



DR-NTU (Data)

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:

Institutional repository

Reviewer Entry

Reviewer 1

Comments:
Accept.

Reviewer 2

Comments:
Accept.

Brief Description of Repository

DR-NTU (Data) is the institutional open access research data repository for Nanyang Technological University (NTU), supported by the NTU Library and Centre for IT Services (CITS) [1]. The repository was launched in November 2017, following the NTU Research Data Policy's [2] requirements for research data management, where final research data produced by research projects conducted at or under the auspices of NTU must be deposited in data repository and made available for sharing. The repository accepts final research data from multidisciplinary research. It is listed in Re3data.org [2] and OpenDOAR [3]. The repository is built on an open source web application called Dataverse [4], which is developed and used by Harvard University.

References:

1. About DR-NTU (Data): <https://libguides.ntu.edu.sg/drntudataguidespolicies/home>
2. NTU Research Data Policy: <https://www.ntu.edu.sg/research/ntu-research-data-policy>
3. Re3data.org – DR-NTU (Data): <https://www.re3data.org/repository/r3d100012440>
4. OpenDOAR - DR-NTU (Data): <https://v2.sherpa.ac.uk/id/repository/3925>
5. Dataverse Project: <https://dataverse.org/>

Reviewer Entry

Reviewer 1

Comments:
Accept.

Reviewer 2

Comments:
Accept.

Brief Description of the Repository's Designated Community.

Faculty, research staff, and students affiliated with Nanyang Technological University (NTU) may deposit their final research data in DR-NTU (Data) in order to make their data discoverable, accessible and reusable [1]. NTU is a research-intensive public university in Singapore. It has 33,000 undergraduate and postgraduate students in the colleges of Engineering, Business, Science, and Humanities, Arts and Social Sciences, Graduate College, and School of Medicine [2].

Data deposited and published in DR-NTU (Data) is open for public, and the metadata is indexed and discoverable in DataCite [3] and Web of Science Data Citation Index [4].

References:

1. About DR-NTU (Data): <https://libguides.ntu.edu.sg/drntudataguidespolicies/home>
2. About NTU: <https://www.ntu.edu.sg/about-us>
3. DataCite Search: <https://datacite.org/search.html>
4. Clarivate Web of Science Data Citation Index:
<https://clarivate.com/webofsciencegroup/solutions/webofscience-data-citation-index/>

Reviewer Entry

Reviewer 1

Comments:

Accept.

Reviewer 2

Comments:

Accept.

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

Reviewer Entry

Reviewer 1

Comments:

Accept.

Reviewer 2

Comments:

Accept.

Comments

DR-NTU (Data) allows depositors, who have relevant permission, to add, publish and edit datasets. However, each dataset is curated upon creation, with DR-NTU (Data) Managers adding extensive metadata to dataset and files [1]. Managers, who are typically Data Librarians, would advise depositor on enhanced curation (e.g. enhancement of documentation) and data-level curation (e.g. additional editing of deposited data files). The depositors and/or dataset's contact person(s) are expected to perform the enhancement and additional editing, if necessary.

Level C Curation (conversion to new formats) only takes place when the object meets the criteria of Preservation Level 3 [2], where Format Migration preservation strategy will be applied.

References:

1. DR-NTU (Data) Curator Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/curator>
2. DR-NTU and DR-NTU (Data) Preservation Plan [intranet only]

Reviewer Entry

Reviewer 1

Comments:
Accept.

Reviewer 2

Comments:
Accept.

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

- o Datacite for DOI minting (through Global Dataverse Community Consortium)

Reviewer Entry

Reviewer 1

Comments:
Accept.

Reviewer 2

Comments:
Accept.

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

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:
Not applicable.

Other Relevant Information.

- o DR-NTU (Data) contributes its metadata to the DataCite search and disseminating services [1].
- o DR-NTU (Data)'s metadata are open for harvesting from discovery platform such as Ex Libris Central Discovery Index [2].
- o DR-NTU (Data) is indexed in Web of Science Data Citation Index as well, providing citation counts for the datasets [3]

References:

1. DataCite Search: <https://datacite.org/search.html>
2. Ex Libris Central Discovery Index: <https://www.exlibrisgroup.com/products/primo-discovery-service/content-index/>
3. Clarivate Web of Science Data Citation Index:
<https://clarivate.com/webofsciencegroup/solutions/webofscience-data-citation-index/>

Reviewer Entry

Reviewer 1

Comments:

Accept.

Reviewer 2

Comments:

Accept.

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

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Accept.

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Accept.

Response:

Mission statement of DR-NTU (Data) is displayed on the repository's homepage [1] and user guides [2]. This is in line with library's role in NTU Research Data Policy, which approved by the university RDM Group [composition of the Group is described in R5 and R6].

Mission:

DR-NTU (Data) curates, stores, preserves, makes available and enables the download of digital data generated by the NTU research community in various disciplines. The repository develops and provides guidance for managing, sharing, and reusing research data to promote responsible data sharing in support of open science and research integrity.

