Assessment Information

[CoreTrustSeal Requirements 2020–2022](#)

Repository: Data Repository for the University of Minnesota
Website: https://conservancy.umn.edu/drum
Certification Date: 29 June 2021

This repository is owned by: **University of Minnesota**

Data Repository for the University of Minnesota

Notes Before Completing the Application

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

True

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:

Institutional repository

Brief Description of Repository

The Data Repository for University of Minnesota (DRUM) is a well-curated subset of a larger general institutional repository called the University of Minnesota Digital Conservancy. DRUM, and associated data curation service, launched November 2014. DRUM only accepts deposits from University of Minnesota affiliates but all data housed in DRUM are publicly available. DRUM is listed in the Registry of Data Repositories re3Data.org under ID:r3d100011393.

Brief Description of the Repository's Designated Community.

As a public land-grant institution and flagship research institution for the state, the Data Repository for the University of Minnesota connects university research with the citizens of the state of Minnesota and the world, in fulfillment of the university's tripartite research, teaching and pubic engagement mission.

Level of Curation Performed. Select all relevant types from:

D. Data-level curation – as in C above; but with additional editing of deposited data for accuracy

Comments

Our team of data curators collaborate with data providers to ensure that the data are as accurate and reusable as possible.

Insource/Outsource Partners. If applicable, please list them.

University of Minnesota Office of Information Technology

Data Curation Network, administered by the University of Minnesota

Summary of Significant Changes Since Last Application (if applicable).

Newly worded answers to all questions. However the technology, infrastructure and other key aspects have mostly stayed the same.

Other Relevant Information.

The Data Repository for the University of Minnesota is used world-wide and the 500+ datasets have garnered 390,654 downloads since launch.

ORGANIZATIONAL INFRASTRUCTURE

1. 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:

The Data Repository for the University of Minnesota (DRUM) provides University of Minnesota affiliates a place to share, publish, and preserve their digital data for long-term access and future use (1). As the institutional data repository, DRUM's mission is to fulfill the University's mission.

The University of Minnesota is the 6th largest public university in the United States and the flagship research university for the state of Minnesota. The University of Minnesota was founded in 1851 and supports over 60,000 students across 5 campuses, geographically distributed throughout the state.

The official University-wide policy "Research Data Management: Archiving, Ownership, Retention, Security, Storage, and Transfer" (2) defines ownership and stewardship responsibilities for "research data that are generated or acquired by faculty, staff, and students through the use of University facilities and resources." Per the policy, research data are owned by the University's Board of Regents (the University's top-level governing body), but stewardship responsibilities fall solely on the principal investigator, or individual leading a research project. According to the policy, the University Librarian is responsible for "ensure[ing] accessibility and preservation of research data through curation, metadata, repositories, and other access and retrieval mechanisms to meet federal, state, sponsor, and University requirements."

By ingesting and preserving research data created at the University, and disseminating and making it accessible to the citizens of Minnesota and the world, DRUM directly supports the University's tri-part mission (3) to "preserve knowledge, understanding, and creativity," to "share that knowledge, understanding, and creativity" and to "extend, apply, and exchange knowledge between the University and society...by making the knowledge and resources created and preserved at the University accessible to the citizens of the state, the nation, and the world."

The University Libraries launched DRUM in 2015 to enable University affiliates to exercise their stewardship role and openly share their data in a fully-curated and openly accessible digital repository. Structurally, DRUM is housed within the larger institutional repository of the University of Minnesota, the University Digital Conservancy (UDC), but supported with additional curation workflows, staff, and policies.

1. University Digital Conservancy. "About the Data Repository": https://conservancy.umn.edu/pages/drum/ (accessed 2020-08-10)
2. University of Minnesota. "Research Data Management: Archiving, Ownership, Retention, Security, Storage, and Transfer." https://policy.umn.edu/research/researchdata (accessed 2020-0810)
3. University of Minnesota Board of Regent Policy: Mission Statement:
https://regents.umn.edu/sites/regents.umn.edu/files/2019-09/policy_mission_statement.pdf (accessed 2020-08-10)

Reviewer Entry
Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

2. 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
Accept

Response:

Access and use for data housed in DRUM are specified by a Terms of Use policy (1), an End User Access Policy (2), and when applicable, any Creative Commons or other open licenses applied to datasets by data producers. Users may select a creative commons license in the DRUM upload workflow (feature built into DSpace (3)), or the depositor can work with DRUM curators to apply another license of their choice, such as MIT or GNU. These are enforced by a University-wide Acceptable Use of Information Technology Resources Policy (4). In addition, the library maintains up-to-date awareness and education of applicable regulations to inform data users about the ethical and legal use of digital data and supporting licenses (5, 6). This section will go into more detail on the three areas of evidence required for this section.

License agreements in use
Each record in DRUM includes a link to the Terms of Use (1) with the following language: "By using these files, users agree to the Terms of Use Policy. Content distributed via the University of Minnesota's Digital Conservancy may be subject to additional license and use restrictions applied by the depositor."

DRUM's Terms of Use is as follows: "Authors who share their data here expect that it will be re-used to some degree. Users are expected to abide by the University of Minnesota Acceptable Use Policy, and other University policies, where applicable. However, by using or downloading the data, you signify your agreement to the conditions of use stated below:

- The user will not make any use of data to identify or otherwise infringe the privacy or confidentiality rights of individuals discovered inadvertently or intentionally in the data.
- The user will give appropriate attribution to the depositor(s) of the data in any publication that employs resources provided by the Data Repository.
- If your use or publication requires permission, you must contact the depositors directly; administrators of the Data Repository cannot respond to requests for permission."

