DASISH

Data Service Infrastructure for the Social Sciences and Humanities (DASISH) brings together all 5 ESFRI research infrastructure initiatives in SSH area. The goal of DASISH is to determine areas of cross-fertilization and synergy in the infrastructure development and to work on concrete joint activities related to data, such as data access, data sharing, data quality, data archiving and legal and ethical aspects.

Comprehensive Policy-Rules for Data Management in SSH*

Background and description

There is a growing awareness among research stakeholders – ranging from international organisations, national science organisations, research funding bodies, to data centres, data archives, universities and researchers – on the needs for proper tools, mechanisms and instruments aimed at providing trustworthy long-term preservation of research data. Crucial means to this end are a clearly defined set of *preservation policy rules*. A written preservation policy is an important tool for data preservation services to achieve open and prolonged access to research data and to create trust among stakeholders.

In this report we provide an overview of the state of the art of preservation policies and describe, compare and analyse the scope of policy-rules and the requirements they set for the SSH domain, in particular in Europe and the US. The overall aim has been to establish and recommend a set of concrete policy-rules that will support the integrity of data and build trust in data preservation services.

The report analyses existing preservation policies and procedure templates in addition to concrete policy examples from a selection of service providers. Each policy model is analysed and described individually before the different policy models and policy-rule elements are compared to identify similarities and differences. Based on the analyses the report concludes by recommending a set of policy-rules for the long-term preservation of research outputs.

Findings

With few exceptions all the assessed policies share some key features:

- Close proximity to the OAIS reference model, both in structure, methodology and concepts.
- Three-layered policy structure: (**1**) high-level *content coverage* (entries such as ‘mission’, ‘scope‘, ‘purpose’, ‘objective’, etc.); (**2**) descriptions of *digital object management* procedures (i.e. data lifecycle approach); (**3**) dedicated segments on *sustainability*, *security*, *risk* and *technical infrastructure*.
- Policies reflect service provider characteristics. If services are more *centralised* and curated through human intervention, their policies are laid down in a single or a limited number of policy document(s). If services are *distributed* or decentralised policies tend to be equally distributed.
- Although there are common elements in many of the models, the decision making and procedure implementation in the organisations are to some extent based on ad-hoc solutions due to a lack of common understanding of how to implement the standardisation models (e.g. the OAIS reference model).

---

* Deliverable 4.4 of WP 4: Data Archiving in the Social Sciences and Humanities. Available at: http://dasish.eu/publications//

Table 1: Summary of recommended Preservation Policy elements

Policy Element	Description
Purpose, objectives, scope	Should describe the purpose and function of the organisation, state the rationale for the preservation policy and clearly show how the policy is grounded in the organisational context by establishing clear connections between goals and implementation.
Glossary, definition of terms	As the audience for a preservation policy may be diverse, it can be useful to define key terms at the outset to ensure common understanding, especially if the policy applies many technical terms or terms specific to the organisation
Preservation standards, requirements, legal and regulatory framework	This segment should list and specify how and under which requirements and legal and regulatory frameworks the policy works.
Roles and Responsibilities, financial responsibilities, cooperation	This segment should clarify and define the different key roles and responsibilities for participants involved in the long-term preservation of data.
Pre-ingest, selection and acquisition	This segment provides the rationale and processes for developing and retaining collections based on specific parameters.
Ingest, communication with the depositor	Describes the processing of the data object before it is entered into the archive.
Preservation strategy	Outlines how the organisation approaches the storage of its data collections (e.g. bit stream preservation, transformation to an open format, rendering, emulation, migration, etc.).
Archival storage, security	Specifies the organization's commitment and approach to ensuring the accuracy, completeness, authenticity, integrity, and long-term protection of the organization's data assets.
Risk management	Describes measures on how the organisation achieves a secure and trustworthy technical infrastructure.
Data management, curation, metadata	Outline of the metadata schema in use. Specifies how the different sections of the schema are structured (e.g. descriptive metadata, structural metadata, administrative metadata, preservation metadata, etc.).
Access, use, re-use	Identifies how end users interact with the archive to find, request and receive data and metadata.
Intellectual property	This element describes how the organisation plans to recognise and deals with copyright issues.
Policy review, certification	Statement on how often a review of the policy is carried out (e.g. annually, biannually). Additionally it should include a section on how the data centre is, or aims to become, formally a trustworthy long-term preservation service