

FORS 

explore.understand.share.

Preservation Policy

FORS Data Archive

Creator: FORS Data Archive
Approved by: FORS Executive Board
Date: December 2021

Versioning: 4.0

1. Introduction

1.1 Purpose and aim of this preservation policy

Our preservation policy is intended to give an overview of the broader archival processes of FORS, available to the public on the FORS and SWISSUbase websites. It provides transparency on the objectives and principles that guide the FORS Data Archive with respect to its handling of digital data. It is addressed to research funders, data producers and users, and other interested audiences. The scope of this preservation policy is limited to the archival services provided by the FORS Data Archive. It describes the digital preservation strategies and principles, as well as the responsibilities and procedures involved in ensuring adequate preservation of data managed by the FORS Data Archive.

The preservation policy is reviewed every three years in order to adapt to an ongoing changing technical and research environment. This policy is issued by FORS Data Archive, which oversees the preservation of the data at FORS.

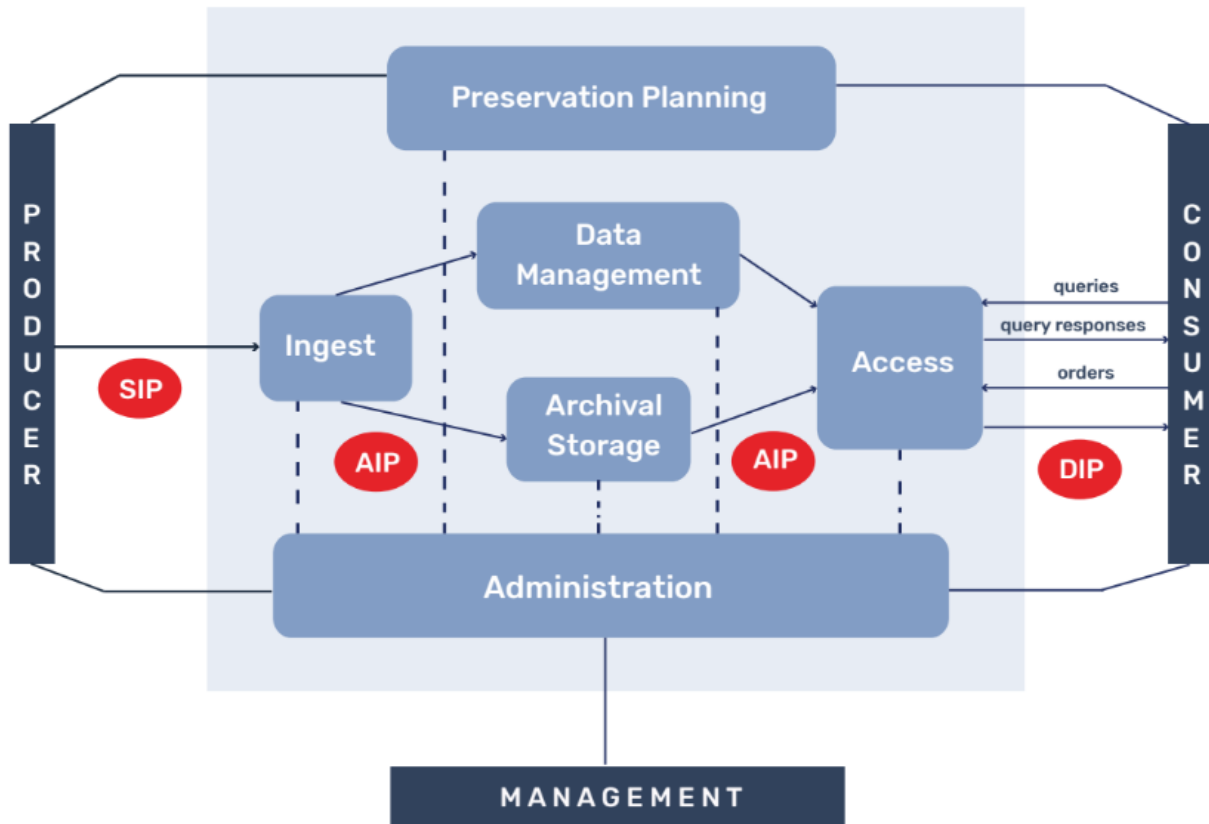
FORS Data Archive is using the online platform SWISSUbase. It is a non-commercial, national, public, cross-disciplinary repository service solution and offers controlled access to research data; ranging from open to restricted. FORS is one of the main partners, alongside the Universities of Lausanne and Zurich. Within SWISSUbase data service units are independently responsible for specific disciplines or partner universities. FORS Data Archive is the data service unit responsible for all social sciences. This policy is specifically for the FORS Data Archive.

1.2 Preservation principles: Definitions and relevant concepts

This preservation policy and all other relevant documentation adhere to the terminology and preservation practises 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. 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). The FORS Data Archive monitors - in coordination with the SWISSUbase/FORS IT team - 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, which is delivered to the system for ingest (comprising data, documentation and related metadata)
- AIP: archival information package, which is stored and preserved within the system (may contain long-term preservation formats)
- DIP: dissemination information package, created to distribute the digital content (which may contain additional formats used by the designated community)



1

1.3 Digital preservation and its challenges

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:

¹ SWISSUbase all rights reserved

- 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 occurrence 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 dataset is complete and unaltered;
- illustrate that we are a **trustworthy** long-term data service 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 may create new expressions (change) of digital objects if appropriate, while keeping track of all performed changes.

2. Principles and purpose of a preservation policy at the FORS Data Archive

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). The FORS Data Archive is the data and research information services department of FORS. The mission of the FORS Data Archive is to support research, teaching, and learning in the social sciences, 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 objective – to lend active support in the long-term to social science research in Switzerland. The FORS Data Archive ensures that preservation is embedded in all its activities, from ingest to access.

The designated community of the FORS Data Archive consists of scientific researchers, students, and instructors; primarily but not limited to the disciplines of the social sciences. The datasets archived by the FORS Data Archive are available via the SWISSUbase catalogue to researchers and students of all disciplines.

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 data owners.

The legal framework within which the FORS Data Archive is operating consists of the following regulations:

- Swiss Federal Act on Data Protection (FADP), 19 June 1992 (Status as of 1 March 2019)
https://www.fedlex.admin.ch/eli/cc/1993/1945_1945_1945/en
- Federal Law for the Promotion of Research and Innovation (FIFG), 14 December 2012 (Status of 1 January 2020)
<https://www.fedlex.admin.ch/eli/cc/2013/786/en>

2.3 Coverage and scope of the collection

The FORS Data Archive solicits quantitative and qualitative data that are collected in Switzerland, financed by a Swiss institution or otherwise affiliated with a Swiss higher education institution, and that fall within the realm of the social sciences (see our acquisition policy).

2.4 Access

Datasets are consulted by users under controlled conditions in accordance with agreed standards and guidelines. With the platform SWISSUbase, access to data is easy and straightforward. Unauthorised access to SWISSUbase is prevented through a system of differential access rights for users and staff. We monitor that the dissemination information package (DIP) 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 might need to be adapted.

SWISSUbase has implemented a system of persistent identifiers for each dataset (DOIs), following the recommendations given by Datacite. Via DOI it is possible to clearly reference and locate digital data and their versions permanently within the SWISSUbase catalogue.

In order to ensure that usage rights, data protection, and intellectual property rights are met, we use standard SWISSUbase contracts for both data depositors and data consumers. The contracts are based upon the principles of open access and relevant national legislation. Data depositors can also choose Creative Commons copyright licences for their data. Customised contracts are possible on request.

3. Sustainability of the preservation policy

3.1 Roles and responsibilities

FORS Data Archive ensures long-term usability and comprehension of its managed data. Efforts are coordinated and supported by the FORS director and executive board members. The FORS Data Archive team closely cooperates with the FORS IT department and SWISSUbase support team. In addition, regular exchange takes place with other Data Archive Units working with SWISSUbase. The technical infrastructure is provided by SWITCH.

The FORS Executive Board is responsible for this preservation policy and guarantees its implementation and its communication throughout the organisation. It is the responsibility of all FORS Data Archive staff members to be up to date with respect to the most recent developments in the digital preservation domain. Toward this goal, they are encouraged to attend training, conferences, and events to further develop their knowledge and skills. All new staff members are introduced to and trained in relevant preservation processes.

FORS works to foster relationships in the social science domain and across disciplines as well as on the level of national and international collaborations with other infrastructures that deal with preservation. As for future development for digital preservation, FORS Data Archive is involved in the development of the SWISSUbase platform in an advisory capacity - together with other data service units - to make sure the platform meets the needs of the designated communities (data depositors as well as data users).

3.2 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 three years. That means that changes in

the designated community or relevant technologies, as well as changes in high-level policies and legal obligations, are taken into account.

3.3 Other related documents 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. Below is a list of those related documents.

- FORS mission statement:
<https://forscenter.ch/about-fors/mission/>
- FORS Governance:
<https://forscenter.ch/about-fors/governance/>
- Acquisition policy
<https://forscenter.ch/about-fors/policies/>
- Preparing your data for deposit in SWISSUbase
<https://forscenter.ch/data-services/help-resources/>
- SWISSUbase User Guide
<https://resources.swissubase.ch/help/user-guide/>

4. Preservation strategy – implementing the policy

In contrast to the preservation *policy*, the preservation *strategy* addresses how the preservation is carried out, focusing on workflows and technical strategies.

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: FORS Data Archive offers help to data producers wishing to deposit data. This can include technical support, but can also consist of data management consultations, guidance and training. Data depositors are given clear instructions on how to properly prepare, document, and deposit their data.

Ingest: We collect data that are in accordance with the provisions of our acquisition policy. If possible we intervene in the (pre-)ingest process as early as possible in order to foster a successful data delivery. All data and their documentation are submitted as a submission information package (SIP) and uploaded to the information system SWISSUbase. In addition to the data, the data depositors create an extensive metadata record in SWISSUbase while establishing an entry. The data submission process is described and specified on the SWISSUbase website through a user guide. Once an SIP is uploaded to SWISSUbase, quality assurance routine checks are carried out for completeness, integrity and validation of the data files, the submitted documentation, and the metadata.

Archival storage: The FORS Data Archive 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 (AIP) and dissemination (DIP). The FORS Data Archive might also enrich the metadata to improve findability and reusability during the archiving process.

Data Management: All data files of the Submission Information Packages (SIPs), the Archival Information Packages (AIPs) and the Dissemination Information Packages (DIPs) are stored as individual files in object storage S3 from SWITCH engines. A strict implementation of checksums ensures the integrity of our digital holdings. The descriptive metadata of an 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. Once saved as an AIP, 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).

The versioning strategy implemented in SWISSUbase automatically triggers new versions when data or metadata are changed. Therefore changes in metadata, documentation or data fields are documented and traceable.

SWISSUbase provides persistent identifiers (DOIs) for all datasets following best practises by Datacite. Each time a version of a dataset is published, a new DOI is assigned to facilitate long-term traceability. Each DOI assigned to a version of a dataset remains active; meaning the page is accessible and the metadata is visible, but only the data from the latest published version can be downloaded directly via the catalogue. Access to old versions is guaranteed via the FORS Data Archive.

Access: The DIP can be downloaded from SWISSUbase according to the access conditions determined by the data depositor. It is possible to set documentation files as “public”; these are thus visible without a contract.

The end user agreement specifies the duration of use. Contracts can be extended. Users are required to inform the archive of any publications based on the data used and to confirm that the data has been deleted after the end of the use of the data. In case of non-compliance with the terms of the contract, user accounts will be blocked.

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.

The FORS Data Archive maintains a list of accepted and preferred file formats that are part of our acquisition policy and are also available in our guide for data depositors. As part of the preservation planning function of the OAIS model, we check periodically to see whether the currently 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

General risk management techniques are carried out by SWITCH – one of the partner organisations in SWISSUbase. SWITCH operates exclusively with servers that are stationed in Switzerland. SWITCH organises the technical and organisational measures based on its security management system (ISMS) in accordance with ISO/IEC 27001.²

Physical security:

- Server rooms (datacentres) are secured by an access control system against unauthorised access (physical, digital);
- Temperatures in the data centres are constantly monitored and provisions are in place against fire.

Backup strategy:

- Daily backups are stored for 30 days;
- Monthly backups are stored for 1 year.

² <https://www.switch.ch/engines/>

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. The FORS Data Archive 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.

The FORS Data Archive migrates 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

CESSDA Training Team (2020). CESSDA Data Archiving Guide Bergen, Norway: CESSDA ERIC.

CoreTrustSeal Trustworthy Data Repositories Requirements (2020-2022):
<https://www.coretrustseal.org/why-certification/requirements/>

Digital Preservation Handbook, 2nd Edition, <https://www.dpconline.org/handbook>, Digital Preservation Coalition © 2015

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

Wilkinson et al. (2016), The FAIR Guiding Principles for scientific data management and stewardship, Scientific Data 3, doi:10.1038/sdata.2016.18

Preservation policy references and further information also available via [CESSDA Resource Directory](#)