

# **ilifu Research Data Management and Open Science Policy**

## **V 2.0 (September 2019)**

### **Preamble**

The ilifu Research Data Management and Open Science Project is part of an overall collaborative venture on the use of data intensive research facilities by researchers at six partner institutions in the Western and Northern Cape regions (University of Cape Town, Stellenbosch University, University of the Western Cape, Cape Peninsula University of Technology, Sol Plaatje University, and the South African Radio Astronomy Observatory (SARAO)).

The goal is to provide researchers and scientists with advanced research data management facilities which enables data processing, storage, curation, storage for re-use and sharing. The project is in response to global trends in Open Science, Open Access and Open Data initiatives, which call for the open sharing of research data to enable acceleration in the advancement of science, innovation and social development.

The project is funded through the Department of Higher Education, Science and Technology's (DHEST) DIRISA (Data Intensive Research Initiative of South Africa) to build a national cyberinfrastructure for researchers and scientists. The ilifu project is one of the first such regional data nodes under the DHEST initiative and will eventually link up with additional data nodes in the national cyberinfrastructure.

The ilifu Research Data Management and Open Science Project is responsible for the development of policies and guidelines on good data management and sharing practices. It is also responsible for a Work Integrated Learning programme for the main project.

These guidelines have been developed to provide support to researchers and scientists working/using the facility wishing to follow best practices and enable data visibility and its sharing. With the accompanying tools and services to be provisioned in the course of the project, such as the ilifu Data Catalogue/s, these generic guidelines provide initial support to aid better data discoverability. Furthermore, ongoing discipline based approaches will provide access to appropriate metadata schemas to enable visibility of the data.

## Document control

<b>Version number</b>	<b>Date</b>	<b>Author</b>	<b>Brief description of change</b>
0.0	27/6/2018	Alecia	Formatting incorporated theme outputs from Google doc in .docx
0.1	28/06/2018	Elisha, Debbie, Alecia	Consolidated, rewrote, edited
0.2	29/06/2018	Elisha, Alecia	Edited
0.3	01/07/2018	Elisha, Alecia	Edited
0.4	02/07/2018	Dale, Alecia	Edited
0.5	02/07/2018	Elisha, Alecia	Edited, formatted
0.5.1	02/07/2018	Elisha, Alecia	Edited
0.5.2	02/07/2018	Alecia	Formatted
0.5.3	25/01/2019	Wouter	Made changes to Section B and added Data Deposit Agreement appendix as Section E
1.0	12/01/2019	Niklas	Policy statement
1.1	21/02/2019	Niklas, Alecia, Debbie	Incorporation of researcher feedback of v.1.0. References
2.0	09/04/- 12/09/19	Nikki, Solomon, Mark Wouter, Ricardo, Debbie, Alfred, Sanjin	Incorporation of User Engagement feedback

## ilifu Research Data Management Policy Statement

1. The ilifu infrastructure is an actively managed space, with dedicated data curation staff, pipelines and workflows in place. This is to ensure consistency of data quality as well as comprehensive reporting on the ilifu grant funded project.
2. The ilifu infrastructure enables the storage and dissemination of research data, software and workflows ‘*as open as possible, as closed as necessary*,’ and in accordance with the FAIR principles<sup>1</sup>. This is in line with an increasing number of funder policy requirements, non-exhaustively listed in **Section E**.
3. The minimum expectation is for data *directly supporting* written research outputs to be shared openly<sup>2</sup>, and as close to the time of publication as possible. Furthermore, users are encouraged to also publish any other original pieces of the research process that are required for reproducibility, such as code or workflows according to the institutional policies of the researcher. Definitions of Openness are referenced in **Section C.8**.
4. It is expected that the data produced are collected and presented with the necessary attention to ethical practice, including, but not limited to: informed consent, data anonymization and protection as referenced in **Section C.6**.
5. It is expected that due consideration is made to enact intellectual property rights. Reuse of research data and code on the ilifu infrastructure requires the same accurate observation of referencing standards and IP stipulations as any other scholarly resource. More detail is provided in **Section A.7**.
6. Discipline-specific best practices and standards in data management planning are to be observed from the proposal stage, through the active research phase, to the managed preservation and publication of the research data expected, created and shared. Data Management Planning is referenced in **Section B.3**.

---

<sup>1</sup> Force11 (2014) The FAIR Data Principles [online] Available from: <https://www.force11.org/group/fairgroup/fairprinciples> [Accessed 12th February 2019]

<sup>2</sup> Open Knowledge International (no date) The Open Definition [online] Available from: <http://opendefinition.org/>

## **Table of Contents**

<b><i>Preamble</i></b>	<b><i>1</i></b>
<b>ilifu Research Data Management Policy Statement</b>	<b>3</b>
<b>Section A. Data Governance, Data policy and Infrastructure Service</b>	<b>5</b>
1. Intent	5
2. Scope	5
3. Infrastructure services	5
4. Data protection	5
5. End User agreements	5
6. Conditions of use	5
7. Intellectual property	5
8. License agreements	5
9. Data security	5
10. Data collection	6
11. Access to data	6
<b>Section B. Standards, interoperability, certification and archiving</b>	<b>7</b>
1. Intent	7
2. Scope	7
3. Research Data Management planning	7
4. Metadata	7
4.1 Requirements	7
4.2 Terms of use	7
5. Data deposit	7
5.1 Data deposit, quality and copyright	8
6. Data assurance	8
7. Storage	8
8. Version control and authenticity	8
9. Archiving	8
<b>Section C. Open Science, Open Access, Ethics, Legal Framework and Reuse</b>	<b>9</b>
1. Intent	9
2. Scope	9
3. Data discoverability	9
4. Data citation	9
5. Authorisation for Reuse	9
6. Ethics	10
7. End user agreement	10
8. Open science and Open access	10
<b>Section D: Glossary of terms</b>	<b>11</b>
<b>Section E: Resources</b>	<b>12</b>
1. Funder Data Policies (Non-exhaustive listing)	12
2. References	12
<b>Appendices</b>	<b>14</b>
A) ilifu Data Deposit Agreement	14
B) The ilifu Consortium Repository End User Agreement	16

## **Section A. Data Governance, Data policy and Infrastructure Service**

### **1. Intent**

The intent is to develop common policies and guidelines to guide researchers working on the ilifu infrastructure [see section A.3]

### **2. Scope**

The following guidelines have been developed to provide good research data management practices for researchers working on the ilifu infrastructure including accessibility, sharing and reuse of data. It includes coverage on Data protection, End-user agreements; Conditions of use; Intellectual Property and License agreements.

### **3. Infrastructure services**

The ilifu infrastructure is a shared and collaborative project which provides processing of big data, and its storage that requires researchers working on the platform to comply with its terms and conditions of use including data policies and guidelines.

### **4. Data protection**

The ilifu infrastructure will ensure the protection of researchers' data with regards to data processing, publication and sharing. Where researchers upload data for processing and storage, where such data may be sensitive as to require anonymizing it is the responsibility of the researcher to ensure such.

Researchers are responsible for the protection of personal data by anonymizing personal identifiers.

The protection of data needs to go beyond processing and sharing of data. The need for redundancy (backup) of the data in a secure environment needs to be included thereby reassuring users that the data is stored in at least one additional facility as opposed to one access site

Data privacy is a shared responsibility which requires that users have a basic understanding of the storage system capabilities as far as data sharing and data protection are concerned. [see Appendix A: ilifu Data Deposit Agreement]

### **5. End User agreements**

The reuse of data and end user agreement will depend on the choice of Creative Commons license selected by the researcher. Where researchers are not sure they may seek clarity from institutional research officers.

The agreement should provide some guidance regarding the CC license prior to researchers reaching out to institutional representatives. License requirements may already have been outlined by the funder / funding agency. Where researchers do not have at least some preliminary information regarding the choice of licences, the risk of selection of an inappropriate license might exist.

### **6. Conditions of use**

Researchers working on the infrastructure (and working within international collaborative projects) will ensure the protection of research subjects (human and animal) and comply with national and international legislative frameworks (e.g. POPI) and institutional IP guidelines.

## **7. Intellectual property**

Researchers using the ilifu facilities are governed by their institutional IP guidelines, as set out in Technology Transfer Office Guidelines, Patent contracts, etc.

## **8. License agreements**

The researcher is required to select the appropriate license agreement for access and reuse of their data. This selection should be in-line with ilifu compliant License agreements.

## **9. Data security**

The ilifu infrastructure implementation follows recognized industry standards in securing data collection and storage systems by selecting tried and tested interface systems such as Figshare. ilifu undertakes to make every effort to ensure the safety and security of both physical infrastructure and data it stores in the cloud infrastructure provided.

## **10. Data collection**

It is expected that researchers collect data in compliance with disciplinary ethical requirements before the data is processed, deposited and archived on the ilifu infrastructure. [see also Data protection]

## **11. Access to data**

The ilifu infrastructure serves the principles of Open Science. To protect the research process, all data published on the data management platform will be made publicly available as specified by the funding agency, not exceeding a period of two years. In the interim period, access to the data may be open (unrestricted), embargoed (available after a specified period of time), restricted (to a specified group) or closed (inaccessible to anyone other than the researcher). Any requests for access to restricted or closed data are to be routed to the researcher.

## **Section B. Standards, interoperability, certification and archiving**

### **1. Intent**

The intent is to ensure that researchers comply with national and international standards on data management, deposit, publishing and sharing. The ilifu infrastructure will ensure interoperability with other external systems. The facilities will be compliant with international certification standards on data certification.

### **2. Scope**

The coverage of this section on standards, interoperability and certification relates to metadata standards, data assurance, submission, deposit, storage, security versions, and authenticity.

### **3. Research Data Management planning**

All ilifu infrastructure users are required to provide a Data Management Plan (DMP) using the template of the approved funder. The DMP should accompany the written application to the ilifu resource allocation committee. An online version of participating institutional DMP tools will be made available for researchers.

### **4. Metadata**

#### **4.1 Requirements**

Sufficient metadata description is needed to aid in the discovery, accessibility and effective re-use of the datasets. The metadata must adequately describe the structure of the data and how it was created, as well as details of the owner and any legal restrictions.

Data that support academic publications must be adequately described according to the metadata standards provided by the data management platform.

The keywords “ilifu” and the name of the preferred domain-specific repository (e.g. IVOA<sup>3</sup>, EGA<sup>4</sup>), must appear as a metadata value in the record to allow for the exchange of metadata between ilifu and domain-specific repositories using the Figshare Open API.

#### **4.2 Terms of use**

Metadata is made publicly accessible on the ilifu infrastructure. The metadata may be re-used in any medium without prior permission for not-for-profit provided the DOI or a link to the original metadata record are given.

### **5. Data deposit**

Data deposits are to be guided by a metadata guide that outlines the preferred metadata schema for a specified field or institution. Uploading of data is accompanied by a data deposit form provided by the data management platform.

Uploaded data should be accompanied by an access license (e.g. Creative Commons license) detailing any restrictions on access. Data licenses may also outline any institutional, legal or regional and national restrictions including intellectual property and copyright restrictions.

The ilifu infrastructure will maintain appropriate data schemas that will enable the interoperability with respective domains of data (e.g. IVOA and EGA) via an Application

---

<sup>3</sup> International Virtual Observatory Alliance (IVOA). <http://www.ivoa.net/>

<sup>4</sup> European Genome-Phenome Archive (EGA). <https://www.ebi.ac.uk/ega/home>

Programming Interface (API) to the ilifu repository system, currently running on *Figshare for Institutions* SaaS. [see Section B. 4.1]

### **5.1 Data deposit, quality and copyright**

See Appendix A: ilifu Data Deposit Agreement

### **6. Data assurance**

No files containing personal data will be accepted. Researchers are responsible for the protection of personal data by anonymizing personal identifiers.

Data curators in the respective partner institutes will moderate deposited data to confirm usability and accuracy. Records corrected or amended in consultation with the depositor will be allocated a new version number. Changes or updates due to data correction will be recorded but previous versions will not be deleted.

File versions will be recorded in the metadata record up loaded with data files and will be reflected in the metadata record.

### **7. Storage**

The ilifu Resource Allocation Committee will allocate storage as per research requests and processing requirements.

### **8. Version control and authenticity**

The ilifu facility makes provision for versioning, control and authenticity of data being processed, stored and archived.

### **9. Archiving**

Users of the ilifu infrastructure are advised to make adequate provision in grant proposals for alternative storage to ensure the long-term archiving of data considered to be of persistent value.

## Section C. Open Science, Open Access, Ethics, Legal Framework and Reuse

### 1. Intent

The intent is to provide guidelines on the overall national open science framework requirements, open access, ethics, legal requirements for data storage and reuse.

### 2. Scope

The coverage of this section of the guideline will relate to data discoverability, data citation, reuse, ethics, end user agreements and access to the data.

### 3. Data discoverability

The facility will be compliant with international standards to enable discoverability e.g. DOI. ilifu realises that not all data that is discoverable can be freely accessed. Data can be subjected to embargoes, access controls, permissions or licences associated with data due to a variety of reasons such as confidentiality, reuse permissions, commercial interests.

### 4. Data citation

The facility encourages that data generated be discoverable and citable by means of the DOI provided by the data management platform. ilifu will adopt the use of data citation guidelines recommended for researchers by the Australian National Data Service (ANDS)<sup>5</sup>

### 5. Authorisation for Reuse

- Data reuse takes place when data which were used for a particular research project are reused in the future by other researchers.
- From the perspective of the future researcher(s), the data would amount to secondary data which have been collected.
- In contrast, the original researcher(s) would consider the data to be reused data.
- Data reuse is facilitated by license agreements which determine the authorisation for reuse.
- A license is a legal document that states how the data should be attributed, and for which purposes it may and may not be used.
- The type of license chosen for a dataset depends on two factors:
  - (a) The amount or proportion of secondary data in the 'new' dataset, and
  - (b) The terms of the original license under which those secondary data were acquired.
- All research data intended for reuse should have licenses included in their metadata records.
- For more information on licensing please see the section on license agreements.
- Data re-users are often not familiar with the secondary data which they come across.
- As such data re-users should take some time to read the metadata records associated with datasets in order to find out what they are allowed to do with the data.
- Below are some easy steps which can be followed by data re-users in order to determine the authorisation for the reuse of datasets:

Step 1: Locate a dataset you desire to reuse

Step 2: Access the dataset's metadata record

---

<sup>5</sup> Australian National Data Service (2015) *Data Citation for Researchers* [online] Available from: <https://www.ands.org.au/working-with-data/citation-and-identifiers/data-citation/data-citation-for-researchers>

- Step 3: Locate the license agreement, used for the dataset in question, within the metadata record. This should appear as part of the “Rights” statement of the metadata record.
- Step 4: Determine the permissions which are granted under the license (this can be done by clicking on the license icon which is linked to the terms of license).

## **6. Ethics**

This guideline aims to promote awareness of and compliance with ethical principles and procedures in the conduct of ilifu research activities, thereby clarifying for researchers their ethical obligations.

All research conducted under ilifu project involving interaction with or observation of human subjects, or information linked to human and animal subjects, or research involving groups of individuals, or organisations must go through a process of ethical screening and clearance.

Researchers working on the ilifu infrastructure are subject to various institutional ethics requirements, to the requirements of any funding agencies and national legislative frameworks, refer to the links below as appropriate.

- [UCT Policy for Responsible Research Conduct](#)
- [UCT Research Integrity Policies](#)
- [UWC Policy on Research Ethics](#)
- [CPUT Research Ethics](#)
- [Stellenbosch Policy for Responsible Research Conduct](#)
- [National Health Research Ethics Council](#)
- [South African Medical Research Committee Research Ethics Policy](#)
- [National Institute of Health Ethics Policy](#)
- [Guidelines on Good Research Practice - Wellcome Foundation](#)

## **7. Data Privacy**

All research conducted under the ilifu project must comply with the Protection of Personal Information Act (PoPIA)(and in case of collaborative international projects researchers must also comply with the necessary legislative frameworks), such that any personal information must be processed (a) lawfully; and (b) in a reasonable manner that does not infringe the privacy of the data subject (PoPIA, Chapter 3, Part A, 1.8, 1.9). Furthermore the processing of any personal must be limited to the scope for which it was intended and for which the subject gave consent. The reuse of any such data is compatible with the PoPIA act “as long as the information is used for historical, statistical or research purposes and the responsible party ensures that the further processing is carried out solely for such purposes...” (PoPIA, Chapter 3, Part A, Condition 4.15).

For the purposes of compliance with the PoPIA act, all researchers of ilifu working with human subjects and personal data are required to store and manage consent forms which include specific conditions around allowing for the possible re-use of data within the ethical practices of open science. These consent forms are to be kept in digital format with controlled access as per the research data management workflow of the relevant project.

Researchers are protected by Intellectual Property Rights from Publicly Financed Research and Development Act (Act No 51 of 2008) against the exploitation of intellectual property in instances where research outputs have a potential commercial value.

Researchers are further guided by the National Archives and Records Service of South Africa Act of 1996 (Act No 43 of 1996) in dealing with the proper management and care of the records of governmental bodies; and the preservation and use of a national archival heritage.

## **8. End user agreement**

[see section A.5 and Appendix B]

## **9. Open Science and Open Access**

The ilifu facility will be guided by the South African Open Science Framework<sup>6</sup>, Department of Science and Technology (2018) Draft White Paper on Science, Technology & Innovation<sup>7</sup> as well as the Copyright Amendment Bill. The Framework seeks to position South Africa to participate competitively and drive innovation in the 21st century big data economy and is strongly supported by the draft 2018 STI White Paper. The Copyright Amendment Bill also proposes a national Open Access policy. Open Access falls within the umbrella of Open Science, a movement encouraging researchers to share their processes, findings, and data in an ethical manner with the research community and beyond. Open Science recognizes the global benefits of researchers learning from and developing from each others work, but also acknowledges that such sharing should be as open as possible and as closed as necessary. This is to ensure the integrity of the research as well as the privacy of all research subjects. Ethical checklists, as practiced by the field of study are required to be completed prior to the sharing of data. With this regard, the Research Data Management and Open Science Policy encourages that all data generated under the umbrella of this project is findable, accessible, interoperable and reusable.

---

<sup>6</sup>S., Hodson, S., Walwyn, D., Wood, J. and Wright, C. (2019). SA-EU Open Science Dialogue Report. [online] Zenodo. Available at: <https://zenodo.org/record/2559469#.XG5kxs8zbR0> [Accessed 27 Feb. 2019].

<sup>7</sup> Department of Science and Technology. (2018) Draft White Paper on Science, Technology and Innovation [online] Available from: [https://www.dst.gov.za/images/2018/Draft-White-paper--on-STI-7\\_09.pdf](https://www.dst.gov.za/images/2018/Draft-White-paper--on-STI-7_09.pdf) [Accessed 27 February 2019]

## Section D: Glossary of terms

**API-** Stands for Application Programming Interface.

**Dataset-** A set of files containing both research data, and documentation.

**DMP-** Data Management Plan is a document which describes how data will be handled throughout the research project and after the completion of the project.

**Documentation-** Any digital files such as a codebook, technical report, methodology, workflows which explain the research data's production or use.

**FAIR-** Stands for Findable, Accessible, Interoperable, Reusable. *'Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals.'*<sup>8</sup>

**Metadata-** Metadata are a description of the data in order to help people to obtain and use the data. The description can be background information on the data, as well as information on how the data can be used. There are standards for describing whole datasets and there are also standards for describing entities within a dataset. Researchers are encouraged to make sure that they understand their own data and that others may find, use and properly cite the data. It is important therefore that researchers help to add metadata to the documents and data sets that they create.

**Open-** *'Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness).'* [source: Open Knowledge International (no date) *The Open Definition* [online] Available from: <http://opendefinition.org/>]

**Research data-** Are defined as factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings. A research data set constitutes a systematic, partial representation of the subject being investigated.

**RDM-** Stands for *Research Data Management*. Concerns the organisation of data, from its entry to the research cycle through to the dissemination and archiving of valuable results. It aims to ensure reliable verification of results, and permits new and innovative research built on existing information.

---

<sup>8</sup> Mark D. Wilkinson, Michel Dumontier [...] Barend Mons (2016) *The FAIR Guiding Principles for scientific data management and stewardship* *Scientific Data* volume 3, Article number: 160018 [online] Available from: <https://www.nature.com/articles/sdata201618>

## Section E: Resources

### 1. Funder Data Policies (Non-exhaustive listing)

Arts and Humanities Research Council (2019) *Research Funding Guide version 4.6*. [online] Available from: <https://ahrc.ukri.org/documents/guides/research-funding-guide1/> [Accessed 12<sup>th</sup> February 2019].

Bill and Melinda Gates Foundation (2019) *Open research data guidelines*. [online] Available from: <https://gatesopenresearch.org/for-authors/data-guidelines#opendata> [Accessed 12<sup>th</sup> February 2019].

Bill and Melinda Gates Foundation (2019) *Open research policies: article policies*. [online] Available from: <https://gatesopenresearch.org/about/policies> [Accessed 12<sup>th</sup> February 2019].

European Commission (2017) *Guidelines to the rules on open access to scientific publications and open access to research data in Horizon 2020 version 3.2* [online] Available from: [ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-pilot-guide\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf) [Accessed 12<sup>th</sup> February 2019]

Medical Research Council. (2016) *Data sharing policy*. [online] Available from: <https://mrc.ukri.org/documents/pdf/mrc-data-sharin-policy/> [Accessed 12<sup>th</sup> February 2019].

Medical Research Council. (2016) *Policy on open research data: clinical trials and public health intervention studies*. [online] Available from: <https://mrc.ukri.org/documents/pdf/mrc-policy-on-open-research-data/> [Accessed 12<sup>th</sup> February 2019].

National Environment Research Council (2019) *Data policy* [online] Available from: <https://nerc.ukri.org/research/sites/data/policy/data-policy/> [Accessed 12<sup>th</sup> February 2019].

National Environment Research Council. (2010) *NERC data policy*. [online] Available from: <http://www.nerc.ac.uk/research/sites/data/policy.asp> [Accessed 12<sup>th</sup> February 2019].

National Science Foundation (2019) *Dissemination and sharing of research results: data sharing policy* [online] Available from: <https://www.nsf.gov/bfa/dias/policy/dmp.jsp> [Accessed 12<sup>th</sup> February 2019].

Wellcome Trust. (2017). *Policy on data, software and materials management and sharing*. [online] Available from: <https://wellcome.ac.uk/funding/gidance/policy-data-software-materials-management-and-sharing> [Accessed 12<sup>th</sup> February 2019].

## Appendices

### A) ilifu Data Deposit Agreement

<sup>9</sup>In order to support a standard and consistent data management practice for research data, depositors are requested to acknowledge understanding of the following principles governing data deposit:

1. Items may only be deposited by accredited members of any of the ilifu partner institutions, or their delegated agents.
2. The Research Data Repository (RDR) administrator will only moderate deposits for the eligibility of authors/depositors, relevance to the scope of the repository, valid layout and format, and the exclusion of spam.
3. The validity and authenticity of the content of submissions is the sole responsibility of the depositor.
4. Content must be academic and may not be fraudulent, offensive or defamatory.
5. Items can be deposited at any time and in any file format, but will not be made publicly visible until the depositors release the data, e.g. when publisher - or funder-specified embargo periods have expired.
6. In making the deposit, the depositor confirms that they are the sole rights holder and/or they have received the right/permission from the rights holder(s) to publish the materials, i.e.:
  - 6.1 if the depositor is not the copyright holder and the material is copyrighted, the depositor acknowledges that they have received permission from the rights holder to publish the copyright material openly and have adhered to the conditions under which the copyright holder has granted publication rights e.g. acknowledged the original creator of the work and/or licensed the work under specified conditions
  - 6.2 where the depositor is not the sole IP creator, that written permission approving the deposit has been received from all the IP creators, and in the case of photographic or video material that the necessary image releases have been obtained from people appearing in the photographs or videos
7. ilifu may take down any publication without notice if any of the items contained within the publication are in breach of any of the above, or:
  - 7.1 If the data includes work that is found to be fraudulent, falsified, plagiarized or against the FAIR principles of Open Science
  - 7.2 The depositor has requested to have the data removed from the repository provided

---

<sup>9</sup> University of Cape Town. (2018). UCT terms of data deposit. [online] Available from: <http://www.digitalservices.lib.uct.ac.za/uct-terms-data-deposit>

that the request is accompanied by a reasonable motivation approved by the Research Data Management Governance Committee in terms of its mandate to approve exceptions to the public release of data for legal, ethical and commercial reasons

**8.** If ilifu or the depositor choose to remove the data from the repository, the metadata may be retained in the repository in order to trace the original DOI and indicate that the removal was deliberate. Furthermore, ilifu may retain a copy of the data for archival and reference purposes

I acknowledge and will adhere to the above principles of data deposit into the ilifu repository.

Signed

Name: \_\_\_\_\_

Date: \_\_\_\_\_

## **B) The ilifu Consortium Repository End User Agreement**

### **Definitions**

**Dataset** – A set or multiple sets of files or a database constituting a discrete and clearly defined set of data and/or the means of generating data related to a research activity or project, or supporting one or more research publications. It may include both data and the means to generate, interpret or validate data, such as computer models and software code. A dataset may include data files (including, where relevant, program files) and contextual supplementary documentation, such as the following: administrative materials, codebooks, user manuals, workflows, protocols, methodologies, laboratory notebooks and Read Me files.

**ilifu Consortium** – An inter-institutional regional data-intensive research facility comprised of the following institutions: Cape Peninsula Institute of Technology, University of Cape Town, University of Stellenbosch, University of the Western Cape, Sol Plaatje University and the South African Radio Astronomy Observatory.

**ilifu Consortium repository** – A searchable and query-able interfacing entity that will be used in order to store, manage and maintain datasets and metadata for the ilifu consortium.

**Metadata** – Accompanying information, either in a separate file or otherwise included in the dataset materials, about a particular dataset, including but not limited to the author's name, publishing date, title of data contents, description of contents, and other such related information.

**Metadata record** – A structured description of a dataset, including information about licensing and conditions of access to the data. Metadata records can refer to datasets held in the repository or elsewhere.

**Repository** – The ilifu consortium repository.

1. Ownership and copyright information for the datasets are stated explicitly in the Rights statement of each metadata record.
2. You agree to use the datasets only in accordance with the ilifu Consortium Repository End User Agreement. You agree to notify the ilifu consortium repository of any breach of its terms or of any infringements of the datasets of which you become aware as quickly as possible.
3. You will abide by the appropriate copyright and license statements applied to datasets and metadata.
4. You will ensure that means of access to data are kept secure and used only for appropriate purposes. In particular, passwords are personal and should not be shared.
5. Whenever you use a dataset you should, where possible, use the bibliographic citation recommended by the Repository, or an equivalent.
6. Use of the data on the ilifu consortium repository is at your sole risk. You agree not to use this Repository for any illegal or unlawful purpose. In particular, you will not use the datasets or metadata in a manner which infringes the law relating to copyright, confidentiality, privacy, data protection, defamation or similar or related doctrines. Moreover, you will not use or attempt to use the datasets or metadata to identify any

individuals from which a study sample may have been selected, nor may you claim to have done so.

7. The repository excludes, to the maximum extent permitted by law, all express or implied warranties of any kind in relation to any dataset or metadata; in particular, the repository shall not be liable for any loss or damage;
  - a. Which may be suffered or incurred by you or a third party in respect of the use by you of any dataset or metadata, or
  - b. Which may arise directly or indirectly in respect of the use by you of any dataset or metadata.
8. Any breach of this ilifu Consortium Repository End User Agreement will lead to the immediate and automatic termination without notice of your access to the services, and could result in legal action against you.
9. This Agreement shall be governed by and interpreted in accordance with the laws of the Republic of South Africa (excluding the conflict of laws rules thereof). All disputes under the ilifu Consortium Repository End User Agreement will be resolved in the relevant South African court(s). You consent to the jurisdiction of such courts and waive any jurisdictional or venue defenses otherwise available.

By ticking "I agree" below you are agreeing to the terms and conditions of use as outlined above.

- I agree to the terms of the End User Agreement.
- I do not agree to the terms of the End User Agreement.