

Qualitative Data Repository

Notes Before Completing the Application

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

True

Reviewer Entry

Reviewer 1

Comments: No comment required.

Reviewer 2

Comments:

CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

Background & General Guidance

Glossary of Terms

BACKGROUND INFORMATION

Context

R0. Please provide context for your repository.

Repository Type. Select all relevant types from:

Domain or subject-based repository

Reviewer Entry Reviewer 1 Comments: No comment Reviewer 2 Comments:

Brief Description of Repository

The Qualitative Data Repository is a curated domain repository for data collected or generated through qualitative and multi-method research in the social sciences and cognate disciplines.

Links: QDR Mission (on homepage towards the bottom): https://qdr.syr.edu

Reviewer Entry Reviewer 1 Comments: No comment

Reviewer 2

Comments:

Brief Description of the Repository's Designated Community.

QDR serves a global social science community and especially those researchers working with qualitative methods and data. We understand "social science" broadly to include other disciplines using social science methodology. In particular, QDR Given the relative novelty of sharing qualitative social science data, especially in the US context, QDR's staff actively engage in training researchers, providing in-person workshops and consultation (e.g., at the annual meeting of the American Political Science Association [1] and the annual Institute for Qualitative and Multi-Method Research [2]) as well as remote instruction via webinars. [3] We have also published a self-guided online version of our course on managing qualitative data.[4]

Links:

1. Short courses directed by QDR staff at the American Political Science Association's Annual Meeting: https://www.maxwell.syr.edu/moynihan/cqrm/Short_Courses_at_APSA/

2. An introductory video on Data Management Plans recorded as part of the Institute for Qualitative and Multi-method research: https://www.youtube.com/watch?v=wIDBNcI9Fbc

3. Webinar: Theory and Practice of Transparent Qualitative Health Research, co-conducted with researchers from the Guttmacher Institute. https://www.youtube.com/watch?v=GQVw9H7Ji6c

4. Online course on managing qualitative data: https://managing-qualitative-data.org

Reviewer Entry

Reviewer 1 Comments: No comment

Reviewer 2

Comments:

Level of Curation Performed. Select all relevant types from:

C. Enhanced curation - e.g. conversion to new formats; enhancement of documentation

Reviewer Entry

Reviewer 1

Comments: No comment

Reviewer 2

Comments:

Comments

During curation, QDR adds extensive metadata to data and files, converts file formats, and advises depositors with regard to respecting relevant ethical and legal limitations on data sharing (de-identification, copyright); curation processes include disclosure-risk review of human-participant data as detailed in R4.

Links:

QDR Curation policy: https://qdr.syr.edu/policies/curation

Reviewer Entry

Reviewer 1

Comments: Version 2: Agreed, this is now addressed

Version 1: Does QDR undertake any deidentification, anonymisation, etc. themselves? This has implications for what data can be published.

Reviewer 2

Comments:

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

QDR's primary storage and computing environment are provided by Amazon Web Services (AWS) (East Coast). We rely on AWS for our main servers, part of our security infrastructure as well as parts of preservation and on-site back-ups. We use Google Cloud (West Coast) as a tertiary storage for preservation copies of published data. Reliance on Google Cloud is minimal as we only rely on its "cold storage" feature. It serves mainly to add diversity in the geographic location of data storage and use or providers.

Syracuse University through the Maxwell Information and Computing Technology Group [3] provides us with local IT support, secure storage for sensitive data during curation, and, in rare cases, for long-term storage, as well as server space for secondary, off-site backups. Syracuse IT also provides us with automated penetration testing via the University of Texas Dorkbot [4].

We mint DOIs through Datacite, membership in which is provided via membership in the Global Dataverse Community Consortium [5].

QDR runs on open-source software. Our homepage runs Drupal [4] and our repository uses the Dataverse software [6]. QDR closely monitors security communication from both projects. QDR is also an active contributor to the Dataverse software [7] and is frequently involved in discussions about new or updated features.

Links:

- 1. Amazon Web Services: https://aws.amazon.com/
- 2. Google Cloud: https://cloud.google.com/
- 3. Syracuse University's Maxwell Information and Computing Technology Group Website: (https://ict.syr.edu/
- 4. University of Texas Dorkbot: https://security.utexas.edu/dorkbot
- 5. Global Dataverse Community Consortium Website: http://dataversecommunity.global/
- 6. Drupal Website: https://www.drupal.org/
- 7. Dataverse Software Website: https://dataverse.org

Reviewer Entry

Reviewer 1

Comments: No comment

Reviewer 2

Comments:

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

- With the failure of the Digital Preservation Network (DPN) [1], we moved QDR to different preservation storage, which combines a standard-based preservation format with external storage in Google Cloud (details below).

- As indicated as planned in the last application, we switched our DOI registration to DataCite.

- Since Fall 2019, we have implemented standardized usage metrics following the standards established by the Make Data Count project (details below).

- We have added several additional revenue sources, in particular, institutional membership and fees for deposits (details below)

- We have updated all relevant software components, including our Drupal and Datacite instances.

Links:

[1] https://duraspace.org/the-digital-preservation-network-dpn-to-cease-operations/?

Reviewer Entry Reviewer 1 Comments: No comment Reviewer 2

Comments:

Other Relevant Information.

QDR is the only repository with a sole focus on qualitative social science data in North America. It is part of Data-PASS, the data preservation alliance for the social sciences (discussed in detail below).

Reviewer Entry

Reviewer 1

Comments: No comment

Reviewer 2

Comments:

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 Agreed.

Reviewer 2

Comments:

4 - The guideline has been fully implemented in the repository

Response:

QDR's mission statement is displayed on our homepage at qdr.syr.edu:

"QDR curates, stores, preserves, publishes, and enables the download of digital data generated through qualitative and multi-method research in the social sciences. The repository develops and disseminates guidance for managing, sharing, citing, and reusing qualitative data, and contributes to the generation of common standards for doing so. QDR's overarching goals are to make sharing qualitative data customary in the social sciences, to broaden access to social science data, and to strengthen qualitative and multi-method research." [1].