References:

1. DR-NTU (Data): <https://researchdata.ntu.edu.sg>
2. DR-NTU (Data) User Guides and Policies: <https://libguides.ntu.edu.sg/drntudataguidespolicies/home>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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
Accept.

Reviewer 2

Comments:

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

Response:

DR-NTU (Data) encourages data sharing, and by default applying CC-BY-NC license to deposited datasets, as stated in the NTU Research Data Policy [1].

If depositors choose to use license which is different from the default, they must describe their license and terms in the dataset's Terms of Use [2], in order to inform others about any conditions for accessing and re-using their data.

Depositors can maintain access control over their dataset by restricting files and granting access to specific people or group of people. Access can be granted to each file, to all files in a dataset, and to all files in a collection of datasets (a dataverse).

For downloaders, they need to abide by the applicable data usage license agreement attached to the dataset [3].

Downloaders are not required to register or having an account in DR-NTU (Data), except for restricted data-files which require depositor's approval.

DR-NTU (Data) does not accept the deposit of sensitive research data, which includes individual's identifiable data [4].

Reference:

1. NTU Research Data Policy: <https://www.ntu.edu.sg/research/ntu-research-data-policy>
2. DR-NTU (Data) Depositor Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/depositor>
3. DR-NTU (Data) General Terms of Use – 7. Downloading User Submissions:
<https://libguides.ntu.edu.sg/drntudataguidespolicies/TermsOfUse>
4. DR-NTU (Data) Collection Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/collection>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

3. Continuity of access

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

Compliance Level:

3 – The repository is in the implementation phase

Reviewer Entry

Reviewer 1

Comments:

3 – The repository is in the implementation phase
Accept.

Reviewer 2

Comments:

3 – The repository is in the implementation phase
Accept.

Response:

In 2016, the Library was given strategic initiative funding by the University for Research Data Management headcounts. These manpower resources were to be part of a newly set up Research Data Management (RDM) Team in the Library then. Part of the RDM Team's key area of responsibilities was to develop the institutional research data repository. After the launch of the data repository in late 2017, the RDM Team has since grown to the current 6 FTE headcounts. The University understands very well the importance of fulfilling its obligations toward the university Research Data Policy as well as funders. It believes that the RDM support must continue in a sustainable approach. Since 2021, the University has shifted the funding source for the RDM support (which includes data repository) that is undertaken by the Library from strategic initiative funding to operational funding. This is a clear demonstration of the University's commitment and undertaking for the university research data repository as part of its operations, and this expectation is clearly articulated in the University Research Data Policy.

DR-NTU (Data)'s mission is to preserve and make available research data deposited in there. All practices and procedures carried out by DR-NTU (Data) are to ensure preservation and continued access to existing and new data deposited, consistent with our mission. DR-NTU (Data) liaises with depositors to ensure adequate data and documentation are provided to ensure usability of data and licence agreements for distribution and preservation. The DR-NTU (Data) Collection Guidelines [2] say that DR-NTU (Data) reserves the right to review, transfer, and remove unpublished datasets, which have been deposited for more than 10 years. The guaranteed preservation period for unpublished data would be 10 years. For published data, DR-NTU (Data) strives to ensure long-term continued access.

As a unit within the NTU, we can rely upon the institutional obligations of The University for both long term continuity and infrastructure support. DR-NTU (Data) does not have agreements currently with other archives or institutions in provision of data in the event of its cessation.

Continuous Improvement:

DR-NTU (Data) still requires a concrete succession plan to ensure continued access to its datasets in the event of cessation of operations. The following are planned improvement items:

- a. Explore possibilities for secondary custodian and/or reciprocal custodian [timeline 2022-24]
- b. Have an agreement with other archive or institution in provision of data access in the event of cessation of operations [timeline 2024-2025]

Reference:

1. NTU Research Data Policy: <https://www.ntu.edu.sg/research/ntu-research-data-policy>
2. DR-NTU (Data) Collection Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/collection>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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
Accept.

Reviewer 2

Comments:

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

Response:

DR-NTU (Data) does not accept sensitive data or research containing personal data or personally identifiable information [1].

DR-NTU (Data) provides the following information to depositors:

a. At the point of account creation:

When Designated Community creates their account with DR-NTU (Data), he/she must accept the General Terms of Use [2]. One of the points in the General Terms of Use highlights that deposited files cannot contain identifiable information:

“User Uploads must be void of all identifiable information, such that re-identification of any subjects from the amalgamation of the information available from all of the materials (across datasets and dataverses) uploaded under any one author and/or User should not be possible. Specifically, User Uploads cannot contain social security numbers; credit card numbers; medical record numbers; health plan numbers; other account numbers of individuals; or biometric identifiers (fingerprints, retina, voice, print, DNA, etc.). The only exceptions for when identifiable information is allowed are when:

- a. the information has been previously released to the public;
- b. the information describes public figures, where the data relates to their public roles or other non-sensitive subjects; or
- c. all identified subjects have given explicit informed consent allowing the public release of the information in the dataset”

b. At the point of dataset creation and publishing:

Automatic email from system will be sent to depositors. The email reiterates that deposited files cannot contain identifiable information.

