



## Assessment Information

[CoreTrustSeal Requirements 2017–2019](#)

Repository:

DataFirst

Website:

<https://www.datafirst.uct.ac.za/>

Certification Date:

05 November 2018

This repository is owned by:

**University of Cape Town**

**CoreTrustSeal Board**

W [www.coretrustseal.org](http://www.coretrustseal.org)

E [info@coretrustseal.org](mailto:info@coretrustseal.org)



# DataFirst

## Notes Before Completing the Application

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

True

### *Reviewer Entry*

#### **Reviewer 1**

Comments:  
Accept

#### **Reviewer 2**

Comments:  
Accept

## CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

### Background & General Guidance

### Glossary of Terms

## BACKGROUND INFORMATION

### Context

*R0. Please provide context for your repository.*

*Repository Type. Select all relevant types from:*

Domain or subject-based repository

*Reviewer Entry*

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

**Comments**

DataFirst is a university-based open research data repository.

*Reviewer Entry*

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

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

Our public access data is available to all, but our target communities are the academic and policy research communities in the social sciences.

*Reviewer Entry*

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

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

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

*Reviewer Entry*

**Reviewer 1**

Comments:  
Accept

**Reviewer 2**

Comments:  
Accept

## ***Comments***

There are many open African data sites, but these sites do not generally comply with open data principles and curation best practice. Firstly, their data are often incomplete and non-primary, and access to them is impermanent. Secondly, because these data are usually aggregated, it is difficult to assess their accuracy.

Our reputation rests on the quality assurance we provide with the data we disseminate. We therefore make extensive data -quality checks on the data we share. Anomalies and errors are either

(i) corrected in consultation with depositors, or

(ii) tagged and noted in the "data quality note" field in the metadata, to help researchers with data reuse

Changes to any of the data files or document files will be denoted by a version change in the file label. Major changes, such as reissuing of the data files, will be denoted by a new version, e.g. version 1.0 to version 2.0. Other changes are denoted by version number increments, e.g. version 2.1. We version at file level, so that researchers don't have to redownload unchanged files. The dataset version number is aligned to the version of the most recently updated file, e.g. if the census household data file is version 2, and the census person data file is version 2.1, the census dataset will be versioned 2.1.

### ***Reviewer Entry***

**Reviewer 1**

Comments:  
Accept

**Reviewer 2**

Comments:  
Accept

## ***Outsource Partners. If applicable, please list them.***

The ICT department at the University of Cape Town maintains our repository infrastructure, including ICT security.

### ***Reviewer Entry***

**Reviewer 1**

Comments:  
Accept

**Reviewer 2**

Comments:  
Accept

## ***Other Relevant Information.***

(i) Certifications

DataFirst was awarded the Data Seal of Approval from 2014 to 2017

(ii) Memberships

(a) Inter-University Consortium of Social and Political Research (since 2011) DataFirst is the official representative for the University of Cape Town, and negotiated for and manages ICPSR membership of other SA universities under a South African national membership

<https://www.icpsr.umich.edu/icpsrweb/membership/administration/institutions/7379>

(b) ICSU World Data System (since 2013) DataFirst is a regular member of the WDS

[https://www.icsu-wds.org/community/membership/regular-members/@@member\\_view?fid=datafirst](https://www.icsu-wds.org/community/membership/regular-members/@@member_view?fid=datafirst)

*Reviewer Entry*

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

## ORGANIZATIONAL INFRASTRUCTURE

### I. Mission/Scope

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

#### *Compliance Level:*

4 – The guideline has been fully implemented in the repository

*Reviewer Entry*

**Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

**Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

#### *Response:*

Our mission to make disaggregated African data easily available to researchers is published on our homepage and has been endorsed by our Governing Board and Internal UCT Advisory Committee:

DataFirst is a research data service dedicated to opening up South African and other African survey and administrative microdata to researchers and policy analysts. We promote high quality research by providing the essential Open Research Data infrastructure for discovering and accessing data and by developing skills among prospective users, particularly in South Africa. We undertake research on the quality and usability of national data and encourage data usage and data sharing." <https://www.datafirst.uct.ac.za/>

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Accept

##### **Reviewer 2**

Comments:  
Accept

## **II. Licenses**

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

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

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

##### **Reviewer 2**

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

### ***Response:***

We share data under three licenses. Our data depositors choose the access level for the data they deposit with us, although we do encourage an open data approach unless the data is potentially disclosive. A middle ground is the non-commercial use option. These access types are clearly indicated by icons alongside the dataset name and on the dataset landing page. Access levels include.