QDR's mission statement is promoted on our webpage as well as various presentations and hand-outs and provides the foundation for our policies and governance.

Links:

1. QDR Mission (on homepage): https://qdr.syr.edu

Reviewer Entry

Reviewer 1

Comments: No comment

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 Agreed

Reviewer 2

Comments:

4 - The guideline has been fully implemented in the repository

Response:

QDR data are governed by a set of agreements between QDR and depositors (Standard and Special Deposit Agreement) and between QDR and downloaders (Standard and Special Download Agreements). Their content is described in detail below. All documentation is published under a CC-BY license. Currently, QDR disseminates no data under an open license (such as CC0 or CC-BY). This decision, based on advice from similar repositories such as UK Data, reflects the particular sensitivity of much qualitative data. We constantly re-evaluate this policy to assure it serves the interest of our depositors, the human participants involved in the data projects that are deposited with us, as well as QDR's stakeholders in the research community; the last such consultation, involving both internal discussions and our research advisory board, confirmed this practice.

Data projects are clearly labeled as either "Standard Access" [e.g. 1] or "Restricted/Controlled Access" [e.g. 2] on the data landing page. Each license leads to a page explaining it in plain language with links to the full text of the license terms. [3, 4] Files are visibly labeled as either "Data" or "Documentation" (with documentation files sorted to the top of the file list). To further clarify the distinction between the applicable licenses for data and documentation files, documentation files are accessible without login, while data files require login.

As part of the standard deposit agreement [5], depositors grant QDR all rights required to perform curation and preservation tasks on the data, specifically (drawn directly from the text of the agreement):

- "To disseminate copies of the data project in a variety of media formats

- To promote and advertise the data project in any forms or media

- To describe, catalog, validate and provide documentation about the data

- To store, translate, transfer, move, copy and re-format the data in any way to ensure its future preservation and accessibility

- To incorporate metadata and documentation for the data project into public access catalogs. - To enhance, transform and/or re-arrange the data project, including the data and metadata, in order to protect respondent confidentiality, improve usability, or to facilitate any task listed [...] above."

As part of our standard download agreement [6], researchers agree to use data only for research and teaching and to not attempt to re-identify individuals in de-identified data. The download agreement also bars the unauthorized redistribution of data, requires proper attribution during its use, and lists possible consequences for violating the agreement, which include possible bans from future use of QDR, reports to both institutional and federal bodies regulating research ethics, and potential legal action.

As part of a special deposit agreement [8], depositors can specify additional access conditions [7] that can regulate who may access the data, and to which users need to explicitly agree (by signing a special download agreement [9]) before downloading. Such access conditions are adjusted to reflect both the sensitivity of the data and the risk for re-identification and are custom-created in close communication with the depositor. Internally, restricted data are handled using specific protocols as specified under R4.

Links:

- 1. Example data project with standard access: https://doi.org/10.5064/F6CBM0MJ
- 2. Example data project with restricted/controlled access: https://doi.org/10.5064/F6GZGCJB
- 3. Plain language page for standard access: https://stage.qdr.org/policies/qdr-standard-access-conditions
- 4. Plain language page for restricted access: https://stage.qdr.org/policies/qdr-restricted-access-conditions
- 5. QDR Standard Deposit Agreement (requires registration, attached): https://qdr.syr.edu/deposit/standarddeposit
- 6. QDR Standard Download Agreement: https://qdr.syr.edu/discover/standarddownload
- 7. QDR Access controls: https://qdr.syr.edu/guidance/human-participants/access-controls
- 8. QDR Special Deposit Agreement (requires registration, attached): https://qdr.syr.edu/deposit/specialdeposit
- 9. QDR Special Download Agreement (requires registration, attached):

Reviewer Entry Reviewer 1 Comments: No comment

Reviewer 2

Comments: Accept

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 Agreed

Reviewer 2

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

Response:

QDR is committed to providing repository services and access to data for the long term. Cognizant that funding and institutional environments can change, QDR has taken various measures to ensure continued access to its materials in the event of cessation of operations, within the scope of its commitment to provide access to data for at least 20 years from the point of deposit. Beyond QDR's strategies for sustainability (R5) and preservation (R14), the repository assures continued access to its material primarily via membership in the Data Preservation Alliance for the Social Sciences [1].

Data-PASS membership (which also includes joint events, statements on policy-related questions, as well as coordination on metadata and preservation-related issues) includes a guaranteed succession rule. If any of Data-PASS's member repositories cease operations, the other members agree to continue hosting its materials. The transfer protocol, as agreed to in the currently active Memorandum of Understanding [2], stipulates that should any Data-PASS member no longer be able to retain materials in its possession, it provides 90-days written notice to the partnership, which then determines a suitable repository to assume custody of the data. As described in the MOU, ICPSR is the most likely host repository for non-federal data, with the National Archives and the Library of Congress serving as last resort options. QDR's deposit agreement [3] explicitly allows QDR to transfer stewardship of deposits to ensure ongoing access.

Links:

- 1. Data-PASS: http://www.data-pass.org/join.html (see first bullet point on "What are the benefits")
- 2. Data-PASS MOU: http://data-pass.org/sites/default/files/Data-PASS_MoU_201504.pdf
- 2. QDR Standard Deposit Agreement (requires registration, attached): https://qdr.syr.edu/deposit/standarddeposit

Reviewer Entry

Reviewer 1

Comments: No comment

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 Version 2: Agreed

Version 1: Agreed, subject to response on comment below.

Reviewer 2

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

Response:

Ethical research and data practices are of concern to all researchers. They can be of particular concern to qualitative researchers who have long-established relationships of trust with research participants. As part of QDR's curation protocol [1], for all data projects that include data gathered from human participants (such as interview transcripts), QDR requests [2] and reviews IRB/ethics board approval [3] and, in particular, the informed consent language used during the research to help the depositor evaluate if the sharing of data is precluded, and/or to aid the depositor in respecting any limits on the sharing of data that result from guarantees made to project participants.

