



ICPSR

SHARING DATA TO ADVANCE SCIENCE

ICPSR and CoreTrustSeal: Repository Certification Experiences and Opportunities

Jared Lyle
IASSIST
Sydney, NSW, Australia
May 29, 2019



Outline

- Overview of ICPSR
- Why assessment is important
- ICPSR's experience with assessment, including effort and resources needed
- Benefits from assessment
- Opportunities to improve the process



- DATA & PUBLICATIONS
- EVERYTHING ELSE

Search Data and Publications

Search

Log In/Create Account



ICPSR

Sharing data to advance science

News



Jun 22, 2017

An ICPSR team gives viewers a spin through our new Data Manager, and a peek at what's next



Jun 19, 2017

ICPSR celebrates launch of new data deposit system!



Jun 16, 2017

Father's Day stats from US Census Bureau

More Headlines

Jun 12, 2017

New Releases through 2017-06-11

May 31, 2017

Due June 19: Call for abstracts for the Jamaica Statistics Symposium 2017, being put on by The Jamaica Statistical Society

May 12, 2017

"Who was the most important person in raising you?" Explore data from the Flint Adolescent Study

[view all news](#)

<http://www.icpsr.umich.edu>

ICPSR



Philip Converse, Warren Miller, and Angus Campbell

- Established 1962
- Originally 22 Members, now consortium of 776 world-wide
- Originally Political Science, now all social and behavioral sciences

ICPSR

- Current holdings
 - 10,000+ studies, quarter million files
 - 1500+ are *restricted studies*, almost always to protect confidentiality
 - Bibliography of Data-related Literature with 80,000 citations
- Approximately 60,000 active MyData (“shopping cart”) accounts
- Thematic collections of data about addiction and HIV, aging, arts and culture, child care and early education, criminal justice, demography, health and medical care, and minorities



Why Assessment is Important

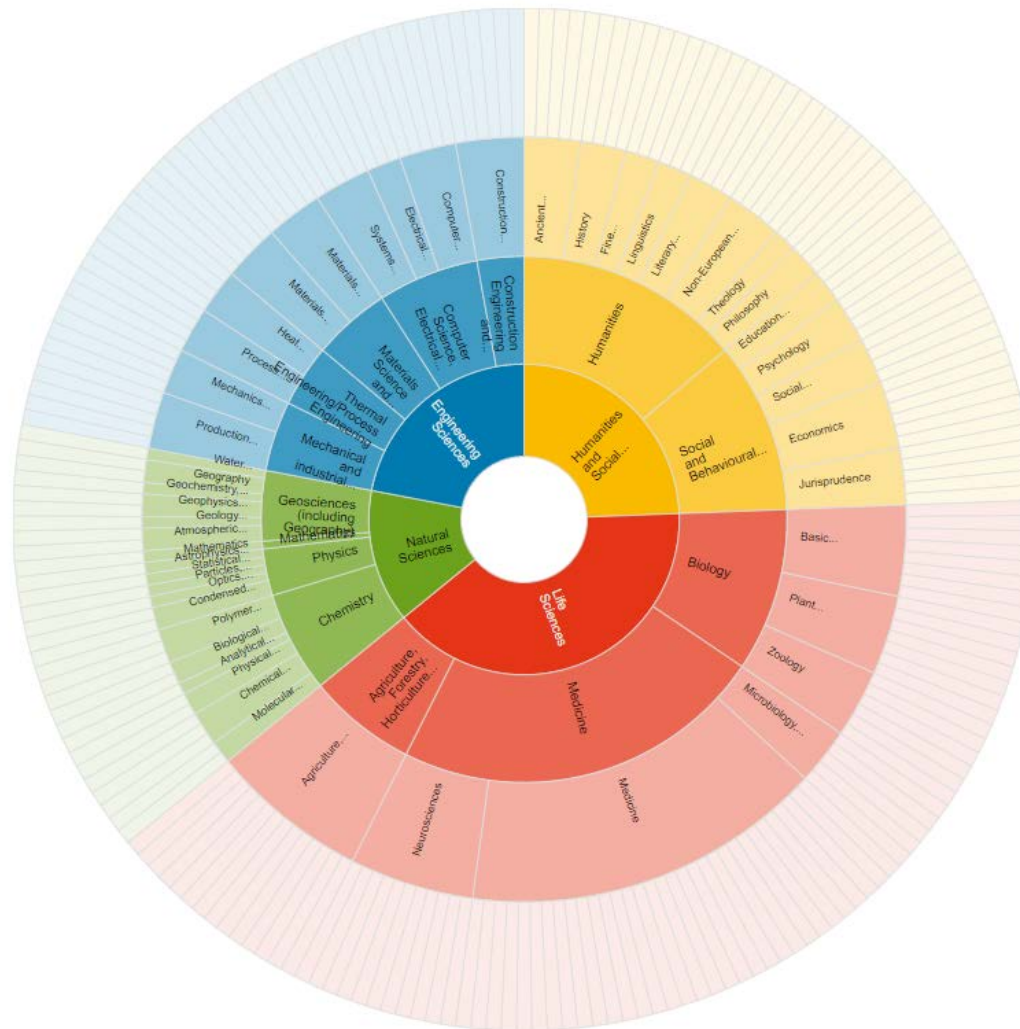
- Provide **transparent** view into the repository
 - **Improve** processes and procedures
 - Measure against a **community standard**
 - Show the **benefits of domain repositories**
-
- Promote **trust** by funding agencies, data producers, and data users that data will be available for the long term

re3data.org

REGISTRY OF RESEARCH DATA REPOSITORIES

Search...

Search



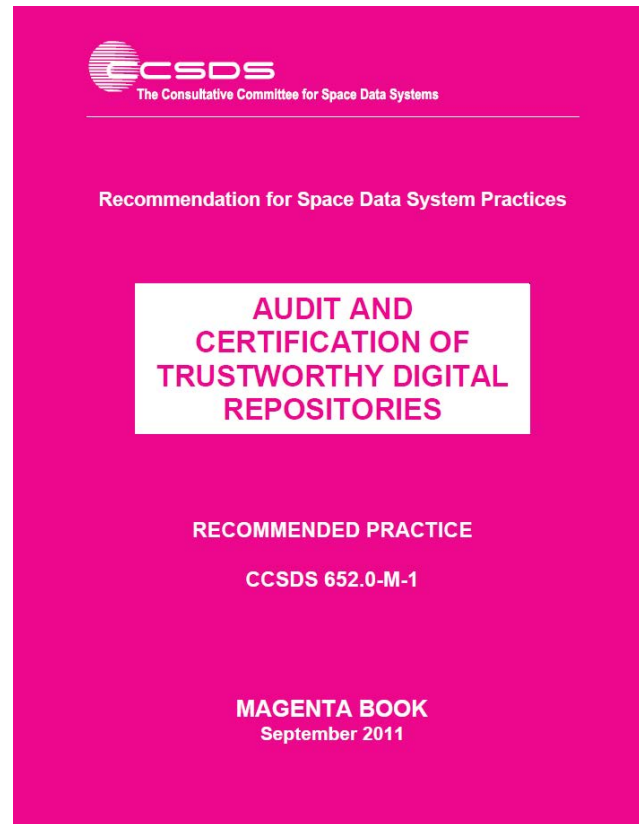


Forever!

Guaranteed!

We promise!

“Claims of trustworthiness are easy to make but are thus far difficult to justify or objectively prove.”



May 12, 2014 by Steve Kolowich

     |  Comments (51)


Hazards of the Cloud: Data-Storage Service's Crash Sets Back Researchers

Dedoose, a cloud-based application for managing research data, suffered a “devastating” technical failure last week that caused academics across the country to lose large amounts of research work, some of which may be gone for good.

SocioCultural Research Consultants, the company that sells Dedoose, is still scrambling to recover as much of its customers' work as possible, and has said in a blog post that “the vast majority” of research data on its platform were not affected.

The crash nonetheless has dealt frustrating setbacks to a number of researchers, highlighting the risks of entrusting data to third-party stewards.

<http://chronicle.com/blogs/wiredcampus/hazards-of-the-cloud-data-storage-services-crash-sets-back-researchers/52571>



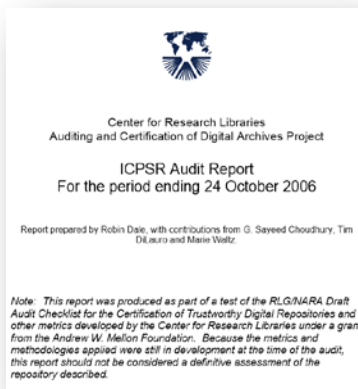
*If we want to be able to share data, we need to store them in a **trustworthy** data repository. Data created and used by scientists should be managed, curated, and archived in such a way to **preserve the initial investment** in collecting them. Researchers must be certain that data held in **archives remain useful and meaningful into the future.***

“An Introduction to the Core Trustworthy Data Repositories Requirements”

https://www.coretrustseal.org/wp-content/uploads/2017/01/Intro_To_Core_Trustworthy_Data_Repositories_Requirements_2016-11.pdf

ICPSR Assessment Experience

2005-2006	CRL test audit (TRAC checklist)
2010-2012	TRAC/ISO 16363 self-assessment
2009-2010	Data Seal of Approval certification
2013	Data Seal of Approval (update)
2013	World Data System certification
2018-2019	CoreTrustSeal



Common Elements of Assessment

- The Organization and its Framework
 - Governance, staffing, policies, finances, etc.
- Technical Infrastructure
 - System design, security, etc.
- Treatment of the Data
 - Access, integrity, process, preservation, etc.

CRL Test Audit, 2005-2006

- Test methodology based on RLG-NARA Checklist for the Certification of Trusted Digital Repositories
- Assessment performed by an external agency (CRL)
- Precursor to current TRAC audit/certification
- ICPSR Test Audit Report:
http://www.crl.edu/sites/default/files/attachments/pages/ICPSR_final.pdf

Effort and Resources Required

- Completion of Audit Checklist
- Gathering of large amounts of data about the organization – staffing, finances, digital assets, process, technology, security, redundancy, etc.
- Weeks of staff time to do the above
- Hosting of audit group for two and a half days with interviews and meetings
- Remediation of problems discovered

Findings

Positive review overall:

*Taken as a whole, ICPSR appears to provide **responsible stewardship** of the valuable research resources in its custody. **Depositors of data to the ICPSR data archives and users of those archives can be confident** about the state of its operation, and the processes, procedures, technologies, and technical infrastructure employed by the organization.*

Findings

Positive review overall, but...

- Succession and disaster plans needed
- Funding uncertainty (grants)
- Acquisition of preservation rights from depositors
- Need for more process and procedural documentation related to preservation
- Machine-room issues noted

Changes Made

- Hired a Digital Preservation Officer
- Created policies, including Digital Preservation Policy Framework, Access Policy Framework, and Disaster Plan
- Changed deposit process to be explicit about ICPSR's right to preserve content
- Continued to diversify funding (ongoing)
- Made changes to machine room

DSA Self-Assessment, 2009-2010

ICPSR
Inter-university Consortium for
Political and Social Research

A partner in social science research

Find & Analyze Data

Membership in ICPSR

Deposit Data

ICPSR Summer Program

Resources for Instructors

Data Management

ICPSR provides leadership and training in data access, curation, and methods of analysis for a diverse and expanding social science research community.

Announcements

[New Releases through 2013-07-14](#)

[12 Weeks until the 2013 ICPSR Meeting!](#)

[ICPSR Posts Guidelines for OSTP Data Access Plans](#)

[ICPSR Updates its Data Management & Curation Site](#)

[More...](#)

Data Seal of Approval 2010

ICSU
WORLD DATA SYSTEM



http://assessment.datasealofapproval.org/assessment_78/seal/pdf
<http://hdl.handle.net/2027.42/144318>

Data Seal of Approval

- Started by DANS in 2009
- The objectives of the DSA are to “safeguard data, to ensure high quality and to guide reliable management of data for the future without requiring the implementation of new standards, regulations, or high costs.”

<http://www.datasealofapproval.org/en/information/about/>

Data Seal of Approval

- 16 guidelines – 3 target the data producer, 3 the data consumer, and 10 the repository
- Example guideline: *(7) The data repository has a plan for long-term preservation of its digital assets.*
- Self-assessments are done online with ratings and then peer-reviewed by a DSA Board member

Procedures Followed

- Digital Preservation Officer and Director of Collection Delivery conducted self-assessment, assembled evidence, completed application
- Provided a URL for each guideline

Effort and Resources Required

- Mainly time of the Digital Preservation Officer and Director of Collection Delivery
- Would estimate two days at most
- Less time required to recertify every two years

Findings and Changes Made

- Recognized need to make policies more public – e.g., static and linkable Terms of Use (previously only dynamic)
- Reinforced work on succession planning – now integrated into Data-PASS partnership agreement
- Underscored need to comply with OAIS – building a new system based on it

CoreTrustSeal, 2018-2019



CoreTrustSeal

- Developed by the DSA-WDS Partnership Working Group on Repository Audit and Certification, a Working Group of the Research Data Alliance
- Merging of the Data Seal of Approval certification and the World Data System certification
- 16 criteria (guidelines)

Requirements

- 16 criteria (guidelines):
 - Organizational Infrastructure (6)
 - Digital Object Management (8)
 - Technology (2)

Example of Evidence – R5

- Guideline Text: R5. The repository has adequate funding and sufficient numbers of qualified staff managed through a clear system of governance to effectively carry out the mission

Example of Evidence – R5

Guidance: The range and depth of expertise of both the organization and its staff, including any relevant affiliations (e.g., national or international bodies), is appropriate to the mission.

ICPSR Response: R5 (one part)

A 12-person Council whose members are elected by the ICPSR membership provides guidance and oversight to ICPSR. Members serve four-year terms, and six new members are elected every two years. The Council acts on administrative, budgetary, and organizational issues on behalf of all the members of ICPSR. [6]

ICPSR's staff of over 100 perform a variety of functions to support ICPSR's archival and training missions. The staff include data curators and managers, librarians, Web developers, communications specialists, user support specialists, administrative staff, and a small team of researchers, as well as software developers, programmers, system administrators, and desktop support specialists. Staff have expertise in digital archiving, data preservation, usability testing, Section 508 review for ADA Section 8 compliance, DOI registration, web traffic analytics, search engine optimization, storage and dissemination of sensitive data, restricted-use data agreements, and researcher credentialing. All staff are required to complete ongoing training related to data security and disclosure risk. [7]

ICPSR Response: R5 (references)

References:

[1] ICPSR Web site, About the Organization:

<https://www.icpsr.umich.edu/icpsrweb/content/about/index.html> (accessed 2018-10-04)

[2] ICPSR 2016-2017 Annual Report, Financial Reports:

<https://www.icpsr.umich.edu/files/ICPSR/about/annualreport/2016-2017.pdf> (accessed 2018-11-08)

[3] ICPSR Web site, Thematic Data Collections:

<https://www.icpsr.umich.edu/icpsrweb/content/about/thematic-collections.html> (accessed 2018-10-04)

[4] ICPSR Web site, List of Member Institutions and Subscribers:

<https://www.icpsr.umich.edu/icpsrweb/membership/administration/institutions> (accessed 2018-11-06)

...

Effort and Resources Required

- 3-5 days of time by the Director of Metadata and Preservation
- Less time required to certify every 3 years

Dear Jared Lyle,

Thank you! We have successfully received your application on behalf of Inter-university Consortium for Political and Social Research to be certified under the CoreTrustSeal. Your application has now been locked, and further editing of its contents is not possible at this time. A URL to access the application for viewing or printing can be found here: <https://amt.coretrustseal.org/inter-university-consortium-for-political-and-social-research/1>

This application has the status of Application Submitted. Please note that if this the first submission of the current application, the review process will begin only after payment has been confirmed of the 1000 EUR administration fee by Inter-university Consortium for Political and Social Research. The legal entity managed by the CoreTrustSeal Secretariat will notify you as to how this fee should be paid. Upon receipt, the application will be sent for review, and we will contact you again with the decision of the CoreTrustSeal Standards and Certification Board, once available. We kindly ask for your patience during this time.

You added the follow additional comments to your application:

Best regards,
CoreTrustSeal Secretariat
(On behalf of the CoreTrustSeal Standards and Certification Board)

CoreTrustSeal Secretariat
The Hague, Netherlands | Tokyo, Japan
+31 6 2386 3243 | +81 4 2327 6395
info@coretrustseal.org

Dear Jared,

On behalf of the CoreTrustSeal Standards and Certification Board, I would like to ask you the following details with respect to the CoreTrustSeal application for Inter-university Consortium for Political and Social Research.

- name of the organization to which the invoice can be addressed
- organization address to which the invoice can be addressed
- email address that can be used for sending the invoice

Upon receipt, the CoreTrustSeal Secretariat will prepare the invoice for the administrative fee, see <https://www.coretrustseal.org/apply/administrative-fee/> .

Thank you in advance.

Should you have any questions, do not hesitate to contact me.

Best wishes,
Lisa de Leeuw
(On behalf of the CoreTrustSeal Standards and Certification Board)



finance@CoreTrustSeal.org | www.CoreTrustSeal.org

Dear Jared,

This is to confirm that Foundation CoreTrustSeal received the administrative fee from ICPSR for 1 CoreTrustSeal Data Repository Certification in good order.

The CoreTrustSeal Secretariat will assign reviewers shortly that will be given a deadline of maximum two months to provide their thoughts to the CoreTrustSeal Board, which will then make a decision at its subsequent meeting. Afterwards, you will be informed on their decision and/or feedback.

Best wishes,

Lisa de Leeuw

(On behalf of the CoreTrustSeal Standards and Certification Board)



finance@CoreTrustSeal.org | www.CoreTrustSeal.org

Still in progress

- All areas marked by reviewers as
4 – The guideline has been fully implemented in the repository

Reviewer feedback:

“This is a well-structured, written and evidenced application. The only suggestion is to briefly review self-assessment statement related to external evidence links to ensure they clarify which aspect of the evidence is relevant. It would also be helpful to clarify which parts of the linked evidence is relevant in the case of longer documents.”

Findings and Changes Made

- Ancillary fine-tuning:

R13. The repository enables users to discover the data and refer to them in a persistent way through proper citation.

- Making older versions of collections available
- File-level persistent identifiers and citations

- Other fine-tuning:

R3. The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.

Comparison of Assessments – Effort and Resources

- Test audit was the most labor- and time-intensive
- TRAC self-assessment involved the time of more people
- CoreTrustSeal (Data Seal of Approval and World Data System) certification least costly

Comparison of Assessments – Benefits

- What did we learn and did the results justify the work required?
 - Test audit was first experience – resulted in greatest number of changes, greatest increase in awareness
 - Fewer changes made as a result of CoreTrustSeal (DSA and WDS); also not as detailed
 - TRAC assessment has surfaced additional issues to address

Benefits continued

- Difficult to quantify
 - Trust of stakeholders
 - Transparency
 - Teaching opportunity for new staff
 - Improvements in processes and procedures
 - Use of community standards and alignment across domains

Other comparisons

- Support by leadership
- Organization-wide involvement
- Interest from community

More opportunities

- IASSIST community of practice for social science CoreTrustSeal applicants and recipients
- Machine-actionable documentation



Thank you!

lyle@umich.edu