This information is also provided in Depositor Guidelines [3] and related FAQ [4]. If depositor deposits deidentified or anonymised datasets, they are requested to notify DR-NTU (Data) Managers.

Guidance on appropriate deidentification and anonymisation techniques is provided in Sensitive Data Guide [5] and Anonymisation Guide [6].

DR-NTU (Data) Managers proactively curate deposited datasets at the earliest possible opportunity for potentially disclosive or sensitive data and reach out to depositors to ensure that the terms of informed consent for data sharing and other contractual/legal requirements allow for open data sharing. Assistance is provided to depositors, on anonymisation/de-identification techniques to ensure data is no longer sensitive and all identifiers removed. De-identified files will be put under access restrictions with relevant terms of access provided by depositors. When deemed necessary, DR-NTU (Data) Managers will deaccession or delete published dataset [1 – Deaccession and Deletion].

References:

1. DR-NTU (Data) Collection Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/collection>
2. DR-NTU (Data) General Terms of Use: <https://libguides.ntu.edu.sg/drntudataguidespolicies/TermsOfUse>
3. DR-NTU (Data) Depositor Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/depositor>
4. FAQ: What kind of data / content can be deposited in DR-NTU (Data)?: <https://libfaq.ntu.edu.sg/faq/261399>
5. Sensitive Data Guide: <https://libguides.ntu.edu.sg/rdm/sensitivedata>
6. Anonymisation Guide: <https://libguides.ntu.edu.sg/anon>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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
Accept.

Reviewer 2

Comments:

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

Response:

DR-NTU (Data) is hosted by the Nanyang Technological University (NTU). As the NTU Research Data Policy requires the sharing of final non-sensitive research data in the repository (or external subject specific open access data repositories), the University would need to ensure that adequate funding and sufficient number of qualified staff are available to manage and provide services for the use of the repository [1]. The repository is currently managed by a team of 3 Full Time Equivalent (FTE) data librarians [2], who are supported by 2 FTE support staff and 1 FTE IT staff. The Data Librarians hold either a master or doctoral degree.

The University Research Data Management Group (RDM Group), whose members include Associate Vice President (Infrastructure and Programmes), Research Integrity Officer, Research Support Office Director, Research Integrity & Ethics Office Director, and Data Librarians, provides advocacy, higher-level governance, and strategic decision making. The Group conducts 3-4 meetings per year, discussing and advising matters related to research data management within the university.

Since the launch of the Research Data Policy in mid-2016 and subsequent launch of the repository DR-NTU (Data) in late 2017, the University has been providing funding to the Library for additional 2-3 new FTE Data Librarians to undertake the

Library's responsibilities in research data management, including the provision of repository support services.

The University has been providing funding for Data Librarians and IT staff to attend conferences/meetings/e-learning/local courses that support them in their research data management work. See Requirement 6 (Expert Guidance) for more details of the conferences/meetings/e-learning/local courses.

Data Librarians are members of the Singapore Alliance of University Librarians (SAUL) Research Services Taskforce (RSTF). We host and conduct data sharing related information sharing meetings and events for which we open the invitation to anyone who could attend the events. A recent example is the RSTF 2020 Webinar (Oct 2020): <https://blogs.ntu.edu.sg/ntulibrary/2020/09/22/rstf-2020-webinar/>

The Data Librarians and support staff liaise with depositors to ensure adequate data and documentation are provided to ensure usability of data and licence agreements for distribution and preservation. In addition, the Data Librarians conduct DR-NTU (Data) workshops for data depositors [3] to demonstrate the use of the repository in an optimal manner.

Reference:

1. NTU Research Data Policy – Roles and responsibilities in Data Management – Library: <https://www.ntu.edu.sg/research/ntu-research-data-policy>
2. Research Data Management: Contact us: <https://libguides.ntu.edu.sg/rdm/contact>
3. Research Data Management: Workshops: <https://libguides.ntu.edu.sg/rdm/workshops/>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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

Reviewer Entry

Reviewer 1

Comments:

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

Reviewer 2

Comments:

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

Response:

DR-NTU (Data) has an ongoing expert guidance and feedback from

1. Internal / in-house, which comprises of

- The University Research Data Management Group, whose members include Associate Vice President (Infrastructure and Programmes), Research Integrity Officer, Research Support Office Director, Research Integrity & Ethics Office Director, and Data Librarians. The Group conducts 3-4 meetings per year, discussing and advising matters related to research data management within the university.

- Library

Data Librarians provide regular workshops and consultations for researchers, which serves as a feedback mechanism for the repository. The communication with researchers also serves as a channel to get advice on disciplinary/subject-related issues.

A dedicated IT staff provides technical support, regular maintenance, and system updates.

- Research Support Office for guidance on research data policy.

- Research Integrity and Ethics Office for expertise concerning confidentiality and ethics.

- Centre for IT Services for expertise on storage and security.

- Legal & Secretarial Office for guidance on legal matters.

- Research Integrity Committee, whose members include research leaders from the University. The Committee conducts two meetings per year, in which DR-NTU (Data) reports its recent developments and solicit advice.

2. External, which includes

- Dataverse Community

The Dataverse community's open source and transparent culture encourages the sharing of administrative and technical expertise, which supplement the expertise of DR-NTU (Data) staff.

DR-NTU (Data) staff actively join the fortnightly Dataverse Community Call [1] and the annual Dataverse Community Meeting [2] to discuss issues and gain insights on trends and latest development in Dataverse.

- Global Dataverse Community Consortium (GDCC)

As member of GDCC, DR-NTU (Data) has access to collaborative venue which supports the registration of persistent identifiers [3].

- Singapore Alliance of University Librarians (SAUL): Research Services Task Force (RSTF)

Data Librarians are members of RSTF. This Task Force serves as a community of practice for Singapore's academic

libraries, where members discuss and share their practices concerning research data management.

3. Expert visits

DR-NTU (Data) taps the expertise of international visitors:

- Talk and workshop by Sonia Barbosa, Harvard Dataverse Manager of Data Curation and Dan Brooke, Dataverse Program Manager (November 2019)

<https://blogs.ntu.edu.sg/ntulibrary/2019/10/16/dataverse2019/>

- Talk by Kevin Ashley, Digital Curation Centre (March 2019)

<https://blogs.ntu.edu.sg/ntulibrary/2019/02/11/datasharingseminar20190306/>

References:

1. Dataverse Community Calls: <https://dataverse.org/community-calls>

2. Dataverse Community Meetings: <https://dataverse.org/events>

3. Dataverse Project Global Dataverse Community Consortium:

<https://dataverse.org/global-dataverse-community-consortium>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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
Accept.

Reviewer 2

Comments:

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

Response:

Upon upload of data files, DR-NTU (Data) performs two integrity checks during the ingest. The Universal Numerical Fingerprint (UNF) is applied to uniquely identify and verify tabular data [1]. The MD5 checksums are calculated for each data file, to verify that the files are not corrupted via upload [2].

Upon deposit of a dataset, DR-NTU (Data) managers review the files and metadata, make changes if required, ask depositor for clarification and further information to ensure the completeness of dataset [3].

DR-NTU (Data) supports version control, and all changes to data files and metadata of published datasets are logged in the version summary and details. Minor changes, which typically occur on metadata, will result in a minor version update. Major changes, such as changes to data files, will result in a major version update. Previous versions of datasets are kept and remain accessible, allowing recovery of data and metadata. Should there be a case of unintentional changes, which resulted in de-accessioning of a version(s) of a dataset, reason for de-accessioning is captured (in the compulsory field) and displayed in the version summary.

To record the origin of data file and any transformation it has been through, provenance information can be uploaded at the data file level. The provenance file must be in JSON format and follow the W3C standard, or as a free text provenance description.

In the version control, the identities of contributors (depositors, curators) are logged automatically for each version. The contributors' accounts are verified by institutional log in. In cases where Depositors need to authorise other users to edit the dataset, the Depositor needs to set the Permission at the dataset or dataverse level. The Permission ensures that only permitted-user-account has access to edit specific dataset.

References:

1. Dataverse Universal Numerical Fingerprint: <http://guides.dataverse.org/en/latest/developers/unf/index.html>
2. Dataverse Configuration: <http://guides.dataverse.org/en/latest/installation/config.html>
3. DR-NTU (Data) Curator Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/depositor>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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
Accept.

Reviewer 2

Comments:

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

Response:

As a generic repository, DR-NTU (Data) does not put any limitations on the field of study represented in the data to be deposit. However, the repository has the Collection Guidelines [1] which guide the selection of data for archiving. If the deposited data is not in line with the Collection Guidelines, the DR-NTU (Data) Managers, who are typically Data Librarians, shall liaise with the depositor to discuss possible action, which may range from moving the data file to other repository or storage, to deaccessioning of dataset. The process of removing dataset is described in the Collection Guidelines – Deaccession and Deletion. A tombstone landing page with basic citation metadata will always be accessible with the existing persistent identifiers [2].

The Collection Guidelines require depositors to provide enough data and metadata/data documentation so that others can understand and (re)use the data.

DR-NTU (Data) requires depositors to fill up mandatory metadata fields necessary for creating dataset citations that follow Force11's data citation principles.

After deposit, each dataset in DR-NTU (Data) is curated by DR-NTU (Data) Manager, following the workflow outlined in Curator Guidelines [3]. The Manager shall check and communicate with the depositors when the metadata is insufficient or incorrect. The Manager also makes suggestions to depositors, and whenever necessary, makes changes to the metadata.

Recommended File Format list is available within the Collection Guidelines [1]. The Manager makes suggestion to depositors to convert their data files if they are not in the recommended format. DR-NTU (Data), however, accepts data in all formats, and makes it clear that the repository cannot guarantee the long-term digital preservation if the data is not in the recommended format.

References:

1. DR-NTU (Data) Collection Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/collection>
2. FAQ: Can I deaccession/unpublish a published dataset? <https://libfaq.ntu.edu.sg/faq/262604>
3. DR-NTU (Data) Curator Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/depositor>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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

Reviewer Entry

Reviewer 1

Comments:

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

Reviewer 2

Comments:

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

Response:

DR-NTU (Data) infrastructure is operated and managed by the NTU Centre for IT Services and has the same level of service quality and operational security as all other application services at the university. NTU Centre for the IT Services is responsible for the infrastructure with sufficient performance and storage space for production and development server.

NTU is committed to taking all necessary precautions to ensure the stability and security of system administration and operations. This includes vulnerability management, patch management, and backup and restoration procedures.

In addition, DR-NTU (Data) maintains on-site back-ups, with monthly full system back-ups, and daily backup with a snapshot of the data and the metadata, utilising both virtual and physical servers. The backups are stored in a secured location and they are only available to authorized employees.

DR-NTU (Data)'s data deposit and curation procedures are documented in Depositor Guidelines [1] and Curator Guidelines [2]. The guidelines serve as standardised workflows which ensure that DR-NTU (Data) managers shall arrive at substantially the same outcome while curating the datasets. DR-NTU (Data)'s data file fixity checks help the managers to ensure data consistency across archival copies.

References:

1. DR-NTU (Data) Depositor Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/depositor>
2. DR-NTU (Data) Curator Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/curator>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

10. Preservation plan

R10. The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.

Compliance Level:

3 – The repository is in the implementation phase

Reviewer Entry

Reviewer 1

Comments:

3 – The repository is in the implementation phase
Accept.

Reviewer 2

Comments:

3 – The repository is in the implementation phase
Accept.

Response:

DR-NTU (Data) is still in the process of implementing the preservation plan. The plan is currently residing on intranet and a copy is shared with Reviewers. Public version of the plan is expected to be ready by end of 2022.

The Preservation Plan [1] covers two repositories:

- DR-NTU, which is the institutional repository for scholarly outputs except research data
- DR-NTU (Data), which is the institutional repository for research data

The Preservation Plan [1] outlines the mandate, roles and responsibilities, operating principles, procedural accountability, preservation strategies, preservation levels, and recommended file format. The Plan draws on Open Archival Information System (OAIS) Reference Model and Library of Congress Digital Collection Management. The Plan emphasises that retention, preservation and access of research data must comply with the national regulations (e.g. Singapore Personal Data Protection Act), as well as institutional policies (e.g. NTU Data Governance).

The preservation strategies include Bit level preservation, Fixity checking, Normalisation, and Format migration.

The preservation levels are as follow:

- Preservation Level 1: applies to all digital object in all formats, where bit level preservation and fixity checking are applied.
- Preservation Level 2: applies to all digital object in all formats, and the object must be published. Normalisation strategy is applied.
- Preservation Level 3: applies to object in recommended formats, and the object must be published. Format migration strategy is applied.

The preservation levels correspond to the level of curation performed (R0), in the following ways:

- For all deposited digital objects, in all formats, not yet published: Preservation Level 1 is automatically applied. Level B Curation (basic curation) is performed. Occasionally, when depositors are highly experienced in applying metadata and supplying adequate documentation, Level A Curation (content distributed as deposited) is adopted.
- For all deposited digital objects, in all formats, and the objects are published: Preservation Level 1 is automatically applied, and Preservation Level 2 is applied as well. Level B Curation (basic curation) is performed. Occasionally, when depositors are highly experienced in applying metadata and supplying adequate documentation, Level A Curation (content distributed as deposited) is adopted.
- For deposited digital objects in recommended formats, and the objects are published: Preservation Level 1 is automatically applied, and Preservation Level 2 is applied as well.

Preservation Level 3 (format migration) shall be applied when the recommended formats have been assessed as having a risk of obsolescence. Dataverse automatically converts certain types of files that contain well-formed tabular data into non-proprietary, tab delimited files. Therefore, the focus of Preservation Level 3 will be on non-tabular data.

Prior to applying format migration: a) there must be evidence that the newer format is able to preserve the object's content, context and structure, b) there must be consultation with depositors and/or community of practice to gauge the

risk of format migration.

The Preservation Level 3 is in line with Level C Curation (Enhance curation, e.g. conversion to new formats). In other words, Level C Curation only takes place when the object meets the criteria of Preservation Level 3.

Currently, DR-NTU (Data) assures long-term access to files and content using strategy as described in Requirement 9.

The General Terms of Use describes the agreement between depositor and DR-NTU (Data) [1]. By agreeing to the Terms, the depositor gives DR-NTU (Data) permission and required licenses to make the deposited content available for archiving, preservation, and access within the repository. The depositor also grants DR-NTU (Data) the right to convert the deposited files and metadata to any medium or format and make multiple copies of the deposited dataset for the purposes of preservation and accessibility.

Depositor must agree and accept the General Terms of Use, upon registration for DR-NTU (Data) account. Salient points on User Uploads, which include repository services, restrictions, licenses and permissions, and submission agreement are auto emailed to depositor, upon creation of each dataset.

Continuous Improvement

The implementation of Preservation Plan still requires improvement, especially in periodic risk-assessment where obsolete file format can be detected, and subsequently migrated. The following are planned improvement items:

- a. Risk assessment via file format identification and validation. The repository's software, Dataverse, recognises and categorises most of file types. However, there are several file types which are not detected by Dataverse yet. Ideally, improvement would include additional detection, by Redetect command, API endpoint, or other methods [timeline 2021-2022]
- b. Risk assessment via community of practice. Workshops and engagement activities may include, in the future, discussion on preservation and file formats widely used in specific subject [timeline 2021-2022]

References:

1. DR-NTU and DR-NTU (Data) Preservation Plan [intranet only - copy is shared with Reviewers]
2. DR-NTU (Data) General Terms of Use: <https://libguides.ntu.edu.sg/drntudataguidespolicies/TermsOfUse>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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:

3 – The repository is in the implementation phase

Reviewer Entry

Reviewer 1

Comments:

3 – The repository is in the implementation phase
Accept.

Reviewer 2

Comments:

3 – The repository is in the implementation phase
Accept.

Response:

DR-NTU (Data) does not judge nor assess the scholarly quality of deposited datasets. However, DR-NU (Data) curators assess the consistency of the dataset with the provided documentation/metadata and request edits and updates from depositors as needed. Data Librarians, who also serve as DR-NTU (Data) managers, hold graduate degrees in science and social sciences.

The workflow, for quality control and check, is outlined in the Curator Guidelines [1] and Depositor Guidelines [2]. As DR-NTU (Data) allows depositors to deposit directly to the repository, issues are typically resolved by asking the depositor to make the edits, or curators perform the metadata edits and communicate this accordingly to the depositor.

DR-NTU (Data) provides metadata fields for depositor to include related publications and related datasets, along with their corresponding persistent identifiers. When appropriate, Curators would advise depositor to add “Replication data for” or “Related data for” as a prefix in the dataset title.

The “Replication data for” prefix is used when depositor has sufficient data documentation that facilitates others to replicate the experiments to derive the same findings as reported in the published paper.

The “Related data for” prefix is used when depositor is sharing less than what is needed for others to replicate.

Metadata customisation and metadata template are supported in DR-NTU (Data) as well. This ensures that dataset from specific subject/discipline can be better described in the custom metadata fields.

DR-NTU (Data) organises dataset’s files in a folder structure [3], allowing end user to have a better clarity to the

relationship between data files.

There is no formalised way for the designated community to comment on, and/or rate data and metadata. However, Data Librarians regularly conduct workshops [4] to engage designated community to better understand their data requirements.

Continuous Improvement:

The following are planned improvement items, to ensure sufficient information about the dataset:

- a. Collaborate with subject experts (e.g. researchers from the university) to develop subject guides on metadata and file format, which are specific to a certain subject or a community-endorsed standard [timeline 2021-2023]
- b. Grow a community of practice, which consists of subject experts, that can be tapped for advice when DR-NTU (Data) Managers require subject-expert-assistance in curating dataset and metadata. The practices can be compiled into online guides and reused in the workshops [4] [timeline 2021-2023]

References:

1. DR-NTU (Data) Curator Guidelines (internal guide): <https://libguides.ntu.edu.sg/drntudataguidespolicies/curator>
2. DR-NTU (Data) Depositor Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/depositor>
3. Finding and using data – Tree view: <http://guides.dataverse.org/en/latest/user/find-use-data.html>
4. DMP and DR-NTU (Data) workshops: <https://libguides.ntu.edu.sg/rdm/workshops>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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

Accept.

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Accept.

Response:

DR-NTU (Data) provides instructions to Depositors on data preparation, deposit and publishing, in the form of Depositor Guidelines [1]. Detailed information on specific steps are provided in the form of FAQs as link outs from Depositor Guidelines. DR-NTU (Data) Managers conduct monthly workshops to the community on DR-NTU (Data) functionalities and deposit workflow [2].

Depositor Guidelines consist of the following sections:

1. General Information:
 - a. What to deposit
 - b. What not to deposit
 - c. When to deposit
2. Prepare your data for deposit
 - a. Prepare your data files
 - b. Prepare related documentation
3. Deposit your files
 - a. Create user account, or log in
 - b. Create a dataverse
 - c. Create a dataset
 - d. Upload files
 - e. Review usage license
 - f. Restrict access (If applicable)
 - g. Share unpublished dataset (Optional)
4. Publish your dataset
5. Edit your published dataset
6. Deaccession

DR-NTU (Data) provides instructions to DR-NTU (Data) Managers on curation workflow, in the form of Curator Guidelines [3].

Curator Guidelines consist of the following sections:

1. Start of curation process
2. Upon creation of a new dataverse
3. Upon creation of a new sub-dataverse
4. Upon creation of a new dataset

- a. Mandatory metadata fields
- b. Additional metadata fields
- c. Files
- d. Terms
- e. Republish datasets (If applicable)

The purpose of these guidelines is to ensure that the resulting dataset meets the collection criteria set out in Collection Guidelines [4] and the specified data and metadata quality standards, in accordance with FAIR principles. In addition, DR-NTU (Data) Managers proactively communicate with Depositors at the earliest possible opportunity (e.g. dataverse/dataset creation and file deposit), and make enhancements/corrections to dataset metadata in consultation with Depositors, where necessary, as stated in Curator Guidelines [3].

These guidelines are published with date and version information. They are reviewed periodically by Data Librarians.

References:

1. DR-NTU (Data) Depositor Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/depositor>
2. Research Data Management: Workshops: <https://libguides.ntu.edu.sg/rdm/workshops/>
3. DR-NTU (Data) Curator Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/curator>
4. DR-NTU (Data) Collection Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/collection>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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

Reviewer Entry

Reviewer 1

Comments:

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

Reviewer 2

Comments:

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

Response:

As a multi-disciplinary, institutional data repository, DR-NTU (Data) is listed in Re3data.org [1] and OpenDOAR [2].

DR-NTU (Data) supports faceted browsing, searching across all metadata fields, and advanced search using specific metadata fields. The Widgets feature in DR-NTU (Data) allows Depositors to list their datasets on their personal websites [3], hence promoting discovery and raising visibility.

DR-NTU (Data) supports the publishing and harvesting of metadata in several standards (e.g. Dublin Core, DDI, and DataCite) over the widely-used OAI-PMH protocol, and exposes collection and dataset-level metadata to search engines. DR-NTU (Data)'s datasets are currently indexed in DataCite [4], Ex Libris Central Discovery Index [5], and Clarivate Web of Science Data Citation Index [6].

DR-NTU (Data) utilises DOI as persistent identifier. The DOI minting service is provided by Datacite, through Global Dataverse Community Consortium membership.

DR-NTU (Data) provides standardised citation of datasets which follows the Force11 Joint Declaration of Data Citation Principles [7]. Citation can be exported as EndNote XML, RIS, and BibTex.

References:

1. Re3data.org – DR-NTU (Data): <https://www.re3data.org/repository/r3d100012440>
2. OpenDOAR – DR-NTU (Data): <https://v2.sherpa.ac.uk/id/repository/3925>
3. Dataverse Management – Widget: <https://guides.dataverse.org/en/latest/user/dataverse-management.html>
4. DataCite Search: <https://datacite.org/search.html>
5. Ex Libris Central Discovery Index: <https://www.exlibrisgroup.com/products/primo-discovery-service/content-index/>
6. Clarivate Web of Science Data Citation Index:
<https://clarivate.com/webofsciencegroup/solutions/webofscience-data-citation-index/>
7. Data Citation by Dataverse Project: <https://dataverse.org/best-practices/data-citation>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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:

3 – The repository is in the implementation phase

Reviewer Entry

Reviewer 1

Comments:

3 – The repository is in the implementation phase

Accept.

Reviewer 2

Comments:

3 – The repository is in the implementation phase

Accept.

Response:

Documentation and metadata are crucial for the reuse of the research data. The Collection Guidelines [1] require depositors to provide enough data and metadata/data documentation so that others can understand and reuse the data.

At the stage of data preparation, depositors are introduced to best practices for file naming and data organisation to reflect the content, status and version of files. Depositors are also recommended to upload README files together with data files. Information on structure/content and format of README files are provided with the goal of allowing others to interpret, understand and reuse the dataset [2].

Upon data deposit, DR-NTU (Data) requires depositors to fill up mandatory metadata fields (e.g. Title, Author, Contact, Description, Subject, Keyword, Kind of Data, Software) necessary for generating dataset citations that follow Force11's data citation principles.

The approach to curation is documented in the Curator Guidelines [3]. As part of the curation process, each deposited dataset is reviewed by DR-NTU (Data) Manager. The Manager shall check and communicate with the depositors when the mandatory metadata is insufficient or incorrect. Whenever applicable, the Manager enriches the metadata by adding

additional metadata information (e.g. Related Publication, Research Grants, Related Dataset) in collaboration with the depositors.

DR-NTU (Data) also provides disciplinary-specific metadata, such as Geospatial metadata, Social Science and Humanities metadata, and Life Sciences metadata. The Manager recommends the disciplinary-specific metadata to depositors when their datasets fit into the specific discipline.

In addition to metadata, the Manager also reviews the deposited files (file name, file size, content, format, documentation, tags, descriptions) and gives suggestions to depositors, with the aim of increasing the understandability, and ultimately the reuse of the data.

Recommended file formats are stipulated in Collection Guidelines [1] and further explained in Requirement 8 (Appraisal). Depositors are advised to deposit files in open file formats to maximise long-term access and reuse [2]. In addition, DR-NTU (Data) converts certain types of files that contain well-formed tabular data into non-proprietary, tab delimited files [4], for ease of reuse.

The default license applies by DR-NTU (Data) is CC-BY-NC, which allows for reuse of data, with proper attribution, in non-commercial context.

Continuous Improvement:

The following are planned improvement items, to ensure that dataset continues to be understood and reused:

- a. Collaborate with subject experts (e.g. researchers from the university) to develop subject guides on metadata and file format, which are specific to a certain subject or a community-endorsed standard [timeline 2021-2023]
- b. Grow a community of practice, which consists of subject experts, that can be tapped for advice when DR-NTU (Data) Managers require subject-expert-assistance in curating dataset and metadata. The practices can be compiled into online guides and reused in the workshops [4] [timeline 2021-2023]

References:

1. DR-NTU (Data) Collection Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/collection>
2. DR-NTU (Data) Depositor Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/depositor>
3. DR-NTU (Data) Curator Guidelines: <https://libguides.ntu.edu.sg/drntudataguidespolicies/curator>
4. Dataverse Tabular Data File Ingest: <https://guides.dataverse.org/en/latest/user/tabulardataingest/index.html>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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
Accept.

Reviewer 2

Comments:

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

Response:

DR-NTU (Data) is built on open-source software, Dataverse. Dataverse software is developed by the Institute for Quantitative social Science at Harvard University [1] and supported by a dedicated team, along with international community of Dataverse users who are also part of several organizations that set standards and best practices for scholarly data sharing, including the Research Data Alliance, Force11 and DataCite.

Development of the Dataverse software is managed using GitHub [2], which allows the community to contribute to the software and track changes. New releases of the software are continuous, about 3-4 updates per year and DR-NTU (Data) keeps up with the software updates.

DR-NTU (Data) adopts the standards which are supported by Dataverse, such as OAI-PMH, SWORD API, Schema.org/JSON-LD for provenance JSON files, DDI, Dublin Core and Shibboleth/SAML.

Hardware and software inventories and configuration information are recorded in a NTU managed application that takes inventory as and when there are changes in the systems.

Bandwidth, connectivity and storage space sufficiently meet the needs of the Designated Community. Storage space is auto monitored, with alert and regular reports sent to Data Librarians.

References:

1. Dataverse Project - About: <https://dataverse.org/about>
2. Dataverse Github: <https://github.com/IQSS/dataverse>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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
Accept.

Reviewer 2

Comments:

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

Response:

DR-NTU (Data) is owned by Nanyang Technological University (NTU). The University Acceptable IT Usage Policy [1] applies to the Designated Community of DR-NTU (Data).

DR-NTU (Data) complies with the NTU Centre for IT Services' security requirements to ensure consistent and systematic

information security. The requirements include system security requirements, operational requirements and regular auditing and review. Computer Security Incident Response Team (CIRT), which is part of Centre for IT services and led by department's information security officer helps to improve the security of DR-NTU (Data) network, reduce the number of security incidents and the potential harm caused, as well as raise awareness of security issues among IT consultants and end users. This includes any incident affecting information security at NTU, incidents that compromise confidentiality and integrity of data, as well as unwanted incidents affecting the availability of data.

Authorization to modify the application and system update is limited to a designated Centre for IT Services staff and Library IT staff.

The technical infrastructure including the servers are located in a secure data center, where only authorized employees have access to the equipment after identification. The systems are all provided with redundant power supplies that are located on separate power groups that are powered by UPS and generators, even if there is a power failure.

Backups are made to tapes and are done on a daily, weekly and monthly basis. The backups are stored in a secured location and they are only available to authorized employees.

References:

1. NTU Acceptable IT Usage Policy: <https://www.ntu.edu.sg/cits/Pages/Acceptable-IT-Usage-Policy.aspx>

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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:

We do not have further feedback at this point.

Reviewer Entry

Reviewer 1

Comments:

This application demonstrates evidence of robust procedures and systems that are mostly in line with the expectations of the CoreTrustSeal criteria. Following review, the repository has recognised where there is opportunity to improve and adjusted their compliance level for criteria R11 and R14 accordingly. They have provided reviewers with a copy of their current preservation plan in support of R10 and clarified the oversight and guidance that is provided by the University leadership. As a result of these changes, it is my view that the application provides a comprehensive and accurate assessment of the repository against the CoreTrustSeal criteria.

CoreTrustSeal board comment: It is expected that the applicant implements their plan to make their preservation policy publicly available for R10 for the next renewal round.

Reviewer 2

Comments:

Overall comments (including any good points)

The repository has satisfied all requirements of CTS. Although it is a young repository, it has a team of experienced librarians and other staff in the field and provide quality services to the community. The repository has a solid infrastructural and technical foundation as a part of the NTU.