

Policy 1 - Data Sharing Policy

This document should be read in conjunction with the HSRC Data Preservation Policy, the HSRC Data Curation Standard Operating Procedure (SOP), the HSRC Project Management Framework, as well as applicable IT related policies and procedures.

1. Introduction

The HSRC supports data sharing and therefore ensures that data is managed, organised and preserved for further reuse by the HSRC, as well as by the wider scientific and stakeholder community.

Apart from being good research practice, the importance of data sharing is emphasised by the HSRC Act (Act No. 17 of 2008) which states that the HSRC is to “...develop and make publicly available new data sets to underpin research...”.

The HSRC also subscribes to the central principles of the Organisation for Economic Co-operation and Development (OECD)’s report “Promoting Access to Public Research Data for Scientific, Economic and Social Development” that advances open access to data yielded by publicly-funded research. According to this initiative, data arising from publicly-funded research is viewed as a public good, produced in the public interest, and should be openly available to the maximum extent possible. This is in line with the requirements of the Intellectual Property Rights from Publicly Financed Research and Development Act (Act No. 51 of 2008) which requires that “...intellectual property emanating from publicly financed research and development is identified, protected, utilised and commercialised for the benefit of the people of the Republic..”.

In addition, the HSRC recognises that an increasing number of donor organisations require that access be granted to data emerging from research funded by them. It is also becoming standard practice for peer reviewed journal publishers to require access to data linked to research outputs for scrutiny to address the risk of data falsification. The HSRC promotes the Responsible Conduct of Research and as such supports the sharing of its research data.

2. Purpose of the Data Sharing Policy

This policy aims to provide a broad framework for the sharing of HSRC-generated data.

3. Scope the Data Sharing Policy

This policy applies to:

- The sharing of HSRC-generated research data – whether qualitative or quantitative in nature – as it stands at project completion;
- Data yielded by publicly-funded research;

- Data that is privately funded, subject to the approval of the data owner ;
- Data, as well as the related documentation that provides the context or provenance of the data.

4. Policy statements

4.1 The majority of HSRC data will be made available for sharing

The HSRC facilitates access to the research data it generates. It acknowledges, however, that circumstances may arise in which data cannot be made available for reasons including intellectual property right and ethical considerations. The HSRC applies mixed levels of access regulation, combining restricted access with open access. These permissions are applied as appropriate for the applicable target audiences.

4.2 The HSRC assumes copyright, unless otherwise negotiated

Intellectual property rights of all data gathered using Parliamentary grant moneys will be handled according to the stipulations of the Intellectual Property Rights from Publicly Financed Research and Development Act (Act No. 51 of 2008). This implies that copyright for all data gathered using Parliamentary grant moneys will be retained by the HSRC.

Copyright to data gathered on behalf of clients will be negotiated as part of project contracts. In these cases joint ownership is preferred, but if this can't be achieved, permission must be sought for the secondary use of the data by HSRC researchers.

If original contracts did not address these issues or negated the HSRC's intellectual property rights, negotiations must be re-opened retrospectively with clients or funders to agree on equitable terms.

4.3 Data will be shared in a timely manner

As an organisation that impacts on policy, the HSRC expects the timely release and sharing of data. However it also accepts that there are a number of variables at play and that what constitutes timely release will depend on the type of data, as well as the nature of the research being undertaken. The HSRC recognises that those researchers who collected the data have a legitimate interest in benefiting from their investment of time and effort. Accordingly, such researchers may benefit from first and continuing use, but not from prolonged exclusive use.

Data must be made available for sharing within 12 to 24 months after the official completion date of the project concerned, acknowledging that the specific time will be influenced by the nature of the data collected. Where data from large epidemiologic or

longitudinal studies is collected over several discrete time periods or waves, the data should be shared in waves as these become available.

HSRC researchers interested in the secondary use of the data before the official release date can apply for access that will be negotiated with the ED and Principal Investigator (PI) / project leader concerned.

Anyone making use of the data within three years of its being disseminated will be required to acknowledge the PI / project leader and/or co-workers (as appropriate) on any resulting outputs should the PI / project leader and/or co-workers so desire. This provision notwithstanding, any output based on HSRC-generated data must acknowledge the original data source, the data owner and the HSRC appropriately.

4.4 Data sharing will follow an official, managed process

Appropriate methods for data sharing must be employed so as to ensure adherence to regulations in relation to security, intellectual property rights and confidentiality. Access must be managed via official end user license agreements. Sharing of data informally is not permissible.

No data will be shared without the knowledge of the Principal Investigator (PI) / Project Leader or a delegated authority within a research programme. In cases where data does not fall under the auspices of an existing research programme, permission will be obtained from the DCEO: Research.

4.5 In order for data to be shared it must be preserved

Preservation ensures that data remains usable and authentic over time. Appropriate preservation actions during and after the research process must be adhered to. Guidelines for this are provided in the HSRC Data Preservation Policy.

4.6 The rights and privacy of human subjects will be protected

The HSRC operates within a strict ethics framework set out by its Research Ethics Committee. Data sharing must comply with the stipulations of the Protection of Personal Information Act (Act No. 4 of 2013) and the Electronic Communications and Transactions Act (Act No. 25 of 2002). Measures must be implemented to protect the rights and confidentiality of participants.

4.7 The HSRC may charge for data related services

Charges may be levied for assistance to users, academic or commercial, where such assistance extends beyond basic support and requires the involvement by research personnel. The charges must compensate for the time spent dealing with queries or providing support. The costs and availability of staff to assist will be negotiated directly with the relevant research unit.

4.8 The sharing of data is a shared responsibility

Various parties are required to participate in the data sharing process, including the DCEO: Research, Executive Directors, researchers / data managers, principal investigators / project leaders, research data curators, the directors of IT and Legal Services, as well as data users. The responsibility of HSRC staff may be transferred to a designated alternative.

5. Support for this policy

- The HSRC Act (Act No. 17 of 2008)
- Intellectual Property Rights from Publicly Financed Research and Development Act (Act No. 51 of 2008)
- Protection of Personal Information Act (Act No. 4 of 2013)
- Electronic Communications and Transactions Act (Act No. 25 of 2002)

6. Definitions

| | |
|------------------------|---|
| Data | Data refers to computerised datasets and may be <ul style="list-style-type: none">• quantitative data files consisting of a matrix of numbers or words and its related metadata, such as variable labels, code labels and missing value definitions, including data with spatial references and maps based on these references, or summary data tables. The final dataset might include both raw data and derived variables which would be described in the documentation associated with the dataset,• qualitative data sets that might include transcripts, thematic coding and conceptual maps. |
| Data curation | Data curation is the active management and appraisal of data over the corresponding lifecycle of scholarly and scientific interest. |
| Data management | Data management occurs primarily within the life cycle of a project and is carried out by the research team. It includes all activities concerned with overseeing, directing and organising the digital objects relevant to the research process. The data, documentation, |

code and command files and metadata need to be managed.

| | |
|---------------------------------|---|
| Data sharing | Data sharing includes the actions to provide access to and facilitate the re-use of data. |
| Confidential information | Confidential information refers to information for which unauthorised access or disclosure could result in an adverse effect, for example personal identifying information or sensitive information. |
| Personal information | Personal information refers to information about an identifiable, living, natural or juristic person individual, but excludes information about an individual who has been dead for more than 20 years. It is regulated in terms of the Protection of Personal Information Act (Act No. 4 of 2013) and the Electronic Communications and Transactions Act (Act No. 25 of 2002). |
| Metadata | Information that describes significant aspects of a resource. It may exist at various levels, typically from that of the data collection through to the individual variables of each data file in that collection. |
| Preservation | Preservation refers to all actions to ensure enduring access to the full content of digital resources over time. Preservation typically occurs after a research project has finished and is usually carried out by a data centre or archive. |

7. Authorisation

This Policy Document was approved by the HSRC Board and the CEO on 20 August 2014 after consultation with interested parties.


Prof O. Shisana
CEO

Date 20 August 2014