Where deposited data are de-identified, curators review all data to help depositors decide whether de-identification has been carried out properly [4]. Where data and related documentation are in a language in which no member of the QDR staff is proficient, curation staff uses automated translation to spot check de-identification and conveys best de-identification practices to depositors. In all cases, QDR's role is advisory. The final responsibility for decisions concerning de-identification remains with depositors. QDR, however, will not publish confidential data that its curators consider insufficiently de-identified or where the informed consent obtained from participants rules out data publication and sharing [e.g. 5]. These decisions are discussed in detail with depositors, and disagreement between curators and depositors is exceedingly rare.

In complex cases, QDR is able to draw, in addition to in-house expertise, on the expertise of several members of its Research Advisory Board, several of whom have relevant experience [6]. QDR is also an active participant in the ongoing FORCE11/COPE working group on data publication ethics [7], which seeks to align practices across data repositories.

During an initial consultation with depositors, QDR staff also helps the depositor to assess the sensitivity of data and potential disclosure risks and aids the depositor in identifying appropriate levels of access controls ranging from access for all registered users to access only on-site. Details of available access conditions are described in the documentation of access conditions [8].

For sensitive data, QDR follows strict protocols for transmitting, handling, and storing data. Depositors are instructed to encrypt data using AES-256 encryption prior to transfer, using SFTP or a Dropbox business folder with multi-factor authentication enforced for all users able to access content. Sensitive data are stored using AES 256 encryption. Where the depositor requests additional safeguards for sensitive data, we help them to decide which access conditions should be imposed so that the data can be downloaded. The data are then distributed under a Special Download

Agreement reflecting those access conditions. The conditions specified in the agreement reflect the nature of the disclosure risk in the data and can contain, for example, requirements for IRB/ethics board approval and/or a data security plan.

Sensitive data requires responsible use. QDR ensures that data is only released to personally identified individuals: access to data is granted following authentication via institutional e-mail and videoconferencing.

Additional requirements for data use are specified in the special download agreement and follow both depositor requests and QDR's assessment of the identifiability and risk for human participants of the data in question. The general requirements, by level of sensitivity, are outlined in QDR's "Handling Sensitive Data" policy [9] under "Access to Restricted Data". For low sensitivity restricted data, authentication, a research plan, and assurances to not distribute the data further are typically sufficient for access. For medium sensitivity data, QDR requires a detailed data security plan as well as IRB approval for the proposed research and, in addition to the depositor's signature, the signature of an authorized institutional representative on the special download agreement. For highly sensitive data, QDR only allows access in person in a monitored room and screens users' notes. (While QDR does have the capacity to provide such access, it does not currently hold any data it classifies as highly sensitive). QDR is continuously exploring additional means of certifying researchers for access to sensitive data and thus facilitating access. We are participating as a pilot institution in ICPSR's "Research Passport" initiative [10] (see the working paper linked below) that will leverage cross-repository collaboration to certify researchers in handling sensitive data.

QDR generally advises depositors against access conditions requiring their approval. Where depositor approval for data access is required, QDR implements an access policy that allows QDR to determine whether access is granted after a set period of time without response from the depositor and based on previously agreed criteria [11].

Through its Terms and Conditions as well as its Standard Download Agreement [11], QDR also requires that researchers agree to use data ethically for data not deemed sensitive. As outlined, this includes giving attribution when using the data, not re-publishing it without explicit consent, and not using it for commercial purposes.

Given these precautions, we expect any misuse QDR's data to be rare. Should it occur, QDR's Download Agreements stipulate a range of sanctions for violation of the agreements, including deletion of user accounts, contacting the QDR institutional representative at the user's home institution (if that institution is a member of QDR) and the IRB at the user's home institution, and in cases endangering human participants, reporting to the federal Office of Human Research Protection.

QDR limits access of QDR staff to sensitive data. All access is overseen by senior staff, who have trained (and published) on the handling of sensitive qualitative data and regularly attend international workshops and conferences in data science and management to remain informed of state-of-the-art practices and technology. We also frequently consult QDR's technical advisory board, which includes several experts on responsible data practices, for guidance.

As of 2021, all new QDR staff working with data deposits, including graduate assistants, receive dedicated security training, use multi-factor authentication (via Yubi keys) to access any relevant accounts, and sign an attestation on the responsible handling of confidential information.

Links:

- 1. QDR Curation Policy : https://qdr.syr.edu/policies/curation
- 2. QDR Standard Deposit Agreement (requires registration, attached): https://qdr.syr.edu/deposit/standarddeposit

3. QDR IRB guidance: https://qdr.syr.edu/guidance/human-participants/irb

4. QDR De-identification guidance: https://qdr.syr.edu/guidance/human-participants/deidentification

5. Blog describing example of rejected data:

https://qdr.syr.edu/qdr-blog/participant-protection-informed-consent-and-data-sharing

6. QDR Research Advisory Board: https://qdr.syr.edu/about/governance

7. FORCE11 Research Data Publishing Ethics Group: https://force11.org/groups/research-data-publishing-ethics/

8. QDR Access controls and conditions: https://qdr.syr.edu/guidance/human-participants/access-controls

9. QDR Handling Sensitive Data Policy: https://qdr.syr.edu/policies/sensitivedata

10. ICPSR Researcher Passport: https://radius.icpsr.umich.edu/radius/passport

11. Depositor-approved access:

https://qdr.syr.edu/guidance/human-participants/access-controls#depositor-approved-access

12. QDR Standard Download Agreement: https://qdr.syr.edu/discover/standarddownload

Reviewer Entry

Reviewer 1

Comments: Review Version 2: This is now addressed

Review Version 1: Question for QDR:

What does QDR do in cases where the data is not sufficiently de-identified? I note from your policy that while QDR has only an advisory role, will you publish data that you believe is not adequately de-identified or does not address sufficiently disclosure risks? Or would you assign a higher level of sensitivity and limit access?

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

Agreed

Reviewer 2

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

Response:

QDR is housed in the Center for Qualitative and Multi-Method Inquiry (CQMI), a unit of the Maxwell School of Citizenship and Public Affairs, a nationally leading public policy school at Syracuse University. CQMI is also the home of the Consortium for Qualitative Research Methods (CQRM), which conducts an annual international Institute for Qualitative and Multi-methods Research with around 180 participants [1].

The repository is led by social scientists at Syracuse University and Georgetown University as well as information scientists at the University of Washington at Seattle [2]:

Director:

Colin Elman, Professor of Political Science, Syracuse University.

Technical Directors:

Carole Palmer, Professor and Associate Dean for Research, Information School, University of Washington, Seattle Nic Weber, Assistant Professor, Information School, University of Washington, Seattle

Director of Research:

Diana Kapiszewski, Associate Professor of Political Science, Georgetown University

Associate Director: Sebastian Karcher, Research Assistant Professor of Political Science, Syracuse University QDR's Associate Director and Curation Specialist assist users and curate deposits with the support of two to three graduate student assistants as well as part-time support from CQMI personnel.

QDR has a small team of developers (one frontend/database, one systems/dev-ops), building on a lightly customized version of the Dataverse open source development software to reduce development costs.

QDR staff regularly attend professional meetings to present their work and benefit from an international community of data specialists. Among conferences attended in the past three years are, the annual meeting of the International Association of Social Science Information Services and Technology (IASSIST), Research Data Alliance (RDA), Dataverse Community Meetings, Research Data Access and Preservation Summit (RDAP), Open Repositories, Force 11, and the Joint Conference on Digital Libraries. QDR and/or its personnel are members and actively participate in international bodies including RDA, RDAP, and IASSIST. QDR's director serves on the Center for Open Science's Transparency and Openness Promotion (TOP) Guidelines Coordinating Committee. QDR's Associate Director serves on the Technical Steering group of DataCite and is a member at large (US) on the IASSIST administrative council.

Funding for the repository currently comes from various sources:

Grant funding for core operations provided by the National Science Foundation (Political Science Program) as well as project-based funding by the Sloan Foundation, the Institute for Museum and Library Services, and internal grants at Syracuse University. Past funders also included the Robert Wood Johnson Foundation. Current grand funding guarantees operations through 2025.

In-kind support by Syracuse University (office space, IT support, administrative support, graduate assistants) Revenues from institutional membership [3]. QDR currently has 28 institutional members [4]. During the first round of renewals for institutional members, 92% of membership institutions renewed their membership. We expect membership to grow and make up a larger proportion of total revenue going forward.

Storage and curation fees for individual deposits. Due to its grant funding, QDR is able to generously award waivers for researchers without access to funding for curation and storage costs and has not refused a deposit for lack of funds. Nevertheless, about 25% of new deposits include payment for storage and curation. As QDR is more frequently written into grant applications, we expect the number, and likely the share of funded deposits to increase.

QDR has a robust, and increasing, user base. As of May 2022, QDR has over 3,700 registered users and the site receives an average of about 2,000 monthly unique users according to google analytics. Most users (ca. 50%) are from the United States, but include researchers from across the globe, spanning five continents.

Links:

- 1. QDR Governance: https://qdr.syr.edu/about/governance
- 2. QDR Contact and personnel: https://qdr.syr.edu/contact
- 3. Institutional Membership: https://qdr.syr.edu/membership/join
- 4. List of institutional members: https://qdr.syr.edu/content/list-qdr-institutional-members

Reviewer Entry

Reviewer 1

Comments: Would be useful to update the usage information which is currently from 2018: Version 2: Agreed, this is now addressed

"As of October 2018, QDR has over 3,100 registered users and the site receives an average of about 1,700 monthly unique users according to google analytics. Most users (ca. 70%) and visitors (ca 60%) are from the United States, but both groups include researchers from across the globe, spanning five continents."

Reviewer 2

Comments: Accept

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 Agreed

Reviewer 2

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

Response:

From its inception, QDR has sought out and received advice from information and data science experts and its designated community of social scientists and other researchers. QDR has two advisory boards [1]:

The Technical Advisory Board, whose members are specialists from libraries and data repositories, advises QDR on technical questions including development, curation, and digital preservation.

The Research Advisory Board, whose members are leading social scientists as well as qualitative researchers in the health and education sciences, assures that QDR serves the interests of its main constituency, practicing qualitative researchers.

Each of these boards meets twice a year via teleconference. QDR also reports to both boards at the end of each quarter and solicits advice and feedback. For specific questions or major decisions, QDR seeks out the advice of board members individually.

In addition, QDR regularly presents its work at international conferences and workshops (see R5). To gain additional insights, QDR also organizes workshops convening diverse groups of international experts to focus on key concerns of archiving and curating qualitative social science data. Pre-COVID workshops were on Annotation for Transparent Inquiry (ATI), an approach to transparency in qualitative research [2]. Jointly with its Data-PASS partners, QDR also frequently organizes workshops for journal editors on topics related to depositing data supplementing journal articles [3]. QDR also frequently consults with Institutional Review Boards (IRBs) and their members on the ethics of data sharing. In collaboration with PRIM&R, the principal organization for IRB professionals, QDR has established a dedicated "hashtag" for questions related to data sharing arising in the review of IRB applications. [4]

Finally, QDR personnel regularly teach data management to researchers, both in seminars and in individual consultations, see R0/Designated community for details. These conversations serve as a constant check that the services offered address the needs and concerns of the repository's designated user community.

Links:

1. QDR Advisory boards: https://qdr.syr.edu/about/governance

2. ATI workshops: https://qdr.syr.edu/qdr-blog/annotation-transparent-inquiry-ati

3. Data-PASS "JEDI" workshops: https://dpjedi.org/events.html

4. Blogpost on the "#SBERdatasharing" on the PRIM&R forums:

https://qdr.syr.edu/qdr-blog/primr-and-qdr-help-irb-professionals-discuss-data-sharing-social-behavioral-and-economic

Reviewer Entry

Reviewer 1 Comments: No comment

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 Agreed

Reviewer 2

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

Response:

QDR follows the OAIS reference model in handling data [1]. On receipt, QDR personnel check data files and metadata for completeness and integrity and, as needed, solicit updated or additional files from depositors. Curators keep an unchanged version of the original deposit data and metadata (submission information package, SIP) throughout the curation process [2].

While there is no formalized identity check in place, QDR's curation team typically communicates directly by phone/Zoom with depositors and encourages the use of institutional emails for registration and communication.

On ingest, the Dataverse software automatically creates an SHA-256 checksum for every ingested file that allows for checking file integrity manually including by users and third parties. Files are stored on AWS S3 [3], where redundant copies of each file are stored on distributed servers and integrity checks at rest are performed using content-MD5 checksums and cyclic redundancy checks (CRCs). AWS also performs integrity checks during data transfer. QDR uses file versioning in its S3 buckets to provide additional redundancy and protection against human and software errors [4]. On publication, a preservation package using an RDA recommended BagIt format with OAI-ORE metadata [5] is automatically stored in Google "coldline" archival storage [6], which provides 99.999999999% durability over a given year. The preservation format has been tested to allow round-trip restoration of datasets between Dataverse instances, and proof-of-concept transfer to other RDA-conformant repositories (e.g. Clowder) has been explored [7].

The Dataverse software automatically enforces version updates on data for every change of published data using a two-digit versioning system (e.g., 2.1). Smaller changes to the metadata are recorded as minor changes, such as 2.1 to 2.2. Updates of data or other major changes receive a new version number (e.g., from 2.1 to 3.0).

Links:

1. QDR Digital preservation policy: https://qdr.syr.edu/policies/digitalpreservation

2. QDR Curation policy: https://qdr.syr.edu/policies/curation

3. AWS S3 storage: https://aws.amazon.com/s3/faqs/

4. Poster on QDR's Preservation Storage workflow: https://doi.org/10.6084/m9.figshare.8304146.v1

5. Poster presented to RDA on interoperability of the OAI-ORE format:

https://www.rd-alliance.org/interoperable-data-archiving-and-migration-using-rdri-working-group-recommendations

6. Google coldline storage: https://cloud.google.com/storage/archival

7. RDA Recommendation for Interoperable Data Archiving:

https://www.rd-alliance.org/interoperable-data-archiving-and-migration-using-rdri-working-group-recommendations

Reviewer Entry

Reviewer 1

Comments: No comment

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 Agreed

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

QDR prioritizes the acquisition and curation of deposits according to our collection development and appraisal policy [1]. Under that policy, and following its mission, QDR prioritizes "qualitative data, or data associated with mixed-method research with a strong qualitative component, that are generated and/or used in the social sciences or cognate disciplines" and/or that hold "great intellectual value and/or that are of high quality." Where data are clearly outside of our collection policy (e.g., are purely quantitative or life-science oriented), we point depositors to suitable alternative archives based on their specifications.

All data are reviewed by QDR's Associate Director or Curation Specialist. QDR largely follows the recommendations of the UK Data Archive [2], augmented with information from the Library of Congress [3], for preparing gualitative data for archiving and has also developed its own recommendations for data preparation and preferred formats [4]. QDR only requires depositors to complete a small number of metadata fields that provide basic bibliographic description of the data [5]. However, QDR works closely with depositors to encourage and help them to provide in-depth documentation about the collection or generation and context of the data. Where metadata initially provided by depositors are too sparse to allow secondary users to make sense of the data (e.g., where no data collection methodology is described), QDR works with depositors to improve documentation. In line with the collection development policy, data that are found too lacking in documentation to be useful are not published. QDR curation staff, in collaboration with depositors, will convert detailed documentation into structured metadata. QDR's metadata application profile closely follows (a subset of) Data Documentation Initiative (DDI) Codebook, the de-facto standard for social science metadata. To the extent possible, metadata categories are linked to more generic vocabularies, specifically Dublin Core and the DataCite Metadata Kernel. QDR staff converts files that are sub-optimal formats into preferred file formats during curation where possible [6]. Where no suitable format for archiving exists, QDR archives files as they are and commits to bit-level preservation. QDR staff proofreads and systematizes documentation provided by depositors to generate rich, standardized metadata (see R9 for more on file formats, conversion, and metadata).

Links:

- 1. QDR Collection development and appraisal policy: https://qdr.syr.edu/policies/collectiondevelopment
- 2. UK Data file format recommendations: https://www.ukdataservice.ac.uk/manage-data/format/recommended-formats
- 3. Library of Congress Digital Preservation Formats: https://www.loc.gov/preservation/digital/formats/intro/intro.shtml
- 4. QDR Data preparation guidance: https://qdr.syr.edu/guidance/managing/preparing-data
- 5. QDR Metadata application profile: https://qdr.syr.edu/policies/metadata
- 6. QDR Recommended file formats: https://qdr.syr.edu/guidance/managing/formatting-data

Reviewer Entry

Reviewer 1

Comments: No comment

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 Agreed

Reviewer 2

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

Response:

QDR's data storage procedures are documented in its preservation and curation policies [1] and follow the OAIS reference model. The main storage facilities of the repository are on AWS S3, which itself has significant protections against data loss such as redundant file storage across multiple data centers. In addition, QDR maintains secondary-site back-ups at Syracuse University, as well as Google cloud "coldline" storage for preservation copies of files. Both AWS and Google perform regular file-integrity checks to guard against the failure of storage media. Full system back-ups are performed on AWS S3 daily and can be used for quick recovery in typical scenarios, with back-ups at Syracuse and Google allowing recovery of data following a catastrophic event [2].

QDR's preservation policy [3] is based on recommendations from the Library of Congress as well as other data repositories with significant holdings of qualitative data such as UK Data and DANS. Following receipt of a data deposit, files are converted to recommended storage formats and ingested into the Dataverse repository system. The file formats and types and file migration follow industry standards and recommendations. All changes are recorded in a readme file accompanying the data. All used file formats are monitored for obsolescence using the Library of Congress's Sustainability of File Formats pages [4]. Files in formats threatened by obsolescence are converted to suitable replacement formats.

Most files currently archived with QDR are not sensitive and do not require special security provisions. Sensitive materials are stored using AES 256 encryption on both AWS and local servers [5]. All access to server software is controlled using virtual private networks.

Links:

- 1. Curation policy: https://qdr.syr.edu/policies/curation
- 2. QDR Security and infrastructure: https://qdr.syr.edu/policies/security
- 3. Digital preservation policy: https://qdr.syr.edu/policies/digitalpreservation
- 4. Library of Congress's Sustainability of File Formats:
- https://www.loc.gov/preservation/digital/formats/fdd/descriptions.shtml
- 5. Sensitive data: https://qdr.syr.edu/policies/sensitivedata

Reviewer Entry

Reviewer 1

Comments: No comment

Reviewer 2

Comments: Accept

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

Reviewer Entry

Reviewer 1

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

Reviewer 2

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

Response:

QDR's preservation policy [1] describes the full preservation framework following the structure of OCLC's "Trusted Digital Repository" framework. As outlined in the policy, preservation of all files is guaranteed for a minimum of 20 years during which all efforts will be made to ensure permanent access to files. QDR assures access to files and content by using a file-format migration strategy as described in R9 and is committed to bit-level preservation where suitable preservation

formats are not available.

The obligations of repository and depositor are clearly laid out in the Standard/Special Deposit agreement, at least one of which is signed by every depositor prior to the publication of data projects, marking the transfer of custody. The deposit agreement explicitly permits QDR to transform, duplicate, and disseminate the data (in the form of a non-exclusive license).

QDR's preservation actions are specified in both preservation and curation policy [2]. An extended version of the curation policy servers as QDR's "Curation Handbook," [3] which has also recently been published as part of a peer reviewed publication. QDR describes best practices for preparing data deposits in a dedicated guidance page on the QDR website [4], and also works with depositors whose initial deposit does not meet our internal standards.

Links:

- 1. Digital preservation policy: https://qdr.syr.edu/policies/digitalpreservation
- 2. Curation policy: https://qdr.syr.edu/policies/curation
- 3. QDR Curation Handbook (article in the Journal of eScience Librarianship) : https://doi.org/10.7191/jeslib.2021.1207
- 4. Data preparation guidance: https://qdr.syr.edu/guidance/preparing-data

Reviewer Entry

Reviewer 1

Comments: No comment

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

Reviewer Entry

Reviewer 1

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

Reviewer 2

Comments:

4 - The guideline has been fully implemented in the repository

Response:

The quality of shared data depends upon their understandability and re-usability. These qualities, in turn, depend upon the organization of the data, and the clarity and completeness of the documentation that accompanies them (i.e., how well they describe the data, the process through which they were collected/generated, and the context of their creation) [1]. QDR encourages depositors to provide all relevant information that allows for well-informed re-use of the data and works closely with them to help them provide the highest possible level of data and documentation quality. This process relies on the subject expertise of QDR's curation staff. The conceptual framework of our curation approach – enabling epistemically responsible curation – has recently been published as a peer-reviewed article in The Qualitative Report [2]. Curation staff also assess the consistency of the data with the provided documentation and request changes, fixes, or updates from depositors as needed [3]. Curation is supervised by senior staff all of whom hold graduate degrees in social science. Metadata are generated in consultation with depositors using the Dataverse input mask, which maps (and exports) to Data Documentation Initiative (DDI) Codebook, the de-facto meta standard for social science, as well as other metadata formats such as DataCite XML and can be harvested using OAI-PMH. The actual cataloging of the metadata is performed by QDR curation staff based on depositor input and is subject to review by the depositor [4].

As part of the curation process, QDR also links to published work that uses or cites the data. QDR was one of the first repositories to fully implement Make Data Count usage data [5] following standardized Project Counter metrics. Usage metrics are displayed on every item landing page [6].

There is no formalized way for the designated community to comment or rate data or metadata. Nonetheless, QDR regularly works with scholars who re-use data in teaching and research in order to better understand their requirements and, if needed, adjust cataloging and curation practices.

Links:

1. Collection development and appraisal policy: https://qdr.syr.edu/policies/collectiondevelopment

2. How Data Curation Enables Epistemically Responsible Reuse of Qualitative Data. The Qualitative Report: https://doi.org/10.46743/2160-3715/2021.5012

3. Curation policy: https://qdr.syr.edu/policies/curation

4. Metadata application profile: https://qdr.syr.edu/policies/metadata

5. Make Data Count: https://makedatacount.org/

6. Sample landing page with Make Data Count Metadata: https://doi.org/10.5064/F6VOIB8H

Reviewer Entry

Reviewer 1

Comments: No comment

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 Agreed

Reviewer 2

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

Response:

QDR's workflows in handling, storing, and preserving data and keeping it secure are described in the following documents:

- Preservation policy [1] (describes conformance to OCLC's trusted digital repository and the OAIS reference model)

- Curation policy [2] (supplements the preservation policy with a specific focus on QDR's activity to increase data and metadata quality and assure ethical sharing of data), detailed in more detail in our Curation Handbook [3]

- Appraisal and Collection Development [4] (describes QDR's criteria for accepting data)
- Sensitive data [5] (describes the handling of different levels of sensitive data)
- Security [6] (describes back-ups and security provisions)

- Standard [7] /Special [8] deposit agreements (formal agreement outlining depositor and repository rights and obligations at a high level of abstraction)

R8 describes QDR's appraisal procedures. Where data are found to not fit QDR's mission or the repository is otherwise unable to accept them, curators will actively assist the relevant researcher(s) in finding an alternative location for the data. Together with QDR's mission statement, the Appraisal policy specifies the types of data stored by QDR, i.e. data generated through and/or used in qualitative and multi-method research The diversity of such data complicates automated checking and analysis, which is why QDR relies heavily on its expert curation staff throughout the data lifecycle. As described in R9, QDR describes its handling of data to depositors in an agreement that they sign (Standard deposit agreement) and provides additional details in its curation policy. The handling of confidential data is described above in R4 and in the "sensitive data" policy. When depositors wish to place restrictions on access to their data, these are specified individually in coordination with the depositor and codified in a set of special deposit/download agreements. Transformation of data for archiving is described in the preservation and curation policies and in R9 above. Security, audit,

and backup procedures are outlined in the Security document and R16.

Links:

- 1. Digital preservation policy: https://qdr.syr.edu/policies/digitalpreservation
- 2. Curation policy: https://qdr.syr.edu/policies/curation
- 3. Curation Handbook: https://doi.org/10.5281/zenodo.4672677
- 4. Appraisal and collection development policy: https://qdr.syr.edu/policies/collectiondevelopment
- 5. Sensitive data: https://qdr.syr.edu/policies/sensitivedata
- 6. Security and infrastructure: https://qdr.syr.edu/policies/security
- 7. Standard deposit agreement (requires registration): https://qdr.syr.edu/deposit/standarddeposit
- 8. Special deposit agreement (requires registration): https://qdr.syr.edu/deposit/specialdeposit

Reviewer Entry Reviewer 1

Comments: No comment

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

Reviewer Entry

Reviewer 1

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

Reviewer 2

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

Response:

Making data findable, accessible, interoperable, and reusable (FAIR, Wilkinson 2016 [1]) is a core mission of a data repository and QDR constantly seeks to improve the discoverability of its holdings. The Dataverse catalog used by QDR offers search, including powerful advanced search options as well as faceted browsing. Crucially for the findability of qualitative data, QDR allows for full-text search in textual data.

QDR's Dataverse catalog also provides harvesting facilities via OAI-PMH as well as a dedicated API, allowing machine-readable access to metadata. Currently, QDR metadata are harvested by Harvard Dataverse [2] as part of the Data-PASS catalog. QDR also optimizes its metadata for the Datacite metadata kernel, which makes it accessible via the DataCite Metadata store [3]. Using JSON-LD/schema.org metadata embedded on item pages, QDR data are also findable through the newly released google dataset search [4].

QDR provides standardized citations as well as bibliographic metadata for reference managers, including Dublin Core and JSON-LD/schema.org metadata embedded on the page, for every project and every file. Every data project and every file within every project are registered with a DOI with DataCite.

QDR is widely recommended by libraries and journals, including PLOS, Nature Scientific Data, eLife, etc [5]. QDR seeks to advance the scholarly web by using and promoting persistent identifiers such as DOIs and ORCIDs as much as possible. We strongly encourage depositors to use their ORCID IDs as part of their records and we link to related data and publications using DOIs whenever they are available. We expect to add support for Research Organisation Registry (ROR) identifiers during Q3/Q4 2022, with support in the Dataverse arriving shortly [6], and are closely following

Links:

1. Wilkinson 2016: https://doi.org/10.1038/sdata.2016.18

2. QDR data findable through the Harvard Dataverse: https://dataverse.harvard.edu/dataverse/qdr

3. QDR data Datacite Search: https://search.datacite.org/repositories/gdcc.syr-qdr

4. Google dataset search: https://datasetsearch.research.google.com/search?query=site%3Adata.qdr.syr.edu&docid=L2c vMTFuMGprbDIndA%3D%3D

5. QDR entry on fairsharing.org, documenting inclusion on various lists of recommended repositories:

https://doi.org/10.25504/FAIRsharing.bmz5ap

efforts to promote PIDs for grants and DMPs.

6. ROR ID support in Dataverse: https://github.com/IQSS/dataverse/issues/6640

Reviewer Entry

Reviewer 1

Comments: No comment

Reviewer 2

Comments: Accept

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 Agreed

Reviewer 2

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

Response:

QDR encourages the reuse of data in its repository by displaying it with rich context and by actively promoting it via social media and other channels. If the archiving of qualitative data is a very new endeavor for social scientists who engage in qualitative research, re-using qualitative data collected or generated by other scholars is even more unfamiliar. As such, QDR is developing a research agenda on the reuse of qualitative data and continuously seeks to adapt its practices to facilitate reuse. Recent highlights include a symposium discussing various approaches to re-using one of QDR's datasets in TQR [1] and a forthcoming article discussing the use of QDR projects that include "Annotation for Transparent Inquiry" in teaching [2]. Our standardized Make Data Count metrics show over 50,000 unique file downloads since October 2019 (when MDC counts began) [3].

Documentation and metadata are crucial prerequisites for the reuse of data by third parties. For metadata, QDR only strictly enforces minimal, Dublin Core requirements on data (title, author, description, subject, deposit date) [4]. However, as part of the curation process, QDR typically develops significantly richer metadata in collaboration with the depositors. QDR's metadata profile is based on the Data Documentation Initiative (DDI) version 2.5 (Codebook), in line with most other social science data repositories. The repository is actively monitoring developments of the DDI standard that would provide better support for qualitative data, but the current "Lifecycle" (3.3) version of the standard holds few advantages for qualitative data. In particular, the use of DDI 2.5 is in line with other repositories with significant qualitative data holdings such as the UK Data Archive. As an XML format, DDI can be converted to updated forms of the standard using XSLT. QDR does not use any additional metadata standards for qualitative data (none of which exist in wide-enough circulation to be useful for findability and re-use). The repository's principal approach to documenting details specific to qualitative studies is to encourage (and, if not provided, specifically request) the inclusion of ample narrative details in the study description.

DDI output is currently automatically generated by the Dataverse software QDR uses. There is significant interest among the Dataverse user and development community to further improve DDI support, so that further developments of the DDI are likely to be incorporated into Dataverse.

QDR ensures the understandability of all deposited data through intensive, manual curation by its subject experts. QDR curators read all documentation and regularly request changes or additions to improve understandability. They also work with depositors on structuring their deposit and naming data files to maximize the ability of others to understand and ultimately re-use the data. The approach to curation is documented in the curation policy [5].

The licenses used by QDR allow for re-use of all data in research and teaching, but generally disallow the re-publication of data elsewhere, i.e. data are not under open licenses. The license terms are specified in QDR's Standard [6] /special [7] download agreements. QDR uses a range of automated monitoring (e.g. google alerts and google scholar alerts) to detect any unauthorized republication of data, but as of today that has been no known case. Should such violations be detected, QDR's primary interest would the takedown of the materials via takedown requests and, if necessary, the threat of legal action. In extreme cases, QDR's terms and agreements allow sanctions against individual researchers as described under R4.

These less permissive licenses are chosen due to the complex nature of some qualitative data, e.g., those under copyright, which limits their sharing, and those gathered from human participants, which can only be shared in a way such that research participants remain protected. QDR's practices are based on the practices and recommendations of comparable repositories such as the UK Data Archive. The repository will consider publishing data under open CC-BY-SA (Creative Commons Attribution-Share-Alike) or CC0 (Public domain waiver) licenses. All documentation is published under a CC-BY-SA license.

Links:

- 1. Introduction to TQR symposium on reuse: https://doi.org/10.46743/2160-3715/2021.5015
- 2. Teaching with ATI article: https://doi.org/10.33774/apsa-2021-1fnv8-v3
- 3. Make Data Count counter at the top of data repository landing page: https://data.qdr.syr.edu
- 4. QDR Metadata application profile: https://qdr.syr.edu/policies/metadata
- 5. QDR Curation policy: https://qdr.syr.edu/policies/curation
- 6. Standard deposit agreement (requires registration, attached): https://qdr.syr.edu/deposit/standarddeposit
- 7. Special deposit agreement (requires registration, attached): https://qdr.syr.edu/deposit/specialdeposit

Reviewer Entry

Reviewer 1

Comments: Version 2: Agreed - this is now addressed

Version 1: It would be useful to highlight any metadata standards or practices that are specific to qualitative data (noting DDI is oriented towards quantitative collections).

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 Agreed

Reviewer 2

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

Response:

QDR follows ISO 14721:2012, section 4.1.1.1 (common services) [1] as a reference model for technical infrastructure development. The technical directors for QDR, in consultation with a Technical Advisory Board, monitor the implementation of services and review emerging standards for qualitative data management on a biannual basis [2]. Infrastructure development activities follow a semi-annual roadmap produced by QDR's technical directors, and approved by the QDR's Technical Advisory Board.

Hardware and software inventories and configuration information are recorded in a QDR managed wiki, and updated quarterly. All software running the production environment of QDR is open-source - this includes operating systems running on EC2 and S3 servers (Linux), a content management system based on Drupal, a repository framework based on Dataverse [3], as well as a suite of configuration management (Chef) [4], continuous integration (Jenkins) and monitoring tools (Nagios) [5]. To further ensure continuous delivery of deployed code, our team also relies upon open-source tools to perform automated tests (Selenium) [6] as well as infrastructure execution and management tool (Terraform) [7]. Each of these tools is well-supported by open-source communities. Our technical directors and system administrator regularly monitor security and vulnerabilities related to this suite of software.

The hardware used to run QDR is provisioned at Amazon Web Services, and managed by our technical development team. Our infrastructure at AWS is configured with a set of Virtual Private Clouds for security (described in detail in R16),

and we ensure proper bandwidth is available by using Elastic Load Balancing which distributes incoming user traffic across multiple EC2 instances [8].

Links:

- 1. ISO 14721: https://www.iso.org/standard/57284.html
- 2. Security and infrastructure: https://qdr.syr.edu/policies/security
- 3. Dataverse: http://dataverse.org
- 4. Chef: https://www.chef.io/solutions/infrastructure-automation/
- 5. Nagios: https://www.nagios.org/
- 6. Selenium: http://www.seleniumhq.org/
- 7. Terraform: https://www.terraform.io/
- 8. Digital preservation policy: https://qdr.syr.edu/policies/digitalpreservation

Reviewer Entry

Reviewer 1

Comments: No comment

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 Agreed

Reviewer 2

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

Response:

Security and risk management are carried out by QDR's technical team, in collaboration with the Syracuse Maxwell School IT department, and a contract with a cloud infrastructure provider AWS. Dedicated instances are purchased from AWS and refreshed every two years [1]. Technical infrastructure is physically located in US-EAST (Ohio) but can be moved relatively quickly through QDR's use of the infrastructure management tool Terraform (as described in R15). QDR created a virtual private cloud (VPC) for different applications deployed to AWS. The VPC is achieved through private IP subnets, as well as a virtual private network (VPN) that secures access to the VPC (this is achieved through authentication).

QDR tracks security announcements from the Dataverse and Drupal communities and applies patches and updates as needed to address identified issues. The team also tracks automated security alerts generated by https://github.org/ [2] related to vulnerabilities found in open-source components used within these applications and will deploy updates in advance of community alerts when warranted. QDR regularly updates operating systems, databases, programming libraries and other components to operate using versions with long-term support and security coverage. Through Syracuse, QDR monitors its infrastructure via penetration testing services (i.e. U.T. Austin's Dorkbot [3]). As described in R9 and R12, QDR creates redundant storage (located both at Syracuse and in the cloud with AWS) that prevents data loss and limits the impact of service outages in the case of a natural disaster. End-user access to data requires registration and agreement to QDR's General Terms and Conditions of Use [4] (described in R2).

Links:

1. Security and infrastructure: https://qdr.syr.edu/policies/security

2. GitHub "Dependabot" security warnings: https://docs.github.com/en/code-security/supply-chain-security/managing-vuln erabilities-in-your-projects-dependencies/configuring-dependabot-security-updates

3. UT Austin "Dorkbot": https://security.utexas.edu/dorkbot

4. QDR Terms and conditions: https://qdr.syr.edu/termsandconditions

Reviewer Entry

Reviewer 1 Comments:

No comment

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:

Reviewer Entry Reviewer 1 Comments: Reviewer 2 Comments: