me all I knew);
Their names are **What** and **Why** and **When**
And **How** and **Where** and **Who.**”



Tools: the big three

- DataONE
(<http://www.dataone.org/>)
- DMPTool (<https://dmp.cdlib.org/>)
- Databib (<http://databib.org/>)

DataONE

“Data Observation Network for Earth (DataONE) is the foundation of new innovative environmental science through a distributed framework and sustainable cyberinfrastructure that meets the needs of science and society for open, persistent, robust, and secure access to well-described and easily discovered Earth observational data.” [what, why, when, how, where, who]

The DataONE Best Practices DB

“The DataONE Best Practices database provides individuals with recommendations [what] on how to effectively work with their data [how] through all stages of the data lifecycle [when]. Users can access best practices within the database by either clicking on a stage of the lifecycle, selecting keywords (under advanced search) or using free search.”

DMPTool

DMPTool helps researchers “create ready-to-use data management plans [**what**] for specific funding agencies; meet requirements for data management plans [**why**]; get step-by-step instructions and guidance for data management plans [**how**]; learn about resources and services available at your institution [**where**] to fulfill the data management requirements [**why, again**] of their grants.”

DMPTool2

“With generous funding from the Alfred P. Sloan Foundation, the DMPTool team is developing the next generation of the DMPTool.”

Additionally, the DMPTool Project received an IMLS grant to develop education and outreach materials for librarians, including a webinar series and a centralized location for resources.

Databib

“Databib is a tool for helping people identify and locate [**where**] online repositories [**what**] of research data.”

(Also: the Directory of Open Access Repositories [[OpenDOAR](#)])

DCC's Disciplinary Metadata Registry

“While data curators, and increasingly researchers [**who**], know that good metadata [**what**] is a key for research data access and re-use [**why**], figuring out precisely what metadata to capture and how to capture it is a complex task... This page provides links to information about these disciplinary metadata standards, including profiles, tools to implement the standards, and use cases of data repositories currently implementing them [**how**].”

Training Resources

- MANTRA
- DataONE
- UK Data Archive
- RDMRose

Tools: local spin-offs

In 2012, ASERL and SURA formed the Research Data Management Group in response to federal data management requirements.

The group produced two documents...

I. [A Step-By-Step Guide to Data Management](#)

The goal: a simple, one-page introduction to data management for faculty and researchers, with links to appropriate tools.

The product: a longer (five pages), more complex document.

Basically, a radically distilled version of the [DataONE Best Practices Primer](#).

The steps

1. Assemble your data management toolkit
2. Plan (at the beginning of the project)
3. Collect and check your data
4. Describe and document your data
5. Select a repository for your data
6. Store and preserve your data

2. ASERL/SURA Model Language for Research Data Management

The goal: “...to assist their member institutions in drafting sound policies to govern the uses and management of research data generated by university faculty and staff.”

Again, a guide or template, not an out-of-the-box solution.

Model policy components

1. Aims and objectives
2. Suggested policy
3. Data ownership
4. Stakeholders
5. Related policies (e.g. university's IP policy)
6. Appendix (roles and responsibilities)
7. Suggested sources

Some questions for discussion

Have you used these tools at your institution/Library?

Do researchers at your institution know about these tools?

Do they use them?

Do they use them effectively?

What's our role as librarians and subject specialists?

Thank you

Lizzy: lizzy.rolando@library.gatech.edu

Aaron: trehuaj@auburn.edu