FORS
explore.understand.share.

Preservation Policy DARIS

Creator:	Renate Kunz, Sybil Krügel
Approved by:	Georg Lutz (director)
Date:	April 2017
Versioning:	3.0

1. Introduction

1.1 Purpose and aim of a preservation policy

The preservation policy[1] is intended to give a succinct overview of the broader archival business processes of FORS, available to the public on the DARIS website. It provides transparency on the objectives and principles that guide DARIS with respect to its handling of digital data. It is addressed to FORS staff as well as research funders, data producers and users, and other interested audiences.

The scope of this preservation policy is limited to the data and research information services department DARIS of FORS. It describes the digital preservation strategies and principles of DARIS, as well as the responsibilities and procedures involved in ensuring adequate preservation of and access to the data held within the archive.

The preservation policy is reviewed every other year in order to adapt to a changing technical and research environment. It is issued by the DARIS data service team, which oversees the preservation of the data at FORS.

1.2 Preservation principles: Definitions and relevant concepts

The DARIS preservation policy and all other relevant documentation adhere to the terminology and preservation practices outlined by the Open Archival Information System (OAIS) Reference Model. The OAIS model provides both a functional model – the specific tasks performed by the archive, such as storage or access – and a corresponding information model, to support long-term maintenance and access to digital material.[2] The functional entity "Preservation Planning" encompasses tasks such as development of preservation strategies and standards, development of packaging designs and migration plans, and monitoring of technology (innovations in storage and access technologies) and the designated community (shifts in scope or expectations). DARIS monitors the technical fitness of its archive, does regular risk assessments of the stored digital objects (which includes technology monitoring for the aforementioned object types), and plans for preservation actions.

An essential element of the OAIS model is the grouping of all information into packages:

- SIP: submission information package
- AIP: archival information package
- DIP: dissemination information package

1.3 Digital preservation and its challenges

---

[1] In the literature, the terms preservation policy, preservation procedure and preservation strategy are often used interchangeably. For our purpose, we understand that a preservation policy describes principles, values, intentions and overall general ways of thinking of the organisation (conceptual level, high-level). The preservation policy is implemented through the preservation strategy. Unlike the general statements in policies, the substance of the strategy is very specific and tangible.

[2] CCSDS (2009). Reference model for an open archival information system (OAIS). Draft recommended standard CCSDS 650.0-P-1.1. The Consultative Committee for Space Data Systems. http://public.ccsds.org/sites/cwe/rids/Lists/CCSDS%206500P11/Attachments/650x0p11.pdf.

The goal of any digital repository is the preservation of data in order to make them accessible for the long term.

The aim of our digital preservation policy is therefore to:

- make sure that data are **usable and accessible** to researchers, in other words, make sure that they can be located, retrieved, and presented;
- ensure the **reliability and authenticity** of the data: reliable data are data whose contents can be trusted as a full and accurate representation of the transactions, activities, or facts to which they attest, and can be depended upon in the course of subsequent transactions or activities. Data should be created at the time of the incident to which they relate, by individuals who have direct knowledge of the facts or by instruments routinely used within the business to conduct the transaction. An authentic record is one that can be proven to be what it purports to be, to have been created or sent by the person purported to have created it, and to have been created at the time purported;
- ensure the **integrity and quality** of the digital objects: to prove that a record is complete and unaltered;
- make sure to become a **trustworthy** long-term archive for the research community;
- ensure **transparency** and demonstrate to funders and users our organisation's long-term commitment to its data collections; and
- create a basis for establishing priorities and guidelines for daily operation and internal business processes**.**

However, there are some challenges to digital preservation that a repository is faced with, notably maintaining the integrity, authenticity, and quality of data while guaranteeing their accessibility. In order to face the problem of hardware and software obsolescence due to technology change, a repository can create new manifestations of digital objects. This usually implies that parts of the content are converted, thus inevitably risking loss of some content. Some losses might be acceptable, while others may threaten the authenticity of the data. If (future) users cannot trust the archival holdings, all preservation efforts are in vain.

In order to ensure the integrity of our digital holdings, we are developing our own solution for applying checksums within our new system FORSbase. Checksums are be applied not only at ingest (SIP) but also during the archival storage of our digital holdings (AIP).

2. Principles and purpose of a preservation policy at DARIS

2.1 Purpose and function of the institutional setting

FORS is a research infrastructure in the social sciences serving researchers from Switzerland and abroad. FORS emerged in 2008 from three existing institutions: SIDOS (data archive), the Swiss Household Panel, and SELECTS (electoral studies). DARIS, again, is the data and research information services department of FORS. The mission of DARIS within FORS is to support high quality research, teaching and learning in the social sciences and humanities, by acquiring, curating, and managing data and related digital resources, and by promoting and disseminating these resources as widely and effectively as possible.

Digital preservation is therefore a crucial aspect of its activities, since all services would not be possible without the guaranteed preservation of data. The preservation policy enables FORS to fulfil its primary founding objective – to lend active support in the long term to social science research in Switzerland. DARIS ensures that preservation is embedded in all activities, from ingest through to access.

The preservation policy of DARIS relates to other organisational policies, such as the mission statement, the collections policy, the disaster recovery plan, and the security and access policy.
The designated community of DARIS consists of scientific researchers, students and instructors in the fields of the social sciences.

2.2 Legal aspects

Digital preservation must also address legal issues such as usage rights, data protection, and intellectual property rights. They are specified in national and international regulatory frameworks as well as managed through contractual agreements between the data archive and the rights holder.

The legal framework within which DARIS is operating consists of the following regulations:

- Swiss federal act on data protection (FADP),19 June 1992 (Status as of 1 January 2014)
- Federal law for the promotion of research and innovation (FIFG), 14 December 2012 (available only in French or German)
- DARIS deposit agreements: http://forscenter.ch/wp-content/uploads/2013/11/deposit-contract_e.pdf
- DARIS end user license: http://forscenter.ch/wp-content/uploads/2013/11/Download-contract-_E.pdf

2.3 Coverage and scope of the collection

DARIS solicits quantitative and qualitative data that are collected in Switzerland and that fall within the realm of the social sciences (see our collections policy). The data should have further analytic potential that can bring new insights to the social sciences. In addition, the projects should be good examples of well-designed, conducted, and documented research. The data and documentation are digital, and could include text, audio, image, and video formats. DARIS assumes long-term responsibility for the preservation and accessibility of the data. In practice, this means that data are preserved beyond the next round of technical change.

2.4 Access

Data are consulted by users under controlled conditions in accordance with agreed standards and guidelines. With our new system FORSbase, access to data is easier and more straightforward for our users. A registered user can download data from FORSbase or from our Nesstar Server. Unauthorised access to FORSbase is prevented through a system

of different access rights for users and staff. We guarantee that the dissemination information package (DIP) that is submitted to the user is suitable for reuse. As with the SIP and AIP, the suitable file formats for DIP may change over time and need to be closely monitored.

We intend to introduce a system of persistent identifiers for our datasets (DOI's) into FORSbase. The implementation of DOIs for our data holdings using da¦ra is planned to be finalised by mid-2017. With the DOI system it will be possible to clearly reference and locate digital data permanently.

In order to ensure that usage rights, data protection and intellectual property rights are met, we use standard contracts for both data depositors and data consumers. The contracts are based upon the principles of open access and relevant national legislations.

3.   Sustainability of the preservation policy

3.1 Roles and responsibilities

The DARIS team guarantees long-term usability and comprehension of its archive's data. Efforts are coordinated and supported by the FORS director and executive board members. The DARIS team closely cooperates with the FORS IT department and the IT department of the University of Lausanne, which is responsible for the technical infrastructure. The director of FORS is responsible for this preservation policy and guarantees its implementation and its communication throughout the organisation. It is the responsibility of all DARIS staff members to be up to date with respect to the most recent developments in the digital preservation domain. Toward this end, they are encouraged to attend trainings, conferences, and events to further develop their knowledge and skills. All new DARIS staff members are introduced to and trained in preservation issues (http://forscenter.ch/en/about-us-2/staff-3/).

FORS recognises the need for collaboration in digital preservation and works to foster relationships both in the social science domain and across disciplines. As a future development for digital preservation, DARIS will focus on ensuring the efficient operation of the FORSbase platform.

3.2 Relationships with other documentation and policies

The preservation policy is supported by relevant documentation and key policies that need to be taken into account when implementing digital preservation activities. The following table provides a list of those related documents.

Document	Availability
DARIS mission statement	Website: http://forscenter.ch/wp-content/uploads/2015/05/Mission-statement_EN_v21.pdf
FORS deed of foundation	Website: http://forscenter.ch/wp-content/uploads/2013/11/Deed_of_foundation.pdf
Legislation	Website: http://www.admin.ch/opc/en/classified-

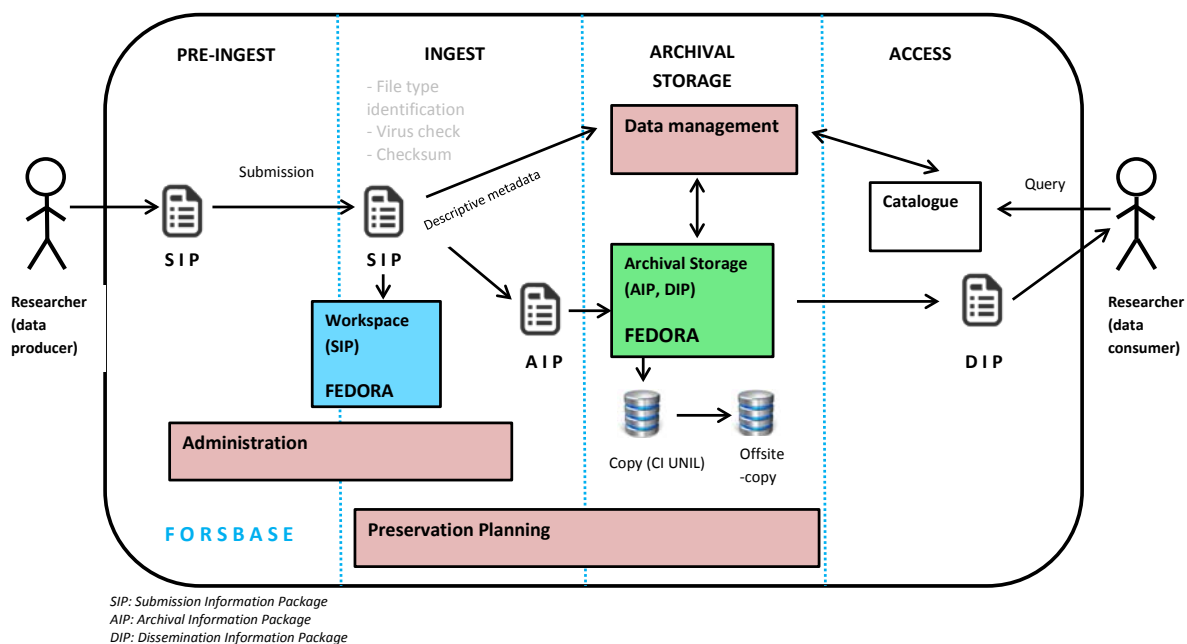|	compilation/19920153/index.html
Collections policy	Website: http://forscenter.ch/wp-content/uploads/2015/05/Collections-Policy_E.pdf
Disaster recovery plan	Internal document (IT department University of Lausanne)
Succession plan	Internal document (to be announced)
Deposit agreement	Website: http://forscenter.ch/wp-content/uploads/2013/11/deposit-contract_e.pdf
End user agreement	Website: http://forscenter.ch/wp-content/uploads/2013/11/Download-contract-_E.pdf
Data protection	Website: http://forscenter.ch/en/data-and-research-information-services/2221-2/deposit-data/data-protection/
Guidelines for depositors	Website: http://forscenter.ch/en/data-and-research-information-services/2221-2/deposit-data/guidelines-for-deposit/

3.3 Monitoring and review

It is necessary to regularly evaluate the preservation policy in order to ensure the continued relevance of the organisation's digital preservation aims and procedures and to detect any weaknesses or required changes. The policy therefore undergoes periodic review every two years. That means that changes in the designated community or relevant technologies, as well as changes in high-level policies or legal obligations, are taken into account. For more information, please contact the data service team at DARIS (http://forscenter.ch/en/about-us-2/staff-3/).

4. Preservation strategy – implementing the policy

In contrast to the preservation policy, the preservation strategy addresses how the preservation is carried out, and focuses on workflows and technical strategies. Four issues are mainly concerned: 1) Preservation Planning 2) Storage, 3) Duplication, and 4) Backup

4.1 The archiving workflow

Our archiving workflow is in close accordance with the OAIS model and uses its terminology (apart from the pre-ingest function, which is not part of the OAIS model).

PRE-INGEST — INGEST — ARCHIVAL STORAGE — ACCESS

- File type identification
- Virus check
- Checksum

Researcher (data producer) → SIP → Submission → SIP → Descriptive metadata → Data management

Workspace (SIP) FEDORA

Administration

Preservation Planning

FORSBASE

AIP → Archival Storage (AIP, DIP) FEDORA

Copy (CI UNIL) → Offsite-copy

Catalogue ← Query ← Researcher (data consumer)

DIP

SIP: Submission Information Package
AIP: Archival Information Package
DIP: Dissemination Information Package

Pre-ingest: DARIS offers guidance and help for data producers wishing to deposit data. This can include technical support, but can also consist of data management training and guidance. Data depositors are given clear instructions on how to properly prepare, document, and deposit their data. In addition to that, solicitation of data is also included in the pre-ingest phase.

Ingest: Irrespective of particular file formats, we collect data that are in accordance with the provisions of our collections policy. If possible we intervene in the (pre)-ingest process as early as possible (prospective) in order to foster a successful data delivery.

All data are submitted as a submission information package (SIP) and uploaded to our information system FORSbase. This means that in addition to the data, the data depositors create an extensive metadata record in FORSbase while establishing an entry. The data submission process is clearly described and specified on our FORSbase website through a series of quick guides.[3] Once a SIP is uploaded to FORSbase, quality assurance routine checks are carried out for completeness, integrity and validation of the data files, the submitted documentation and the metadata. DARIS staff reserves the right to adapt the form of the received digital objects (e.g. the formats) of the SIP in order to prepare them for archival storage.

In FORSbase data and documentation files are screened for format at ingest and checked on a regular basis for format readability. All data files (of the Submission Information Packages (SIPs), the Archival Information Packages (AIPs) and the Dissemination Information Packages (DIPs) are stored as a csv file in Fedora. However, the uploaded files (SIP) on the one hand and the AIP and DIP on the other hand are stored for security reasons on two different instances. In order to ensure the integrity of the digital holdings, checksums are integrated in FORSbase (Fedora) and are applied not only at ingest (SIP) but also during the archival storage of the digital holdings (AIP).

---

[3] https://forsbase.unil.ch/

Archival storage: The descriptive metadata of a SIP are treated separately from the data themselves. These are, however, interlinked with an identification number. This means that the data management function occurs independent of primary data within FORSbase. Approved storage technology is provided and guaranteed by the IT department of the University of Lausanne (see section 4.3 on IT infrastructure). The IT centre of the University of Lausanne also uses redundant servers at the EPFL Neuchâtel in case of inundation, fire, malicious actions and technical failure.

Once saved as an AIP in Fedora, the data are not deleted from this instance, and adequate preservation strategies are employed in order to guarantee their long term preservation (see 4.4, Migration as priority).

Our current versioning strategy implemented in FORSbase automatically triggers new versions when data or metadata are changed. Minor or major changes in metadata, documentation or data fields are recorded in FORSbase. We are currently in the process of implementing a system of persistent identifiers (DOI) provided by da¦ra. The versioning, handled so far by FORSbase, would follow the recommendations of da¦ra.

Access: DARIS guarantees that its digital holdings are identifiable and traceable at all times through our catalogue. The data and their associated metadata and documentation are made available to the user as a dissemination information package (DIP). The DIP can be downloaded from FORSbase according to the access conditions determined by the data producer.

4.2 Accepted and preferred file formats

In order to make the data and documentation accessible for other users and to be able to process and use them over time, data files should be in a form that best ensures their longevity. Digital data are constantly at risk from changes in the hard- and software environment. Those risks can be levelled off by using file formats that are non-proprietary, openly documented, unencrypted, and uncompressed. In addition to that, the frequency of migration and the costs of preservation are lower.

DARIS maintains a list of accepted and preferred file formats that are part of our collections policy and are available on our website. As part of the preservation planning function of the OAIS model, we check periodically to see whether the current accepted and preferred file formats are still considered to be suitable for preservation, and we update the list when necessary. The adoption of new file formats and the consequent replacement of old ones are planned in advance and communicated to stakeholders accordingly.

4.3 IT infrastructure

Data recovery provisions, backup strategies, and risk management techniques are carried out by the IT department of the University of Lausanne – the host institution of FORS.

The following backup and risk management strategies are in place:

Physical security

- Server rooms (datacentres) are secured by an access control system against unauthorised access;
- Temperatures in the datacentres are constantly monitored and provisions are in place against fire, water, power failure, or any other major disaster.

Redundant server infrastructure

- Redundant servers are located at two physically different locations on the University campus;
- There are frequent incremental and full backups to onsite disk storage.

Storage media

- Diversity of storage media (disk);
- Frequent media refreshment.

The backup and storage measures of the University of Lausanne therefore guarantee fast and complete recovery of our archive holdings in case of an incident.

4.4 Migration as a priority

In contrast to a non-electronic object, a digital object always needs an environment in order to render its content. Since this environment is constantly evolving at a rapid pace, digital objects may become unreadable or obsolete. DARIS has adopted a migration based approach to digital preservation. We migrate file formats that have come close to obsolescence to new file formats that are more sustainable and guarantee future usability. The potential risk of information loss will be mitigated by testing of migration pathways and validation of migrated files.

At DARIS we migrate files where needed, but will always maintain the original manifestation of the data and all subsequently generated manifestations of the original files. In this case, we adhere to the principle of reversibility: being able to revert to an earlier version of a digital file after migration. We also fully document the migration process in the form of a detailed migration history as part of the metadata associated with the data file.

5. Resources

Becker, Christoph et al. Systematic planning for digital preservation: evaluating potential strategies and building preservation plans, in: International Journal on Digital Libraries, 4/2009, S. 133-157. http://www.ifs.tuwien.ac.at/~becker/pubs/becker-ijdl2009.pdf

Core Trustworthy Data Repositories Requirements (2016): ICSU World Data System (WDS) and Data Seal of Approval (DSA) https://drive.google.com/file/d/0B4qnUFYMgSc-eDRSTE53bDUwd28/view

Data Seal of Approval: http://www.datasealofapproval.orgFriese, Yvonne. How to develop a preservation policy; Guidelines form the nestor working group; IASSIST Quarterly Fall/Winter 2012, p. 17-22. http://www.iassistdata.org/sites/default/files/iqvol36_34_friese.pdf

Guidelines for the creation of an institutional policy on digital preservation (2014) (nestor-materials 18). http://files.d-nb.de/nestor/materialien/nestor_mat_18-eng.pdf

Lavoie, Brian. 2014. The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition). Great Britain: Digital Preservation Coalition. Accessed: http://dx.doi.org/10.7207/twr14-02

Schumann, Natascha, and Jonas A. Recker. 2013. "De-mystifying OAIS compliance: benefits and challenges of mapping the OAIS reference model to the GESIS Data Archive." IASSIST Quarterly Jg. 36, H. 2 6-11. http://www.iassistdata.org/iq/de-mystifying-oais-compliance-benefits-and-challenges-mapping-oais-reference-model-gesis-data-arc

Preservation policy references:

- GESIS: http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/DAS_Preservation_Policy_eng.pdf
- DANS: http://www.dans.knaw.nl/nl/over/organisatie-beleid/juridische-informatie/dans-preservation-policy-uk.pdf
- Digital Curation Coalition (DCC): Preservation policy template for repositories http://www.dcc.ac.uk/sites/default/files/documents/Preservation%20policy%20template.pdf
- UKDA: http://www.data-archive.ac.uk/media/54776/ukda062-dps-preservationpolicy.pdf
- The National Archives: http://www.nationalarchives.gov.uk/documents/information-management/digital-preservation-policies-guidance-draft-v4.2.pdf
- Swiss Federal Archives: http://www.bar.admin.ch/dokumentation/00445/00527/index.html?lang=enhttps://www.bar.admin.ch/bar/en/home/services-publications/publications/archival-publications.html