



**UNIVERSITY of the
WESTERN CAPE**

**RESEARCH DATA MANAGEMENT POLICY
Draft 1.6 November 2020
Approved Version - Approval Date**

Policy name	University of the Western Cape Research Data Management Policy
Executive responsible	Deputy Vice Chancellor: Research & Innovation
Policy drivers	Director: eResearch Director: Research Development Director: Library Services
Governance	Senate Research Committee
Policy life-cycle	Every three years
Relevant Legislation	Data Protection Act of 1998 Patents Act 57 of 1978 Copyright Act 98 of 1978 Competition Act 89 of 1998 Electronic communication and transaction act 25 of 2002 Intellectual Property Rights from Publicly Financed Research and Development Act 51 of 2008 Protection of Personal Information Act 4 of 2013
Institutional policies	USAf POPIA Industry Code of Conduct: Public Universities Intellectual Property Policy Research Ethics Policy Open Access Policy Records and Archives Policy Institutional Operating Plan (IOP)

Table of contents

1. Introduction	3
2. Purpose Statement	3
3. Policy Scope	4
4. Principles	4
5. University Responsibilities	5
6. Researcher Responsibilities	5
7. Appendices	6
A: Data Governance and Standards	6
FAIR Data	6
Data Management Plan (DMP)	6
Data Depositing	6
Metadata Standards	6
Storage	7
Version Control	7
Data Discovery and Citation	7
Ethics and Re-use	7
B: Glossary of terms	8
Data Management Plans (DMPs)	8
Dataset	8
Data Repository	8
De-Identification (Anonymisation)	8
Digital Object Identifier (DOI)	8
Documentation	8
FAIR Data	8
Metadata	8
Open Access (OA)	9
Open Data	9
ORCID	9
Research Data	9
Research Data Management (RDM)	9
C: Resources	9
8. Policy Formulation Process	10

1. Introduction

Research data underpins research outputs, and sharing and preserving this data within a coherent and functional research data management (RDM) strategy is integral to research and institutional integrity. Recent advances in information technology and digital science have changed the broader research perspective, and the amount of data being generated has increased, and continues to increase radically. This has necessitated a need to ensure that research data is managed appropriately to guarantee accessibility and security.

The University of the Western Cape recognises that good practice in research data management and open access to research data is integral to quality research. Moreover, good RDM practices coupled with openness help protect intellectual and financial investments; support verification of research results and enable innovative research.

Access to research data from UWC will not only highlight the excellence of this institution's research output, but also potentially foster collaborations and enhance public engagement. Making research data accessible maximises its impact, whether academic, cultural, social or economic.

Researcher concerns arising from intellectual property rights and authenticity and ownership of data, have underlined the need for data storage and preservation. Thus, research data, including software that was developed as part of the research project (if necessary to validate research findings), must be stored in a secure environment, with sufficient detail (metadata) and documentation. Metadata will enable discoverability and re-use of research data. All stakeholders, national and international funders, academic publishers and other related organisations, require research data to be managed in a transparent manner.

In South Africa, the National Research Foundation (NRF) requires that data supporting publications be deposited in an accredited open access (OA) data repository with a registered Digital Object Identifier (DOI) for citation and referencing purposes. This was mandated in March 2015. The general expectation is that publicly funded research data constitute a public good, and should be made openly available.

This policy must be read together with other sections of the UWC Research Policy.

2. Purpose Statement

The purpose of this policy is to facilitate the management of research data at the University of the Western Cape, and to ensure that all research data generated at the University are managed and curated to support the institution's aspirations of a research-intensive university.

This policy aims to encourage good practice in RDM across the institution via a framework which details processes, roles and responsibilities, support services and facilities available for this purpose.

3. Policy Scope

The policy is applicable to all employees and students associated with the University of the Western Cape by formal contract, full-time or part-time, who are conducting research/ projects. This policy is applicable to all research data collected in the course of formal and informal data collection during the lifespan of the research.

This policy provides a framework for the University to provide research data management (RDM) regime based on the principles of reproducible research; data availability/openness; data re-use as well as recognition of legitimate or acceptable limitations of data availability/openness.

4. Principles

The following principles are adopted as best practices in RDM.

All data collected is subject to the POPI Act of 2013, particularly Section 19 (Use of Identifiable Personal Information in Scientific Research).

All research conducted at the University of the Western Cape must comply with this Policy, to ensure that:

4.1 Research data generated through public funding are a public good, produced in the public interest and should therefore be made openly available, with as few restrictions as possible, and in a timely and responsible manner. Research data should therefore be 'as open as possible, as closed as necessary'. Openness aids in accelerating research and fostering the re-usability of data, but at the same time, it is recognised that being closed is sometimes necessary in order to safeguard privacy. This is of particular importance with regard to confidentiality and protection of personal information (POPI Act);

4.2 Identifiable personal data will be anonymised or de-identified in order to remove direct or indirect identification;

4.3 Data management plans (DMPs) are created at the onset of a research project, and RDM happens throughout the life-cycle of research projects. DMPs are to be duly declared in accordance with funder requirements and/or faculty or departmental guidelines;

4.4 Data underpinning Masters and PhD research projects should be deposited in the Institutional Research Data Repository at graduation;

4.5 Data with acknowledged long-term value should be appropriately preserved and remain accessible and usable for future research;

4.6 Research data is owned by the University unless otherwise agreed to by the University in terms of a research contract concluded with a funder;

4.7 Sufficient metadata should be recorded and made openly available to enable other researchers to discover, understand, re-use and cite;

4.8 Published research results should include information on how to access the supporting data;

4.9 UWC recognises that there are legal, ethical and commercial constraints on the release of research data. Confidential research data may be stored in the Institutional Research Data Repository, but will not be made public until cleared by the appropriate authority **and, in the case of personal information, with explicit consent from participants;**

4.10 Researchers may be entitled to a limited period of privileged use of the data collected to enable them to publish;

4.11 All users of research data should acknowledge the sources of their data and abide by the terms and conditions under which they are accessed;

4.12 Research data collected as part of a joint research initiative may be stored in the Institutional Research Data Repository, if the external party does not have access to an appropriate facility. Where externally owned or stored data is required for reproducible research or verification of the results and findings, the metadata is to be stored in the Institutional Research Data Repository;

4.13 **Research datasets are required to be stored for a minimum of five (5) years after completion of the original project. Where intellectual property is involved,** or if there are particular statutory or contractual requirements, a longer period may be required. In some cases, and in particular where research involving human subjects is concerned, funding bodies may require that all raw data be kept indefinitely.

5. University Responsibilities

The University of the Western Cape via the Deputy Vice-Chancellor: Research & Innovation is responsible for policy, implementation, compliance with funder requirements; advocacy & awareness as well as providing an enabling environment for research data stewardship and governance.

UWC Senate Research Committee is responsible for ensuring that research data management procedures are in place to protect the University against risks as well as monitoring compliance by the researchers in their respective faculties and research units.

6. Researcher Responsibilities

Researchers are responsible for documenting data storage and management requirements in their data management plans; comply with funder policies and deposit data timeously with the Institutional Research Data Repository or domain-specific data infrastructures or other trustworthy repositories.

Principal Investigators (PIs) are accountable for ensuring that:

- All participants register for and use an ORCID when depositing data;
- Research groups implement good data practices;
- Data is handled in accordance with UWC policies, legal, ethical, regulatory and contractual obligations;

- A data management plan is in place for the life-cycle of the research/project.
- **Data management practices are compliant with the POPI Act.**

The Research Development, Library and eResearch Units are responsible for providing guidance, advocacy, disposal and capacity building to support good practices in data management and open science in general.

7. Appendices

A: Data Governance and Standards

FAIR Data

UWC is committed to the FAIR open data principles which state that data should be 'Findable, Accessible, Interoperable and Reusable'. This international standard for Open Data and Open Science informs the RDM Policy.

Data Management Plan (DMP)

Researchers are encouraged to submit a data management plan (DMP) that will outline the life-cycle of the research data for each project. This is required at the inception of the research project. UWC researchers will be provided with a DMP tool or use the funder's DMP template if prescribed.

Data Depositing

Researchers and principal investigators (PI's) are responsible for depositing their research data, validity and authenticity of the content submitted. Researchers can deposit data at any time but only make it publicly visible once it is permissible to do so (or when publisher or embargo periods have expired).

By depositing their data, researchers confirm their sole rights to the content, or that they have acquired permission from the owners to deposit the content. They must acknowledge having received permission and adherence to copyright if not the sole owner of the content. Deposited research data with Intellectual Property attached to it will not be made public and will be subject to the policies of the UWC Technology Transfer Office (TTO).

Metadata Standards

Data deposits are guided by the metadata schema used in the Institutional Research Data Repository. Researchers depositing to discipline-specific data repositories, metadata schemas for such repositories and disciplines will apply.

Appropriate content description should be made to allow discovery, accessibility and re-use of datasets. Content on the subject, ownership and any legal restrictions must be clear. Data must be described adequately following the metadata standards provided by the data management platform and / or discipline specific metadata schema used.

Storage

It is the researcher's responsibility to ensure that materials generated and collected from their research activities are stored securely in a durable and accessible format and in a manner that ensures its authenticity and integrity as well as meeting all legal and confidentiality requirements.

Non-digital research data must be retained securely within the department; in the researcher's own office or the laboratory in which they were generated. In cases where the data is confidential it must be secured appropriately in compliance with ethical and legal requirements.

Version Control

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

UWC researchers can also ensure that digital files are well-structured internally. Appropriate file naming convention should be applied by simply adding document titles, authors and their contact details, dates, version control information, and column and row labels for tables and spreadsheets increase the ability of research data to be discovered, managed and interpreted over time.

Data Discovery and Citation

The research data management platform will comply with international standards in using persistent digital identifiers (such as the DOI and the ORCID researcher identifier) to make data discoverable and citable. UWC will adopt the use of data citation guidelines recommended for researchers by the Australian National Data Service (ANDS).

Ethics and Re-use

All research conducted at UWC 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. UWC Researchers are subject to institutional ethics requirements, to the requirements of any funding agencies and to national legislative frameworks.

Data re-use takes place when data which were used for a particular research project are re-used in the future by other researchers. Data re-use is facilitated by license agreements which determine the authorisation for re-use. 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.

B: Glossary of terms

Data Management Plans (DMPs)

DMPs are organic documents that describe how data will be managed during the life-cycle of the research project. DMPs are often requested by funders at the bid stage and should state what data will be created and how, as well as outlining plans for sharing and preservation. Any restrictions on access to the data should also be noted along with mechanisms to protect unauthorized access.

Dataset

A structured collection of data associated with a research project.

Data Repository

A central secure digital warehouse, where data and associated metadata are stored and maintained.

De-Identification (Anonymisation)

The process of severing the link between information and a data subject (natural persons). In the case of de-identification (or anonymisation), the POPI Act no longer applies. Information is considered de-identified if the information cannot be re-identified by any reasonably foreseeable method or through linking it with other information that identifies the data subject. In this way, de-identification is distinguished from masking or pseudonymisation.

Digital Object Identifier (DOI)

A DOI is a persistent link to a digital object. This can include various aspects of research data, including a research paper, or a dataset or only metadata. All research data that is published is assigned a DOI. UWC's IDR, Kikapu, assigns DOIs to published research data.

Documentation

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

FAIR Data

For data to meet the international Open Access guidelines of FAIR, it needs to be findable, accessible, interoperable and reusable.

Metadata

Enables and enhances the discovery and re-use of data through structured human and machine-readable information about the data. It is a description of the data that helps users to discover and use the data. Metadata includes background information on why and how the data was collected, as well as software used and how it can be used and shared. There are standards for describing whole datasets and there are also standards for describing entities within a dataset. Metadata is essential in making research data discoverable.

Open Access (OA)

Immediate, online, freely available and accessible research outputs. Open Access refers to a publishing model for scholarly communication and publishing that makes research information available online at no cost and free from restrictions like copyright.

Open Data

Data that is freely accessible to everyone to use and share without restrictions or mechanisms of control.

ORCID

ORCID (Open Researcher and Contributor ID) is a non-proprietary and platform-agnostic alphanumeric code that uniquely identifies individual academic researchers and contributors.

Research Data

Data collected, securely stored and used as a primary source for research in scholarship across all academic disciplines. Research data can be quantitative or qualitative, and include any information created, collected or generated to validate original research findings. In this context, Research Data is understood to include Research Software adopted and/or developed as part of the research project if this is necessary to validate research findings, such as those published in a journal paper. It is thus good practice to make available research software and adequate documentation to enable others to more easily validate your research findings.

Research Data Management (RDM)

RDM encompasses all aspects of creating, collecting, organising, housing, delivering, maintaining, archiving, preserving, sharing and managing research data.

C: Resources

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

Australian National Data Service (ANDS). *Data Citation*. [online] Available from: https://www.ands.org.au/_data/assets/pdf_file/0005/724334/Data-citation.pdf

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

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

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

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

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. *The FAIR Guiding Principles for scientific data management and stewardship*. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

8. Policy Formulation Process

	Contributors	Details	Date
1	A Nqotole, M Snyders, C Hart and G Davis	Initial draft	18/01/2019
2	A Nqotole, P Zimani & S Neerputh	Revised draft	28/04/2019
3	M Vaccari and B Fielding	Comments & Ownership	06/05/2019
	Reviewed by Senate Library Committee	Comments/Feedback	14/05/2019
4	Review by Senate Research Committee	Comments/Feedback	13/09/2019
5	Review by Senate Executive Committee	Referred back for wider consultation	15/10/2019
6	Reviewed by REC Members	Comments/Feedback	04/11/2019
7	Deans, Deputy Deans of Research, SARChI Chairs, DVC: Academic, REC Chairs	Comments/Feedback	11/03/2020
8; 9	M Vaccari, M Snyders, A Nqotole & S Schafer	Comments/Feedback	20/07/2020; 30/09/2020
10	Senate Meeting	Discussion	05/11/2020
11	All	Comments/Feedback	25/11/2020
12	Senate		2021