(i) Public Access (shown with a green arrow icon)

Users log in and complete an online form stating how they will use the data, and can then immediately download the data and documentation

(ii) Licensed Access Data (shown with an orange figure icon)

This is non-commercial use data. Each online data request is checked and approved. Users receive an email with a link to download the data (within 2 days)

(iii) Data available in a data enclave (our secure research data centre at the university) (shown with a blue padlock icon)

This is for sensitive or potentially disclosive data (e.g. survey data with household GPS coordinates, or firm-level data).

The metadata is available online and researchers can apply for access to the Centre online at

<https://www.datafirst.uct.ac.za/services/secure-data-services>

We also publish metadata for open research data from other sites (shown with a blue arrow icon) to assist data discovery

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

Accept

##### **Reviewer 2**

Comments:

Accept

### **III. Continuity of access**

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

#### ***Compliance Level:***

3 – The repository is in the implementation phase

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

3 – The repository is in the implementation phase

##### **Reviewer 2**

Comments:

3 – The repository is in the implementation phase

#### ***Response:***

Our parent body, the University of Cape Town funds our repository infrastructure and pays the data service administrator's salary. Funding for additional data-related activities such as data quality work and data research is from sponsors or partner institutions. Long-term preservation would include migration to an alternative, university-supported repository in the event of DataFirst being unable to continue the work. We have operated since 2001, and this scenario is unlikely. However, our Internal Advisory Committee, representing senior university academics, is working with the University administration to formally document the University's ongoing support for our repository.

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Accept

##### **Reviewer 2**

Comments:  
Accept

## **IV. Confidentiality/Ethics**

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

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

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

##### **Reviewer 2**

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

### ***Response:***

(i) Ethics approval for our repository

We obtained approval to share data from our Faculty Ethics Committee, and are specifically mentioned in their guide <http://www.commerce.uct.ac.za/Downloads/Commerce%20Ethics%20in%20Research%20Handbook%20June%202018.pdf>

We sign MOUs with depositors for curatorship and sharing permissions

<https://www.datafirst.uct.ac.za/files/20180412-datafirst-depositor-mou-v5.docx>



(ii) Ethics approval for depositors' data collection

Since 2017 all depositors are required to provide us with a copy of documented ethics approval for their data collection. We publish these documents online with the data files. We have also added a field in our DDI-compliant metadata for ethics approval information.

(iii) Anonymising data

Most of the data deposited with DataFirst is anonymised. When this is not the case, we follow anonymisation procedures documented at <https://www.datafirst.uct.ac.za/images/docs/datafirst-disclosure-control-flowchart.pdf>

(iv) Sharing sensitive/personal data

Our secure centre at the University holds sensitive data <https://www.datafirst.uct.ac.za/services/secure-data-services> The Centre was set up according to international standards for safe centres, in consultation with IT experts and experts from the UK Data Archive, who have extensive experience managing a safe Centre <https://www.ukdataservice.ac.uk/get-data/how-to-access/accesssecurelab>

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Accept

##### **Reviewer 2**

Comments:  
Accept

## **V. Organizational infrastructure**

*R5. The repository has adequate funding and sufficient numbers of qualified staff managed through a clear system of governance to effectively carry out the mission.*

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

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

##### **Reviewer 2**

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

## ***Response:***

### **(i) Funding**

The repository infrastructure and repository administrator's salary are university funded. Data research work and other data projects are project- or donor-funded.

### **(ii) Governance**

Our governance structure includes (a) a Governing Board representing international data repositories, government data producers, and South African universities and policy research institutes, and (b) an Internal Advisory Committee made up of UCT academics and senior DataFirst staff

### **(iii) Skills-sets**

We are a small but highly-skilled team including (a) the data service/repository Manager (Mphil Information Science UCT covering African government data, 23 year's experience curating research data), and (b) two Data Analysts, a micro-economist and a demographer, and (c) a Web/Media Manager, as well as administrative staff.

## ***Reviewer Entry***

### **Reviewer 1**

Comments:  
Accept

### **Reviewer 2**

Comments:  
Accept

## **VI. Expert guidance**

***R6. The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either inhouse or external, including scientific guidance, if relevant).***

## ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

## ***Reviewer Entry***

### **Reviewer 1**

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

### **Reviewer 2**

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

## ***Response:***

Expert input is received from our Governing Board. Board members include:

(a) Internationally respected data experts, including Directors of the UK Data Archive (founded in 1967)

[www.data-archive.ac.uk](http://www.data-archive.ac.uk) and the Inter-University Consortium for Political and Social Research (founded in 1962)

<https://www.icpsr.umich.edu/icpsrweb/>

(b) Established social scientists, including the directors of influential research units at 7 South African universities and the African Economic Research Consortium, an African research-policy think-tank

(c) Government data producers (Statistics agency, other government departments)

This team provides invaluable expert input to our work. We report back to the Board on follow-up to any service improvements we make on their recommendation. Examples are: implementing DOIs, and working with SA journal editors to encourage more complete data citation by SA authors.

### **Helpdesk**

Our online helpdesk is a tool for communicating with our data users, and for soliciting feedback on the quality and usability of our data. As a university-based repository, we work closely with academics, and have built up excellent relationships with senior staff who provide feedback, directly or via our helpdesk <http://support.data1st.org/support/home>

We self-assign response and follow-up times on our helpdesk regarding user suggestions and work within these to improve the quality of the data, working with depositors.

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

Accept

#### **Reviewer 2**

Comments:

Accept

## **DIGITAL OBJECT MANAGEMENT**

### **VII. Data integrity and authenticity**

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

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

**Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

***Response:***

Standards

(a) Our processes for digital object management are modelled on the Open Archival Information System ISO standard <https://www.datafirst.uct.ac.za/images/docs/20150618-microdata-service-model.pdf>

(b) We use DDI and Dublin Core metadata standards

(c) Depositor credentials - We sign a standard MOU with depositors, which confirms roles and responsibilities, and their right to deposit the data <https://www.datafirst.uct.ac.za/files/20180412-datafirst-depositor-mou-v5.docx>

(d) Data Integrity and reliability - The original data deposit and all iterations of the data are housed on a secure server at UCT, and we use checksums and versioning to ensure data integrity and reliability. Our data platform keeps a record of all changes and by whom these were made.

(e) Metadata - We create detailed DDI-compliant metadata for the data we share, and for the data we link to, using Nesstar publisher, which is a free mark-up tool for creating metadata exportable in xml

(f) Versioning - We include version notes in the metadata for information on changes to each version of the data. We use file-level versioning, so that researchers only need to download new versions of files, rather than whole datasets. We only publish the most up-to-date version, but do provide access to earlier versions and tombstone citations on request.

(g) Persistent identifiers - DataFirst joined DataCite at the end of 2017 to obtain DOIs and have minted DOIs for over 60 of our datasets and will have minted DOIs for all datasets by the end of 2018.

***Reviewer Entry***

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

## VIII. Appraisal

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

***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

#### ***Response:***

We've modelled our research data life-cycle at

<https://www.datafirst.uct.ac.za/images/docs/20150618-microdata-service-model.pdf>

This model is based on the Open Archival Information System ISO standard for open repositories

[https://en.wikipedia.org/wiki/Open\\_Archival\\_Information\\_System](https://en.wikipedia.org/wiki/Open_Archival_Information_System)

The data deposit stage ("data ingest" in OAIS) is stage A in our model. We are working with our software developers to add an online self-deposit module to our data platform.

##### (i) Policy

We accept all social science and health data, in common with social science data services worldwide. We accept both qualitative and quantitative data. We do not take aggregated data or data we cannot share in some form. We don't refuse deposits of shareable microdata. We assess each deposit, and don't make the data available if it is of dubious quality. But that is very rare. Most data has some worth, and we prefer to provide data quality notes to ensure users know the limitations of the data. Information for data depositors can be found on our site at

<https://www.datafirst.uct.ac.za/services/deposit-data>

##### (ii) Documentation

We encourage depositors to provide all documents used in data collection and reporting on the data. We use this, often in consultation with depositors, to create useful metadata

##### (iii) Data formats

We do specify preferred data formats, but we take any usable format, and convert these to standard software formats for data analysis, such as SPSS or Stata. We also comply with open data principles by providing the data as .csv files so the data can be analysed by researchers who do not have access to proprietary software.

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

Accept

##### **Reviewer 2**

Comments:

Accept

## **IX. Documented storage procedures**

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

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

### ***Response:***

The preservation dataset is represented by digital object B on our data life-cycle model at

<https://www.datafirst.uct.ac.za/images/docs/20150618-microdata-service-model.pdf>

(i) Preservation dataset

The original deposit copy and all iterations of the dataset are housed on a server managed by the ICT department at the university. The restricted-access data is housed on a secure server, isolated from external servers. Servers have uninterruptable power supplies, and are backed-up weekly, and replaced every 5 years, or when necessary.

(ii) Storage infrastructure

Storage, backup, migration, preservation and security requirements are managed by UCT's ICT department, in accordance with UCT's IT infrastructure policies [http://www.icts.uct.ac.za/policies\\_guidelines](http://www.icts.uct.ac.za/policies_guidelines) Data is backed up to servers housed in a secure server room. The secure server is backed up to a second secure server, which is part of an intranet not connected to external networks. Backups are of the data holdings on the server, not the software.

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

Accept

##### **Reviewer 2**

Comments:

Accept

## **X. Preservation plan**

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

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

### ***Response:***

(i) Responsibilities for preservation

Our depositor MOU formalises the rights and responsibilities of data owners and DataFirst as the data distributor, including data owners' right to remove DataFirst's curatorship permissions

<https://www.datafirst.uct.ac.za/services/deposit-data>

(ii) Data Ownership

Ownership does not pass to us, but data owners agree to allocate curatorship and dissemination rights to us.

(iii) Retention period

We preserve and publish data in the long-term, as we believe there is no cut-off date for the utility of data. As an illustration, we recently digitised data from a survey carried out in 1948. These data are now being used by historians, sociologists, and other social scientists to re-explore South Africa's Apartheid past.

(iv) Persistent Access

Our datasets have persistent identifiers to ensure their continued availability to the research community. In 2017 we became members of DataCite to obtain DOIs for our datasets.

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

Accept

##### **Reviewer 2**

Comments:

Accept

## **XI. Data quality**

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

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

### ***Response:***

Data assurance is stage 3 on our data life-cycle model

<https://www.datafirst.uct.ac.za/images/docs/20150618-microdata-service-model.pdf>

Our reputation as a trusted repository depends on the quality assurance we provide for the data we distribute. This has made us the repository of choice among academics for South African data.

We comply with data curation standards and best practice, to ensure the quality of the data. Quality attributes we take into account are:

#### **(i) Accessibility**

We provide infrastructure and metadata to ensure data is discoverable. We comply with open data requirements by sharing our public access data free and online and for any purpose, and in non-proprietary and well as proprietary formats. We do require user registration, but this is for reporting purposes, and is not related to data access. Through our secure centre, we make data accessible which would otherwise not be available to researchers

#### **(ii) Accuracy and reliability**

Through an iterative process of consultation with data owners and data users, we improve the accuracy and reliability of the data

#### **(iii) Comparability**

We distribute standard file formats and standardised (DDI compliant), machine-readable metadata to ensure data can be compared across space and time

#### **(iv) Security**

We use a combination of user agreements, permissions, disclosure control, technology, and safe spaces to ensure data confidentiality

#### **(v) Interpretability**

Our metadata provides usage and provenance information to help researchers to re-analyse the data. Our online



helpdesk provides ongoing user support <http://support.data1st.org/helpdesk>

(vi) Data quality feedback

While we undertake data quality checks before publishing the data, we also crowd-source among academics for feedback on the quality and usability of the data. We receive feedback via our online helpdesk <http://support.data1st.org/helpdesk>.

At best, we work with owners and users to improve the data. At the least we provide data quality notes in our metadata.

We highlight this "virtual cycle of data re-use" on our model

<https://www.datafirst.uct.ac.za/images/docs/20150618-microdata-service-model.pdf>. We use this to demonstrate to data owners the value of sharing their data for further research.

### *Reviewer Entry*

#### **Reviewer 1**

Comments:

Accept

#### **Reviewer 2**

Comments:

Accept

## **XII. Workflows**

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

### *Compliance Level:*

4 – The guideline has been fully implemented in the repository

### *Reviewer Entry*

#### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

#### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

### *Response:*

Our repository workflow is modelled at

<https://www.datafirst.uct.ac.za/images/docs/20150618-microdata-service-model.pdf>

Summary of processes: Deposits are negotiated with depositors via our online support site or in person, and MOUs are signed to clarify roles and responsibilities of depositors and DataFirst as curator. Data checking involves quality checks

and data anonymisation where necessary. This stage feeds into the metadata creation stage, and in this stage we often liaise with depositors to ensure we create accurate and helpful metadata. Data files are uploaded with any useful data collection documents and any analytic records. Datasets are published when ready or, for funded projects, on data release dates. We report regularly to depositors on who is using their data and for what purpose, for their impact assessments. Finally, we source and cite publications based on the data, which is an on-going task.

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Accept

##### **Reviewer 2**

Comments:  
Accept

## **XIII. Data discovery and identification**

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

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

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

##### **Reviewer 2**

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

### ***Response:***

#### (i) Data discovery

Our repository has a searchable online metadata catalogue which we maintain according to international data curation standards (DDI for metadata at dataset level, Dublin Core for metadata on documents)

#### (ii) Metadata harvesting

Our Repository software was developed by the World Bank's Development Data Group, who harvest our metadata for the World Bank's Central Microdata Catalog

[http://microdata.worldbank.org/index.php/catalog?sort\\_by=rank&sort\\_order=desc&sk=](http://microdata.worldbank.org/index.php/catalog?sort_by=rank&sort_order=desc&sk=)

We are also listed in the DataCite Registry of Research Data Repositories

<https://www.re3data.org/search?query=DataFirst>, amongst others

(iii) Data attribution

We use the DataCite data citation standard. We include a recommended citation field in the metadata for each dataset.

We provide information on data citation best practice <https://www.datafirst.uct.ac.za/services/citations>. Each dataset on our catalogue has links to publications that cite the data

Our datasets have persistent identifiers and we are currently minting DOIs for each datasets (60+ datasets have DOIs to date) and plan to allocate these persistent identifiers to all our datasets by the end of 2018.

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

Accept

##### **Reviewer 2**

Comments:

Accept

## **XIV. Data reuse**

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

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

### ***Response:***

Metadata creation is depicted by stages 4 and 5 in our curation model

<https://www.datafirst.uct.ac.za/images/docs/20150618-microdata-service-model.pdf>

#### (i) Deposit metadata

We don't require depositors to provide standardised metadata, but we do request all documents used during the data collection and reporting process. From these we create standardised, searchable metadata for each dataset, using the Nesstar Published markup software which is based on the DDI metadata standard. Through experience and interaction with academics, we have learnt which metadata fields are most relevant for data-intensive social research, such as information on the lowest level of geographic aggregation of the data

#### (ii) File formats

We provide data in formats used by the social science research community, e.g. SPSS, Stata, but also in open formats such as .csv and R, to comply with open data standards.

#### (iii) Flexibility

Metadata creation is a dynamic process, and we regularly update our metadata, e.g. when new versions are released after error correction, or when new information on historical data comes to light. We also update quality notes after user feedback. Being in a university setting enables us to keep updated on the tools and information required by academics and students who undertake data analysis

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Accept

##### **Reviewer 2**

Comments:  
Accept

## TECHNOLOGY

### XV. Technical infrastructure

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

#### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

**Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

**Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

***Response:***

(i) Standards

Our repository work-flow and infrastructure are based on the ISO 14721:2012 (CCSDS 650.0-P-1.1) Open archival information system (OAIS) standard

Infrastructure

We benefit from the preferential bandwidth and access to the national internet backbone given to universities in South Africa. This means fast connectivity and minimal down-time for our site. The ICT department at the University maintains our infrastructure, including our webserver for our website and open data platform, and our secure server for the secure centre. The ICT department also keeps an inventory of our software, and system documentation.

Our open data platform software is developed by the World Bank's Development Data Group. We were initially the test site for the software and have contributed to additions such as the citations module. The Group is responsible for updates to the software.

***Reviewer Entry*****Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

**XVI. Security**

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

***Compliance Level:***

4 – The guideline has been fully implemented in the repository

***Reviewer Entry*****Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

**Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

***Response:***

(i) Open data infrastructure

Our technical infrastructure is managed by the IT department at UCT, with adherence to international cyber-security standards <http://www.icts.uct.ac.za/security>

(ii) Secure Centre

Because this is a facility unique to the institution, we drew up two documents which relate to information security for the Centre:

(a) a procedures manual <https://www.datafirst.uct.ac.za/files/2018-srdc02-procedures-manual-v2.pdf>

(b) an ISO information security compliance document <https://www.datafirst.uct.ac.za/services/secure-data-services>

***Reviewer Entry***

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

## **APPLICANT FEEDBACK**

### **Comments/feedback**

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

***Response:***

Applying for your certification is an excellent means of assessing our service, because it highlights service gaps, in particular, gaps in documenting what we do. It is a valuable exercise for repositories who want to adhere to international curation best practice.

*Reviewer Entry*

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept