Implementation of the Data Seal of Approval

The Data Seal of Approval board hereby confirms that the Trusted Digital repository Oxford Research Archive for Data (ORA-Data) complies with the guidelines version 2014-2017 set by the Data Seal of Approval Board.
The afore-mentioned repository has therefore acquired the Data Seal of Approval of 2013 on November 23, 2015.

The Trusted Digital repository is allowed to place an image of the Data Seal of Approval logo corresponding to the guidelines version date on their website. This image must link to this file which is hosted on the Data Seal of Approval website.

Yours sincerely,

The Data Seal of Approval Board

Assessment Information

Guidelines Version: 2014-2017 | July 19, 2013
Guidelines Information Booklet: DSA-booklet_2014-2017.pdf
All Guidelines Documentation: Documentation

Repository: Oxford Research Archive for Data (ORA-Data)
Seal Acquiry Date: Nov. 23, 2015

For the latest version of the awarded DSA   http://assessment.datasealofapproval.org/seals/
for this repository please visit our website:

Previously Acquired Seals: None

This repository is owned by: **Bodleian Digital Library (BDLSS)**

United Kingdom

T 01865280040
E ora@bodleian.ox.ac.uk
W http://www.bodleian.ox.ac.uk/bdlss

Assessment

0. Repository Context

Applicant Entry

Self-assessment statement:

The Oxford Research Archive for Data (ORA-Data) is a multidisciplinary archival store for digital data produced as a result of research by Oxford University academics. It is provided by Bodleian Digital Library Systems and Services (http://www.bodleian.ox.ac.uk/bdlss accessed 09/06/2015) on behalf of the University of Oxford, and forms part of the wider Oxford University Research Archive (http://www.bodleian.ox.ac.uk/ora accessed 09/06/2015) which provides a permanent and secure archive of the University's scholarly research output. ORA-Data complements other data archives by providing a local archive for researchers who are not able or do not want to deposit their data elsewhere, though it is not intended to replace national, subject, or other established data collections. The service can be used to store data that underpins scholarly publications, so that the data can be cited and accessed (if applicable). ORA-Data was launched as a pilot service on 1 December 2014 and has subsequently continued to be developed, particularly so as to comply with the recommendations of UK research councils and other funding bodies.

ORA-Data is administered from within the University of Oxford and has no outsource partners. As such, it fully complies with the University's policy on the management of research data and records (http://researchdata.ox.ac.uk/university-of-oxford-policy-on-the-management-of-research-data-and-records/ accessed 09/06/2015).

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

1. The data producer deposits the data in a data repository with sufficient information for others to assess the quality of the data, and compliance with disciplinary and ethical norms.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Those depositing in ORA-Data are required to be members of the University of Oxford (access is restricted to those with an Oxford University Single-Sign-On), or are required to have been members of the University of Oxford at the point of creating their research outputs and research data (in which case retrospective deposit is facilitated by ORA-Data staff). As such, they are required to comply with the legal and ethical criteria contained in the University's Policy on the Management of Research Data and Records:

http://www.admin.ox.ac.uk/media/global/wwwadminoxacuk/localsites/researchdatamanagement/documents/Policy_on_the_Management (accessed 25/06/2015)

Basic advice for data producers is provided at http://www.bodleian.ox.ac.uk/bdlss/digital-services/data-archiving (accessed 09/06/2015) and more detailed guidance is provided from the ORA-Data LibGuide at http://ox.libguides.com/ora-data (accessed 09/06/2015).

The ORA-Data Deposit Conditions are available from the 'About ORA-Data' section of the LibGuide (http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0019/190009/ORA-Data-Deposit-Agreement.pdf accessed 24/06/2015) along with ORA-Data policies relating to submission, retention, preservation, withdrawal, complaints and take-down (http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0012/190011/ORA-Data-Submission-Preservation-and-Withdrawal-Policies.pdf accessed 24/06/2015).

Datasets and metadata are deposited using a clearly presented 6-step deposit form, with a combination of mandatory and optional fields (selected entries from both being available to the data consumer) and clear and concise pop-up guidance for each; specific guidance is provided in the ORA-Data LibGuide (http://ox.libguides.com/how-to-deposit and http://ox.libguides.com/ora-data-checklist both accessed 09/06/2015). Step 1 requires the depositor to agree to the ORA-Data Deposit Conditions referred to above. There is the option of citing related publications (i.e. those based on the data) where applicable. All deposited datasets and metadata are subject to review by ORA-Data staff whose role is also to guide depositors through the process and answer

questions as required; where mandatory fields (marked by asterisk) are not completed or contextual information (particularly with regard to a mandatory 'Documentation about your dataset and processes' field) is deemed to be insufficient, the data producer will be contacted and asked to supply further information.

Meanwhile, further information is presented on the University's Research Data Oxford website (http://researchdata.ox.ac.uk/ accessed 09/06/2015), which is co-managed by the University's IT Services and the Bodleian Libraries; as well as general information regarding research data, this site provides an overview and entry point for ORA-Data (http://researchdata.ox.ac.uk/preserving-your-data/ora-data/ accessed 09/06/2015) and specific information relating to the expectations of UK research councils and other funding bodies (http://researchdata.ox.ac.uk/funder-requirements/ accessed 09/06/2015).

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

2. The data producer provides the data in formats recommended by the data repository.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

3. In progress: We are in the implementation phase.

Self-assessment statement:

ORA-Data currently accepts any file type for preservation, but its LibGuide (http://libguides.bodleian.ox.ac.uk/ora-data-checklist accessed 10/06/2015) advises depositors to consider which formats will ensure the broadest possible accessibility by others, both now and in the future (and if in doubt to consider depositing more than one format of the same item to help future-proof the data). It advises that plain text files (such as .txt; .csv; .html; .xml) are both human and machine readable, and can be opened in any operating system by a wide range of applications (unlike some proprietary software formats) and also provides links to a number of external sites which provide advice on file formats:

UK Data Archive guide to file formats (http://www.data-archive.ac.uk/create-manage/format/formats-table accessed 10/06/2015)

Digital Curation Centre guide to file formats (http://www.dcc.ac.uk/resources/curation-reference-manual/completed-chapters/file-formats accessed 10/06/2015)

Jisc Digital Media Infokit: digital file formats (http://www.jiscdigitalmedia.ac.uk/infokit/file_formats/digital-file-formats accessed 10/06/2015)

Depositors are required to document (either within the metadata or in a deposited Readme file) how, why and when their data was created and whether any specific software was used to create, edit or process the files (e.g. a note indicating that a particular software package is required to read and work with the files); they are advised to include sufficient information about the dataset to make it easily intelligible to others. Meanwhile, ORA-Data staff monitor deposits for outmoded (or soon-to-be outmoded) file formats at both the review and migration stages of the deposit process, and contact depositors to advise as necessary.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

Advice for your DSA renewal submission: your repository earns to provide a formal list of recommended file formats - and not (just) links to external references...

3. The data producer provides the data together with the metadata requested by the data repository.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Data producers have the option to deposit a dataset and a metadata record, or just a metadata record (where the dataset is deposited elsewhere).

Basic advice for data producers is provided at http://www.bodleian.ox.ac.uk/bdlss/digital-services/data-archiving (accessed 09/06/2015) and more detailed guidance is provided from the ORA-Data LibGuide at http://ox.libguides.com/ora-data (accessed 09/06/2015). Clear and concise pop-up guidance for each field on each step, and detailed guidance is also provided in the ORA-Data LibGuide (http://ox.libguides.com/how-to-deposit and http://ox.libguides.com/ora-data-checklist both accessed 09/06/2015).

The 6-step deposit form presents a combination of mandatory and optional fields relating to descriptive, structural and administrative metadata:

1. Upload files (agree to deposit licence; browse and select files; upload files)

2. About your data (Title information; documentation about dataset and processes; publisher; publication year; plus optional abstract, subjects, keywords, and language, temporal and geographical information) NB: includes option to enter existing DOI when depositing metadata record only, or request a new DOI if depositing dataset with metadata record. NB: in addition to completing the mandatory 'Documentation about dataset and processes' field, depositors are advised to explain if specific software is required to read and work with the files, and are encouraged to include a standard ReadMe file with their dataset.

3. About you (creator information, with names/email/affiliation linked from the University's core user directory as appropriate, and a selection of 16 different creator roles)

4. Related publications (title, relationship, abstract, DOI, URL and/or citation)

5. Data archiving and access (payment options, nominate data steward, select access conditions for metadata and or dataset, including embargoes)

6. Funders and licences (details of funding awards, options for licensing data) NB: Pop-up help provides guidance on how to license research data with reference to Ball, A. (2012). 'How to License Research Data'. DCC How-to Guides. Edinburgh: Digital Curation Centre (http://www.dcc.ac.uk/resources/how-guides/license-research-data accessed 25/06/2015)

If any mandatory fields have not been completed upon submission, the data producer is automatically referred back to the field in question.

Following submission, ORA-Data staff review each metadata record and will contact the data producer if metadata in any mandatory fields are insufficient or incorrect, or if they believe additional metadata in either mandatory or optional fields will make the record/dataset substantially more discoverable (this is particularly the case where research has been funded by UK research councils which have specific rules or requirements regarding research data).

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

4. The data repository has an explicit mission in the area of digital archiving and promulgates it.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

"As the University of Oxford aims to lead the world in research and education, so the Bodleian Libraries aim to support the University in its mission by providing a world-leading library service. The mission of the Bodleian Libraries is to provide an excellent service to support the learning, teaching and research objectives of the University of Oxford; and to develop and maintain access to Oxford's unique collections for the benefit of scholarship and society." (http://www.bodleian.ox.ac.uk/bodley/about-us/policies accessed 25/06/2015)

The key role of digital services in achieving these objectives is reflected in the Libraries' specifically digital policies regarding metadata, digitisation, architecture and preservation, the latter being available at http://www.bodleian.ox.ac.uk/bodley/about-us/policies/preservation (accessed 25/06/2015) and complemented by the Bodleian Libraries' policy on the management and preservation of research data and records (http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0015/190014/Bodleian-RDM-Policy.pdf accessed 25/06/2015).

Meanwhile, ORA-Data's policy concerning the storage and preservation of data (http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0012/190011/ORA-Data-Submission-Preservation-and-Withdrawal-Policies.pdf accessed 25/06/2015) states the following:

1. Every reasonable effort will be made to retain items indefinitely.

2. The repository will try to ensure continued readability and accessibility.

• Items will be migrated to new file formats where possible and if deemed necessary.

- It may not be possible to guarantee the ongoing readability of all file formats.

3. The repository regularly backs up its files according to current best practice.

4. The original bit stream is retained for all items, in addition to any upgraded formats.

5. In the event of the repository being closed down, the database will be transferred to another appropriate archive in the control or management of the University of Oxford.

Furthermore, the Bodleian works closely with agencies such as the Digital Preservation Coalition (http://www.dpconline.org/ accessed 25/06/2015) and the Data Curation Centre (http://www.dcc.ac.uk/ accessed 25/06/2015).

Promotional activities are varied, and include presentations to University divisions, faculties and departments, meetings with potential depositors (especially as new research data requirements are issued by UK research councils), and dissemination through the University's research data website (http://researchdata.ox.ac.uk/ accessed 25/06/2015) and its blog (http://researchdata.ox.ac.uk/home/research-data-blog/ accessed 25/06/2015).

Information in the form of printed leaflets and posters is also widely distributed around the libraries and academic departments of the University. All front-facing Library staff are made aware of ORA and ORA-Data both at the induction training offered to new staff, and by frequent presentations at Staff Development events. ORA and ORA-Data staff also conduct an active programme of outreach events within the departments and faculties of the University.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

5. The data repository uses due diligence to ensure compliance with legal regulations and contracts including, when applicable, regulations governing the protection of human subjects.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

ORA-Data is a service provided by Bodleian Digital Library Systems and Services (BDLSS), part of the Bodleian Libraries of the University of Oxford. The University is a civil corporation established under common law, which was formally incorporated by the Act for Incorporation of Both Universities 1571 under the name of 'The Chancellor Masters and Scholars of the University of Oxford' (see http://www.admin.ox.ac.uk/statutes/375-092.shtml#_Toc30485125 accessed 25/06/2015).

Those depositing in ORA-Data are required to be members of the University of Oxford (access is restricted to those with an Oxford University Single-Sign-On), or are required to have been members of the University of Oxford at the point of creating their research outputs and research data (in which case retrospective deposit is facilitated by ORA-Data staff). As such, they are required to comply with the legal and ethical criteria contained in the University's Policy on the Management of Research Data and Records:

http://www.admin.ox.ac.uk/media/global/wwwadminoxacuk/localsites/researchdatamanagement/documents/Policy_on_the_Management (accessed 25/06/2015)

ORA-Data has a number of service-specific policies and conditions, openly available from http://www.bodleian.ox.ac.uk/bdlss/digital-services/data-archiving (accessed 25/06/2015) and elsewhere, including:

ORA User Licence (http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0011/190010/ORA-Data-User-Agreement.pdf accessed 25/06/2015). NB: data deposited in ORA-Data is accessed via the ORA interface (http://ora.ox.ac.uk/ - accessed 25/06/2015) and so the ORA User Licence covers both research publications and research data.

ORA-Data Submission, Preservation and Withdrawal Policies
(http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0012/190011/ORA-Data-Submission-Preservation-and-Withdrawal-Policies.pdf
 accessed 25/06/2015)

ORA-Data Deposit Conditions
(http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0019/190009/ORA-Data-Deposit-Agreement.pdf accessed
25/06/2015), which state that data does not:

- contain any material which is unlawful

- infringe any intellectual property rights (e.g. copyright or database right) of any other person

- contain any material which is defamatory of any person

- contain any confidential information

- contain any information that relates to an individual who can be identified

For clarification of anonymity, depositors are referred to the guidance of the UK Information Commissioner's
Office (https://ico.org.uk/media/1061/anonymisation-code.pdf accessed 25/06/2015).

The ORA-Data Deposit Conditions are available to, and must be accepted by (via a tick-box), a depositor at Step 1
of ORA-Data's 6-Step deposit process before a dataset can be deposited.

Guidance for depositors is provided by the University's research data website (http://researchdata.ox.ac.uk/home/managing-your-data-at-oxford/ethical-legal-commercial/ accessed 25/06/2015), and the Central University Research Ethics Committee (CUREC) (http://www.admin.ox.ac.uk/curec/ accessed 25/06/2015). Meanwhile, ORA-Data reviewers are familiar with will the guidance of the UK Information Commissioner's Office (https://ico.org.uk/media/1061/anonymisation-code.pdf accessed 25/06/2015) and will check deposited data to ensure anonymization should the associated metadata suggest there may be risk of disclosure.

Staff who operate ORA-Data and its underlying systems are trained in the handling and disclosure of sensitive information and, when necessary, are able to call upon the wider services of the Bodleian Libraries and the University of Oxford should they need to seek further advice.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

6. The data repository applies documented processes and procedures for managing data storage.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

ORA-Data is governed by the Bodleian Libraries' policy on the management and preservation of research data and records (http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0015/190014/Bodleian-RDM-Policy.pdf accessed 25/06/2015), though it also has a service-specific policy (http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0012/190011/ORA-Data-Submission-Preservation-and-Withdrawal-Policies.pdf accessed 25/06/2015) which states the following:

1. Every reasonable effort will be made to retain items indefinitely.

2. The repository will try to ensure continued readability and accessibility.

- Items will be migrated to new file formats where possible and if deemed necessary.

- It may not be possible to guarantee the ongoing readability of all file formats.

3. The repository regularly backs up its files according to current best practice.

4. The original bit stream is retained for all items, in addition to any upgraded formats.

5. In the event of the repository being closed down, the database will be transferred to another appropriate archive in the control or management of the University of Oxford.

More broadly, the policies which underpin the Bodleian's digital activities are available at http://www.bodleian.ox.ac.uk/bodley/about-us/policies/digital-policies-overview (accessed 26/06/2015) and its policy on architecture (http://www.bodleian.ox.ac.uk/bodley/about-us/policies/architecture accessed 26/06/2015) outlines: infrastructure; virtualization; scalability; duality (two disk copies plus tape back-up); storage capacity; network capacity; back-up regime; disaster recovery; security; commonality; and services alignment.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

7. The data repository has a plan for long-term preservation of its digital assets.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

ORA-Data is agnostic with respect to file types and metadata. Any file type can be deposited and stored, as can any metadata in XML format. In addition to individual files, ORA-Data accepts data packages, i.e., data objects comprising many mixed file types such as processed data spreadsheets, raw data in any form, text, and audio files. Data objects can also be in the form of zipped files that can be automatically unzipped if required. Data in ORA-Data are managed and curated so that they remain findable, accessible and reusable in the long term. ORA-Data provides bit stream preservation of research data; it cannot retain all the structural information of a relational database, or the user interface features of a website. By storing and preserving the data in ORA-Data, however, researchers are able to retrieve their data and use it to answer new research questions and to query the data in different ways to build new services on top of the data. Data producers are advised, however, that if datasets are created using proprietary software then continued access and reuse may be problematic (as some manufacturers protect the internal structure of their files and some do not make future upgrades back-compatible) and they are advised to contact the appropriate ORA-Data team member for technical advice. (See http://www.bodleian.ox.ac.uk/bdlss/digital-services/data-archiving accessed 26/06/2015)

Meanwhile, the ORA-Data LibGuide (http://libguides.bodleian.ox.ac.uk/ora-data-checklist accessed 26/06/2015) advises depositors to consider which formats will ensure the broadest possible accessibility by others, both now and in the future (and if in doubt to consider depositing more than one format of the same item to help future-proof the data). It advises that plain text files (such as .txt; .csv; .html; .xml) are both human and machine readable, and can be opened in any operating system by a wide range of applications (unlike some proprietary software formats) and also provides links to a number of external sites which provide advice on file formats:

UK Data Archive guide to file formats (http://www.data-archive.ac.uk/create-manage/format/formats-table accessed 10/06/2015)

Digital Curation Centre guide to file formats (http://www.dcc.ac.uk/resources/curation-reference-manual/completed-chapters/file-formats accessed 10/06/2015)

Jisc Digital Media Infokit: digital file formats (http://www.jiscdigitalmedia.ac.uk/infokit/file_formats/digital-file-formats accessed 10/06/2015)

Depositors are required to document (either within the metadata or in a deposited Readme file) how, why and when their data was created and whether any specific software was used to create, edit or process the files (e.g. a note indicating that a particular software package is required to read and work with the files); they are advised to include sufficient information about the dataset to make it easily intelligible to others. ORA-Data staff monitor deposits for outmoded (or soon-to-be outmoded) file formats at both the review and migration stages of the deposit process, and contact depositors to advise as necessary. Meanwhile, the University's central IT Services department, as well as the network of IT Support Officers, can also provide depositors with extensive advice and assistance on converting their data to more robust and preservable formats.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

8. Archiving takes place according to explicit work flows across the data life cycle.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Comprehensive procedural documentation is maintained and shared internally using an ORA-Data-specific area of proprietary team collaboration software (Atlassian Confluence), and this details the following:

- Policies and licences

- Funders' requirements

- Staff training

- Technical architecture

- Data deposit workflows

- Reviewer workflows

- File upload procedure (automated for single files up to 2GB and/or for multiple files up to 5GB, but by alternative means of file transfer for larger files)

- Optimal file formats

- Metadata fields (mandatory and optional, with appropriate pop-up help text and input restrictions)

- DOI assignment workflow and syntax

- Guidelines and other training materials

- Frequently asked questions

All data which underpins the research outputs of the University of Oxford are accepted but, where potential deposits are not deemed appropriate (e.g. where a deposit might reasonably be considered to constitute the research paper), ORA-Data staff will open a dialogue with the depositor.

Meanwhile, ORA-Data reviewers are familiar with will the guidance of the UK Information Commissioner's Office (https://ico.org.uk/media/1061/anonymisation-code.pdf accessed 25/06/2015) and will check deposited data to ensure anonymization should the associated metadata suggest there may be risk of subjects' privacy being compromised. Where necessary, reviewers will contact depositors and refer them to the guidance of the UK Information Commissioner's Office (https://ico.org.uk/media/1061/anonymisation-code.pdf accessed 25/06/2015), putting publication of the dataset in ORA-Data 'on hold' until it has been sufficiently anonymized.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

9. The data repository assumes responsibility from the data producers for access and availability of the digital objects.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

ORA-Data's Deposit Licence for data outlines conditions of deposit (http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0019/190009/ORA-Data-Deposit-Agreement.pdf accessed 25/06/2015) and must be accepted (via a tick-box) by the data producer at Step 1 of ORA-Data's 6-Step deposit process before a dataset can be deposited. All deposits are subject to review by ORA-Data staff and the dataset is checked if the metadata suggests that the conditions of the licence are in danger of being contravened (e.g. to ensure anonymization should the associated metadata suggest there may be risk of disclosure).

Meanwhile, a disaster recovery policy is included in the Bodleian's digital policy on architecture (http://www.bodleian.ox.ac.uk/bodley/about-us/policies/architecture accessed 26/06/2015).

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

10. The data repository enables the users to discover and use the data and refer to them in a persistent way.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

3. In progress: We are in the implementation phase.

Self-assessment statement:

ORA-Data currently accepts any file type for preservation, but its LibGuide (http://libguides.bodleian.ox.ac.uk/ora-data-checklist accessed 10/06/2015) advises depositors to consider which formats will ensure the broadest possible accessibility by others, both now and in the future (and if in doubt to consider depositing more than one format of the same item to help future-proof the data). It advises that plain text files (such as .txt; .csv; .html; .xml) are both human and machine readable, and can be opened in any operating system by a wide range of applications (unlike some proprietary software formats) and also provides links to a number of external sites which provide advice on file formats:

UK Data Archive guide to file formats (http://www.data-archive.ac.uk/create-manage/format/formats-table accessed 10/06/2015)

Digital Curation Centre guide to file formats (http://www.dcc.ac.uk/resources/curation-reference-manual/completed-chapters/file-formats accessed 10/06/2015)

Jisc Digital Media Infokit: digital file formats (http://www.jiscdigitalmedia.ac.uk/infokit/file_formats/digital-file-formats accessed 10/06/2015)

Depositors are required to document (either within the metadata or in a deposited Readme file) how, why and when their data was created and whether any specific software was used to create, edit or process the files (e.g. a note indicating that a particular software package is required to read and work with the files); they are advised to include sufficient information about the dataset to make it easily intelligible to others. Meanwhile, ORA-Data staff monitor deposits for outmoded (or soon-to-be outmoded) file formats at both the review and migration stages of the deposit process, and contact depositors to advise as necessary.

Datasets deposited in ORA-Data may be searched for and retrieved via the Oxford Research Archive interface (http://ora.ox.ac.uk/ - accessed 25/06/2015) which facilitates keyword searching of descriptive metadata (though not deep searching of dataset content) or a browse option which allows faceted filtering within five broad categories: type of work; subject; college; item date; faculty).

OAI harvesting is facilitated (indeed, ORA-Data is participating in the Jisc-funded UK Research Data Discovery Service which aims to aggregate metadata for research data held within UK universities and national, discipline specific data centres - https://www.jisc.ac.uk/rd/projects/uk-research-data-discovery accessed 01/07/2015) and ORA-Data's metadata records are crawled by major search engines such as Google, thereby maximising the visibility and impact of its content.

All records submitted in ORA-Data are assigned a universally unique identifier (UUID) and this can be used to create a persistent URL upon request. Digital Object Identifiers (DOIs) can be issued for all datasets deposited in ORA-Data, or for datasets deposited in secure data stores elsewhere in Oxford University according to the Bodleian's policy for assigning DOIs to datasets (http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0014/190013/Oxford-University-DOI-Policy-ORA-Data.pdf accessed 01/07/2015). Once a DOI has been issued, the terms of the University's DOI Allocation Agreement with the British Library (acting as the UK allocation agent of DataCite - https://www.datacite.org/ accessed 01/07/2015) do not allow either the dataset itself or five key metadata fields (identifier; creator; title; publisher; year published) to be amended.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

11. The data repository ensures the integrity of the digital objects and the metadata.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

3. In progress: We are in the implementation phase.

Self-assessment statement:

The Bodleian Digital Library (BDLSS) is in the process of commissioning a new disk storage platform (for ORA-Data and other services) which will support integrity checking (e.g. checksums, SHA-1). At present, data integrity checking is undertaken on the version of the data backed up on tape rather than on disks, though cross-checking of multiple tape-disk copies is planned.

While the underlying storage platform (Databank) does allow for multiple versioning of data, this has not yet been implemented for all workflows. ORA-Data currently provides only basic version control for multiple versions of the same data. Because a dataset which has been provided with a Digital Object Identifier (DOI) cannot be amended or added to under the terms of the University's DOI Allocation Agreement with the British Library (acting as the UK allocation agent of DataCite - https://www.datacite.org/ accessed 01/07/2015), any subsequent version of that dataset needs to be considered as a new deposit (with a new DOI if desired), though step 2 ('About your data') of the 6-step deposit process provides an optional field to provide the version number or name, and step 4 ('Related publications') allows the depositor to describe the relationship between the new and earlier datasets and link to them using DOIs or URLs. At present, this necessitates duplication of metadata entry in each new record for a new version of the dataset. However, it is anticipated that new functionality will be incorporated which will allow multiple versions of the same dataset to be assigned to the original data record in ORA-Data, with newer versions of the dataset retaining the same DOI string but each with a different version suffix (the DOI syntax is based on the Dryad model at https://www.nescent.org/wg_dryad/DOI_Usage - accessed 03/07/2015).

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

12. The data repository ensures the authenticity of the digital objects and the metadata.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Procedural guidelines are followed regarding review of datasets and associated metadata, and these are available to ORA-Data staff in a document in the ORA-Data-specific area of proprietary team collaboration software (Atlassian Confluence). The guidelines are as follows:

1. File checks

- Automatic virus checks are run on all files uploaded to the system, and this alerts ORA-Data staff to those that fail.

- ORA-Data staff check that files can be opened without problems. In the case of datasets containing a very large number of files (e.g. upwards of 100) it may be too time-consuming to check each individual file, in which case a 5% sampling technique is recommended. It may not be possible for reviewers to open some files produced using specialist or proprietary software, and in these cases the depositor may be contacted to suggest that they also deposit a copy of their data in a plain file format (e.g. .txt; .csv; .html; .xml), which will be more 'future-proofed' and accessible/useable for other researchers.

- File names are checked for any obvious disclosures of identifiable information about participants in the research. This may apply especially to research in the Social Sciences (which often uses interviews, either in transcript or audio-visual form) and to medical and psychology research (e.g. which may include information about patients or volunteers in studies).

2. Metadata checks

- ORA-Data staff check that all mandatory fields have been completed adequately, and contact the depositor for further details as required. The mandatory data fields are:

1. Title

2. Documentation about your dataset and processes

3. Publisher name [should auto-fill as 'University of Oxford' but depositors can change as they require]

4. Year published [should auto-fill as the current year, but depositors can change as they require]

5. Creator information [there must be at least one named person listed]

6. Archiving service payment [an option must be selected here, and further details provided]

7. Nominate a data steward [at least one data steward should be named, and s/he cannot be a creator]

8. Can this file be made freely available? [default is 'yes', but if 'after a certain period' or 'no' are selected, reviewers check that the depositor has then filled in the embargo details]

9. Can this ORA catalogue record be made freely available? [default is 'yes', but if 'after a certain period' or 'no' are selected, check that the depositor has then filled in the embargo details]

10. Is your research the result of a funding award? [if 'yes' is selected, the reviewers check that full details of the award have been given]

- Reviewers check that information is entered and spelled correctly. Titles (e.g. in the fields 'Title' and 'Title of related publication') should be entered in sentence case, with capitalisation reserved for the initial letter and proper nouns.

- Reviewers check any URL links cited are not broken or inappropriate.

- Where appropriate, reviewers may expand some of the metadata based on the description given by the depositor and their submitted data files (e.g., adding a broad subject heading or keyword).

3. Documentation checks

- ORA-Data staff check that all datasets deposited have some form of accompanying data documentation. Depositors may either have entered text directly in the 'Documentation about your dataset and processes' field in step 2 ('About your data') of the 6-step deposit process, or may instead have uploaded supporting documentation with their data files. Either way, this documentation should briefly explain the context of the data, including whether any specific software was used to create, edit or process the files, and should explain such things as acronyms, abbreviations or names of variables used in the data. Depositors are asked to give sufficient information to make the dataset intelligible and reusable for another researcher. Reviewers look out for instances where a depositor has obviously not given enough information and, if necessary, they will get in touch with them for clarification.

- ORA-Data staff check to see if any other supporting file names are referenced in the deposit form, and that any such files have been uploaded.

4. Linked publication checks

- If details of a related publication are given in the 'Related publications' section, reviewers check that any bibliographic details, URLs, DOIs or citations given appear to be correct.

- ORA-Data staff check whether there is a copy of the publication already in ORA itself (i.e. a publication) and, if not, consider asking the depositor if they would like to deposit a copy that can then be linked to the dataset.

5. Copyright, access & embargo checks

- ORA-Data staff check that access conditions (i.e. embargos) and any additional licence chosen for the data are appropriate with regard to the ownership and/or confidentiality of the data.

6. Identifiable information checks

- Uploaded files may be briefly checked for any obvious disclosures of identifiable information about participants in the research. This may apply especially to research in the Social Sciences (which often uses interviews, either in transcript or audio-visual form) and to medical and psychology research (which may include, e.g., information about patients or volunteers in studies).

- Reviewers look out for potentially controversial areas of research (e.g. medical testing involving animals) where the data and catalogue record may benefit from an embargo, even if the depositor hasn't requested or realised this. In such cases, reviewers may contact the depositor to check as appropriate.

In most cases, the identity of depositors is confirmed through restriction to the ORA-Data service to members of the University of Oxford (access is restricted to those with an Oxford University Single-Sign-On). Those who were members of the University of Oxford at the point of creating their research outputs and research data, but are no longer so and therefore do not have an Oxford University Single-Sign-On, must contact ORA-Data staff to discuss their qualification for depositing in ORA-Data. Furthermore, all 'person' fields in the data form are linked to Oxford's Core University Directory Service (https://help.it.ox.ac.uk/internal/sld/cud - accessed 03/07/2015) and this allows depositors' contact details and affiliation (college, department etc.) to be auto-completed in the data form; it is anticipated that ORA-Data will also investigate the potential for using the recently implemented 'ORCID at Oxford' service for this purpose (http://openaccess.ox.ac.uk/orcid-at-oxford-goes-live/ - accessed 03/07/2015).

There is an audit trail for each deposit in ORA-Data which logs the date, person (depositor, reviewer etc.), status (draft; submission; claimed for review; approved; system verified; data migrated; published; DOI registered; referred, etc.) along with information about all processes carried out.

Depositors must contact ORA-Data staff to request post-publication changes (the reviewer can 'refer' a completed deposit form back to the depositor for a given period of time where it is justified), though this is limited if a Digital

Object Identifiers (DOI) had previously been issued, in which case the terms of the University's DOI Allocation Agreement with the British Library (acting as the UK allocation agent of DataCite - https://www.datacite.org/ - accessed 01/07/2015) do not allow either the dataset itself or five key metadata fields (identifier; creator; title; publisher; year published) to be amended.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

13. The technical infrastructure explicitly supports the tasks and functions described in internationally accepted archival standards like OAIS.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

3. In progress: We are in the implementation phase.

Self-assessment statement:

The ORA-Data service is an implementation of the OAIS reference model. It complies with metadata schema standards including Dublin Core (see http://dublincore.org/ - accessed 30/09/2015), MODS (see http://www.loc.gov/standards/mods/ - accessed 17/08/2015) for legacy data, and the RDF (see http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/ - accessed 30/09/2015). It is also working toward compliance with RIOXX (see http://rioxx.net/ - accessed 17/08/2015) for research data, so as to ensure compliance with the RCUK policy on open access. Meanwhile, it has non-public infrastructure plans (for ORA-Data and other services) which are reviewed and developed on an on-going basis; it is currently implementing a new infrastructure under the repository (VIPR) and is preparing for the transfer of the ingest forms/repository to the VIPR infrastructure in order to ensure stable infrastructure in the long-term. Meanwhile, a market review is under way to help inform development of future technology platforms for the Oxford Research Archive as a whole.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

14. The data consumer complies with access regulations set by the data repository.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Data deposited in ORA-Data is accessed via the ORA interface (http://ora.ox.ac.uk/ - accessed 25/06/2015) and so the ORA User Licence covers both research publications and research data (http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0011/190010/ORA-Data-User-Agreement.pdf – accessed 03/07/2015). The licence states that "If you use any content in breach of the terms of this user licence, your right to use the ORA website will cease immediately and you must, at the University's option, return or destroy any copies of the content you have made". Each record delivered through the ORA interface also contains a link to a summary of the terms of use (http://ora.ox.ac.uk/information/termsOfUse - accessed 03/07/2015).

Depositors may embargo their datasets but, where applicable, are expected to comply with funding bodies' guidelines on making their data freely available. Requests to view restricted use (confidential) data may be made to ORA-Data staff who will open dialogue with the data producer (or the nominated data steward if the data producer can no longer be contacted).

Data depositors can assign a licence of their choice from a drop-down list in step 6 of the 6-step deposit process, or choose 'bespoke licence' and enter the text of any bespoke licence, access terms and conditions. For guidance on how to license research data, they are referred to Ball, A. (2012) 'How to License Research Data' DCC How-to Guides. Edinburgh: Digital Curation Centre (http://www.dcc.ac.uk/resources/how-guides/license-research-data – accessed 03/07/2015).

We rely upon service users and members of the various stakeholder communities to identify and report any potential breeches of the licence under which data has been made available. In the first instance we would attempt to resolve any licence breaches via informal contact with the infringing party, but if necessary we are able to call upon the services of the University's Legal Services Department for advice and support on taking more robust steps.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

15. The data consumer conforms to and agrees with any codes of conduct that are generally accepted in the relevant sector for the exchange and proper use of knowledge and information.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Data consumers using ORA-Data are obliged to conform to the terms outlined in theORA User Licence (http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0011/190010/ORA-Data-User-Agreement.pdf accessed 25/06/2015). NB: data deposited in ORA-Data is accessed via the ORA interface (http://ora.ox.ac.uk/ - accessed 25/06/2015) and so the ORA User Licence covers both research publications and research data.

However, because data depositors in ORA-Data are obliged to comply with the legal and ethical criteria outlined in Section 5 above (confidential, sensitive or personal information is not to be included, or must be sufficiently anonymised, etc.), and are encouraged to declare the name/s of the rights holder/s (in step 6 of the 6-step deposit process), and are given a choice of licences (or indeed the option to identify a bespoke licence) to apply to their data (also in step 6 of the 6-step deposit process), the onus is very much on the data depositor not to provide data which might be misused in any way by the data consumer. ORA-Data does not therefore 'police' the data that is consumed through its open access interface (other than offering the facility for applying embargoes to datasets) because its deposit criteria is already rigorous in this regard and, where there is likelihood that data records/deposits may include confidential, sensitive or personal information, they will already have been rejected or referred back to the depositor as part of the initial review procedure.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

16. The data consumer respects the applicable licences of the data repository regarding the use of the data.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Data depositors can assign a licence of their choice from a drop-down list in step 6 of the 6-step deposit process, or choose 'bespoke licence' and enter the text of any bespoke licence, access terms and conditions. For guidance on how to license research data, they are referred to Ball, A. (2012) 'How to License Research Data' DCC How-to Guides. Edinburgh: Digital Curation Centre (http://www.dcc.ac.uk/resources/how-guides/license-research-data – accessed 13/07/2015).

Oxford University's Research Data Oxford website (http://researchdata.ox.ac.uk/ accessed 03/07/2015) provides specific advice about IP rights, sharing and licensing (http://researchdata.ox.ac.uk/home/sharing-your-data/rights-licensing/ - accessed 13/07/2015), and directs users to the University's intellectual property policy available at http://www.admin.ox.ac.uk/statutes/790-121.shtml#_Toc28143157 – accessed 13/07/2015).

The University of Oxford complies with United Kingdom's Data Protection Act of 1998 (https://www.gov.uk/data-protection/the-data-protection-act accessed 13/07/2015) and government guidelines on IP and copyright (https://www.gov.uk/government/organisations/intellectual-property-office accessed 13/07/2015) and the data consumer is also obliged to comply with the terms of the ORA User Licence (http://www.bodleian.ox.ac.uk/__data/assets/pdf_file/0011/190010/ORA-Data-User-Agreement.pdf).

Meanwhile, in conjunction with a number of other leading UK universities, the University of Oxford has responded to the RCUK's recent call for consultation on its draft 'Concordat On Open Research Data' (see Version 10 at http://www.rcuk.ac.uk/research/opendata/ - accessed 30/09/2015) which seeks to establish a set of expectations of good practice for both the creators and consumers of open research data.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments: