Research Data Management Guidelines
version 1.0

“Africa Multiple” Cluster of Excellence (EXC 2052/1 – 390713894)
Contents

1. Introduction ...................................................................................................................... 1-2
2. Opportunities and Chances .............................................................................................. 2
3. Roles and Responsibilities ............................................................................................... 2-3
4. Legal and Ethical Considerations .................................................................................... 3-7
   4.1. Privacy - Personal data ............................................................................................... 3-5
   4.2 Sensitive and confidential data .................................................................................... 5
   4.3. Copyright ...................................................................................................................... 5-7
      4.3.1. Copyright of Non-Cluster Data and Databases ......................................................... 6
      4.3.2. Copyright of Researchers' Own Data ................................................................... 6-7
   4.4. Further Legal Considerations .................................................................................... 7
5. Research Data Management Procedures .......................................................................... 7-11
   5.1. Metadata: Data Organisation and Documentation ...................................................... 8-9
   5.2. Data Evaluation .......................................................................................................... 9
   5.3. Data Selection ............................................................................................................. 9
   5.4. Access & Sharing ......................................................................................................... 10
   5.5. Storage / Archiving .................................................................................................... 10-11
   5.6. Data Erasure ............................................................................................................... 11
6. Entry into Force; Reviews and Revisions .......................................................................... 11
1. Introduction

The “Africa Multiple” Cluster of Excellence, funded by the German Research Foundation (DFG) under Germany’s Excellence Strategy (EXC 2052/1 – 390713894), seeks new interdisciplinary, theoretical and methodological approaches to the study Africa and African diasporas in their multifaceted intercontinental and global entanglements. The research data generated and processed in the various research projects of the Cluster will be managed in the Cluster’s Digital Research Environment (DRE). The aim is to develop a platform that enables researchers to organize, store, retrieve and share their data both within and outside of the Cluster, while also laying the foundation for new inter- and transdisciplinary perspectives through the linking of diverse datasets from a variety of disciplinary fields. The Digital Research Environment (DRE) will also lay the foundations for future research approaches. Data management and usage is subject to the Cluster’s overall principle of reflexivity.

The establishment of the DRE constitutes one of the strategic goals of the Cluster of Excellence and involves the professional digital support and curation of the Cluster’s research. It will enable the management of the Cluster’s data through the research data life cycle (storage, organization, structuring, sharing and reuse) aiming at solutions that conform to the FAIR (findable, accessible, interoperable, researchable) principles of Research Data Management and the needs of each individual scholar involved in the Cluster’s research projects.

The data policy guidelines\(^1\) formulated here are intended to inform all Cluster members about its data management procedures, give orientation to current and newly arriving researchers and to ensure that data management practices are carried out in accordance with relevant legal and ethical standards. Crucially the guidelines outline the responsibilities of the researchers in meeting relevant data management standards, policies and procedures.

A sound research data management policy and its implementation are key to the overall functioning of the Cluster of Excellence. They constitute the cornerstone of the cluster’s research infrastructure while also facilitating the evaluation of research processes and outcomes. The digital solutions offered through this data management policy will also be of great value for external cluster partners, especially those based at the African Cluster Centres and elsewhere in Africa, while also enhancing and augmenting the field of knowledge on African and Africa-related research and its digital presence. The DRE will be a core instrument to enable cooperation at eye level with our African partners in order to fulfil the Cluster’s ambitious aim to reconfiguring African Studies.

These data policy guidelines have been formulated in accordance with the University’s research data management guidelines (2016)\(^2\), the white paper “Safeguarding Good Scientific Practice” of the German Research Foundation (DFG)\(^3\) as well as the “Principles for the Handling

---

of Research Data” of the German Research Foundation (DFG) as part of the “Alliance of German Science Organisations”⁴. Our guidelines also incorporate relevant subject specific requirements from the different fields, disciplines and research foci represented in the Cluster.

The data management guidelines presented in this document apply to all Cluster members including researchers from partner institutions and universities for their activities financed by the Cluster. All Cluster researcher are committed to comply with the below specified regulations.

2. Opportunities and Chances

Through the joint data management measure in the Cluster and a common metadata structure developed with the participants of the Cluster, we aim to ensure a structured organization of research data. The participation of all Cluster projects as well as its partners in this measure will guarantee equal data rights for all Cluster members. The open access strategy will also allow the sharing of data with scholars outside of the Cluster. This strategy is one of our aims to overcome inequalities in academic research. It enables us to use data across disciplines, enhances the visibility of the research, creates possibilities for collaboration between the participating scholars, and consequently also contributes to increasing the research competence. Apart from opening up new perspectives, we also seek to stimulate reflexive discussions on research data, which is another important component in the Cluster. At the same time, we need to do justice to the exigencies of the freedom of research. The producers of the data will keep the full right of data ownership, which is protected by security mechanisms in the data management system and the archives where we ensure a secure back-up of the research data.

3. Roles and Responsibilities

The data management guidelines are binding for all researchers active within the “Africa Multiple” Cluster. Each researcher is responsible for ensuring that the data management practices comply with the applicable laws protecting individuals and intellectual property, the university regulations and specific research funding arrangements with third parties. This applies in particular to ensuring "good scientific practice" and to compliance with the standards and requirements relevant in specific academic disciplines.

In accordance with the “Guidelines for Safeguarding Good Scientific Practice” and the “DFG Guidelines on the Handling of Research Data”⁵ of the German Research Foundation (DFG), all cluster researchers (project leaders, research associates, research assistants, doctoral students, postdocs etc.) are obliged to protect research data from loss, prepare it for sustainable use, document it and save it for long term access and use. Likewise, sensible data will be protected from unauthorized access in order to prevent any potentially harmfully misuse of research data generated by the Cluster as outlined in detail below.

---

The “Africa Multiple” Cluster of Excellence calls on its researchers to store, archive and publish research data in specialist repositories or in the university’s existing institutional infrastructures and comply with the Open Access Resolution of the University of Bayreuth (UBT).

Institutions and projects within the cluster are encouraged to develop their own internal guidelines, which may contain more detailed rules for dealing with research data and for the systematic set-up and use of data management plans in the context of project planning.

Considerations regarding the confidentiality as well as the availability and usability of data have direct effects on the storage and required infrastructure for managing data. It is important to plan data processing and usage scenarios, while identifying priorities, necessities and resources needed for research data management (cf. Chapter 5)

The Cluster’s researchers are required to liaise with the data curators to identify the appropriate times and methods for the data to be archived, shared, deleted and supplied for reuse. The Digital Solutions team supports the management of data through advice, guidance and the provision/procurement of relevant tools. The Digital Solutions team (which includes the data curators and the software specialists undertaking the development of the digital solutions) is available for consultation depending on each individual project member’s needs for their data management, and the available resources within the Cluster.

4. Legal and Ethical Considerations

Compliance with the EU General Data Protection Regulation (EU-Datenschutz-Grundverordnung, DGSVO)\(^6\), the Bavarian Data Privacy Act (BayDSG)\(^7\) and with the fundamentals of research ethics constitutes a mandatory component of all research projects. There are two important stipulations that follow from the legal aspects of research data management. First, data protection requires researchers to protect personal data against unauthorized access. Second, with regard to copyright issues two aspects have to be taken into consideration: (1) the access to and subsequent use of works protected by copyright and (2) the creation of new copyrighted objects and the control over the rights resulting from the creation of these.

4.1. Privacy - Personal Data

Research data is subject to a number of German and international laws including the above mentioned EU General Data Protection Regulation (GDPR / DSGVO) and the Bavarian Data Privacy Act.

The purpose of both regulatory frameworks is laying down “rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data”\(^8\), the protection of “fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data”\(^9\) and the movement

\(^6\) «https://gdpr-info.eu/» (last accessed 17.01.2020).
\(^8\) Art. 1 GDPR Subject-matter and objectives, «https://gdpr-info.eu/art-1-gdpr/» (last accessed 17.01.2020).
and processing of data within the European Union. They protect all natural persons from misuse of personal data processed “wholly or partly by automated means”. The regulation applies to all data processors in the Union, “regardless of whether the processing takes place in the Union or not”. Personal Data is defined as any information that leads to the direct or indirect identification of a natural person by so called “identifiers” such as name, date of birth, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person. Persons engaged in the collection or processing of personal data are prohibited from obtaining, using or processing such data without authorization, and they are obliged to maintain data confidentiality even after the end of their activities.

Personal data (e.g. address, age) can only be lawfully collected if the data collector or data processor has obtained a written consent by the person in question. For underage children the parents have to give consent. As a rule, this consent must be provided in written form. The consent statement must be comprehensible and obtained without coercion. The consent form must include a certain set of information to be in accordance with data privacy law:

- the names and contact details of the responsible person / institution /organization
- the name and contact details of a contact person
- which kind of data is collected, the purpose of the data collected and how it is going to be used
- the criteria for the retention period of the data
- the right to withdraw consent on the part of the informant or participant

As far as possible, the researcher should inform the respondents about the objectives of the research project and the purpose of the data collection. This does not necessarily mean to disclose the exact formulation of the research questions, but the aim should be the creation of a high degree of transparency towards the participants or respondents. The data curators within the cluster are available to offer advice regarding the required procedures and documentation associated with obtaining consent from research participants.

There are varying perspectives regarding the obtaining informed consent across different academic and research fields (e.g. anthropology, sociology, linguistics). However, the legal requirements outlined within formal legislation for the handling of personal data supersede any guidelines developed for particular disciplines or projects.

---

12 Art. 3 GDPR Territorial scope, «https://gdpr-info.eu/art-3-gdpr/» (last accessed 17.01.2020).
4.2. Sensitive and Confidential Data

The BDSG contains special provisions on the use of personal data in research. It states that personal data collected or stored for research purposes may only be processed or used within the scope of those purposes. The data collected must be anonymized as soon as the research process allows it. Until that time, data associated with a particular person must be kept separate and may only be combined with other data for research purposes. Academic institutions may only publish personal data if they have the consent from the person concerned and the publication of the personal data is indispensable for the presentation of the research findings on current events.

In addition to the above, there are ethical guidelines within each academic discipline which, while not legally binding are no less important. It is the responsibility of each researcher to investigate whether a specialized association for their scholarly field has developed a code of ethics and discipline-specific approaches for the handling of confidential and sensitive data.

The anonymization of personal data should be done at the earliest possible time (for example, directly during transcription) by replacing names and location with abbreviations and pseudonyms. Researchers should only collect the personal information they need, and they should store personal data on separate storage media with password protection, so that anonymized data cannot be found together with signed declarations of consent or the original audio, film and image files.

4.3. Copyright

Copyright law is a very complex and specialized legal matter. Hence, only the most important aspects of some copyright issues can be outlined here. In cases of doubt, advice from a lawyer should be sought by the researchers of the “Africa Multiple” Cluster of Excellence.

Copyright law protects principally any “author’s own intellectual creations” according to the meaning of the German Copyright Act. Protected works according to the act are:

- Literary works, such as written works, speeches and computer programs;
- Musical works;
- Pantomimic works, including works of dance;
- Artistic works, including works of architecture and of applied art and drafts of such works;
- Photographic works, including works produced by processes similar to photography;
- Cinematographic works, including works produced by processes similar to cinematography;
- Illustrations of a scientific or technical nature, such as drawings, plans, maps, sketches, tables and three-dimensional representations.

16 An overview of subject specific recommendations concerning research data management is provided by the German Research Foundation, cf. «https://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/index.html#anker62194854» (Fachspezifische Empfehlungen zum Umgang mit Forschungsdaten; last accessed 17.01.2020)
Copyright licenses can allow various degrees of distribution, remixing, adapting data, also commercially, as long as the author of the original work is credited, e.g. the various degrees of the Creative Commons Licenses\textsuperscript{19}.

The author’s copyright is no longer subject to copyright protection 70 years after its author’s death, thus becoming public domain. It should be noted that the statutory periods for copyright protection vary within and outside the EU, and that special legal rights apply to archive materials.

4.3.1. Copyright of Non-Cluster Data and Databases

In principle, academic research and data management has to take intellectual property rights into consideration. If a work is protected by copyright, the consent of the author for its duplication or retransmission is essential. In this context, it should be noted that databases under German law are subject to specific protection, which grants the database creators the sole right to disseminate and reproduce them.

Only the sharing/reproduction of an insignificant part of the database is permitted without the consent of the database manufacturer after consultation. In addition, consent for the duplication of a substantial part of a database is not required if such reproduction is only for personal or illustrative and teaching purposes. In addition, copies of both databases and copyrighted works may be used for personal research purposes if the reproduction is clearly not meant for commercial purposes. On the other hand, the retention of copyrighted data by research organizations is not normally covered by this rule, as it usually seeks to make data available to more than one person i.e. sharing it among researchers. In this respect, the management of research data has to consider which external data and programs have been used and which restrictions are associated with their use. In cases of doubt, clearances should be sought with the right holders.

Concerning the reuse of material stored in national archives, the reuse of images or other information material from institutions like museums or private foundations, researchers should be confidential with the exploitation rights regulations of these institutions. Especially in the case of national and/or federal archives with their specific regulations laid down in Archival Law Regulations, issues of data protection and reuse of their data might fall under special restrictions to consider. In the field of contemporary artists and musicians, researchers should be aware that artists’ copyrights might be represented by a collecting society (e.g. VG Bild-Kunst in Germany/Bonn, DALRO in South Africa).

4.3.2. Copyright of Researchers’ Own Data

Researchers should consider the issue of legal protection and usage rights associated with their own data. Of particular importance are the ways in which usage rights can be controlled. There is a wide spectrum of possible licensing models. If no or only specific restrictions are desired, then usage rights are most easily provided with a suitable open license (e.g. GPL or

\textsuperscript{19} See the website of the non-profit organization: «https://creativecommons.org/»
Creative Commons Licence). If academic, technical or methodological patents protect generated data, an experienced patent attorney should be consulted in order to draft a license agreement.

4.4. Further Legal Considerations

In order to avoid legal ambiguities in the management of research data, contract and license drafts should be reviewed by a lawyer. DRE can provide assistance, if needed. In particular, the different data protection and copyright provisions in international contexts and the resulting complexity of legal claims should not be underestimated. This issue is particularly relevant in the context of the “Africa Multiple” Excellence Cluster where research is likely to be conducted across several geographical locations. Due to the still relatively new and constantly changing legal situation, a system for regular legal consultation should also be considered.

5. Research Data Management Procedures

The management of the research data within the Cluster sets out to follow the FAIR Principles in order to ensure its organization in ways that makes it findable, accessible, interoperable and, reusable (depending on the legal and ethical regulations and principles).

The Research Data Management procedures have the dual importance of creating and organizing data in a way that makes it legible and actionable by humans and computer/software. It acknowledges the need to link data with its authors/owners and provide the basis for the citation of data at the same level of importance as publication citations. It therefore follows the Data Citation Principles laid down by the Data Citation Synthesis Group in 2014.\(^{20}\)

The members of the Digital Solutions team within the “Africa Multiple” Excellence Cluster (including the data curators, software specialists and the Vice Dean) are available to provide service, advice and consultation on data management tasks. Further data management services and advice are available from existing data management services at the University of Bayreuth, namely the University Library, the University’s Research Data Management Team (FDM@UBT) and the IT Service Centre.

The next section outlines the responsibilities of both the individual researchers/project leaders and the research data management team within the Cluster. The following tasks will be discussed with reference to the various responsibilities associated with each activity:

- Metadata: Data Organisation and Documentation
- Data Evaluation
- Data Selection
- Access & Sharing
- Storage / Archiving
- Data Erasure

\(^{20}\) Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014 https://doi.org/10.25490/a97f-egyk
5.1. Metadata: Data Organisation and Documentation

Every research project is required to comprehensively document the creation and processing of research data in ways that the origin of the data itself as well as the associated identifiers and classifications can be understood at any time. This is typically done by using a metadata scheme. Metadata is data that describes data. The descriptive documentation associated with the research data, renders it identifiable, “categorisable”, discoverable and usable over time. Metadata is often structured according to specific standards, which are developed and used by particular disciplines and user communities to standardise descriptions for particular resource or types (see Table 1).

Table 1: Examples of metadata standards across different disciplines

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Metadata Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic</td>
<td>• Dublin Core</td>
</tr>
<tr>
<td>Humanities Data</td>
<td>• The Text Encoding Initiative</td>
</tr>
<tr>
<td></td>
<td>• LIDO (Light Weight Description of Objects)</td>
</tr>
<tr>
<td>Social Sciences Data</td>
<td>• Data Documentation Initiative</td>
</tr>
<tr>
<td>Academic Data</td>
<td>• CCLRC Academic Data Model</td>
</tr>
<tr>
<td>Multimedia</td>
<td>• NISO Z39.87-2002 Technical Metadata for Digital Still Images</td>
</tr>
</tbody>
</table>

Metadata standards are often expressed in machine readable Markup Languages (e.g. XML) that facilitate the use and interaction of different computer software to identify, interpret and link data across multiple sources. The use of an appropriate metadata structure enhances the visibility of the data by rendering the data easy to process, identify, index and recall. For example, a metadata description of a sociological data set that does not specify its association with the domain of social sciences may still be classified by a human reader as belonging to the subject field of social sciences. However, without a descriptive classification it would not be possible for software to identify its association with any specific field. Furthermore, missing descriptive metadata relating to the origin of the data may prevent data from being linked to its author and greatly diminish the discoverability of both the author and his/her work.

Overall, metadata assists the process of research data management and scholarly dissemination by facilitating:

- Discoverability of data
- Data identification
- Data interoperability
- Data association with publications and related datasets; and
- Validation and quality assurance of data.

Metadata is typically stored and managed within a file or a database system. This requires a data model flexible and elastic enough to hold all the information required. This is usually decided upon by researchers within individual projects.
It is recommended that researchers develop a good understanding of metadata standards relevant to their discipline and field of research. In addition, they are required to liaise with the data curators in their efforts to devise a metadata schema specific to the research undertaken within the “Africa Multiple” Excellence Cluster.

5.2. Data Evaluation

Data management planning is a crucial step in the process of identifying the requirements for handling, storing and sharing of data gathered during the research process. In order to ensure the effective management of research data, all researchers are required to prepare and submit data management plans outlining the anticipated data requirements for each project and research activity. The data curators will support the researchers in the process of drafting their data management plans. In particular, they will obtain information from each cluster member about their data and needs through the an online questionnaire.

Data management plans may be reviewed and updated at regular intervals throughout the lifecycle of each individual project. This will be conducted through follow up interviews or online surveys which may take place at regular intervals during the funding period of the Cluster.

The data management plans put together by the Cluster’s data curators are structured along the lines of the Digital Curation Centre (DCC)\(^{21}\) checklist and Horizon 2020 project\(^{22}\) guidelines.

5.3. Data Selection

Each researcher is responsible for identifying and selecting the data to be managed within the Digital Research Infrastructure of the Cluster. In cases where data is generated by more than one scholar, decisions regarding the data management within the project should be made collectively accepted by all researchers involved in the project.

Due to the diversity of research themes and academic disciplines involved, it is not possible to make general statement about what data should be handed over to the Cluster’s research data infrastructure. The quality of the data as well as the data’s authenticity and anonymity (if required) is solely dependent on the individual researcher who submits, uploads or hands over to the Digital Solutions team and its data curators.

Submitted data must include explanatory notes and statements regarding their context, ownership, provenance and other relevant details associated with the data. In addition, technical details regarding the format and accessibility of the data must be supplied as part of the data documentation, particularly in cases where specific software is needed to access or read the data.

\(^{21}\) [http://www.dcc.ac.uk/resources/data-management-plans/checklist](http://www.dcc.ac.uk/resources/data-management-plans/checklist)

5.4. Access & Sharing

While encouraging the sharing of data and compliance with the Open Access policy for academic publications\(^{23}\) it is firmly acknowledged that not all research data generated within each project can be published or shared with others. Certain types of data may be too sensitive for open or limited access due to legal or ethical restrictions or other reasons, which make the data unsuitable for broader distribution.

Therefore, the amount and the way in which research data is shared through the Cluster’s data-sharing infrastructure must be assessed and decided upon by the researchers in charge of the respective individual projects.

The Research Data that is made accessible will be given a persistent identifier, e.g. a DOI, which will allow for the identification of the data if cited or used.

Research Data not made accessible can still be searchable through the recording of the data’s descriptive metadata. In other words, the research data can be described in general terms and the contact details of the primary researcher or contact person can be listed to provide information for those wishing to obtain further information about the data or the project in question.

The metadata scheme used to describe the Cluster data will be developed in a collaborative process by the data curators and the Cluster’s software developers in consultation with the participating researchers.

5.5. Storage / Archiving

The storage of data in the DRE infrastructure will grant scaled access rights to ensure that data usage and its sharing occurs in accordance with individual research project requirements. The data access levels and their associated restrictions are based on the 5-Level-Data Classification schemed proposed by Harvard University\(^{24}\) (cf. Table 2).

The data will be archived for the entire funding period of the Cluster (i.e. at least until 12/2025; if the funding is extended for a second period, the data will remain available until 12/2032). Thereafter the Institute of African Studies (IAS) at UBT will continue to manage Africa related research data with new adjusted data management guidelines, that will also address the requirements of partners of the IAS.

Possibilities of long-term archiving of Cluster data might be available through the University’s cooperation in the RADAR (Research Data Repository) services.\(^ {25}\)

\(^{23}\) cf. the University’s statement on its Open Access strategy: https://www.ub.uni-bayreuth.de/de/download/openaccess-strategie.pdf

\(^{24}\) See https://security.harvard.edu/dct, last accessed 19.08.2019

\(^{25}\) https://www.radar-service.eu/de/ueber-uns
Table 2: Data Access Levels

<table>
<thead>
<tr>
<th>Access Level</th>
<th>Short description</th>
<th>Data handling restrictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 5</td>
<td>Information that would cause severe harm to individuals or the University if disclosed</td>
<td>e.g. no access, no DRE storage, stays, consultation with data management team about safe and secure storage</td>
</tr>
<tr>
<td>Level 4</td>
<td>Information that would likely cause serious harm to individuals or the University if disclosed</td>
<td>e.g. only researchable project descriptive metadata in the DRE. Retrieval only for UBT based scholars; no accessible data</td>
</tr>
<tr>
<td>Level 3</td>
<td>Information that could cause risk of material harm to individuals or the University if disclosed</td>
<td>e.g. access limited to UBT based Cluster members</td>
</tr>
<tr>
<td>Level 2</td>
<td>Information the disclosure of which would not cause material harm, but the University has chosen to keep confidential</td>
<td>e.g. shall be confined to Cluster members and the ACC’s</td>
</tr>
<tr>
<td>Level 1</td>
<td>Public Information</td>
<td>No restrictions</td>
</tr>
</tbody>
</table>

5.6. Data Erasure

Erasure of data stored within the Cluster’s digital research infrastructure should only be effected by staff with appropriate authorization and access. (e.g. data administrators, data curators). Deletion and erasure must be preceded by a consultation with the Cluster’s data curators, although it is possible that systems will be developed to facilitate this process to enable erasure to be conducted independently by individual researchers, data owners and project administrators.

6. Entry into Force; Reviews and Revisions

These “Research Data Management Guidelines v 1.0” have been approved and released by the Academic Committee of the “Africa Multiple” Excellence Cluster, funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) under Germany’s Excellence Strategy – EXC 2052/1 – 390713894, on xx/xx/2020.

The Guidelines enter into force with retrospective effect from 01/01/2020. All cluster members will receive an electronic copy. A downloadable file will also be available on the clusters website.

The Guidelines will be subject to review in order to address changes and developments occurring within the cluster. The review might result in new or amended versions of the guidelines, to be approved by the Cluster’s Academic Committee and distributed among the Cluster members.

For further information, please contact the Digital Solutions team: africamultiple-digital@unibayreuth.de