Conditions of use

DRUM's End User Access Policy (4) stipulates the two mechanisms by which users may access the data: open and by request. The policy is as follows: "The Data Repository is an open access repository and makes collection holdings freely available, worldwide. Data authors may choose to make their data available in two ways:
- Open Data (default): These data are available for immediate download. Users may contact the author with questions regarding the data. Authors may choose to apply a Creative Commons license to their data, which will give users certainty that they do not need permission for any uses allowed by the license. However, even without a Creative Commons license, users will be able to download and use data - subject to the DRUM Terms of Use.
- By Request: In some cases, an author may choose to control access to their work, for up to 2 years, by which end-user access is moderated through the authors' permission (via email). If this is the case, the restriction and request form are clearly indicated in the record of the data. If you have trouble requesting access to data or do not get a response via email, please contact the Data Repository staff."

Noncompliance measures

The University of Minnesota's Acceptable Use Policy (4) states: This University-wide policy stipulates that the University of Minnesota has assigned responsibilities to various entities and personnels to monitor and investigate improper use of digital data. Individuals who are found to have violated the policies may be subject to limitation or termination of user privileges, and other disciplinary or legal actions (4).

1. University Digital Conservancy. "DRUM Terms of Use."
https://conservancy.umn.edu/pages/drum/policies/#terms-of-use (accessed 2020-08-10)
2. DRUM End-User Access Policy: https://conservancy.umn.edu/pages/drum/policies/#end-user-access-policy (accessed 2020-09-10)
3. DSpace, "Creative Commons Support for DSpace Items" https://wiki.lyrasis.org/display/DSDOC5x/Functional+Overview#FunctionalOverview-CreativeCommonsSupportforDSpaceItems (accessed 2021-04-13)
4. University of Minnesota Acceptable Use of Information Technology Resources Policy:
https://policy.umn.edu/it/itresources (accessed 2020-08-10)
5. University Libraries. "Copyright Services." https://www.lib.umn.edu/copyright (accessed 2020-08-10)
6. University Libraries. "Copyright and data licenses." https://www.lib.umn.edu/datamanagement/copyright (accessed 2020-08-10)

3. 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
Accept

Response:

Data deposited into DRUM is considered an institutional asset that is the University Libraries' responsibility to preserve for the long-term. Data producers sign a deposit agreement (1) upon ingest granting the rights, not to the repository, but to the University of Minnesota's Board of Regents. It is through this formal, high-level University entity that long-term stewardship takes place. Therefore, succession planning would only be required in the result of the University of Minnesota's dissolution. As the state's only land-grant institution of higher education (established since 1851) and as the only major public research university in the state of Minnesota, the University of Minnesota is unlikely to cease operation or substantially change its scope or mission. As a libraries-based repository, the succession plan for DRUM content is similar to all content managed and held by the library system. In the event of disaster, the disaster recovery procedure, outlined in section 7.2 of the University of Minnesota Libraries Digital Preservation Framework (2), follows the disaster recovery plans of University data centers (3) and related policy on recovery (4), all subject to regular review.

Specifically, DRUM is supported by the University's digital preservation unit, which supports the long-term usability of data files held in DRUM in several ways. Regular data monitoring and secure storage maintains the integrity of data. A

migration plan and provenance records make sure that data are fit for reuse over time (5). Currently, DRUM commits to preserving data for a minimum of ten years (6). The University Digital Conservancy (UDC) is committed to "provid[ing] free, public access and long-term preservation to work created at the [University of Minnesota]." [6]

1. University of Minnesota Libraries. "DRUM Deposit Agreement"

http://conservancy.umn.edu/pages/drum/policies/#deposit-license (accessed 2020-08-19)

2. University of Minnesota Libraries. "Digital Preservation Framework"

https://www.lib.umn.edu/dp/digital-preservation-framework (accessed 2021-04-13)

3. University of Minnesota. "Data Center Security Standard and Process"

https://it.umn.edu/resources-it-staff-partners/information-security-standards-guidelines/data-center-security (accessed 2021-04-13)

4. University of Minnesota. "Data Storage and Backup & Recovery Standard" https://policy.umn.edu/it/securedata-appdsbr (accessed 2021-04-13)

5. University of Minnesota Libraries. "UDC Preservation Policy" http://conservancy.umn.edu/pages/policies/#preservation (accessed 2020-08-19)

6. University of Minnesota Libraries. "DRUM Policies and Terms of Use"

https://conservancy.umn.edu/pages/drum/policies/ (accessed 2020-08-19)

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

4. 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
Accept

Response:

DRUM manages disclosure risk and adheres to disciplinary norms and ethics in a number of ways via policy, education, curation procedure, and staff collaboration with experts.

Policy: DRUM staff communicate to potential depositors via DRUM's data collection policy (1) that "Data must be non-restricted data that DO NOT contain any private, confidential, or other legally protected information (e.g., personal identifiable information)." To confirm their adherence to policy, all depositors, as a part of the deposit agreement (2), must agree at deposit that "The Content contains no restricted, private, confidential, or otherwise protected data or information that should not be publicly shared."

Education: Users are guided to the University Security Classifications of Data for appropriate categorization of the inherent risks in the data they create. (3) DRUM staff also help users manage and de-identify sensitive data via visual inspection for indirect identifiers and with guidance on the University Library's "Sensitive Data" web page (4) which links to University Medical School's procedure for de-identifying health data (5) regulated by United States federal laws (e.g., HIPAA Privacy Rule).

Curation procedure: All datasets deposited to DRUM undergo curatorial review to ensure that accepted datasets are in compliance with DRUM policy, as well as disciplinary and ethical norms. Special procedures are applied in the following instances:
If the data contain any information derived from human subjects, DRUM curators request a copy of the participant consent form that the IRB approved to ensure that all de-identified data are shared in accordance with this approved plan. If not, curators work with the depositors to either seek further approval from the IRB or to re-consent their participants. In some cases staff must reject their data from DRUM and help them find an appropriate restricted-access repository. (6)
If the data contain any information derived from the invasive use of animals DRUM curators request a copy of the Institutional Animal Care and Use Committee (IACUC) approval demonstrating the benefit of the research and that pain and discomfort was minimized. (7)
If the data contain sensitive information that does not fall into the above categories, curators often request a consultation prior to acceptance and may work with University offices to ensure all ethical considerations are followed. Some examples include data with specific geographic locations of artifacts and dig sites and data derived from or of importance to local tribal nations.

Finally, DRUM staff are engaged with regular data ethics training and outreach activities on campus (both as participants and as instructors) and work closely with the Office of Research and the IRB education committee to ensure that data sharing plans involving confidential data are worded with long-term data archiving and sharing in mind. (see for example,

8)

1. University Digital Conservancy. "Data Collection Policy."

https://conservancy.umn.edu/pages/drum/policies/#data-collection-policy. (accessed 2020-08-10)

2. University Digital Conservancy. "Deposit License." https://conservancy.umn.edu/pages/drum/policies/#deposit-license (accessed 2020-08-10)

3. University Policy Library. "Data Security Classifications by Type." https://policy.umn.edu/it/dataclassification-appa. (accessed 2020-08-10)

4. University Libraries Research Data Services. "Sensitive data." http://www.lib.umn.edu/datamanagement/sensitive. (accessed 2020-08-10)

5. University of Minnesota Health Information Privacy & Compliance Office. "Creating a De-Identified Data Set." https://www.healthprivacy.umn.edu/policies-procedures/creating-de-identified-data-set (accessed 2020-08-10)

6. University Libraries Research Data Services. "IRB considerations." https://www.lib.umn.edu/datamanagement/irb. (accessed 2020-08-10)

7. University of Minnesota. "Institutional Animal Care and Use Committee" https://research.umn.edu/units/iacuc/submit-maintain-protocols/applying/criteria-review. (accessed 2020-08-10)

8. University of Minnesota. "Annual Research Ethics Day Conference — The Power and Perils of Research Data: Generating, Storing & Sharing Data Responsibly." https://consortium.umn.edu/conference/annual-research-ethics-day-conference-power-and-perils-research-data-generating-storing-sharing-data-responsibly (accessed 2020-08-10)

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

5. 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

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
Accept

Response:

DRUM is hosted by the University Libraries, a centrally-funded unit of the University of Minnesota. The University has an overall annual operating budget of $4.2 Billion (for fiscal year 2020) that is state-funded with additional stable revenue streams from student tuition, research grants, and other sources (1).

End-user operations are managed by a data curation team based in the Libraries. DRUM staffs between 4-6 curators that have a portion of their time dedicated to curation work for the data repository. To utilize their expertise and skills, curators are assigned datasets for curation based on the file formats and discipline of the data depositor, such as GIS data, social science data, scientific data, digital humanities data, and public health data. Each curator has working knowledge of the field and understands the common data types and software used in the field. All staff are supported by annual professional development funds. The DRUM staff includes:

DRUM director

DRUM coordinator and digital repository archivist

Scientific data curator, biological science

Scientific data curator, physical and computer sciences

Health sciences and human subjects data curator

Social sciences and human subjects data curator

Spatial/GIS data curator

Curators have the following job responsibilities:

Contribute to the development of data-related repository activities, workflows, and policies.

Execute technical processes involved in managing the lifecycle of digital datasets in order to maintain the integrity of and access to archived datasets; e.g., data transformation projects.

Serve as functional expert for consultations with University of Minnesota researchers on data issues and services (e.g., data management planning) and the development and delivery of training and instructional materials on data curation.

Process submissions for depositing and archiving datasets in the digital repository.

Develop strong connections with faculty and students to determine and address service needs.

Seek opportunities for collaborative partnerships with designated departments and research centers.

Provide research lifecycle support and leadership, including shared service development in areas such as data management planning, research networking, and personal information management.

Contribute to the knowledge base of the profession through research, publication, and professional engagement.

DRUM is supported by a number of staff for technological development and digital preservation who split their time across

several systems.

Digital preservation and repository technology director

Digital repository lead developer

Digital preservation analyst

Digital repository staff ensure that the underlying digital software is well-managed and up to date. Additionally DRUM's lead developer makes custom enhancements to the administrative and user interface to meet specific staff and user needs. Digital preservation and technology staff consult on policy and service needs as needed.

DRUM is advised and provided oversight from a number of University groups. These currently include a Research Data Services Team (responsible for campus data management training and services); DSpace Management Team (repository development/road mapping); University Digital Conservancy Sponsors (governance for DRUM); and other library staff such as metadata strategy.

1. University of Minnesota. "University Budget." http://finance.umn.edu/budget.html (accessed 2020-08-25)

6. 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

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

Response:

As an institutional repository for the University, DRUM is regularly presented to University committees for feedback and evaluation (eg., University Senate committees such as the Senate Research Committee and the Council of Research Associate Deans). These groups include faculty, researchers, post-docs, IT staff (from colleges across campus including science and engineering and the medical school), and administrators from various research offices such as the Office for the Vice President for Research, the Minnesota Supercomputing Institute, and the University of Minnesota Informatics Institute.

Additionally, the Director of DRUM is a member of the University Storage Council (1) and the Storage Champion Network (2). These groups provide guidance and oversight for a number of campus infrastructure facilities, and in particular, provide a platform for regular communication between research staff and researchers (3). For example, the Storage Council recently developed the Storage Selection Tool (4) which defines DRUM in the context of other data storage solutions on campus and ensures that the appropriate types of data are deposited to the correct location on campus.

A large multidisciplinary institution generates a wide range of file formats with significant disciplinary diversity. Noting this curation challenge as the University of Minnesota institutional data repository, DRUM sought expertise beyond one institution. In 2016, DRUM became the founding member and lead institution in the national Data Curation Network (DCN) (5). The DCN brings together the collective expertise and guidance from (currently) 12 partner data repositories that share nearly 50 expert staff who discuss common challenges, research and develop solutions, and (most significantly) curate each other's data in a shared staffing model. The DCN employs a full-time Coordinator who reviews incoming datasets and assigns them to Network curators based on file format and domain expertise. The curators in the DCN perform standardized checks and deliver recommendations for how the local institution may collaborate with the data depositors to enrich and improve the findability, accessibility, interoperability, and reuse (FAIRness) of the data prior to publication. See DCN's workflow details (6). DCN team members meet for weekly stand ups and participate in special interest groups around shared topics, teach workshops, and collaborate to research curation practices for common data types (e.g., R, Microsoft Access, Python), which often result in publications called Data Curation Primers thereby sharing expertise publicly for all curators (7).

1. University of Minnesota. University Storage Council. https://it.umn.edu/experts-communities/university-storage-council (accessed 2020-08-25).
2. University of Minnesota. Storage Champion Network.
https://it.umn.edu/experts-communities/storage-champion-network. (accessed 2020-08-25).
3. Wilgenbusch, J. C., Bates, C., Johnston, L., Baller, J. A., Bollig, E., & Neuhauser, C. (2019). A Comprehensive Campus-based Approach to Address the Opportunities and Challenges Posed by Data Intensive Research. In PEARC '19: Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (learning). July 2019 Article No.: 89 Pages. 1–3. https://doi.org/10.1145/3332186.3333162.

4. University of Minnesota. Storage Selection Tool.

https://it.umn.edu/services-technologies/comparisons/select-digital-storage-options . (accessed 2020-08-25).

5. Data Curation Network. "A shared staffing model for digital data repositories". https://datacurationnetwork.org/ (accessed 2020-08-25).

6. Data Curation Network. "The DCN Curation Workflow". https://datacurationnetwork.org/resources/workflows/ (accessed 2020-08-25).

7. Data Curation Network. "Data Curation Primers". https://datacurationnetwork.org/resources/data-curation-primers/. (accessed 2020-08-25).

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

DIGITAL OBJECT MANAGEMENT

7. 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
Accept

Response:

DRUM uses DSpace for its technology platform, which is based on the OAIS model for tracking the integrity and authenticity of all digital objects housed in the system. Measures are in place to capture the authenticity, provenance, and relations within and between datasets, including versions of the same data. Specifically:

Fixity: A MD5 checksum is issued for each file on ingest into DRUM. Checksums are routinely validated to ensure that the digital objects do not change overtime (1).

Provenance metadata: DRUM records all provenance metadata in the form of authenticated user submission credentials, contact information, timestamps, and a list of the original files submitted to DRUM. This ensures that staff can verify who submitted data to DRUM and when.

Depositor Authentication: The identity of depositors to DRUM must be authenticated according to DRUM policy that "Data must be authored by at least one University of Minnesota researcher with an active [University of Minnesota] Internet ID," and this check is done using the Shibboleth Internet protocol (2)

Documentation of completeness: All DRUM submissions receive curatorial review. As part of this review, data files and metadata that are incomplete will be augmented in collaboration with the data depositor. The DRUM collection policy states that "Data must include adequate documentation describing the nature of the data at an appropriate level for purposes of reuse and discovery. All data receive curatorial review and data that are incomplete or not ready for reuse may not be accepted into the repository" (2).

Curation log: DRUM curators keep record of the original submission metadata and create a working copy of the files. Curators document all the changes made during the curation process in a curation log, which is preserved as part of the metadata (but not available for public download). If additional changes beyond minor metadata updates are requested by the data depositor after curation, these changes are documented in a change log that is stored alongside the curation log.

Version control: DRUM welcomes new versions of a dataset, but versioning is closely mediated. The new version is screened for protected information and goes through a curatorial process just like a new submission of a dataset. Each version gets its own handle. The relationship between the versions is reflected in the URL of the handle and in the metadata fields of the data record. Data users can access all versions of the datasets in DRUM. The record of each version of the dataset includes the persistent identifiers to the other versions (3, for example).

File format validation: File extensions are automatically checked against the DSpace bitstream format registry (4). Then as part of the DRUM curation workflow, curators open and review the files included in the submission. If errors are encountered when trying to open files, curators work with the depositors to confirm the file formats.

1. University Digital Conservancy. "Policies and Guidelines, Preservation Policy"
https://conservancy.umn.edu/pages/policies/#preservation (accessed 2020-08-19)
2. University Digital Conservancy. "DRUM Policies and Terms of Use, Data Collection Policy"
http://conservancy.umn.edu/pages/drum/policies/#data-collection-policy (accessed 2020-08-19)
3. Example for Versioning: http://hdl.handle.net/11299/177631l (accessed 2020-09-08)
4. DSpace, "Metadata and Bitstream Format Registry,"
https://wiki.lyrasis.org/display/DSDOC5x/Metadata+and+Bitstream+Format+Registries (accessed 2021-04-13)

Reviewer Entry
Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

8. 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

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
Accept

Response:

DRUM has a comprehensive data collection policy (1) that is followed for all data submitted to DRUM. All data are appraised by a DRUM coordinator in a review process to ensure that data meet the local collection policy and do not violate the deposit agreement (2). DRUM's collection scope is exclusive to University of Minnesota affiliates and University of Minnesota related content. DRUM avoids deposits that fall outside of scope due to the required secure login process using Shibboleth authentication. The collection is relevant and useful by doing the following:

1. Quality Control of Data: It is DRUM policy that "Data must include adequate documentation describing the nature of the data at an appropriate level for purposes of reuse and discovery. All data receive curatorial review and data that are incomplete or not ready for reuse may not be accepted into the repository." (1) To ensure compliance with this policy, DRUM supports the data producer before (educational workshops, before-you-submit guides, consultations) and during (mediate all deposits and partner with the submitter) the submission. For example, DRUM collects required metadata during the initial submission process where data depositors are instructed to upload documentation files and/or incorporate information about methodology in the metadata fields "abstract" and "description" in the DRUM Deposit Form. When documentation delivered is not sufficient, curators use a readme.txt template (3) to send to the producer for completing. Additionally, curators ensure that all files collected adhere to legal and ethical requirements. In data

management education with potential depositors, DRUM staff recognize and recommend discipline specific repositories for those data that fall out of DRUM's scope.

2. Quality Control of Metadata:The metadata are based on Dublin Core Element Set and incorporate all the applicable Dublin Core elements. Automated assessment is in the form of required fields (author, title, date, data type, agreement to terms of deposit); DRUM metadata schema is published (4). The metadata helps potential data users find and understand the data. DRUM metadata are exposed to online search engines, such as Google and Google Scholar, and are brokered through external indexing systems, such as DataCite and the Web of Science Data Citation Index. This exposure helps with the discovery of the dataset and helps potential data users determine the relevancy of the dataset for their use. Additionally, persistent identifiers are added to each record's metadata in the form of a DOI (via DataCite.org) and handles (via DSpace). Finally, all metadata are available for (open) ingest by third parties (XML of via API).

3. File Formats: DRUM has published a preservation policy that links to a list of preferred formats for which the University of Minnesota provides full support in preservation and accessibility (5). The policy is based on the University of Minnesota Libraries Digital Preservation Framework (6). Additionally, the DRUM coordinator reviews file formats as they are deposited.If applicable, data files are converted into machine-readable and non-proprietary formats (preferably formats fully supported by the institutional repository preservation policy [5]) and the original files are retained. For example, all Microsoft Excel files are transformed to non-proprietary formats using the freely available "Excel Archival Tool" developed by a former DRUM staff member (7); however, both the Microsoft Excel files and non-proprietary formats are kept in the repository.

Rigorous attention to detail during ingestion appraisal and curation is one of the measures used to prevent published datasets from being removed. DRUM supports tombstones for any withdrawn items allowing the DOI to resolve even when items are removed [8].

1. University Digital Conservancy. "DRUM Policies and Terms of Use"
https://conservancy.umn.edu/pages/drum/policies/#data-collection-policy (accessed 2020-08-28)
2. University Digital Conservancy. "DRUM Policies and Terms of Use, Deposit License"
http://conservancy.umn.edu/pages/drum/policies/#deposit-license (accessed 2020-08-25)
3. DRUM Readme Template: http://z.umn.edu/readme (accessed 2020-08-28)
4. University of Minnesota Libraries. (2015). The Supporting Documentation for Implementing the Data Repository for the University of Minnesota (DRUM): A Business Model, Functional Requirements, and Metadata Schema. Retrieved from the University of Minnesota Digital Conservancy, http://hdl.handle.net/11299/171761.
5. University Digital Conservancy. "Policies and Guidelines, Preservation"
https://conservancy.umn.edu/pages/policies/#preservation (accessed 2020-08-28)
6. University of Minnesota Libraries. "Digital Preservation Framework"
https://www.lib.umn.edu/dp/digital-preservation-framework (accessed 2020-08-28)
7. McGory, John. Excel Archival Tool: http://z.umn.edu/exceltool (accessed 2020-08-28)
8. Lyrasis. "DSpace 5.x Documentation" https://wiki.lyrasis.org/display/DSDOC5x (accessed 2020-08-28)

9. 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

Response:

The Operating Principles of the University Libraries' Digital Preservation Framework states that the Libraries will "employ appropriate storage management technologies for digital resources, utilizing on-line, near-line, and off-line storage as appropriate" and will "ensure that hardware, software, and storage media containing archival copies of digital content is managed in accordance with environmental, quality control, security, and other standards and requirements" (1). Data files in DRUM are written to an Isilon storage system that replicates all changes every 8 hours to a geographically separated University of Minnesota Data Center Isilon cluster. The local Isilon cluster stores the data in such a way that the data can survive the loss of any two disks or any one node of the cluster. Snapshots are taken three times a day, and kept for 30 days. The 2nd cluster employs the same protections as the local cluster, and both verify each block via an MD5 checksum as it's read into memory, and again when it's written. The DSpace software also performs fixity checks on materials on a regular basis ensuring the integrity of the files over time.

1. University of Minnesota Libraries. "Digital Preservation Framework, Operating Principles"

https://www.lib.umn.edu/dp/digital-preservation-framework#4.2 (accessed 2020-08-25)

2. University of Minnesota Libraries. "Digital Preservation Implementation Activities"

https://www.lib.umn.edu/dp/implementation-activities (accessed 2020-08-25)

10. 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

Response:

The Digital Preservation Framework outlines a plan for the long-term preservation of its digital assets and "the Libraries intend to align its policy, procedures, and practices with the Trustworthy Repositories Audit & Certification (TRAC) standard (ISO/DIS 16363)" (1). Specifically, the UDC that houses DRUM, supports the long-term usability of data files held in DRUM in several ways. Regular data monitoring and secure storage maintains the integrity of data. Fixity checking is done on a daily basis. Files are stored in a location in which regular snapshots are being taken and maintained for 7 days after a file is modified or deleted. In addition, a full backup of the database is made every two hours and stored in two separate locations. The Libraries has two preservation support levels: fundamental support and advanced support (2). More detailed information on the approach to preservation including migrations, preferred file formats, and procedures can

be found on the Digital Preservation Practices section of the Libraries' website (3).

Contract and transfer of custody

DRUM Deposit Agreement (4) states that "I understand that the Digital Conservancy will do its best to provide perpetual access to my Content. In order to support these efforts, I grant the Regents of the University of Minnesota ("University"), through its Digital Conservancy, the following non-exclusive, perpetual, royalty-free, world-wide rights and licenses:

to access, reproduce, distribute and publicly display the Content, in whole or in part, in order to secure, preserve and make it publicly available, and

to make derivative works based upon the Content in order to migrate the Content to other media or formats, or to preserve its public access.

These terms do not transfer ownership of the copyright(s) in the Content. These terms only grant to the University the limited license outlined above."

1. University of Minnesota Libraries. "Digital Preservation Framework"
https://www.lib.umn.edu/dp/digital-preservation-framework (accessed 2020-08-25)
2. University of Minnesota Libraries. "Digital Preservation Framework, Technical Support Levels"
https://www.lib.umn.edu/dp/digital-preservation-framework#6 (accessed 2020-08-25)
3. University of Minnesota Libraries. "Digital Preservation, Digital Preservation Practices"
https://www.lib.umn.edu/dp/digital-preservation-practices (accessed 2020-08-25)
4. University Digital Conservancy. "DRUM Policies and Terms of Use, Deposit Agreement"
http://conservancy.umn.edu/pages/drum/policies/#deposit-license (accessed 2020-08-25)

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

11. 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

Response:

It is DRUM policy (1) that "Data must include adequate documentation describing the nature of the data at an appropriate level for purposes of reuse and discovery. All data receive curatorial review and data that are incomplete or not ready for reuse may not be accepted into the repository." DRUM's quality control during curation is conducted by adhering to the CURATED steps which are briefly listed below (more detail defined in the full checklist, see 2).

Check – Create an inventory of the files and review received metadata

Understand – Run the data/code, read documentation, assess for QA/QC red flags

Request – Work with the depositor to address any missing information or changes needed

Augment – Enhance metadata for discoverability and contextualize data with appropriate linkages (e.g., PUID for paper or published code, etc.)

Transform – Convert files to non-proprietary formats, if appropriate

Evaluate – Review overall data package for FAIRness

Document – Record all curation activities in a log file.

Issues that arise, following the Request step of CURATED, the curator returns data to the depositor for rectification and requests missing information or changes that will help the designated community understand the data. Any changes made to metadata, files, and documentation are noted in the dataset's curation log which is preserved as part of the metadata (but not available for public download).

Users and community members are able to contact the Deposit depositors whose email is listed in a separate, required field at time of ingestion and made public on the dataset's record page. DRUM also has a feedback form (3) that is open to accepting feedback from members of the community. These questions are answered by the University of Minnesota Archives staff who are well equipped to respond to needs and concerns from the user community. Feedback on the quality of a dataset would be directed to the depositor, who would then notify us of any changes that should be made. Otherwise, direct communication between users and depositors is encouraged by DRUM's requirement that each dataset include an "Author Contact" name and email displayed prominently in the metadata.

Additionally, curators ensure that all files collected adhere to legal and ethical requirements (more information provided in other sections of this application) and that data are contextualized by linking out to related publications and source

material.

1. University Digital Conservancy. "DRUM Policies and Terms of Use, Data Collection Policy"
https://conservancy.umn.edu/pages/drum/policies/#data-collection-policy (accessed 2020-08-28)

2. Data Curation Network. Checklist of CURATED Steps Performed by the DCN: http://z.umn.edu/curate (accessed 2020-08-28)

3. University Digital Conservancy. "Contact Us" https://conservancy.umn.edu/contact/ (accessed 2020-08-28)

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

12. 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:
4 – The guideline has been fully implemented in the repository
Accept

Response:

DRUM's data handling workflow is articulated in two ways.

First, a user-oriented description appears on DRUM"s public website. [1] The workflow handling of data in DRUM involves:
Curation: All data in DRUM are assigned to a subject-expert curator who performs data-level curation procedures

pioneered by the University of Minnesota Libraries, the lead institution in the Data Curation Network project. Curation includes:

Documentation: Work with depositors to create additional data documentation as necessary to describe and make the data re-usable, e.g., readme template

Metadata creation: Identify and apply standardized Dublin Core metadata schemas (see DRUM's open OAI metadata feed)

Quality control/Quality assurance: Open files, run code, detect/request missing files and/or information (e.g., what do "blank" cells mean?)

Chain of custody: Maintains the chain of custody for datasets via the repository technology system, DSpace, which is based on the OAIS standard for tracking the integrity and authenticity of all digital objects housed in the system

File format transformations: Ensure that data locked in proprietary formats can be preserved and migrated for reuse in the future

Creative Commons Licenses: Assist depositors in choosing an appropriate license

DataCite DOIs: Mint persistent URLs that ensure long-term citability

Discovery and access: DRUM is a publicly available collection of digital research data generated by University of Minnesota researchers, students, and staff. Anyone can search and download the data housed in the repository, instantly or by request. All datasets are fully indexed and search engine optimized for broad findability in Google, Google Scholar, Web of Science and others.

OAI Feed: DRUM metadata may be harvested using the open OAI protocol.

RSS Feed: Reader friendly updates to DRUM are available via the RSS protocol.

Preservation: DRUM provides long-term preservation of digital data files using services such as migration (limited format types), secure backup, bit-level checksums, and disaster recovery mechanisms in place.

Second, a much more detailed life-cycle curation workflow is maintained by DRUM curator staff and includes step by step procedures for ingest, appraisal, selection, curator processing. This document describes staff hand-offs and the specific steps taken in instances of PII, human subjects, and a range of other use cases. The unpublished DRUM Curator's Handbook may be accessed online (2).

1. University of Minnesota Digital Conservancy. "DRUM Services." (accessed 2020-08-25). https://conservancy.umn.edu/pages/drum/services/

2. University of Minnesota Libraries. DRUM curation procedures "Curator Handbook". (accessed 2020-08-25). https://docs.google.com/document/d/1XtZaxqo4pgmbHJiJQK6LmfdUYzadEQNaaoXiqOEp7og/edit#

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

13. 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

Response:

DRUM provides data discoverability and identification in a number of ways. DRUM has an advanced search tool powered by Solr full-text indexing of all text-based files in the repository as well as faceted browsing of the repository metadata (1). In addition, the full-text index and all associated metadata are exposed to online search tools such as Google and Google Scholar. DRUM provides an OAI feed for metadata harvesting: DRUM metadata is indexed by Thomson Reuters' Data Citation Index, a service for discovery of datasets, and SHARE, an American Research Libraries program. DRUM uses Dublin Core as its metadata standard. In addition to metadata dissemination, DRUM provides two kinds of persistent identifiers: a handle and a Digital Object Identifier (DOI). DRUM uses the DataCite Fabrica service to generate a DOI for each dataset (2). In addition, to help facilitate proper attribution, each dataset has a suggested citation, which includes the DOI (see example, 3).

1. University of Minnesota Digital Conservancy. "Data Repository for U of M": http://hdl.handle.net/11299/166578 (accessed 2020-08-27)
2. Datacite Fabrica. https://doi.datacite.org/ (accessed 2020-08-27)
3. Example on Suggested Citation: http://dx.doi.org/10.13020/D6V88Z (accessed 2020-08-27)

14. 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:
4 – The guideline has been fully implemented in the repository

Reviewer 2

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

Response:

All datasets accepted into DRUM are intended for reuse. The collection policy (1) states "Data must be deposited for open access -- this means that visitors to the Data Repository site may download and reuse your data. Data that is not suitable for reuse should not be shared in the Data Repository. Authors will have the option of restricting access for a maximum of two years (see End-user Access Policy)." Curators help facilitate reuse in the following ways:

Metadata: The metadata are based on Dublin Core Element Set and incorporate all the applicable Dublin Core elements. DRUM's metadata schema is published online (2). The metadata helps potential data users find and understand the data. DRUM metadata are exposed to online search engines, such as Google and Google Scholar, and are brokered through external indexing systems, such as DataCite and the Web of Science Data Citation Index. This exposure helps with the discovery of the dataset and helps potential data producers determine the relevancy of the dataset for their use. Additionally, persistent identifiers are added to each record metadata in the form of a DOI (via DataCite.org) and handles (via DSpace). Finally, all metadata are available for (open) ingest by third parties (XML of via API).

Data formats: The University Libraries' digital preservation department maintains a list of recommended file formats (3). During the curation process, curators work with data producers to convert files into the recommended file formats that

ensure that the file can be migrated, when needed, to the contemporary file format for use. These migrated formats are kept alongside the original file. DRUM can accept any format for inclusion in the repository, so while some disciplinary-specific formats may not have a migration path to an open format, DRUM accepts these formats as-is, and includes information about what software is needed to open the file in the data documentation.

Evolution of file format and migration: Regular data monitoring and secure storage maintains the integrity of data. As appropriate, files may be migrated to ensure that data files will be fit for re-use over time (4)

Ensuring understandability of data: DRUM strives to ensure that the files it maintains are usable and understandable. For example, DRUM collects information on methodology during the initial submission process where data depositors are instructed to upload documentation files and/or incorporate information about methodology in the metadata fields "abstract" and "description" in the DRUM Deposit Form. Metadata regarding the methodology, such as the instrument used to produce data and any codes or acronyms used in the data collection, is essential for data users to access and conduct secondary analysis and, if not included in the submission, DRUM curators work closely with the data producers to add a readme file using DRUM's template (5).

Attribution: The DRUM end-user policy (6) states "The user will give appropriate attribution to the author(s) of the data in any publication that employs resources provided by the Data Repository." To aid in this policy, each record in DRUM displays a "Suggested Citation" helping end users appropriately cite the dataset as well as a link to the policy on terms of use.

1. University of Minnesota Digital Conservancy. "DRUM Policies and Terms of Use, DRUM Data Collection Policy" https://conservancy.umn.edu/pages/drum/policies/#data-collection-policy (accessed 2020-09-08)
2. University of Minnesota Libraries. (2015). The Supporting Documentation for Implementing the Data Repository for the University of Minnesota (DRUM): A Business Model, Functional Requirements, and Metadata Schema. Retrieved from the University of Minnesota Digital Conservancy, http://hdl.handle.net/11299/171761.
3. University of Minnesota Libraries. "Digital Preservation Preferred File Formats" https://www.lib.umn.edu/dp/preferred-file-formats (accessed 2020-09-08)
4. University of Minnesota Digital Conservancy. "Policies and Guidelines, Preservation Policy" http://conservancy.umn.edu/pages/policies/#preservation (accessed 2020-09-08)
5. University of Minnesota Libraries. Readme Template: http://z.umn.edu/readme (accessed 2020-09-08)
6. University of Minnesota Digital Conservancy. "DRUM Policies and Terms of Use, Conditions." https://conservancy.umn.edu/pages/drum/policies/#terms-of-use (accessed 2021-04-13)

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

TECHNOLOGY

15. 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
Accept

Response:

DRUM is locally hosted using DSpace 5.5, which is a fully documented, open source software (1), with a strong world-wide user community and active governance (one of our staff is on the Leadership Team and our institution is a "gold" member). DSpace was originally built on the standards of the OAIS reference model specifications and has an up to date roadmap (2) and strategic plan (3). The DRUM workflow encompasses processes from data ingest to data dissemination. Four stages of curation (Receive, Appraise and Select, Processing, Access) ensures that DRUM datasets are independently understandable and available for the data users. DSpace is fully documented, open source software, with a user community [1]

A full-time Libraries developer maintains the current state of the repository software using up-to-date operating systems and software updates and documents all changes in GitHub. With the assistance of this developer and the bandwidth available to a major research institution, DRUM is able to meet the availability needs of the designated community.

Drawing from our disaster recovery plan detailed in section R3, backups of the DRUM database are done every hour which reduces possible data loss. Six virtual servers are used for various stages for development and production. Using virtual machines allows any of them to be available for any reason as needed as there is high availability of these virtual

machines which can be put into place if needed for various purposes. The process of doing so is tested all the time; for example the staging server is often a very close copy of the production server, verifying that the production server could be moved as needed and redeployed.

1. DSpace. https://duraspace.org/dspace/ (accessed September 10, 2020)
2. RoadMap. https://wiki.lyrasis.org/display/DSPACE/RoadMap (accessed April 13, 2021)
3. Strategic Planning. https://wiki.lyrasis.org/display/DSPACE/Strategic+Planning (accessed April 13, 2021)

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

16. 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
Accept

Response:

The University of Minnesota takes the necessary precautions to ensure that data housed in the data repository is protected and secure, including authentication, access, and account management (1), physical data center security (2) network firewalls (3), log management to help detect unauthorized activities on the system (4), and protection against malicious activity such as malware and viruses (5). Access to the back end of the system is controlled using Shibboleth

and restricted to the system administrator and the University's linux storage administrators. Data curators have access to edit DRUM records. View access is open to all. Access logs are reviewed looking for suspicious activities such as bots or crawlers, when found access to the materials is cut off to those "users" or "IP addresses." Materials receive a virus check upon ingest to ensure the data is not infected. Drawing from our disaster recovery plan detailed in section R3, materials are replicated to a second location. Snapshots are also taken on a regular basis, and maintained for 30 days. Both allow for recovery of the data if needed.

1. University of Minnesota. "Policy: Authentication, Access, and Account Management"

https://it.umn.edu/resources-it-staff-partners/information-security-standards/authentication-access-account-management (accessed April 13, 2021)

2. University of Minnesota. "Policy: Data Center Security"

https://it.umn.edu/resources-it-staff-partners/information-security-standards-guidelines/data-center-security (accessed April 13, 2021)

3. University of Minnesota. "Network Firewall"

https://it.umn.edu/resources-it-staff-partners/information-security-standards/network-firewall (accessed April 13, 2021)

4. University of Minnesota. "Log Management"

https://it.umn.edu/resources-it-staff-partners/information-security-standards-guidelines/log-management

5. University of Minnesota. "Virus/Malware Protection Management"

https://it.umn.edu/resources-it-staff-partners/information-security-standards-guidelines/virusmalware-protection (accessed April 13, 2021)

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

APPLICANT FEEDBACK

Comments/feedback

These Requirements are not seen as final, and we value your input to improve the CoreTrustSeal certification procedure. Any comments on the quality of the Requirements, their relevance to your organization, or any other contribution, will be considered as part of future iterations.

Response: