Assessment Information

CoreTrustSeal Requirements 2017–2019

Repository: Inter-university Consortium for Political and Social Research
Website: https://www.icpsr.umich.edu/
Certification Date: 28 August 2019

This repository is owned by: University of Michigan
Inter-university Consortium for Political and Social Research

Notes Before Completing the Application

We have read and understood the notes concerning our application submission.

True

Reviewer Entry
Reviewer 1
Comments: 

Reviewer 2
Comments: Accept

CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

Background & General Guidance

Glossary of Terms

BACKGROUND INFORMATION

Context

R0. Please provide context for your repository.
Repository Type. Select all relevant types from:

- Domain or subject-based repository

Reviewer Entry

Reviewer 1
Comments: Accept

Reviewer 2
Comments: Accept

Brief Description of Repository

An international consortium of more than 780 academic institutions and research organizations, the Inter-university Consortium for Political and Social Research (ICPSR) provides leadership and training in data access, curation, and methods of analysis for the social science research community. ICPSR curates, preserves, and disseminates original social science data for research, instruction, and policy evaluation. The organization archives over 10,000 data collections comprising more than 250,000 files of data and documentation. It hosts 21 specialized collections of data in education, aging, criminal justice, substance abuse, terrorism, and other fields. ICPSR collaborates with a number of funders, including U.S. statistical agencies and foundations, to create thematic data collections and data stewardship and research projects. ICPSR's educational activities include the Summer Program in Quantitative Methods of Social Research, a comprehensive curriculum of intensive courses in research design, statistics, data analysis, and social methodology. ICPSR also leads several initiatives that encourage use of data in teaching, particularly in undergraduate instruction. ICPSR-sponsored research focuses on the emerging challenges of digital curation and data science. ICPSR leads or takes part in many policy initiatives and grant-funded activities that result in publications that address issues related to data stewardship. ICPSR researchers also examine substantive issues related to our collections. ICPSR is a unit within the Institute for Social Research at the University of Michigan and maintains its office in Ann Arbor. [1]

ICPSR was a recipient of the 2019 National Medal for Museum and Library Service, the United States’ highest museum and library honor. [2]

References:

**Brief Description of the Repository’s Designated Community.**

ICPSR’s Designated Community includes: researchers, students, practitioners, and other users who need to obtain access to the ICPSR archive of social and behavioral science data. [1]

References:


**Level of Curation Performed. Select all relevant types from:**

A. Content distributed as deposited, B. Basic curation – e.g. brief checking; addition of basic metadata or documentation, C. Enhanced curation – e.g. conversion to new formats; enhancement of documentation, D. Data-level curation – as in C above; but with additional editing of deposited data for accuracy

**Comments**

ICPSR offers both an option of making data available to the user community in the condition deposited (no curation) along with an option for member-funded curation, which involves review, enhancement, and quality checking of the data to
ensure usability and findability. [1]

The membership-funded General Archive, as well as the thematic collections, generally follow ICPSR’s standard curation workflow. The specific steps depend on the unique characteristics of each dataset, but in general, ICPSR data processors always perform the following procedures: review data for confidentiality issues, convert hard copy documentation to electronic form, convert software-specific documentation to PDF/A, generate multiple data formats for dissemination and preservation, create documentation compliant with the DDI specification, create a metadata record, assign a Digital Object Identifier (DOI). They may also do the following: recode variables to address confidentiality concerns, check for undocumented/out of range codes, add question text to variables, create variable labels, create value labels, identify and address foreign language characters, adjust format widths, optimize file size, standardize missing values, check for consistency and skip patterns, make online analysis version with question text, add variables to the Social Science Variables Database, gather citations to related publications for the Bibliography of Data-Related Literature. [2]

openICPSR is an alternative to the standard ICPSR deposit and curation process, providing immediate, self-controlled distribution of data and metadata, albeit without the rich curation and preservation features; data in openICPSR are distributed and preserved as-is, exactly as they arrive and without ICPSR-added curation.

References:
[1] ICPSR Collection Development Policy, Curation Levels:
https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/policies/colldev.html (accessed 2018-10-04)

[2] ICPSR Web site, Data Enhancement:
https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/lifecycle/ingest/enhance.html (accessed 2019-08-12)


Reviewer Entry
Reviewer 1
Comments:
Accept

Reviewer 2
Comments:
Accept

Outsource Partners. If applicable, please list them.

Reviewer Entry
Reviewer 1
Comments:

Reviewer 2
Other Relevant Information.

ICPSR has participated in several repository audits and certifications. In 2006, it was a test case for Trusted Repositories Audit & Certification (TRAC, now ISO 16363). ICPSR was one of the first six data repositories to earn the Data Seal of Approval in 2011. ICPSR earned the World Data System certification in 2013. ICPSR has also achieved United States Federal Information Security Management Act (FISMA) Moderate compliance and received Authority-To-Operate (ATO) from multiple federal agencies under FISMA [1]. The Consortium of European Social Science Data Archives (CESSDA) has chosen ICPSR as an “industry best” or “gold standard” organization by which to benchmark the levels and qualities of service that their national data archives provide to the research community. [2] [3]

ICPSR partners with several United States federal statistical agencies and foundations to curate, preserve, and provide access to their data resources. These thematic collections [4] generally follow ICPSR’s standard workflow, although for purposes of this certification, ICPSR is certifying just the membership-funded General Archive [5]. Divisions within ICPSR that conduct limited or no curation of data, such as openICPSR [6], a self-archiving public access research data-sharing service, are also not covered under this application.

References:

https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/preservation/trust.html (accessed 2018-11-08)


[4] ICPSR Web site, Thematic Data Collections:
https://www.icpsr.umich.edu/icpsrweb/content/about/thematic-collections.html (accessed 2018-10-04)


Reviewer Entry
Reviewer 1
Comments:
Accept
ORGANIZATIONAL INFRASTRUCTURE

I. Mission/Scope

R1. The repository has an explicit mission to provide access to and preserve data in its domain.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Response:

ICPSR has a mission statement describing its commitment to digital archiving: “ICPSR advances and expands social and behavioral research, acting as a global leader in data stewardship and providing rich data resources and responsive educational opportunities for present and future generations.” [1]

The organization integrates the mission into its policies (e.g., ICPSR Digital Preservation Policy Framework [2], procedures (e.g., ICPSR deposit guidelines and form [3] [4], and practices (e.g., data management and curation activities [5]). A Strategic Plan provides further guidance in mission areas [6].

As an organization acting for its member institutions, funding bodies, and depositors, ICPSR has accepted responsibility for providing access to and preserving its digital assets. Within ICPSR, the Director, the Associate Director, the Metadata and Preservation unit, the Information and Technology unit, the Curation unit, and the project directors and project managers all contribute to the lifecycle of digital content at ICPSR. [7]

References:
II. Licenses

R2. The repository maintains all applicable licenses covering data access and use and monitors compliance.

Compliance Level:

4 – The guideline has been fully implemented in the repository
Reviewer Entry
Reviewer 1
Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2
Comments:
4 – The guideline has been fully implemented in the repository

Response:

The applicable licenses are found in the ICPSR membership contract [1], the Terms of Use [2], and the Restricted-use Data Agreements [3]. The Terms of Use require the data consumer to comply with access regulations imposed both by law and by the data repository, and to conform to codes of conduct that are generally accepted in higher education and scientific research for the exchange of knowledge and information. This includes agreeing to use datasets solely for research or statistical purposes and not for investigation of specific research subject, except when identification is authorized in writing by ICPSR. It also includes agreeing to make no use of the identity of any research subject discovered inadvertently, and to advise ICPSR of any such discovery. [2]

ICPSR informs and guides data consumers about relevant data access and use issues and determines what are considered breaches of the license terms.

References:
[1] ICPSR Membership Application, FY2019-2010:

[2] ICPSR Web site, What are ICPSR's terms of use?:
https://www.icpsr.umich.edu/icpsrweb/content/shared/ICPSR/faqs/terms-of-use.html (accessed 2018-11-06)

[3] ICPSR Web site, Restricted-Use Data Management at ICPSR:
http://www.icpsr.umich.edu/icpsrweb/content/ICPSR/access/restricted/index.html (accessed 2018-11-06)

Reviewer Entry
Reviewer 1
Comments:
Accept.

Reviewer 2
Comments:
Accept

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

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1
Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2
Comments:
4 – The guideline has been fully implemented in the repository

Response:

Upon deposit, the depositor gives ICPSR permission “To store, translate, copy or re-format the Data Collection in any way to ensure its future preservation and accessibility” [1]. ICPSR commits to the long-term preservation of that content [2].

ICPSR has a long history (since 1962) of managing, curating, and preserving data [3]. ICPSR’s diversified funding model offers stability and reliability [4].

ICPSR commits to finding a successor to manage the data holdings in the event that ICPSR can no longer manage them. The Articles of Collaboration [5] of the Data Preservation Alliance for the Social Sciences (Data-PASS) [6], of which ICPSR is a partner, specify “Any partner that can no longer retain their material will provide the Partnership with ninety (90) days written notice to provide for a reasonable transition. A determination regarding the most suitable partner to receive the material will be based on the subject matter and content of the materials as well as any economic encumbrances that may be involved in the transfer.” Consideration will also be given to the potential successor’s storage, preservation, and access services, especially as they align with ICPSR’s capabilities. Therefore, it is possible that the Data-PASS partners would share responsibility for any of the partners’ holdings should it be necessary.

References:


[3] ICPSR Web site, ICPSR: The Founding and Early Years:
IV. Confidentiality/Ethics

**R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.**

**Compliance Level:**

4 – The guideline has been fully implemented in the repository

**Reviewer Entry**

**Reviewer 1**
Comments: 4 – The guideline has been fully implemented in the repository

**Reviewer 2**
Comments: 4 – The guideline has been fully implemented in the repository

**Response:**

ICPSR accepts data with identifying information under conditions consistent with the informed consent of the study participants and the relevant Institutional Review Board (IRB) approval. ICPSR staff work with data depositors to address
Disclosure risks. Once data are deposited with ICPSR, staff employ stringent procedures to protect the confidentiality of individuals and organizations whose personal information may be part of the archived data collection. Steps ICPSR staff take to maintain data confidentiality include: (1) Data are deposited and processed in a secure non-networked environment. (2) Completing a detailed review of all datasets to assess disclosure risk, (3) If necessary, modifying data to reduce disclosure risk, (4) Limiting access to datasets for which modifying the data would substantially limit their utility or the risk of disclosure remains high, (5) Training staff and consulting with data producers in methods of disclosure risk assessment and mitigation. [1] [2]

For data that cannot be modified to protect confidentiality without significantly compromising the research potential of the data, ICPSR provides access for researchers and students to restricted-use versions of the data that retain confidential or sensitive data. For these datasets, ICPSR imposes strict legal requirements and employs a graduated system of data access to ensure the protection of human subjects, with the stringency of the access rules reflecting the degree of risk. Generally, the method of dissemination is related to the level of disclosure risk or the sensitivity of the data. Each dissemination method has a different application process and requirements. [3]

ICPSR has established several mechanisms through which restricted data can be distributed: (1) Secure Download: Upon approval, researchers will receive an encrypted file via e-mail which they may download to the secure location specified in the application. (2) Virtual Data Enclave (VDE): Rather than receiving a downloadable copy of the data, researchers are approved to access the data, stored on ICPSR's servers, via a remote desktop connection. Researchers are unable to move files from the remote desktop to their own desktop or to the Internet. To receive output from the VDE, researchers must request that ICPSR conduct a disclosure review on the desired files. (3) Physical Data Enclave: The physical data enclave is located onsite at ICPSR's offices in Ann Arbor, MI. Data in the physical enclave typically contain highly sensitive personal information or data from protected populations. To analyze data located in the Physical Enclave, researchers must travel to ICPSR. To receive output, researchers must request that ICPSR conduct a disclosure review on the desired files. (4) Restricted Survey Documentation and Analysis (SDA): Researchers apply to analyze the data online via the Survey Documentation and Analysis (SDA) system. The online analysis system has built-in disclosure protection parameters for restricted data. [4]

References:
V. Organizational infrastructure

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.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Response:

With more than 55 years of service to the social sciences, ICPSR is the largest archive of digital social and behavioral science data in the world. ICPSR is a unit within the Institute for Social Research at the University of Michigan and maintains its office in Ann Arbor. [1]

ICPSR’s diversified funding model offers stability and reliability. The three primary sources of revenue include grants and contracts, membership dues, and tuition [2]. ICPSR provides data archiving and dissemination services for more than 20 government agencies and foundations, including the Bureau of Justice Statistics, the National Science Foundation, the National Institutes of Health, the Alfred P. Sloan Foundation, the Laura and John Arnold Foundation, the Bill & Melinda Gates Foundation, and the Robert Wood Johnson Foundation [3]. Some of these partnerships have been in place for decades. Membership dues from ICPSR’s over 780 member institutions [4] and tuition from the Summer Program in Quantitative Methods [5] make up other revenue streams.
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. [7] 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. [8]

ICPSR operates in accord with three organizational documents: a Constitution [9], Bylaws [10], and a Memorandum of Agreement with the University of Michigan and the Institute for Social Research [11]. The organization also maintains several policies that inform and guide its work as an archive, including an overarching Strategic Plan [12] that lays out the organization’s priorities for coming years. Other policies cover areas such as digital preservation [13], data access [14], collection development [15], and disaster planning [16].

References:


[7] ICPSR Web site, Staff: https://www.icpsr.umich.edu/icpsrweb/content/about/contact-people.html (accessed 2018-11-06)
VI. Expert guidance
R6. The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either inhouse or external, including scientific guidance, if relevant).

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1
Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2
Comments:
4 – The guideline has been fully implemented in the repository

Response:

ICPSR secures ongoing and regular guidance and feedback, including from staff and its Designated Community.

Senior staff [1] are subject experts in their respective disciplines and research in a variety of areas related to digital archiving, including human subjects protection and disclosure risk, digital preservation, and effective practices across the research data life cycle. ICPSR Council members are experts in their fields and represent key substantive areas of social and behavioral science research [2]. The Council’s work is supplemented by advisory committees of leading scholars who help ICPSR address needs in special fields. Advisory committees have been formed to guide the Consortium in census data, American politics, comparative politics, gerontology, economics, and many other fields. Advisory committees have helped establish policies in data processing priorities, Summer Program curriculum development, and technology planning. [3]

ICPSR regularly communicates with its Designated Community for feedback. We frequently communicate with our over 780 consortial official representatives (ORs) [4], who work directly with data users in libraries or department settings. We organize biennial meetings for these same ORs [5]. We conduct formal surveys of all ORs, also biennially. We have regular contact with our sponsoring agencies [6], who fund research and help us prioritize archiving and dissemination activities. We attend over 30 professional conferences each year, where we meet with and respond to our community, including students and researchers. We also conduct outreach through webinars [7] and social media [8], as well as pop-up surveys on our web site.

As a center in the Institute for Social Research, ICPSR undergoes periodic, mandated self-assessments and organizational reviews to ensure efficient and appropriate functioning, including five-year situational reviews that assess the organization and its director. Criteria for these assessments include financial stability and sustainability, contributions
to social science research, responsiveness to the mission, depth and breadth of the holdings and their relevance, educational strength, overall efficiency and organizational capacity, etc. Comprehensive interviews with ICPSR stakeholders are part of the situational reviews. [9]

References:
[1] ICPSR Web site, Staff: https://www.icpsr.umich.edu/icpsrweb/content/about/contact-people.html (accessed 2018-10-04)


Reviewer Entry
Reviewer 1
Comments: Accept

Reviewer 2
Comments: Accept

DIGITAL OBJECT MANAGEMENT
VII. Data integrity and authenticity

R7. The repository guarantees the integrity and authenticity of the data.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1
Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2
Comments:
4 – The guideline has been fully implemented in the repository

Response:

ICPSR runs checksums at every step of the life cycle from original deposit through data processing to creation of the Archival Information Package for long-term preservation. ICPSR uses a regimen of checksum utilities to routinely confirm the integrity of deposited content that ensures the completeness of the collection [1]. Storage is managed according to strict criteria regarding media, redundancy, etc. The “data pipeline” at ICPSR parallels the OAIS Reference Model and the research data life cycle model [2].

Staff who curate and manage data have a set of instructions that they adhere to and they document data transformations [3]. All changes to data values during the curation workflow are done using syntax and processing history files containing commands used to make adjustments to the data; for example, adjustments found in a syntax file may include adding variable labels, changing value labels, revising missing value designations, and/or inserting recodes. Syntax files allow tracking and replication of changes to data.

Once data are distributed to the public, versions numbers are incremented whenever a file is replaced (version number of the resource is part of its citation). Version history is included as part of the study metadata. ICPSR provides Web access to the most recent version(s) of files. For older versions, users may need to contact ICPSR directly. [4]

References:

[1] ICPSR Digital Preservation Policy Framework:
https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/preservation/policies/dpp-framework.html (accessed 2018-11-06)

VIII. Appraisal

R8. The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for data users.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Response:

A Collection Development Policy informs the selection of data for archiving. The policy sets forth a description of the characteristics of data that ICPSR has interest in adding to the collection. ICPSR intentionally casts a broad net in order to add a wide range of data that would be of interest to the diverse fields representing the social and behavioral sciences.
However, at the same time the organization applies additional appraisal criteria to determine the appropriate level of curatorial investment that ICPSR will make to ensure long-term and effective use of the data. [1]

ICPSR publishes a list of preferred formats and encourages the community to comply with them; these include SAS, SPSS, and Stata files, as well as ASCII files accompanied with setup files [2] [3]. When non-recommended formats are deposited, ICPSR attempts to normalize them to preservable formats [4]. In terms of preservation of files submitted in non-recommended formats unable to be normalized into preservable formats, ICPSR commits to bit-level only preservation.

Once data are submitted in a submission information package, the data pass through a "pipeline" for quality control, processing, and enhancement [5]. Curation staff follow established and documented internal procedures to review and curate data. Staff communicate and work directly with depositors, as needed, to curate data.

References:


Reviewer Entry
Reviewer 1
Comments: Accept

Reviewer 2
Comments: Accept

IX. Documented storage procedures
R9. The repository applies documented processes and procedures in managing archival storage of the data.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1
Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2
Comments:
4 – The guideline has been fully implemented in the repository

Response:

ICPSR pursues several strategies for managing archival storage, including: (1) Redundancy in several off-site locations, including on spinning disk and removable media [1]. (2) “Warm backups” of the Web server in two locations enabling the Web site to “fail over” to another server in the event that the ICPSR server is not available [2]. (3) Checksums are run on a routine schedule to ensure the integrity of the data in the repository [3]. (4) Storage media are monitored for deterioration and new storage media are evaluated on an ongoing basis as part of a Technology Watch program.

ICPSR has implemented policies and procedures dealing with data recovery in the event of disaster [4] [5]. These documents also deal with the human side, ensuring that systems for communication are in place should a disaster occur [6]. ICPSR has defined a specific disaster plan for archival storage and tested the restoration of content from backup copies. Periodic drills are conducted to make sure that staff are familiar with how to respond in the event of emergencies or disasters [7].

For older archival content, outdated file formats have been updated to current formats to avoid the risk of loss through obsolescence, and stored formats are routinely monitored to identify potential preservation issues [8] [9].

ICPSR has considered appropriate security levels for each phase of life cycle management and as confidential data are increasingly acquired, it is adjusting existing procedures accordingly. Currently, ICPSR has a system with a layered design to support a private non-networked area, a semi-private area, and a public area [10]. When new technologies for managing security become available, ICPSR evaluates each for possible integration into the current security approach.

References:
[1] ICPSR Web site, Archival Storage:
X. Preservation plan
R10. The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.

**Compliance Level:**

4 – The guideline has been fully implemented in the repository

**Reviewer Entry**

**Reviewer 1**

Comments: 4 – The guideline has been fully implemented in the repository

**Reviewer 2**

Comments: 4 – The guideline has been fully implemented in the repository

**Response:**

The ICPSR Digital Preservation Policy Framework is the highest level digital preservation policy document at ICPSR. It makes explicit ICPSR's commitment to preserving the digital assets in its collections through the development and evolution of a comprehensive digital preservation program. The Digital Preservation Policy Framework documents that ICPSR has adopted normalization and migration as its primary preservation strategies and refers to our disaster planning program [1]. The possible obsolescence and potential vulnerabilities of current file formats are monitored as one component of an ongoing technology responsiveness program [2].

Upon deposit, the depositor gives ICPSR permission “To store, translate, copy or re-format the Data Collection in any way to ensure its future preservation and accessibility” [3]. ICPSR commits to the long-term preservation of that content.

**References:**

[1] ICPSR Digital Preservation Policy Framework:
https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/preservation/policies/dpp-framework.html (accessed 2018-10-04)

[2] Data-PASS Web site, Data Security Standards: Integrity and Availability:

XI. Data quality

R11. The repository has appropriate expertise to address technical data and metadata quality and ensures that sufficient information is available for end users to make quality-related evaluations.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1
Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2
Comments:
4 – The guideline has been fully implemented in the repository

Response:

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 [1]. All staff are required to complete ongoing training related to data security and disclosure risk [2].

Staff who manage data have a set of instructions that they adhere to when reviewing and curating data and metadata [3]. Staff have decades of experience resolving technical data and metadata quality issues and are active in international data curation and preservation efforts. For instance, ICPSR helped to establish and now plays a leadership role in the Data Documentation Initiative, which is a structured metadata standard for the social and behavioral sciences [4]. DDI has been widely adopted by the international social science data archiving community, and its use is spreading among national statistical agencies and other data providers.

ICPSR provides an extensive Bibliography of Data-Related Literature that links over 75,000 research publications to the ICPSR data on which they are based. The Bibliography is one of ICPSR’s unique resources and helps users, PIs, and
funders understand in a deeper way than page views or downloads how data are re-used. [5]

References:
[1] ICPSR Web site, Staff: https://www.icpsr.umich.edu/icpsrweb/content/about/contact-people.html (accessed 2018-11-08)


Reviewer Entry
Reviewer 1
Comments:
Accept

Reviewer 2
Comments:
Accept

XII. Workflows

R12. Archiving takes place according to defined workflows from ingest to dissemination.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry
Reviewer 1
Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2
Comments:
ICPSR’s workflow from acquisition to dissemination is based on the OAIS Reference Model [1]. Data producers deposit files (SIPs) with ICPSR through an online form [2]. Deposited files undergo a selection and appraisal process [3]. Files are then cleaned, normalized, described, and assembled (as AIPs) for archival storage [4]. Metadata about the files are captured throughout ICPSR’s curation activities [5]. Data are available (as DIPs) through the ICPSR web site, either by downloading files or through online analysis in a Web browser [6].

References:
[1] ICPSR Web site, OAIS-Based Processes:
https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/lifecycle/oais.html (accessed 2018-11-06)


[3] ICPSR Collection Development Policy:
https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/policies/colldev.html (accessed 2018-11-06)

[4] ICPSR Web site, Data Enhancement:
https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/lifecycle/ingest/enhance.html (accessed 2018-11-06)


[6] ICPSR Web site, Access and Dissemination:
https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/lifecycle/access.html (accessed 2018-11-06)

Reviewer Entry
Reviewer 1
Comments:
Accept

Reviewer 2
Comments:
Accept

XIII. Data discovery and identification

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

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

ICPSR ensures that all of its data are well-described, discoverable, and uniquely identified. ICPSR provides Digital Object Identifiers (DOIs) at the study level as well as downloadable data citations for use in publications based on analyses of the data [1] [2]. Guidance on data citation practice is provided in multiple locations on the ICPSR Web site [3].

ICPSR offers a robust search of its holdings based on Solr/Lucene so that users can find data efficiently [4] [5].

In documenting its data holdings, ICPSR uses the Data Documentation Initiative metadata specification, a standard for the content, presentation, transport, and preservation of documentation. Study-level records may be exported from study home pages as XML compliant with these standards: DDI Codebook, DDI Lifecycle, Dublin Core, and DAta Tag Suite (DATS) [6]. Batches of ICPSR study-level metadata are available via an OAI-PMH service. ICPSR metadata records are licensed under a Creative Commons Attribution-Noncommercial 3.0 United States License. [7]

ICPSR is indexed in several generic registries of resources, including re3data.org [8] and FAIRsharing [9].

References:


XIV. Data reuse

R14. The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data.

**Compliance Level:**

4 – The guideline has been fully implemented in the repository

**Reviewer Entry**

**Reviewer 1**

Comments:
Accept

**Reviewer 2**

Comments:
Accept
Reviewer 2
Comments:
4 – The guideline has been fully implemented in the repository

Response:

ICPSR has ensured long-term reuse of its data for over 55 years. ICPSR creates study-level metadata records in DDI format using information supplied by data depositors and other sources [1].

ICPSR archives the original digital content received, the normalized versions of processed data, and superseded versions of data that have been distributed. The archived files will enable ICPSR to retain the ability to regenerate distribution formats over time. ICPSR has adopted normalization and migration as its primary digital preservation strategies. [2]

Data are distributed in formats that the social science research community typically uses: SPSS, SAS, Stata, R, tab-delimited (for use in spreadsheets like Excel), and online analysis [3]. Data use tutorials describing how to use these formats are available [4].

References:


Reviewer Entry

Reviewer 1
Comments:
Accept

Reviewer 2
Comments:
Accept

TECHNOLOGY
XV. Technical infrastructure

R15. The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services it provides to its Designated Community.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1
Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2
Comments:
4 – The guideline has been fully implemented in the repository

Response:

To manage the repository and its data resources, ICPSR has designed and implemented a custom end-to-end data management platform called Archonnex. Based on an Enterprise Software Architecture (ESA) approach and open source technologies, Archonnex consists of a core infrastructure of system, network, and architectural software components that are connected and assembled to build applications to meet the complex needs of the ICPSR mission. This component approach improves scalability both horizontally and vertically. The Archonnex platform, which is implemented using a Java technology stack, follows the Open Archival Information System (OAIS) reference model [https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/lifecycle/oais.html (accessed 2018-11-08)] and leverages open source products such as Fedora repository software from DuraSpace, Apache Foundation’s Tomcat, Jena, and Solr. It is designed for extensibility to integrate with other systems that support a Services Oriented Architecture (SOA) approach. Archonnex complies with the U.S. Government’s Section 508 Standards for Web-based Intranet and Internet Information and Applications, and with the Web Content Accessibility Guidelines 2.0, Level AA, to serve a diverse population of users. [1]

“Warm backups” of the Web server are in two locations and enable the Web site to “fail over” to another server in the event that the ICPSR server is not available [2]. The local network is connected to the University of Michigan (UM) backbone which has redundant connections to the Merit nationwide network providing reliable and fast access to the Internet. The UM monitors ICPSR’s production computing systems 24 hours a day, seven days a week, notifying an on-call systems engineer in the event of system degradation. The network has experienced a better than 99-percent
uptime over the last decade.

References:


Reviewer Entry
Reviewer 1
Comments:
Accept

Reviewer 2
Comments:
Accept

XVI. Security

R16. The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry
Reviewer 1
Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2
Comments:
4 – The guideline has been fully implemented in the repository

Response:

ICPSR has implemented policies and procedures dealing with data recovery in the event of disaster [1] [2]. These documents also deal with the human side, ensuring that systems for communication are in place should a disaster occur [3]. ICPSR has defined a specific disaster plan for archival storage and tested the restoration of content from backup copies. Periodic drills are conducted to make sure that staff are familiar with how to respond in the event of emergencies.
“Warm backups” of the Web server are in two locations and enable the Web site to “fail over” to another server in the event that the ICPSR server is not available [5]. The local network is connected to the University of Michigan (UM) backbone which has redundant connections to the Merit nationwide network providing reliable and fast access to the Internet. The UM monitors ICPSR’s production computing systems 24 hours a day, seven days a week, notifying an on-call systems engineer in the event of system degradation. The network has experienced a better than 99-percent uptime over the last decade.

ICPSR has considered appropriate security levels for each phase of life cycle management and as confidential data are increasingly acquired, it is adjusting existing procedures accordingly. Currently, ICPSR has a system with a layered design to support a private non-networked area, a semi-private area, and a public area [6]. When new technologies for managing security become available, ICPSR evaluates each for possible integration into the current security approach.

The ICPSR network is protected from intrusion in a variety of ways, including network-based firewalls that block unneeded ports and protocols, a Windows Host firewall with custom ICPSR settings, and current antivirus software and malware protection that run on all client computers, which are protected by boundary firewalls. Additional firewalls protect specific areas of the network based on project and contractual commitments. Intrusion detection systems are employed to monitor network traffic moving in and out of the main UM network, and remote logging is used to monitor systems and network traffic for unusual patterns. ICPSR pursues ongoing risk management efforts in accord with established policies. The organization has achieved FISMA Moderate compliance and received Authority-To-Operate (ATO) from multiple federal agencies under FISMA.

References:
APPLICANT FEEDBACK

Comments/feedback

These requirements are not seen as final, and we value your input to improve the core certification procedure. To this end, please leave any comments you wish to make on both the quality of the Catalogue and its relevance to your organization, as well as any other related thoughts.

Response:

Thanks for the opportunity to go through the new CoreTrustSeal certification process. One suggestion about the process is to make it possible to print off the application (or view as PDF) while an applicant is completing the process. Currently, individual responses are captured in the text boxes, which aren't easily viewed or printed, and which makes it difficult to do a comprehensive review and edit of the submission. Another suggestion is to make it easier to capture references within the documents, such as footnotes or endnotes. A WYSIWYG interface also would be nice.