**MEMORANDUM OF UNDERSTANDING**

Between:

DataFirst, University of Cape Town, and:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ [Insert name of Party B]

**1. Purpose and Scope**

The purpose of this MOU is to clearly identify the roles and responsibilities of each party as they relate to DataFirst curating and publishing research micro-datasets deposited by [Insert name of Party B]

**2. Background**

DataFirst’s research data service, based at the University of Cape Town, gives provides long-term open access to anonymised research data on South Africa and other African countries. Our data holdings include survey and administrative microdata. Ours is the only research data repository awarded the [CoreTrustSeal](https://www.coretrustseal.org/why-certification/certified-repositories) international Trusted Data Repository Certification.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ [Insert name of Party B]

[Provide a brief description of Party B] wishes to collaborate with DataFirst to make their data available to the research community as [select all that apply]:

Public use data (Creative Commons –attribution only license)

Research use data (Creative Commons – non-commercial use only license)

Restricted-access use data (Secure Research Data Centre data[[1]](#footnote-1))

**3. Responsibilities under this MOU:** DataFirst

DataFirst shall undertake the following activities related to data deposited by [Insert name of Party B]:

1. Quality check the data and ensure the data is anonymised for public access
2. Disseminate the data in compliance with national data legislation[[2]](#footnote-2) and research ethics requirements
3. Make the data and documents available in the long term
4. Publish informative metadata with the data
5. Provide [Insert name of Party B] with data quality feedback from data users
6. Provide [Insert name of Party B] with regular statistics on the use of their data

**4. Responsibilities under this MOU**: [Name of Party B]

[Party B] shall undertake the following activities:

1. Deposit data with DataFirst to be shared with the wider research community, with no embargo period/ after an embargo period of [Insert number of months] (up to 18 months)
2. Confirm they are the data owners or have permission from the data owners to deposit the data with DataFirst
3. Provide all data collection and data analysis documentation
4. Accept DataFirst’s advice on data that may not be suitable for sharing

**5. It is mutually understood and agreed by and between the Parties that**:

1. Ownership of the data remains with the depositor

2. This MOU may be adapted when necessary with the agreement of both parties.

3. This agreement may be terminated at the request of either party.

4. This MOU includes the reimbursement of funds between the two parties

6. **Effective Date and Signature**

This MOU shall be effective upon the signature of authorized officials of DataFirst and [Party B].

The Parties indicate agreement with this MOU by their signatures.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Name: DataFirst’s Representative

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Signature: DataFirst’s Representative

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Date

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Name: Representative of Party B

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Signature: Representative of Party B

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Date

**Addendum 1. How we Collaborate to Share Data**

**What we do**

DataFirst has responded to the demand for African data by establishing an indigenous data service, based at the University of Cape Town. The data service gives researchers access to anonymised disaggregated data from African governments and research projects via our [online data repository](https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/central/about).

We handle the entire data life-cycle, including data transfer, quality checks, disclosure control (anonymisation), and data description, and publish data under clear access and use licenses, including [Creative Commons](https://creativecommons.org/licenses/) licensing. Our work-flows can be viewed online on our [data life-cycle model](https://www.datafirst.uct.ac.za/images/docs/20150618-microdata-service-model.pdf).

**Our Trustworthiness**

DataFirst’s online data repository is certified as a trusted data repository by the [CoreTrustSeal](https://www.coretrustseal.org/) certification. It is the only Africa-based repository with this international certification.

We preserve and share data in compliance with national data legislation, such as [POPI](https://www.justice.gov.za/inforeg/docs/InfoRegSA-POPIA-act2013-004.pdf). We assist depositors to comply with funders’ data sharing stipulations, as well as research ethics and data privacy norms. Our data is shared according to [open data principles](https://sunlightfoundation.com/policy/documents/ten-open-data-principle) (free, online, complete, primary) or as non-commercial use data, in line with international data curation standards and best practice.

Where data is too sensitive or disaggregated to share online, we assist research teams to share their data in our [Secure Research Data Centre](https://www.datafirst.uct.ac.za/services/secure-data-services) at the University of Cape Town. Researchers need clearance to access the Centre. Research from the Centre investigates phenomena at the local level, which is vital information for policymakers.

**Our Depositors**

We work in partnership with government ministries and other data producers to preserve and share their data in the long term, with academics, students, and government research teams. We currently have data sharing agreements with, among others:

* City of Cape Town Municipality
* Department of Basic Education
* Statistics South Africa

We also partner with South African research institutions to rescue and publish their historical South African data that is at risk of being destroyed, because the data is paper-based and/or its stewardship is under threat.

Data depositors get the following benefits from partnering with us:

* Dissemination according to international standards
* A branded page on our data site
* Long-term preservation of their data
* Background and usage metadata
* Regular data usage citations and statistics
* Support to users of their data via [our support site](http://support.data1st.org/).
* Data quality feedback from DataFirst and data users

To partner with DataFirst to preserve and publish your data, contact [lynn.woolfrey@uct.ac.za](mailto:lynn.woolfrey@uct.ac.za)

**Addendum 2. How our Repository Complies with POPIA**

**Sharing Anonymised Data**

DataFirst complies with South African legislation, specifically South Africa’s [Protection of Personal Information Act](https://www.justice.gov.za/inforeg/docs/InfoRegSA-POPIA-act2013-004.pdf) (POPIA) of 2013 (Information Regulator, South Africa, 2013). POPIA governs the retention and further processing of personal information on data subjects. The Act defines ‘‘personal information’’ as “information relating to an identifiable, living, natural person, and where it is applicable, an identifiable, existing juristic person (Information Regulator, South Africa, 2013, p. 15).” That is, it governs the holding and processing of personally identifiable information. Thus, the terms of POPIA do not apply to distribution of anonymised information, where a data subject cannot be identified in the data. DataFirst complies with POPIA by sharing only anonymised data online.

Data deposited with us for sharing is generally already anonymised by the data producers. DataFirst always undertakes disclosure control on the data prior to dissemination, to confirm that the data has been confidentialised. DataFirst does also at times receive data files that contain personally identifiable information. The personal information must be held by DataFirst and further processed to anonymise and otherwise prepare the data for research re-use. The retention and processing of this information is done in according with POPIA 2013 legislation. Details follow on our compliance with POPIA when holding and processing the information.

**Data Retention**

POPIA 2013 restricts the retention period for the holding of information (to only as long as “necessary for achieving the purpose for which the information was collected”) (Condition 3 Clause 14 (1)) (Information Regulator, South Africa, 2013, pp. 31-32). However, the Act makes provision for long-term retention of personal information “for historical, statistical or research purposes” provided the responsible party holding the information puts safeguards in place to ensure the information is used just for these purposes (Condition 3 Clause 14 (2)) (Information Regulator, South Africa, 2013, p. 32). DataFirst has such safeguards in place. The information is held on our secure server with information security protocols in place and with controlled access for staff only.

**Data Processing**

POPIA requires that further processing of personal information must be compatible with the purpose for which the data was collected (POPIA 2013 Condition 4 Clause 15 (1) pp 33-34. But it waives this requirement when information is used for research purposes and the responsible party ensures that “further processing is carried out solely for such purposes and will not be published in an identifiable form” (Condition 4 Section 15 (3) (e)) (Information Regulator, South Africa, 2013, p. 34). POPIA makes the same exceptions for the prohibition of processing of personal information on children (Part C Clause 35 (1) (d) (ii) and (e)) (Information Regulator, South Africa, 2013, p. 48). This type of Personal information deposited with DataFirst is only held by DataFirst for the purposes of creating de-identified data for public distribution. The identifiable information is not published or shared beyond DataFirst staff for the purposes of data cleaning and anonymisation.

**Processing of Special Personal Information**

POPIA also requires consent from data subjects to process special personal information[[3]](#footnote-3) about them, but also waives this requirement if the information is processed “for historical, statistical or research purposes”, and “where seeking consent from data subjects would involve disproportionate effort” and where “guarantees have been put in place to ensure information processing does not negatively affect the privacy of data subjects (POPIA 2013 Part B Clause 27 (d) (i) (ii)) (Information Regulator, South Africa, 2013, pp. 42-43).” DataFirst has the same safeguards in place for special personal information as for other personal information, which is that we house the information securely and process the data only to create anonymised versions for public distribution, and do not publish the personal information.

**Addendum 3. How Our Secure Research Data Centre Complies with POPIA**

DataFirst complies with South African data legislation, specifically South Africa’s [Protection of Personal Information Act](https://www.justice.gov.za/inforeg/docs/InfoRegSA-POPIA-act2013-004.pdf) (POPIA) of 2013 (Information Regulator, South Africa, 2013). POPIA governs the retention and further processing of personal information on data subjects. The Act defines ‘‘personal information’’ as “information relating to an identifiable, living, natural person, and where it is applicable, an identifiable, existing juristic person (Information Regulator, South Africa, 2013, p. 15).” That is, it governs the holding and processing of personally identifiable information. This addendum describes our compliance with POPIA when sharing potentially disclosive data in our Secure Research Data Centre at the University of Cape Town

**Background**

DataFirst has responded to demands from researchers for highly disaggregated data by following the example of data services worldwide and creating a “safe room” at the university for this type of data, our Secure Research Data Centre. The data shared in this safe room does not include direct identifiers. However, because it has low-level geographic variables, there is a possibility that individuals may be identified in the data. The service has therefore been set up to prevent the risks of respondents’ identities being revealed and published. Procedures in the Centre are compliant with POPIA restrictions on the handling of personal data, as explained below.

**Data Retention**

POPIA 2013 restricts the retention period for the holding of information (to only as long as “necessary for achieving the purpose for which the information was collected”) (Condition 3 Clause 14 (1)) (Information Regulator, South Africa, 2013, pp. 31-32). However, the Act makes provision for long-term retention of personal information “for historical, statistical or research purposes” provided the responsible party holding the information puts safeguards in place to ensure the information is used just for these purposes (Condition 3 Clause 14 (2)) (Information Regulator, South Africa, 2013, p. 32).

**Data Processing**

POPIA requires that further processing of personal information must be compatible with the purpose for which the data was collected (POPIA 2013 Condition 4 Clause 15 (1) pp 33-34. But it waives this requirement when information is used for research purposes and the responsible party ensures that “further processing is carried out solely for such purposes and will not be published in an identifiable form” (Condition 4 Section 15 (3) (e)) (Information Regulator, South Africa, 2013, p. 34). POPIA makes the same exceptions for the prohibition of processing of personal information on children (Part C Clause 35 (1) (d) (ii) and (e)) (Information Regulator, South Africa, 2013, p. 48).

POPIA also requires consent from data subjects to process special personal information[[4]](#footnote-4) about them, but also waives this requirement if the information is processed “for historical, statistical or research purposes”, and “where seeking consent from data subjects would involve disproportionate effort” and where “guarantees have been put in place to ensure information processing does not negatively affect the privacy of data subjects (POPIA 2013 Part B Clause 27 (d) (i) (ii)) (Information Regulator, South Africa, 2013, pp. 42-43).” DataFirst has the same safeguards in place for special personal information as for other personal information,

DataFirst’s Secure Centre has all safeguards in place to protect the personal information described above, during processing of this information to create partially de-identified versions for analysis in the Secure Centre. Safeguards include designing a safe space for the use of this data, hosting the data securely off-site and accrediting researchers to access the Centre.

**The Centre**

The design of the Centre is in line with international standards for secure research data facilities. The physical infrastructure includes boundary protection devices which control access with a biometric (fingerprint) reader. Surveillance and safety are covered by a CCTV camera, an intercom and a break-glass fire system.

**Data Protection**

The restricted data is hosted on a secure server in a monitored server room, with access control and a security gate. The server is set up to have two networks, one for the Centre and the other for Intranet access for backups. A UPS protects the server from power outages. The PCs in the Centre are locked down with software: Their BIOS’s are password secured, their external ports are disabled, and network properties shut off using Group Policies. All cabling is hidden and the switch for cabling to the server secured within a locked cabinet within the Centre. Mobile devices and laptops are required to secured in lockers in the Centre. Researchers will be required to authenticate at login with a username and password. These will not be able to be changed, and will be deleted once their access period ends.

**Researcher Accreditation**

In line with international best practice, the Centre administrator accredits users with “safe” researcher status. Researchers at research and educational institutions are eligible to apply to access data in the Centre. When applying, the researchers must provide information on themselves and their institution, and their proposed research in the Centre, including outputs. They must also provide reasons why their research cannot be undertaken with the public access versions of the data. They also sign a confidentiality agreement to protect respondents’ identities. Further information can be found on our [Secure Research Data Centre webpage](https://www.datafirst.uct.ac.za/services/secure-data-services).

1. This option is for data that is potentially disclosive because it has secondary identifiers or detailed geography [↑](#footnote-ref-1)
2. In SA, the Protection of Personal Information Act, 2013 makes provision for long-term retention and processing of personal information “for historical, statistical or research purposes” provided the responsible party holding the information puts safeguards in place to ensure the information is anonymised or shared for research purposes only. DataFirst has such safeguards in place. Addendum 1 elaborates on these safeguards. [↑](#footnote-ref-2)
3. Special personal information is defined by the Act as information on the someone’s religious/philosophical beliefs, race/ethnic original, trade union membership, political persuasion, health/sex life, and criminal behaviour, as well as biometric information (Information Regulator, South Africa, 2013, p. 42). [↑](#footnote-ref-3)
4. Special personal information is defined by the Act as information on the someone’s religious/philosophical beliefs, race/ethnic original, trade union membership, political persuasion, health/sex life, and criminal behaviour, as well as biometric information (Information Regulator, South Africa, 2013, p. 42). [↑](#footnote-ref-4)