cessda saw

Research and Innovation Action

CESSDA Strengthening and Widening

Project Number: 674939 Start Date of Project: 01/08/2015 Duration: 27 months

Deliverable 4.3 – Report Overview of Data Management Policies in Social Science Data Archives

Dissemination Level	PU
Due Date of Deliverable	31/04/17
Actual Submission Date	24/08/2017
Work Package	WP4
Task	4.2
Type	Report
EC Approval Status	16 November 2017
Version	V1.0
Number of Pages	p.1 – p.90

Abstract: The CESSDA Data Management Policy Framework presented in this report consists of a set of policy components and clauses that provide the foundations and organisational arrangements for designing, implementing, monitoring, reviewing and improving Data Management processes throughout the organisations whose main objective is to preserve and provide access to digital research data. Although the framework is primarily aimed at CESSDA service providers and social science data archives, it is general in scope and design and can be applied by many different organisations whose main objective is to preserve digital objects.

Project funded by the EU Horizon 2020 Research and Innovation Programme under the agreement No.674939

History

Version	Date	Reason	Revised by
0.0	09/10/2015	Template for policy framework/template, first draft	Task leader and partners
0.1	16/11/2015	Template for policy framework/template, second draft. Representative case studies selected.	Task leader and partners
0.2	13/04/2016	Template for policy framework/template, third draft. First draft outline of deliverable report. First draft template for case studies / country reports.	Task leader and partners
0.3	05/08/2016	Template for case studies / country reports. Second draft outline of report.	Task leader and partners
0.4	15/10/2016	All case studies finalised.	Task leader and partners.
0.5	15/05/2017	First draft policy framework and template.	Task leader and partners. WP-leader.
0.5	19/06/2017	First draft final deliverable report.	Task leader and partners.
0.6	03/07/2017	Second draft final deliverable report.	Task leader and WP-leader.
1.0	21/07/2017	Final draft to be submitted to the CESSDA AS.	Task leader

Author List

Organisation	Name	Contact Information
NSD	Trond Kvamme	trond.kvamme@nsd.no
FSD	Mari Kleemola	mari.kleemola@staff.uta.fi
ADP	Janez Štebe	janez.stebe@fdv.uni-lj.si
FORS	Sybil Krügel	sybil.kruegel@fors.unil.ch

Time Schedule before Delivery

Next Action	Deadline	Care of
All cases studies finalised	15/10/2017	All T4.2 partners
Policy framework and template finalised	15/05/2017	All T4.2 partners
Final integral version of deliverable report	19/06/2017	NSD
Review by task partners	03/07/2017	All T4.2 partners
Review by the WP leader	19/07/2017	DANS
Review by the Chair of the Delivery Committee	21/08/2017	CSDA
Review by the Project Coordinator	21/08/2017	CESSDA AS
Approval and Submission by the Project Coordinator to the European Commission	24/08/2017	CESSDA AS

EXECUTIVE SUMMARY

This Report Overview of Data Management Policies in Social Science Data Archives provides resources and guidelines for the development of internal service provider policies. It presents an implementable framework and a model that can be used by preservation organisations as a reference when building or renewing their organisational and operational data management policy framework.

When developing the template we have recognised that developing a comprehensive policy framework will only be worthwhile if it is linked to core organisational business drivers and strategies: it cannot be effective in isolation as a separate entity outside of the higher level organisational aspects. Therefore we have included mission statements, wider strategies and organisational infrastructure in the policy framework model.

The model is based on an examination of existing reference models for archives and repositories, high-quality preservation policy templates, and existing state-of-the-art real-world policies from selected service providers. Based on these findings the CESSDA Data Management Policy Framework presents a series of policy elements and structuring advice that can be useful when setting up or refining an organisational policy framework. The template/model provides descriptions and rationale for each policy element, along with samples and examples from real-world policies.

The report maps the current status of CESSDA service provider policies through a selection of policy framework case studies. UK Data Archive, FSD (Finnish Social Science Data Archive), ADP (Slovenian Social Science Data Archive) and FORS/DARIS (the Swiss Centre of Expertise in the Social Sciences) are presented through in-depth descriptions, highlighting their respective strategic and operational policy frameworks.

Abbreviations and Acronyms

ADP	Slovenian Social Science Data Archive
CCSDS	Consultative Committee for Space Data Systems
CSDA	Czech Social Science Data Archive
DAMA	Data Management Body of Knowledge
DARIS	Data and research information services (part of FORS)
DDI	Data Documentation Initiative
DOI	Digital Object Identifier
DSA	Data Seal of Approval
FORS	Swiss Centre of Expertise in the Social Sciences
FSD	Finnish Social Science Data Archive
GESIS	German Social Science Infrastructure Services
ICPSR	Inter-university Consortium for Political and Social Research
IPR	Intellectual Property Rights
OAIS	Reference Model for an Open Archive Information System (OAIS)
PAIMAS	Producer-Archive Interface – Methodology Abstract Standard
TDR-EU	Trusted Digital Repository EU
UKDA	UK Data Archive
WDS	World Data System

TABLE OF CONTENT

Executive Summary **4**
Table of Content **6**
1. Introduction **7**
1.1. Aims and Objectives 7
1.2. Background 7
1.3. Terminology – Nuances, Definitions and Clarifications of Concepts 10
1.4. Approach and Methodology 11
2. The CESSDA Data Management Policy Framework **13**
2.1. Background 13
2.2. Components of the CESSDA Data Management Policy Framework 15
2.3. Policy Element Description Template 18
2.4. Policy Clauses 19
2.4.1. Policy Context Clauses 19
2.4.2. Organisational Infrastructure Policy Clauses 22
2.4.3. Data Management Policy Clauses 35
3. Case Studies **53**
3.1. UK Data Archive 53
3.1.1. Organisational Infrastructure 53
3.1.2. Digital Object Management 58
3.2. FSD 65
3.2.1. Organisational Infrastructure 65
3.2.2. Digital Object Management 67
3.3. ADP 70
3.3.1. Organisational Infrastructure 70
3.3.2. Digital Object Management 72
3.4. FORS / DARIS 78
3.4.1. Organisational Infrastructure 78
3.4.2. Digital Object Management 79
3.4.3. Future Plans 82
4. Conclusion **82**
References **83**
List of trust frameworks 83
List of guidelines and sources 83
List of Policies 85
CESSDA Archives 85
Other Policies 85
List of Figures **86**
List of Tables **86**
Appendix 1: How to prepare a Service Provider policy **87**

1. INTRODUCTION

1.1. AIMS AND OBJECTIVES

This report was made in the context of the CESSDA SaW project (CESSDA Strengthening and Widening), task 4.2 "Development Support: Development of the necessary administrative, technical, and methodological support needed to establish and develop trustworthy data archives". In general the aim of the task is to provide resources and guidelines for the development of service provider policies for research data archives and data centres. It aims to assist new and developing service providers to prepare and upgrade their policies and to integrate them into their internal workflows and processes.

The report maps the current status of CESSDA service provider policies through a selection of policy framework case studies; UK Data Archive, FSD (Finnish Social Science Data Archive), ADP (Slovenian Social Science Data Archive) and FORS/DARIS (the Swiss Centre of Expertise in the Social Sciences) are presented through in-depth descriptions, highlighting their respective strategic and operational policy frameworks.

Secondly, building on existing (preservation) policy templates and guidelines the main part of the report provides a policy model, the CESSDA Data Management Policy Framework, that can be used as a template and resources for the development and refinement of data archive policies.

1.2. BACKGROUND

Recent years have seen an explosive growth in the amount of research data. Data constitutes the raw material of scientific output and understanding, and the assembling, scrutinizing, organizing and disseminating of data serve an important purpose for the scientific community and the general public.

At the same time there has been a heightened awareness of the importance of openness and usability of data, across research disciplines, across technical platforms and across borders. This in turn has led to significant challenges with respect to data management, curation, access, and long-term data preservation. Key challenges were summed up by the High Level Expert Group on Scientific Data in 2010[1]:

- "How will we preserve the data? What will be the point of storing all this scientific data if, a century from now, it has degraded, been corrupted, or is simply too difficult for anyone but a well-equipped expert to use? Over time non-maintainability of essential hardware, software or support environment may make the information inaccessible and/or users may become unable to understand or use the data.

- How will we protect the integrity of the data? As the 'data tide rises' higher, how will we detect unauthorized alterations? Should every researcher, and every citizen, have access to the data repositories? Should there be different levels of access allowed
- How will we convey the context and provenance of the data? Given the emerging trend to make all publicly funded research data publicly available, just how will users from a wide

[1] European Union: Riding the wave. How Europe can gain from the rising tide of scientific data. Final report of the High-Level Expert Group on Scientific Data. October 2010. http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf

range of backgrounds understand and query the data they are accessing, and recognise the special circumstances under which it was collected?

- What new funding and business models will we need, so that everyone – researchers, enterprises, citizens – has adequate incentive to contribute to the data infrastructure? What kinds of data, under what circumstances, should be free?

- How will we protect the privacy of individuals linked to the data on the one hand, while providing researchers access to vital data on the other hand?"

A later report, by the High Level Expert Group on the European Open science Cloud, praises the long traditions and relatively healthy research infrastructures in Europe, but at the same time emphasise that "..a fundamental shift nevertheless is needed to match the potential to generate ever increasing amounts of data and to turn these data into knowledge as renewable, sustainable fuel for innovation in turn to meet global challenges"[2]

This fundamental shift in a sustainable research infrastructure is dependent on trustworthy and functional research data archives and data centres which can provide long-term data preservation and access services. The archives and data centres in turn must be able to define and refine their services to adapt to the constant changes and evolvement of new types of research data on the one hand, and new technologies and research tools on the other hand. There is a need to define and refine the extent and content of data curation services, and to identify rules for data management and processing that are designed for use across different disciplines and technological platforms.

Crucial means to this end are service provider policies that are rooted in organisational strategies that together give clear and unambiguous guidance and orientation on the operational level. Clear statements, agreed at the highest level in an organisation, about what needs to be preserved, why, and for how long, are essential for long-term management of data. Further, well-defined and publicly available policies and strategies contribute to consistency, accountability and transparency, and create trust among stakeholders. Policies do not exist in isolation; they are a part of a wider process that involves both internal and external actors. A comprehensive policy framework supports the shorter-term management of the institutional activities while also taking into account the longer-term vision of operational activities[3].

There are several benefits of a consistent and comprehensive policy framework. Some examples are mentioned in a report from the Preservation Advisory Centre[4]. According to the report a strong preservation policy will:

- clarify the relationship between the organisation's mission and preservation activity;
- clarify the scope of preservation activity by identifying the collections to be preserved, their significance and the desired retention period;
- act as a focal point for collaborative working across organisations and in some cases between organisations;

[2] "Realising the European Open Science Cloud: First report and recommendations of the Commission High Level Expert Group on the European Open Science Cloud": http://ec.europa.eu/newsroom/document.cfm?doc_id=18851
[3] DASISH, D4.4 "Comprehensive Policy-Rules for Data - Management in SSH": http://dasish.eu/publications/projectreports/DASISH_D_4.4-desember2014.pdf
[4] Preservation Advisory Centre / British Library: "Building a preservation policy": https://www.bl.uk/aboutus/stratpolprog/collectioncare/publications/booklets/building_a_preservation_policy.pdf

- clarify relationships with other aspects of collections management such as collections acquisition, access and security;
- provide a statement of accountability against which performance can be monitored;
- demonstrate the organisation's long-term commitment to its collections to funders and users, internal and external;
- act as a communication tool, internally and externally;
- provide a basis for the development of preservation strategy and preservation programmes;
- provide a basis for establishing priorities and justifying investment;
- demonstrate responsible stewardship for the benefit of current and future users;
- explain to users why certain actions are taken and others are not.

1.3. TERMINOLOGY – NUANCES, DEFINITIONS AND CLARIFICATIONS OF CONCEPTS

The name of this report, "Report overview of data management policies in social science data archives", present several terms that needs to be delimited and contextually understood. The terms "Data", "Data Management", "Policies" and "Archive" all require some conceptual reflection to clarify boundaries and to be defined in this specific context.

Data: data can mean very different things in different settings and contexts. Ask any researchers of their conception of "data" and you will get a different interpretation and understanding from each individual researcher. In our context, which is the service provision of long-term preservation and accessibility of data, data can be understood, on a highly abstract level, as "...a reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing"[5]. This highly abstract definition is taken from the OAIS model; examples mentioned are broad in range and can include "...a sequence of bits, a table of numbers, the characters on a page, the recording of sounds made by a speaking, or a moon rock specimen".

Data are basically materials and products generated or collected during the course of conducting research. But data are not only products in and for itself; data are also determined by the community of interest through process of peer review and project management. Hence, data has to be contextualized within the academic disciplines they stem from[6]. The definition of data can therefore often only be understood and contextualised within a specific research community.

Data Management: the Data Management Body of Knowledge (DAMA)[7] defines data (resource) management as "...the development, execution and supervision of plans, policies, programs and practices that control, protect, deliver and enhance the value of data and information assets". Hence Data Management is something to take into consideration both at the level of the individual or group, e.g. the active researcher who is handling and processing data in her work; and at the level of the institution or organisation that in some form is responsible for the management of data from several individuals or groups of people/researchers. It is in this second meaning that we use the term "data management" in this report.

Data Management is closely related to the term "*data (or digital) curation*", which involves maintaining, preserving and adding value to digital research data throughout its lifecycle. The active management of research data reduces threats to their long-term research value and mitigates the risk of digital obsolescence, as well as providing data sharing among the wider research community. Further it reduces duplication of effort in research data creation; curation enhances the long-term value of existing data by making it available for further high quality research[8].

The term "Preservation" should also be considered when discussing the concept of "Data Management". In the archival sciences and the field of curated long-term preservation, the term "preservation" is often understood very broadly. In many ways it overlaps with the definition of "Data Management" as it is understood in this report. Broadly defined "preservation" can comprise "…the whole of the principles, policies, rules and strategies aimed at prolonging the existence of an object by maintaining it in a condition suitable for use, either in its original format or in a more persistent format, while leaving intact the object's intellectual form"[9].

---

[5] Reference Model for an Open Archive Information System (OAIS): https://public.ccsds.org/pubs/650x0m2.pdf
[6] NIH Data Sharing Policy: https://grants.nih.gov/grants/policy/data_sharing
[7] https://www.dama.org/
[8] Digital Curation Centre: "What is digital curation?": http://www.dcc.ac.uk/digital-curation/what-digital-curation
[9] InterPARES 3 - The InterPARES Project: www.interpares.org/

Policy: a dictionary definition of policy (e.g. Merriam-Webster[10]) tells us that a policy in its most general form can be defined as "…prudence or wisdom in the management of affairs". Hence, a concept that deals with the ability to govern and discipline oneself (either as person or an institution) by the use of reason, skill and good judgement in the general use of available resources; and a caution or cautiousness towards possible dangers or risks. Hence a policy can assist an organisation in effective application of resources, and minimise possible dangers or risk to its operation(s).

More specifically, a second definition tells us that a policy can involve "…a definite course or method of action selected among alternatives and in light of given conditions to guide and determine present and future decisions". This involves a selection of paths or directions for the organisation or individual at hand. It also implies a set of clear and explicit goals and aims, which the methods or course of actions stem from.

If we apply these definitions into the context of data archives and data centres (i.e. providers of long-term preservation services), one could say that the "prudence" and "wisdom" represents the ideas, the objectives and the guiding principles of the organisation: it is where rational intention and purpose is defined before implementation and operation. The actual course or methods of action would represent the implementation of the ideas and intentions and purposes – i.e. an operationalization of wisdom into concrete processes and activities.

To summarise, in this report and in the policy model/template that follows, a "policy" is understood as a clear statement of intent that supports the data archives activities, adds to their legitimacy and trust and allows the institutions to capture the purposes of their operations.

Archive: the terms "archive", "data centre", and "service provider" as they are applied in this report refer to any organisations "…that intends to preserve information for access and use by a Designated Community" for the long term, where the Designated Community is an identified group of potential consumers, or users, who should be able to understand a particular set of information.

1.4. Approach and Methodology

The methods we have used when developing the policy framework template consist of two different approaches:

***Top-down approach*:**

Step #1: by analysing relevant literature and assessing existing policy recommendations, guidelines and templates, we compared content and policy elements and created a shortlist of possible elements to include in the comprehensive CESSDA Policy Framework Template.

Step #2: we assessed the content and elements of the different trust frameworks that constitutes the European Framework for Audit and Certification (explained in more detail below, in the policy template chapter) and compared these with the findings from **step #1**. A preliminary model and template based on these 'first principles' was then created.

Bottom-up approach:

The bottom-up approach consisted of analyses and assessments of the content of state-of-the-art,

[10] https://www.merriam-webster.com/dictionary/policy. Retrieved on March 2, 2017.

real-world data management policies and strategies. Elements and statements from individual organisational policies were compiled, and some were selected for in-depth case studies. The selected cases were the UK Data Archive, FSD (Finnish Social Science Data Archive), ADP (Slovenian Social Science Data Archive) and FORS/DARIS (the Swiss Centre of Expertise in the Social Sciences). These are presented through in-depth descriptions, highlighting their respective strategic and operational policy frameworks, in a separate segment in the report.

Finally, findings from the bottom-up approach, including the case studies, were aligned and compared with the preliminary model that was developed under the top-down approach. The preliminary model was then adjusted and refined to capture any relevant findings from the bottom-up approach, before it was finalised as an implementable policy template.

It should be noted that the model presented in this report may be refined through future iterations; a comprehensive policy framework template must be able to adapt and absorb ongoing changes and developments in the field in which it operates.

2. THE CESSDA DATA MANAGEMENT POLICY FRAMEWORK

Based on the processes and methodologies described above we have developed the CESSDA Data Management Policy Framework. It consists of a set of policy components that provide the foundations and organisational arrangements for designing, implementing, monitoring, reviewing and improving Data Management processes throughout an organisation whose main objective is to preserve and provide access to digital research data[11]. Although the framework is primarily aimed at CESSDA service providers, it is general in scope and design and can be applied by many different preservation organisations.

2.1. BACKGROUND

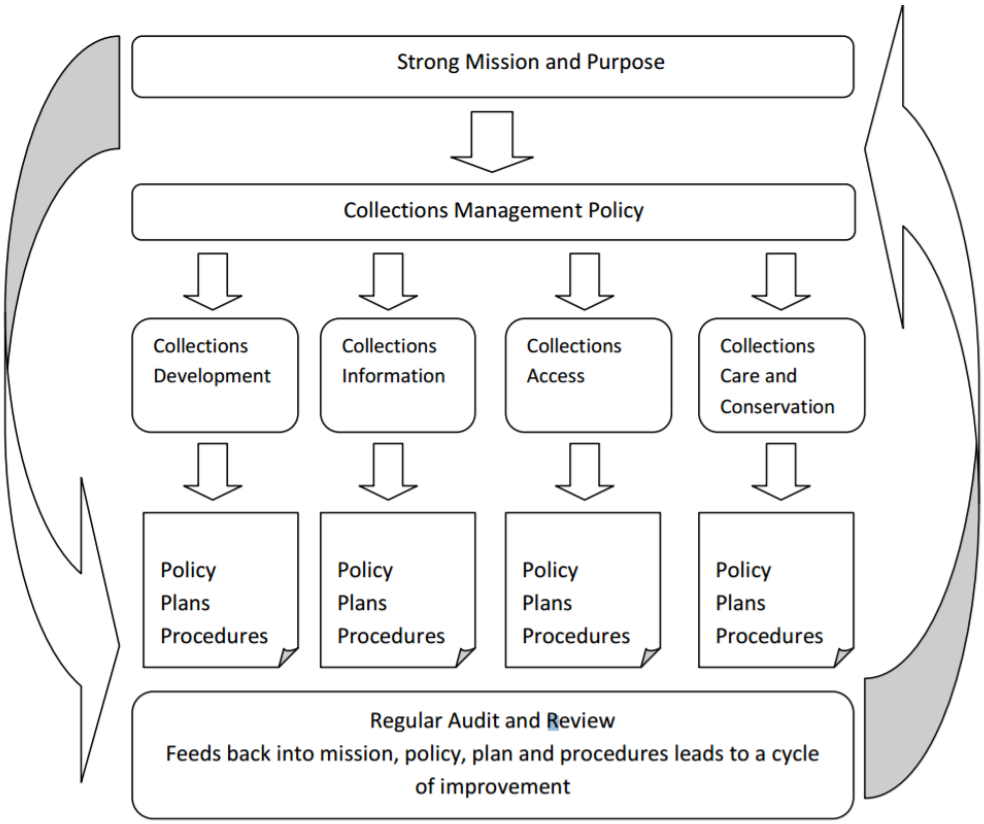
The Framework builds on existing guidelines and policy templates, real-world polices and former infrastructure projects with similar aims and objectives (see appendix for full list of sources).

There is a close connection between policies and trust, so the template also takes cues from the European Framework for Audit and Certification (also known as Trusted Digital Repository EU (TDR-EU)[12]. TDR-EU is a collaboration between Data Seal of Approval, the Repository Audit and Certification Working Group of the Consultative Committee for Space Data Systems (CCSDS) and the DIN Working Group "Trustworthy Archives – Certification". The framework consists of three trust/certification models: the DSA (Data Seal of Approval)[13], the DIN 31644 (the nestor seal for trustworthy digital archives) and the ISO 16363 (audit and certification for trustworthy digital repositories). Well-defined policies that are closely connected to established trust frameworks are essential to trustworthiness.

A possible way of arranging the policy clauses is to operationalise them into management entities that cover different aspects of the curation and management lifecycle of data. The *PAS 197 Collections Management Framework*[14] identifies four sub-components of the overarching Collections Management Policy, namely *Collections Development, Collections Information, Collections Access*, and *Collections Care and Conservation*. Each of the sub-components has its own set of policy clauses, plans and procedures (see figure 1).

[11] This definition is closely aligned to the definition of Collections Management Framework presented in *PAS197:2009 Code of Practice for Cultural Collections Management*. See also the Collection Trust's Accreditation Guidance sheet 1.
[12] http://www.trusteddigitalrepository.eu
[13] The ICSU World Data System (WDS) and the Data Seal of Approval (DSA) are in the process of aligning their procedures and organisations, and becoming a new standard and certification entity. See the DSA homepage for more information.
[14] *PAS197:2009 Code of Practice for Cultural Collections Management*. See also the Collection Trust's Accreditation Guidance sheet 1.

Figure 1: The PAS 197 Collections Management Framework



A similar approach is taken by the Audit and Certification of Trustworthy Digital Repositories (ISO 16363). The document provide a set of normative metrics against which a digital repository or archive may be judged. These metrics are grouped into three different subject areas: Organizational Infrastructure; Digital Object Management; and Infrastructure and Security Risk Management. Further it identifies three possible policy sets in which an organisation can address these metrics:

- The ***collection policy*** elements can be used to understand what the archive holds, what it does not hold, and why. The collection policy supports the broader mission of the archive. Without such a policy the archive is likely to collect in a haphazard manner, or store large amounts of low-value digital content. The collection policy helps the organization to identify what digital content it will and will not accept for ingestion. In an organization with a broader mission than preservation of digital content the collection policy helps to define the role of the archival function within the larger organizational context.

- The ***preservation policy*** elements are written statements, authorized by the archival management, which describes the different approaches to be taken by the archive for the preservation of objects accessioned into its holdings.

- The ***access policy*** elements are written statements, authorized by the archive management, which describes the approach to be taken by the archive for providing access to objects accessioned into its holdings.

Below these policy sets there is in effect also a *procedural activity level*, namely the ***actual services that are carried out***. These are specific, and in some cases measurable processes and process descriptions that applies to specific parts of the organisation; they are procedures that specifies the actions required to complete a service or to achieve a specific state or condition[15]. The relationship between the different levels, terms and concepts that are presented in the Audit and Certification of Trustworthy Digital Repositories can be summarised as follows: an archive is assumed to have a set of Organisational Infrastructure policy clauses, including an overall mission statement and a strategic plan. The mission statement and the organisational clauses identify business drivers and the general conditions and framework the organisation operates within. The strategic plan states how the mission will be achieved, in general terms with goals and objectives. Digital Object Management and Infrastructure and Security Risk Management declare the range of approaches the archive will employ to ensure that the goals and objectives are achieved. The range of approaches can in turn be addressed in a Collection policy, a Preservation policy and an Access policy. Finally the processes and procedures translate the policies into actions that the archive must carry out.

2.2. COMPONENTS OF THE CESSDA DATA MANAGEMENT POLICY FRAMEWORK

The CESSDA Data Management Policy Framework builds on both the *PAS 197 Collections Management Framework* and the *Audit and Certification of Trustworthy Digital Repositories*. In addition it aims to cover the Mandatory Responsibilities and most of the functions of the *Reference Model for an Open Archival Information System (OAIS)*.

A comprehensive policy framework will only be worthwhile if it captures the core business drivers of the organisation on the one hand, AND the operational level policy clauses on the other hand. Given the variety of activities involved in managing and preserving research data and ensuring their continued availability, a preservation policy must relate to other organisational policies. Hence the CESSDA Data Management Policy Framework consists of two policy clause levels:

Organisational Infrastructure Policy Clauses: are policies clauses and policy statements on a high organisational level which applies to all parts of the organisation. These clauses aim to capture the general business drivers, i.e. the conditions, resources and processes that are vital for the existence and continuation of an organisation. In many cases some of these clauses, especially the mission statement and the strategic plan, can be said to constitute distinct elements that exist on a higher level than the organisational clauses themselves, since these clauses defines the reason and rationale behind the existence of the archive / data centre.

- ***Mission Statement***: A written statement, authorized by the management of the repository, which, among other things, describes the commitment of the organization for the stewardship of digital objects in its custody.

- ***Strategic Plan***: A written statement, authorized by the management of the repository, that states the goals and objectives for achieving that part of the mission of the archive concerned with preservation. Preservation Strategic Plans may include long-term and short-term plans.

These in turn sets the premises for the rest of the organisational clauses and the implementation

[15] See CCSDS 652.0-M-1 *Audit and Certification of Trustworthy Digital Repositories*: https://public.ccsds.org/pubs/652x0m1.pdf

policy clauses on the level below.

Data Management Policy Clauses: or the *digital object management* policy clauses are the *implementation* of the organisational clauses. Data Management Policy Clauses describe the approaches taken to fulfil the organisational clauses. These implementations can apply to specific parts of the organisation, to specific processes, and, in the archival context, to specific parts of the data holdings or data collections. Implementation clauses define how institutional activities should 'behave' and is an implementation of purposes and properties defined at the organisational level.

When designing a policy framework it should be taken into consideration that a short, high-level document is much more likely to be agreed, read and acted upon than a very lengthy or complex document. For these reasons, it has been suggested that it may be more effective to limit the policy to short 2-3 statements of intent that are agreed at a senior level in the organisation (e.g. a mission statement or a general 'preservation policy'), and to highlight other relevant organisational policies or sub-policies to provide greater levels of detail as needed[16].

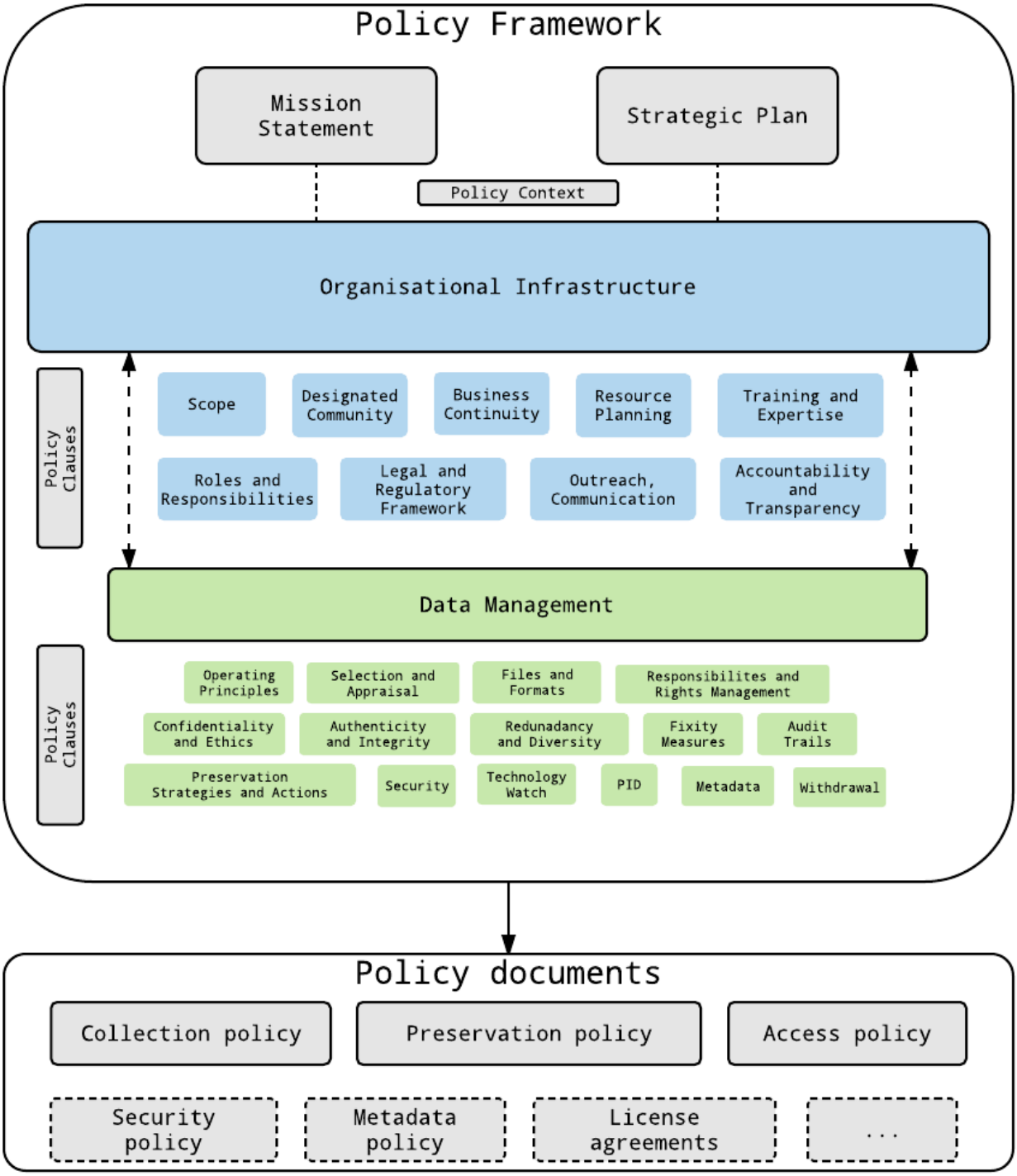
The CESSDA Data Management Policy Framework suggests following the approach taken by the Audit and Certification of Trustworthy Digital Repositories, namely to arrange the different policy clauses into three sets of policies: Collection policy, Preservation policy and Access policy (see definitions and explanations above).

It is important however, to underline that the policy framework and the distinction between Organisational Infrastructure and Data Management with a sub-set of policy clauses, is an *abstract* model. It can result in different policy documents, a different distribution of subjects between policy documents, different document names, etc.; it can be concentrated or 'compounded' into one or a few documents or it can be modulated and distributed among several different and specific policy documents and policy elements. Hence, the clauses are in the subsequent segments presented independent of the suggested arrangement of sub-set policies.

Figure 2: a suggested Data Management Policy Framework for CESSDA Service Providers.

The Organisational Infrastructure policy clauses, including an overall mission statement and a strategic plan identify business drivers of, and the general conditions and framework for, the organisational operations. Digital Object Management clauses declare the range of approaches the archive will employ to ensure that the goals and objectives are achieved. The policy clauses can result in different policy documents, a different distribution of subjects between policy documents, different document names, etc.

[16] Preservation Advisory Centre / British Library: "Building a preservation policy": https://www.bl.uk/aboutus/stratpolprog/collectioncare/publications/booklets/building_a_preservation_policy.pdf


Policy Framework
Mission Statement
Strategic Plan
Policy Context
Organisational Infrastructure
Policy Clauses
Scope
Designated Community
Business Continuity
Resource Planning
Training and Expertise
Roles and Responsibilities
Legal and Regulatory Framework
Outreach, Communication
Accountability and Transparency
Data Management
Policy Clauses
Operating Principles
Selection and Appraisal
Files and Formats
Responsibilites and Rights Management
Confidentiality and Ethics
Authenticity and Integrity
Redunadancy and Diversity
Fixity Measures
Audit Trails
Preservation Strategies and Actions
Security
Technology Watch
PID
Metadata
Withdrawal
Policy documents
Collection policy
Preservation policy
Access policy
Security policy
Metadata policy
License agreements
...

2.3. POLICY ELEMENT DESCRIPTION TEMPLATE

Each policy area will have a set of policy elements that will be described in a Policy Element Description Template. The content of the template is described in the table below.

Table 1: Policy description template

Policy element name	Content
Obligation	Whether the element is *recommended* or *encouraged*. A *recommended* element is considered as providing higher value to the policy framework than an *encouraged* element.
Description	Provides a description and/or a definition of the policy element.
Rationale	Rationale for, and importance of, the policy element.
Related policy elements	No policy element exists in isolation; this segment lists or describes how an element relates to other relevant policy elements.
Examples	One or more policy extracts/samples from existing policies.
Document(s)	List of suggested (names of) organisational documents that can communicate the policy element.
Issues to consider	Short list or description of possible issues and questions the policy element may raise in the archive.

2.4. POLICY CLAUSES

2.4.1. POLICY CONTEXT CLAUSES

The *Policy Context elements* provide bibliographic and general information about the policy framework; they frame the policy elements within the larger organisational context and provide basic information on the creation of the policies.

2.4.1.1. SCOPE OF POLICY

Obligation	Encouraged
Description	Sets the scope of the policy; defines what it addresses and to whom it is addressed.
Rationale	It is important for readers/audience of the policy to be able to identify what aspects of the archival activities it covers.
Related elements	Principle Statement; Contextual links
Examples	**ICPSR**: "*The audience for the framework includes ICPSR members, staff, digital content depositors, funders, and users*". **UKDA Preservation Policy [Scope]**: "*The scope of this policy is limited to the Archive's data collections. It deals with all aspects of preservation and applies to all materials held for long-term digital preservation by the Archive on behalf of the University of Essex and others. The policy only covers the preservation of data collections for which the Archive is the primary custodian. It does not consider preservation of other materials such as the Archive web pages, internal administrative documents and correspondence, and the Archive's intranet. Some of these materials including third-party standards, legislation, policies and procedures have a direct impact on the preservation process, but are governed by the Archive's Records Management Policy. The Archive's preservation strategy is not formulated in a single document but rather is embedded across a range of business information related to data curation and preservation*".
Issues to consider	Limitations / restrictions of the policy: what it does *not* cover; Description of the general policy framework / approach of the organisation (i.e. whether it is formulated in a single document or not.

2.4.1.2. CONTEXTUAL LINKS

Obligation	Encouraged
Description	Addresses how the policy integrates into the organisation and how it relates to other strategies and policies. These could include (but are not limited to) mission statements, strategic plans, a legal mandate, or other guidelines or documents that constitutes the parameters for the policy framework. The segment could include a sample or key section of the organisation's mission statement or mandate so that it relates the policy / policy framework to the organisation's mission and the communities it serves.
Rationale	The section can list other institutional documentation that has a relationship to digital preservation and/or this policy itself. This is an important part of contextualising and positioning the policy in the wider policy framework of the organisation.
Related elements	Scope of the policy; Legal and regulatory framework; Mission statement.
Examples	**ICPSR Digital Preservation Policy Framework**: "*ICPSR fulfils its role as a trusted steward of the heritage of the social sciences by capturing the results of past and current social science research for future researchers. The Digital Preservation Policy Framework supports that mission and is the highest-level digital preservation policy document at ICPSR. It makes explicit ICPSR's commitment to preserving the digital assets in its collections through the development and evolution of a comprehensive digital preservation program. The framework reflects the goals defined in the ICPSR Strategic Plan and contains references to other relevant ICPSR policies and procedure*." **UKDA Preservation Policy [Requirements]**: "*As a core activity in the Archive, preservation does not exist in isolation. It needs to take account of: the aims and objectives of the Archive; its strategic and operational plans; the UK Data Service Collections Development Policy and referenced requirements; the UK Data Archive's Information Security Policy and referenced requirements; the needs of the users of the Archive; archival theory and practice; and the place of the Archive within local, national and international frameworks*". See also **ADP Preservation Policy**, Appendix on "List of internal guidelines and instructions"; and **GESIS Preservation Policy**, "Related documents".
Issues to consider	References to the wider mission of the organisation; Inclusion of a key section of the mission statement (or similar strategic statements); Why the policy has been created, for example, for long term research prospects or a statutory obligation to meet archival requirements.

2.4.1.3. POLICY RELEASE

Obligation	Encouraged
Description	This is more of a policy consideration than an explicit element that should be addressed in the policy. Policy release deals with issues like selecting an appropriate language (i.e. wording) and appropriate communication channels for sharing and distributing the policy clause and/or the policy framework.
Rationale	Policy documents support communication. Appropriate communication channels should be used, and the language used should be balanced between general understandability and the need to employ specialist terms. A short glossary and list of definitions may be useful at the end or at the beginning of the policy document or framework, particularly if it is likely to be read by anyone unfamiliar with some of the terms and concepts used.
Related elements	Affects the full policy framework.
Examples	For examples of glossaries, see **UKDA Preservation Policy**, "Appendix on the Definitions of Terms"; and **ADP Preservation Policy**: "Terminology".
Issues to consider	Should there be different language versions? Through what channels should the policy to be released? Are form and content of the policy appropriate for these channels? Have the (knowledge level) of the audience been adequately considered? Is inclusion of a glossary necessary? Are there ways of obtaining feedback to the content of the policy? And are these feedback mechanisms clearly communicated?

2.4.1.4. BIBLIOGRAPHIC INFORMATION, UPDATING AND VERSIONING

Obligation	Encouraged
Description	This segment should include the history and bibliographic details of the policy version. It can provide the date of last revision, its intended duration and review process, and contact information for the authors.
Rationale	In principle, a policy should form the basis for the work of an institution and not be changed too frequently. On the other hand, a policy is not a static document but should be able to be adjusted to fundamental developments and changes in the organisation in question. And it is important to include version control information in the policy itself so that it is clearly communicated to both internal and external audiences of the policy.
Related elements	-
Examples	The **GESIS Preservation policy** contains several policy elements on bibliographic information and versioning: *Document created by; Document translated by; Date created; Version; Document name; Status (e.g. 'Published' / 'Working version'); Responsible (may be other than creator/author); and Review (e.g. 'Annually')*.
Issues to consider	How are necessary changes and updates to the policy organised? Who is responsible for the evaluation and changes? How can regular checking of the policy be ensured? How often will you commit to review of the policy (i.e. set timetable or on demand) and how will you alert people that it has been updated?

2.4.2. ORGANISATIONAL INFRASTRUCTURE POLICY CLAUSES

The **Organisational policy clauses** are policy clauses and policy statements on a high organisational level which applies to all parts of the organisation. These clauses aim to capture the general business drivers, i.e. the conditions, resources and processes that are vital for the existence and continuation of the organisation. In many cases some of these clauses, especially the mission statement and the strategic plan, can be said to constitute distinct elements that exist on a higher level than the organisational clauses themselves, since these clauses defines the reason and rationale behind the existence of the organisation.

2.4.2.1. MISSION STATEMENT AND PURPOSE OF THE ARCHIVE

Obligation	Recommended
Description	A mission statement that reflects the archives mission and commitment to the preservation of, long term retention of, management of, and access to digital information.
Rationale	It must be clear to all stakeholders (funders, depositors, users) that preservation of, and continued access to, the data is the responsibility, and an explicit role, of the archive.
Related elements	Scope of the archive; Strategic Plan; Designated Community
Examples	**ICPSR** Strategic Plan **[Mission Statement]**: "*ICPSR advances and expands social and behavioral research, acting as a global leader in data stewardship and providing rich data resources and responsive educational opportunities for present and future generations*." **UKDA** Preservation Policy **[Purpose]**: *"The UK Data Archive exists to support high quality research, learning and teaching in the social sciences and humanities by acquiring, developing and managing data and related digital resources, and by promoting and disseminating these resources as widely and effectively as possible. In order to achieve these activities, the primary function of the Archive is to provide long-term preservation activities, and the service-based activities which it carries out on behalf of others, including the provision of access to data which could not occur without the primary preservation activities carried out by the Archive*".
Documents	Mission statement; Charter; Legal, statutory, or government regulatory mandate; Principle Statement in Preservation Policy; Strategic plan.
Issues to consider	What will your organisation do? What will it not do? Consider high level synergies or links with other organisations. May also include information about the level of approval within the organisation that the mission statement has received (e.g., approved public statement, roles mandated by funders, policy statement signed off by governing board, etc.).

2.4.2.2. STRATEGIC PLAN

Obligation	Recommended
Description	A written statement that states the goals and objectives for achieving that part of the mission of the organisation concerned with preservation. It operationalises the mission into specific goals and objectives for achieving the mission. The Strategic Plans can consist of long-term and/or short-term plans. Short-term planning looks at the characteristics of the organisation in the present and sets out strategies for improving them. Long-term planning addresses the situation of the organisation in its social, economic and political environment and develops strategies for adapting and influencing its position to achieve long-term goals. It can suggest both its general direction and the prerequisite conditions for heading in that direction.
Rationale	The plan is intended to act as a tool for development and change. It can provide guidelines for setting the organisation´s priorities, for its allocation of resources to different purposes and for assessments of what it is the service provider wants to achieve in specific areas. This section can also be used to acknowledge the challenges that exists in the field of long-term preservation of digital objects.
Related to	Mission statement
Examples	UKDA Strategic Plan, 2010-2015 ICPSR Strategic Plan
Documents	Strategic Plan (should be a separate document, distinguished from the policy framework).
Issues	Address possible challenges of preserving and providing access to digital objects; consider whether the organisation wants to address and set periodical goals; clearly define objectives and goals, and measures to achieve them; consider addressing issues like resource prioritisation and resource allocation.

2.4.2.3. SCOPE OF THE ARCHIVE

Obligation	Recommended
Description	This is a statement that defines the type of content and information the archive will preserve, retain, manage, and provide access to. It identifies and describes the content, subjects, languages, etc., which the archive covers.
Rationale	It is important to understand, both internally and externally, what the repository holds and accepts, what it does not hold and rejects, and why. The scope supports the collection policy clauses and the broader mission of the repository. Without a clearly defined scope the archive is likely to collect in a haphazard manner, or store large amounts of low-value or "irrelevant" content. In an organization with a broader mission than long-term preservation research data the collection policy helps to define the role of the archive within the larger organizational context (source: Audit and Certification of Trustworthy Digital Repositories).
Related to	Mission Statement; Strategic Plan; Designated Community; Selection and Appraisal
Examples	From the **ICPSR Collection Development Policy [What data are in scope?]**: "*Building on broad, inclusive collection development policies from the past and also acknowledging the increasing importance of research infrastructure that supports cross-disciplinary research, ICPSR seeks data from many disciplines, in support of many methods, and about wide-ranging population groups as described below* [the lists are arranged in three categories: Disciplines, Data generation techniques, and Population groups]. *These lists are not intended to be exhaustive of what ICPSR is interested in collecting, but rather to show the wide range of data that are considered in scope for ICPSR. Also, as user demand for data broadens, and to better anticipate future trends in research, ICPSR is willing to consider additional kinds of data not appearing below. See also Out of Scope and High-Priority Areas*". **GESIS: Digital Preservation Policy; Selection and Acquisition**: "*The Data Archive primarily collects digital data from empirical social research. GESIS statutes determine the character and content of the service and this accordingly affects data offers. The latter are designed to serve the purpose of investigating "social change in national and international comparative and historical perspective" (Statutes §2, subpar. 2b) and support "international comparative research." The subject of "social change" is the main focus of GESIS's services and particular importance is assigned to comparability of data across space and time. However, in principle the GESIS Data Archive curates quantitative data sets relevant to social science research questions as long as data is well-prepared and documented*". **UKDA** Collections Development Selections and Appraisal Criteria states that UKDA acquires data to meet three central purposes, namely "…Potential secondary use and analysis for research; Teaching and learning use; and Replication and validation of research". There is also a segment on Appraisal criteria for submission and data discovery: "A core set of criteria is used for judging strategic high-level value, a second set for assessing user need and analytic value and a third set relating to usability and accessibility. These are classified into three stages. Value criteria are used as a first stage for appraising data for selection. If data do not pass the first stage, that is they are out-of-scope and do meet the core Collections Development remit, they are unlikely to pass to Stage 3 to be assessed for their accessibility and/or usability […]".
Documents	Collection Policy; Collection Development Policy

Issues	What subject areas will be included or excluded? Are there language considerations? Will translations be included or required? (Will text within data files, metadata or other documentation in other languages be translated into English, for example?)

2.4.2.4. DESIGNATED COMMUNITY

Obligation	Recommended
Description	This element identifies and defines the knowledge base of the archive. According to the OAIS model the Designated Community consists of "...potential Consumers who should be able to understand a particular set of information". Note that the Designated Community may be composed of multiple user communities. The Designated Community is defined by the archive and this definition may change over time.
Rationale	It is crucial for an archive to identify a Designated Community in order to be able to serve them optimally and make the preserved information useable and understandable in the long term. Note that sometimes it may be necessary to change the definition of the Designated Community; information originally intended for a narrowly defined community may need to be made more widely understandable at some future date.
Related to	Mission Statement; Scope of the Archive; Selection and Appraisal
Examples	**ICPSR Preservation Policy Framework [Access and Use]**: "*The designated community at ICPSR, as described by OAIS, includes traditional users, i.e., social science researchers and graduate students at member institutions; and newer categories of users, e.g., undergraduates, policymakers, practitioners, and journalists*". **UKDA Collections Development Policy**: "*The UK Data Service's designated user community is made up of social science and related data users within HE and FE in the UK, though best efforts are made for all users. (For the purpose of this definition, quantitative and qualitative data collections created by or for historians are considered to be 'social scientific'.) All users are expected to have a basic understanding of social science methods and techniques relevant to the data collections being accessed*".
Documents	Collection Development Policy; Collection Policy, etc.
Issues	Are the archive`s holding open to all or just a defined group of people? Are there different Designated Communities for different parts of the archive / data holdings?

2.4.2.5. BUSINESS CONTINUITY PLAN

Obligation	Recommended
Description	A statement that defines or refers to approaches the organisation will take to ensure the continued availability and accessibility of data in case the organisation ceases to operate. For example in the case of cessation of funding, either through an unexpected withdrawal of funding, a planned ending of funding for a time-limited project, or a shift of host institution interests.
Rationale	The failure or ceasing of an organisation threatens the long-term sustainability of its information content. A formal plan with identified procedures needs to be in place in case the organisation substantially changes its priorities, or the governing or funding institution substantially changes its scope, or the organisation ceases to operate (source: Audit and Certification of Trustworthy Digital Repositories).
Related to	Strategic Plan; Redundancy and Diversity; Security, etc.
Examples	**UKDA Preservation Policy** [**Funding and Resource Planning**]: "*The UK Data Archive is partially dependent on funding from the Economic and Social Research Council (and others) to ensure the longevity of the resources which it holds. If there were an indication that this funding were to cease succession planning activities (as required by the Succession Plan described in Business Continuity Incident Management Procedures) would be completed and made operational*". **UK National Archives, Business Recovery Plan**: "*In the event of a disaster affecting the daily business operations of The National Archives, the Business Recovery Plan will be activated along with Departmental Contingency Plans for those departments which are affected by the disaster*".
Documents	Preservation Policy; Business Continuity and Incident Management Plan; Business Recovery Plan
Issues	To what extent has the archive ensured that its holdings are preserved even after the archive itself has ceased to exist? What constitutes a crisis which would necessitate the transfer of tasks to third parties, and how is this decided?

2.4.2.6. FUNDING, STAFF AND RESOURCE PLANNING

Obligation	Recommended
Description	Explains how the archive is funded and staffed and how this is sustainable over time. The policy clause could include information on whether/how the organisation is hosted by a recognized institution that ensures ensuring long-term stability and sustainability and which is appropriate to its Designated Community. There should also be information and evidence on funding and available resources, including staff resources, IT resources, and a general budget.
Rationale	Service providers need funding to carry out their responsibilities, along with a competent staff that have the necessary expertise. These are resources that are necessary in order to ensure the viability of the repository over the period of time it has promised to provide access to its contents for its Designated Community. An organisation that undertakes the long term preservation of research data will need dedicated and qualified staff, either in house or contracted, to handle this.
Related to	Strategic Plan; Business Continuity Plan; Staff Training and Expertise
Examples	**ICPSR** Digital Preservation Policy Framework **[Financial Sustainability]**: "*ICPSR has identified specific resources to support and enhance its digital preservation function.* […] To sustain its digital preservation function, ICPSR has allocated a portion of its membership support to digital preservation services. In addition, ICPSR continually seeks external research funding to extend its digital preservation scope and capabilities and has secured contracts to fund specific initiatives. Detailed information about digital preservation funding is available in the ICPSR Annual Report and in the annual budget of ICPSR". **DANS** Preservation Policy **[Sustainability plans and funding]**: "*To fulfil its mission the Archive receives structural lump sum financing from both the KNAW and Netherlands Organisation for Scientific Research (NWO). Should a situation arise which threatens the continued existence of the Archive, these organisations are committed to taking responsibility for the future availability of the data entrusted to the Archive. Institutional depositors, as opposed to individual researchers, constitute another source of funding, as does participation in (international) research projects. This follows from goals in the DANS Strategy Policy*".
Documents	Annual reports; Strategic Plan; Annual budget document; Preservation Policy
Issues	Which cost model is used? Which budget planning documents exist? Which long-term funding plans exist for the digital archive?

2.4.2.7. STAFF TRAINING AND EXPERTISE

Obligation	Encouraged
Description	Describes any professional development program(s) that provides staff with skills and expertise development opportunities.
Rationale	Issues connected to technology, research methodology, research techniques, and data types will continue to change. As will the general practices for digital preservation and requirements of the service provider's Designated Community. The archive must ensure that its staff's skill sets evolve. Ideally the archive strives to meet this requirement through a lifelong learning approach to developing and retaining staff.
Related to	Funding, Staff and Resource Planning; Roles and responsibilities
Examples	ICPSR **Preservation Policy Framework [Challenges]**: "*Training and awareness: All of the ICPSR staff contribute directly and indirectly to the digital preservation function, though the majority of staff members do not have digital preservation as an explicit or significant portion of their responsibilities. ICPSR is committed to providing appropriate training for and raising awareness about digital preservation issues and developments both for its internal staff and for the broader community of data producers, data archivists, and data users*". See also the **ADP Preservation Policy**, which has an extensive coverage of training. See specifically *Competences and development of staff*.
Documents	Preservation policy; Specific training programs;
Issues	Consider including: staff professional development plans; certificates of training and accreditation; and evidence that the organisation reviews and maintains these documents as requirements evolve.

2.4.2.8. ROLES AND RESPONSIBILITIES

Obligation	Recommended
Description	Defines roles and responsibilities within the organisation. Roles and responsibilities should be clear to the employees in the organisation and written down in processes and procedures that are regularly updated. This could be updated job descriptions which sets out the required qualifications of the archive personnel, an organisational chart and/or a staff development plan based on the tasks and objectives of the archive.
Rationale	The archive should be able to document through development plans, organizational charts, job descriptions, etc. that the organisation is defining and maintaining the skills and roles that are required for the sustained operation of the archive. It is important that everyone in the organisation is aware of who is responsible for what. It is important for achieving the goals that the organisation has a clear view who is involved and who is entitled to make decisions. It may also be important for external stakeholders to identify the responsibilities within the archive.
Related to	Funding, staff and resource planning; Strategic Plan
Examples	**ICPSR Preservation Policy Framework [Roles and responsibilities]**: "*As an organization acting for its member institutions, funding bodies, and depositors, ICPSR has accepted responsibility for preserving its digital assets. Within ICPSR, the Director, the Digital Preservation Officer, the Computer and Network Services unit, the Collection Development unit, the topical archive managers, and the Collection Delivery unit all contribute to the management of the digital preservation function and the lifecycle of digital content at ICPSR. The ICPSR Council, an elected advisory board, evaluates high-level policy documents and reviews programmatic plans and progress. The roles and responsibilities within ICPSR for long-term management have been explicitly defined as part of the File-Level Archival Management Engine (FLAME) project*". See also **UK Data Archive Preservation Policy**, list of *Roles and responsibilities*; and **ADP Preservation Policy**: list of *Roles and responsibilities.*
Documents	List of staff roles and responsibilities; Titles and descriptions of work; Organisational charts; Preservation policy
Issues	Consider including staffing plan, competency definitions, job descriptions and evidence that the organisation reviews and maintains these documents as requirements evolve.

2.4.2.9. LEGAL AND REGULATORY FRAMEWORK

Obligation	Recommended
Description	An organisation that provides long-term preservation of research data should know and document the national and international (archival) legislations and regulation that applies to its activities and holdings. Any relevant legal act / legislation concerning digital preservation and the provision of access and re-use of digital material should be referred to.
Rationale	It is important for the service provider to know exactly which regulations apply to it and to ensure that relevant legal regulations are followed and that individual contracts exist for all relevant areas.
Related to	All activities and policies of the archive.
Examples	**FORS Preservation Policy**: "Preservation 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; DARIS end user license" See also **UK Data Archive Preservation Policy**, *Legal and regulatory framework*; and **ADP Preservation Policy**, *Legal framework and responsibilities*.
Documents	Preservation Policy Framework
Issues	Consider addressing the following issues: Legal status of archive/repository; Legal responsibility of the archive; List of national laws that the archive operates under; IPR; Service Level Agreements / Licenses; Data protection / Privacy issues.

2.4.2.10. OUTREACH, COMMUNICATION AND EXTERNAL TRAINING

Obligation	Encouraged
Description	A policy clause that gives an overview of any outreach, outsourcing, communication and education activities undertaken by/for the organisation, either with/for partner organisations, with/for the parent/host institution, or for the broader community. This aspect is about the interaction and communication with relevant stakeholders. It can deal with professional cooperation, outsourcing, training and general communication and outreach efforts to the public.
Rationale	The policy element can bring attention to any cooperation or outsourcing that exists with third-parties / other organisations. Outreach and communication efforts and strategies can ensure the continued communication about the services of the archive to the Designated Community and to the general public. It can also be about providing training and expert guidance to relevant users, as well as accommodating feedback mechanisms that can raise the knowledge and awareness of archival staff.
Related to	Strategic Plan; Designated Community; Funding, Staff and Resource Planning.
Examples	**ADP Preservation Policy [Outsource partners and expert guidance]**: "*The ADP cooperates with various external service providers, who perform certain tasks in the name of the archive. Cooperation with the ARNES takes place on the level of management of the network infrastructure, used by the ADP. The ADP cooperates also with the National and University Library (NUK) in developing software for long-term digital storage, based on Fedora Commons repository platform, as well as with certain other external service providers (maintenance of servers and IT support, programming). Signed agreements on cooperation, which are regularly updated, regulate the cooperation with individual outsource partners"*. See also section 3.2.1 [Submission Information Package] for more on ADPs training activities. When it comes to the provision of training services it may be sufficient to refer to available material, see ICPSR: "*The Guide to Social Science Data Preparation and Archiving provides guidance and templates for depositors to encourage complete and well-documented deposits*".
Documents	Preservation Policy Framework; Strategic Plan
Issues	Communication profile (presence on web and other relevant arenas); Provision of training material; Customer Service.

2.4.2.11. ACCOUNTABILITY AND TRANSPARENCY

Obligation	Recommended
Description	Statement that refers to any audits or certifications that the archive currently has passed, and/or future plans or efforts to undergo audit or to achieve certification. May also include statement on any self-assessments or other 'low-threshold' services or tools that the archive uses to continually quality check and control its activities.
Rationale	Increases trustworthiness towards all stakeholders (e.g. funders, users, owners, etc.). Obtaining accreditation or certification to appropriate standards is a way for ensuring both the quality of data archives and of the quality assurance process.
Related to	Authenticity and Integrity; Legal and Regulatory Framework; Roles and Responsibilities.
Examples	ICPSR Preservation Policy Framework **[Audit and Transparency]**: "*ICPSR is continued to an ongoing self-assessment and improvement process that aligns policies and practice at ICPSR with the Trustworthy Repositories Audit and Certification (TRAC) requirements that were revised and incorporated into the ISO/DIS 16363 (CCSDS 652-R-1). In 2006, ICPSR participated as a test audit in the Certification of Digital Archives research project conducted by the Center for Research Libraries. ICPSR is committed to a two-year cycle of self-assessment and a five-year audit cycle to evaluate, measure, and adjust the policies, procedures, preservation approaches, and practices of the digital preservation function. Current ICPSR policies are available on the ICPSR website and or may be made available upon request*".
Documents	Preservation Policy; Mission and Purpose; Strategic Plan.
Issues	If no certifications or audits have been carried out, has the organisation decided whether to be certified or not? Has your organisation decided on a standard to use in the certification process? What is the main reason to get audited/certified?

2.4.3. DATA MANAGEMENT POLICY CLAUSES

The policies under data management (or digital object management) declares the range of approaches the archive will employ to ensure good practice in creating and managing digital materials, and to achieve the mission and purpose of the archive. Data Management clauses are an implementation of the mission, purpose and strategy of the archive. Comprehensive policies for data management raise awareness of factors which need to be considered when acquiring, ingesting digital materials, and are crucial to the long-term preservation and continued viability and accessibility of digital materials.

The policy clauses presented in the following segment can result in different documents, a different distribution of subjects and policy clauses between documents, different document names, etc. They can be addressed in a general *Preservation policy framework* document, or it can be distributed among several different policy documents, with names like *Collection policy, Collection Development policy, Content policy, Submission policy, Preservation policy*, *Information Security Policy*, and *Access / Reuse policies, etc*.

Alternately, they can be aggregated into broader categories or broader statements. These statements/categories may have names like *content creation, content integrity and authenticity,* and *content maintenance*.

2.4.3.1. OPERATING PRINCIPLES

Obligation	Recommended
Description	This section provides overview of methodologies and philosophies supporting preservation activities and the management of digital objects. It should describe the general approach taken by the organisation to achieve its goals (of being a provider of long-term preservation and access to research material). It can consist of an explicit statement of the intent of the digital preservation program to comply with the Open Archival Information System (OAIS) Reference Model (ISO 14721).
Rationale	Clarifies the general approach taken by the archive in handling and preserving data.
Related to	Mission statement and purpose; Strategic Plan; all succeeding policy clauses under Data Management
Examples	**UKDA Preservation Policy**: "The Archive follows the broad guidance given in the OAIS reference model. The primary value to the UK Data Archive of the OAIS reference model is that it provides a framework on which its activities can be based. The Archive recognises the benefits of the OAIS model. When, in 2005, the Archive assessed its conformance with the OAIS model, the main divergence between model and practice was the strict separation of Archival Information Packages (AIPs) from Dissemination Information Packages (DIPs), and there were a number of activities within the data management function relating to monitoring and management which were not appropriate to the stated objectives of the Archive. This holds true today". **ICPSR Digital Preservation Policy Framework**: "*In achieving its digital preservation objectives, ICPSR recognizes the need to comply with the prevailing standards and practice of the digital preservation community. ICPSR is committed to developing its digital preservation policies, repository, and strategies in accordance with the Open Archival Information System (OAIS) Reference Model (2012). ICPSR tracks and responds to related OAIS initiatives, including developments in digital archives certification, persistent identifiers, preservation metadata, and the producer-archive interface. The mapping of ICPSR's preservation process to OAIS is synthesized in Digital Preservation Requirements Applied to ICPSR*".
Documents	General policy framework; Preservation policy;
Issues	Does the archive follow OAIS or other frameworks or standards for the archiving of digital content? Are there other operating principles or philosophies that support or guide the general activities of the archive?

2.4.3.2. SELECTION AND APPRAISAL

Obligation	Recommended
Description	This policy clause is an extension of the "Scope of the archive"; it deals with the identification of significant properties of information objects that are to be accepted into the archive, by for example delimiting accepted content into different curation categories or levels.
Rationale	In determining the properties to be preserved, the archive have to take into account the balance between the archive's overarching goals and targets, technical possibilities, and the costs of long-term preservation on the one hand and the needs of the designated community on the other hand. A robust appraisal process is required because acquisition decisions will be based on explicit procedures which can be justified; ingest activities can be prioritised and data collections will follow an appropriate ingest and access pathway; reporting can be conducted on collections development activity [Source: UKDA Collections Development Selection and Appraisal Criteria]
Related to	Mission Statement; Scope of the archive; Designated Community; Files and Formats
Examples	**UKDA** Collections Development Selections and Appraisal Criteria lays out five discrete 'Curation Categories' for which it treats all data collections: "*CURCAT1: Data collections selected for long-term curation. These data collections are made available for download, or accessible via online access tools; CURCAT2: Data collections selected for "short-term" management. These data collections will not (initially) be retained for long-term preservation, rather they will be backed-up (i.e., bit-level preservation only), made accessible and discoverable through online access tools (including Nesstar, InFuse, UKDS.Stat, etc.) or via in-house repository software (ReShare); CURCAT3: Data collections selected for 'delivery' only, e.g., where data from third parties are accessed via APIs/web services and delivered to end users via a UK Data Service interface. Issues such as level of trust in owner, what documentation/metadata are required, and how rights/registration are handled need to be agreed; CURCAT4: Data collections selected for "discovery" only. These collections will not be brought formally into the holdings of the UK Data Service, they will exist only in other (institutional) repositories, but the UK Data Service will harvest (or in exceptional circumstances, create) metadata records to allow these data collections to be found more easily; A fifth category (CURCAT5) relates to preservation-only which falls outside the scope of the UK Data Service Appraisal, and which is handled by the UK Data Archive. Data collections may be moved into higher or lower categories over time if the need arises*". **ICPSR** Collection Development Policy: "*ICPSR maintains a broad policy of inclusion based on two levels of curation services. ICPSR offers both an option of making data available to the user community in the condition deposited (no curation) along with an option for member-funded curation, which involves review, enhancement, and quality checking of the data to ensure usability and findability. The selection criteria employed for the two levels of curation services are: **No Curation**: The least restrictive stream of data entering ICPSR is data that receive no curation. Any depositor with data meeting the terms of ICPSR's broad Collection Development Policy may deposit and publish data in openICPSR, an open-access repository. Fees may be collected from non-member institutions for this service. Confidential data in openICPSR requiring restricted access will be supported with fees collected from users. Curation services, paid for by the contributor, are available; **Member-Funded Curation**: ICPSR also accepts and curates data that are considered to be valuable (either in the present or future) to the membership of ICPSR. There are additional*

	selection criteria placed on data that are curated for the ICPSR membership". See also **ADP Preservation policy**: Evaluating the quality of studies and adherence to the criteria of ingest.
Documents	Collection policy; Collection Development policy.
Issues	Does the archive use curation categories/levels? If so, what is the justification for these criteria?

2.4.3.3. FILES AND FORMATS

Obligation	Recommended
Description	This is a sibling element of "Scope of the Archive"; it specifies the types of content that are accepted into the Archive. It can consist of lists of accepted and/or preferred data formats, metadata formats and standards, size and volume limitations (if any), etc. It is a more specific clarification of methods, strategies and processes for acquiring data into the archive.
Rationale	This is a necessary step in defining the (format of) the information which is needed from the information producers or depositors. This process begins in general with the archive's mission statement, purpose and scope, and may be further specified in pre-accessioning agreements with producers or depositors (e.g., producer-archive agreements) and made very specific in deposit or transfer agreements for specific data and their related documentation.
Related to	Mission Statement; Scope of the Archive; Selection and appraisal; Security.
Examples	**ICPSR Collection Development Policy** [Data formats Accepted]: "*ICPSR prefers data in a readily useable format (see the Library of Congress' Recommended Format Specifications), accessible in a variety of computing and technological settings; ICPSR prefers data formats that promote easy access and use without compromising research value; ICPSR prefers that data files deposited in a raw format be transformable or convertible into formats usable by a variety of statistical or analytical software; ICPSR prefers data files unaccompanied by value-added software; Data in obsolete, proprietary, or hard-to-use formats may still be accepted by ICPSR, although these characteristics may compromise any future use of the data other than as-is, bit-level access*". See also **UKDA**, which presents a table of file formats that are accepted in the archive. UKDA distinguish between different types of data (e.g. quantitative, qualitative, video, image, etc.) and different preservation purposes (formats for sharing, reuse and preservation; and formats for preservation only). See also **ADP Preservation policy** **[Recommended formats]**. Similar to UKDA, ADP distinguishes between quantitative and qualitative data (formats).
Documents	Collection policy; Collection development policy; Format tables;
Issues	Will the archive only accept data/metadata in certain formats, or will you be flexible?

2.4.3.4. RESPONSIBILITIES AND RIGHTS MANAGEMENT

Obligation	Recommended
Description	A statement that makes clear how the organisation deals with and manages intellectual property rights / copyrights, (re)use and access conditions, and roles and responsibilities between archive and depositors and users of data. These issues can typically be addressed in a Deposit or License Agreement which is a formal agreement between depositor of the data and the service provider. The policy element should also address the set of tools and guidelines that are used within the organisation with the intent to control and manage rights issues (intellectual property, copyright, and conditions for access and use).
Rationale	A formal agreement enables the depositor and the service provider clarify and make explicit the roles and responsibilities for both parties. In addition the agreement can specify requirements such as access conditions, property rights, confidentiality and disclosure, embargo, etc. Correct handling of IPR-issues is necessary in order to prevent future mistakes or uncertainty concerning rights. If the organisation does not manage IPR-issues correctly it may violate copyright laws and risk law suits. An explicit policy statement on how this is handled will build trust among relevant stakeholders and reduce likelihood of infringement.
Related to	Legal and Regulatory Framework; Roles and responsibilities; Metadata and documentation.
Examples	**UKDA** Licence Agreement provides the UKDA with the necessary permissions to provide access to data under specified conditions of use: "*I grant a non-exclusive, royalty-free licence to the University of Essex acting by its UK Data Archive of Wivenhoe Park, Colchester, CO4 3SQ (the "Data Service Provider") to hold, make copies of, and provide access to the Data Collection, in accordance with the specified access condition below. In the event of the University of Essex ceasing to be a legal entity, this licence will be transferred to the Economic and Social Research Council (ESRC) or its successors*". When it comes to copyright and other legal issues connected to the deposited data, all depositors must agree to the following: "*I confirm that the Data Collection (i) is not and shall be in no way a violation or infringement of any copyright, trademark, patent or intellectual property right whatsoever of any person(s) or organisation (ii) does not contravene any laws currently in force, including but not limited to the law relating to defamation and obscenity*". Further the depositor must confirm whether he/she owner or joint owner of the data. If the depositor is not the owner of data, the depositor must provide names of those copyright holders he/she acts on behalf of. See also **GESIS** Archive Agreement **(*Property and usage rights*)**; and **DANS** License Agreement **(*Copyright and Related Rights*)**.
Documents	Collection Policy; Collection Development policy; Archive agreement; Depositor agreement; License agreement; Access Policy
Issues	In some cases, the archive accepts no responsibility for mistakes, omissions, or legal infringements within the deposited material. There may be situations in which the depositor deposits data that can possibly infringe the rights of other people or institutions. One way to mitigate against legal risks is to have a 'take-down' policy for

removal of objectionable items.

2.4.3.5. CONFIDENTIALITY AND ETHICS

Obligation	Recommended
Description	Data ingested into the archive for long-term preservation may contain confidential and sensitive information that must be protected to ensure they are not accessed by non-authorised users. The clause could contain information on the security of computer systems, including authorised access to and manipulation of data.
Rationale	Information security is important to protect sensitive information that is ingested into, preserved by, and provided access to, by the archive. In many cases there may be legal or regulatory obligations on the organisation, so a dedicated policy clause on these issues is often a requirement; the organisation must make sure that it follows all relevant legal, ethical and/or regulatory obligations.
Related to	Legal and Regulatory Framework; Security; Withdrawal and Take-Down; Responsibilities and Rights Management.
Examples	**UKDA** Preservation Policy [**Legal and Regulatory Framework**]: "*In terms of national standards for the management of information security, the Archive is certificated to BS ISO/IEC 27001: 2013 - Information technology -- Security techniques -- Information security management systems – Requirements and follows the Cross Government Actions: Mandatory Minimum Measures. Depositors are responsible for reviewing for any ethical issues relating to data collections they wish to deposit including those surrounding the potential for risk of harm to any participants in making data available to third parties. The ESRC's Research Ethics Framework (REF) provides guidance to both the Archive and to its researchers*". **ICPSR** Collection Development Policy **[Security, Privacy and Confidentiality Considerations]**: "*1. ICPSR requires that data deposited in the archive meet recognized standards for privacy and confidentiality of subjects studied. (For information on these standards, see the University of Michigan's Human Research Protection Program information). 2. ICPSR prefers to acquire data that can reside in the public domain. 3. ICPSR requires that data intended for public use be formatted so that identifiers inadvertently included in the data can be removed using standard practices without reducing the research value of the original data. 4. Any access limitations that ICPSR might apply to specific data collections (e.g., a requirement that restricted-use agreements must be signed) should be legally justified and manageable given ICPSR's resources, goals, and mission*". See also **ICPSR`s Approach to Confidentiality**.
Documents	Preservation Policy; Information Security Policy; Access Policy, etc.
Issues	Does the archive follow security procedures and standards that can be codified and monitored (ISO, 2013a; ISO, 2013b; ISO 27002; ISO 27001)? Are there sufficient access controls in place? Does the archive use encryption, redaction (the process of analysing a digital resource, identifying confidential or sensitive information, and removing or replacing it) or other anonymisation techniques?

2.4.3.6. AUTHENTICITY AND INTEGRITY

Obligation	Recommended
Description	Policy statement that addresses the ways in which the archive operates a data and metadata management system suitable for ensuring integrity and authenticity during the processes of ingest, storage, and provision of access. Data Seal of Approval defines *integrity* as any measures that ensure that changes to data and metadata are documented and can be traced to the rationale and originator of the change. Integrity checking covers approaches like encryption, digital signatures, fixity checks etc. *Authenticity* covers the degree of reliability of the original deposited data and its provenance, including the relationship between the original data and that disseminated, and whether or not existing relationships between datasets and/or metadata are maintained.
Rationale	One of the primary objectives of long-term preservation is that the preserved digital objects, once ingested into the archive, are not changed without intent.
Related to	Fixity measures; Accountability and Transparency; Audit Trails and Documentation; Roles and responsibilities; Legal and regulatory framework; Responsibilities and rights management; Metadata and documentation. Understood broadly, authenticity and integrity covers the entire data lifecycle and data object management processes within the archive. As such, authenticity and integrity can be considered as an underlying philosophy, mindset or operating principle, which affects the entire organisation.
Examples	ICPSR Digital Preservation Policy Framework [**System Security**]: "*The processing procedures for digital content at ICPSR actively address the need for ensuring the accuracy and completeness of digital content through the careful comparison of documentation and data submitted and the generation of metadata and documentation for data. The implementation at ICPSR of an automated deposit form addressed the need for ensuring the authenticity of digital assets by requesting detailed information and signatures for submission. ICPSR ensures the authenticity and integrity of its digital content through the active and ongoing use of checksums from receipt of the digital content onward. In addition, ICPSR conducts periodic reviews and audits of its digital content in archival storage*". See also the **UKDA Preservation Policy**, where "authenticity" is a recurring theme throughout the policy, and "Integrity measures" are treated in a separate segment.
Documents	Collection Policy; Preservation Policy; Access Policy, etc.
Issues	Fixity checks / checksums; Documentation of the completeness of the data and metadata; Details of how all changes to the data and metadata are logged; Description of version control strategy and logging of changes; Does the repository check the identities of depositors?

2.4.3.7. REDUNDANCY AND DIVERSITY

Obligation	Recommended
Description	Addresses the approach taken by the archive when it comes to the physical storage of digital objects. It presents the principles that are being employed in the design and/or selection of storage systems for preservation. These systems can still become obsolete over time so the archive should have policies and procedures for the regular migration of digital materials between storage systems as they become obsolete. Note that *migration* between storage systems is a different subject to the *migration* between file formats (see policy clause Preservation Strategies and Actions).
Rationale	Maintaining a systematic and sustainable process for the physical storage of digital objects remains a fundamental requirement in ensuring long term digital preservation. A resilient IT storage system should include onsite storage and/or remote cloud storage and automated replication of digital materials across multiple sites and systems. Different types of storage technology and storage systems to spread risk and achieve a balance of data safety, easy access and manageable cost.
Related to	Fixity checks; Technology watch; Audit trails
Examples	UKDA Preservation Policy [**Physical data preservation and storage**]: "*In order to best safeguard long-term preservation, the Archive follows a policy of multiple copy resilience. Five versions of the complete preservation system are held. A main near-line copy and a shadow copy on two separate preservation servers, only accessible using a dedicated preservation account, the access online copy (on the mirror preservation server), and copies are generated for user access and dissemination. There are also a near-site online copy on a server located in another location within the University of Essex, and an off-site online copy. An encrypted hard disc-based offline copy is also maintained. The Archive follows best practice in the storage and housing of magnetic and optical media. In particular, for environmental conditions for storage media BS ISO 18925:2002) and for the storage of archival materials (BS 5454)*". Harvard Dataverse Preservation Policy [**Backup schedule**]: "*HUIT* (Harvard University Information Technology) *backs up all of the application/system files and databases nightly. It is stored off-site in Carlstadt, New Jersey for 45 days. All research data files in the repository are replicated every 4 hours to a second off-site storage array at 1 Summer St, Boston, MA. Since March 2013, HUIT incorporated the data content of the Harvard Dataverse repository into the DRS Storage Infrastructure. This makes use of the storage management software to create a tape copy of the data to be stored for the long-term at the Harvard Depository*".
Documents	Preservation policy; Security; Authenticity and Integrity, etc.
Issues	Multiple independent copies of digital material stored in different geographic locations; Employing a combination of online storage systems and offline media.

2.4.3.8. FIXITY MEASURES

Obligation	Recommended
Description	Addresses the archives strategy / approach towards integrity or fixity check. A fixity check is a "...process of verifying that a File or Bitstream has not been changed during a given period. A common Fixity Check method is to compute a Message Digest ("hash") at one point and recalculate the Message Digest at a later point; if the digests are identical, the object has not been altered" [Source: PREMIS Data Dictionary]. A fixity check is basically a mechanism to verify that a digital object has not been altered in an undocumented manner. Checksums, message digests and digital signatures are examples of tools to run fixity checks.
Rationale	Fixity measures help avoiding unintended alterations of data. Fixity information, the information created by these fixity checks, provides evidence for the integrity and authenticity of the digital objects and is essential to enabling trust.
Related to	Authenticity and Integrity; Accountability and Transparency; Audit Trails and Documentation.
Examples	**ICPSR Digital Preservation Policy Framework** [**System security**]: "*ICPSR ensures the authenticity and integrity of its digital content through the active and ongoing use of checksums from receipt of the digital content onward. In addition, ICPSR conducts periodic reviews and audits of its digital content in archival storage*". **figshare Preservation Policies** [**Fixity and authenticity**]: "*All data files are stored along with a MD5 checksum of the file content. Files are regularly checked against their checksums to assure that file content remains constant. From infrastructure to development processes, we follow industry standards in order to build a highly secure, adaptable and maintainable product*".
Documents	Preservation policy;
Issues	The use fixity measures such as checksums to record and regularly monitor the integrity of each copy of the digital material; If corruption or loss is detected, how are other copies used to create a replacement?; Storage of fixity information alongside the digital materials and also in separate databases or systems.

2.4.3.9. AUDIT TRAILS AND DOCUMENTATION

Obligation	Encouraged
Description	An audit trail is a chronological record of system activities to enable the reconstruction and examination of the sequence of events and/or changes in an event (Source: CNSS National Information Assurance Glossary). Maintaining a full audit trail of all preservation actions performed on a representation of a record ensures that the actions applied to that representation are documented in sufficient detail for present and future users to understand their nature and consequences. In the OAIS model it overlaps with the *Provenance Information* which documents the history of the Content Information. The Provenance Information tells the origin or source of the Content Information, any changes that may have taken place since it was originated, and who has had custody of it since it was originated, providing an audit trail for the Content Information.
Rationale	Audit trails give future users some assurance as to the likely reliability of the Content Information as it contributes to evidence supporting Authenticity. Audit trails document how digital materials have been acquired and transferred into the storage systems as well as how the storage systems are set-up and operated. This information can be used to provide audit information on data authenticity.
Related to	Authenticity and integrity; Fixity measures.
Examples	UKDA [**Integrity Measures**]: "The UK Data Archive takes its role as custodian of data collections seriously. To this end the complete chain of custody of all data collections is documented through metadata. All actions are explicit, complete, correct and current. However, only the 'original' version can be said to be an integral copy of the version deposited with the Archive. The preservation and dissemination versions are considered to be authentic and there is an audit trail of all alterations in the preservation and dissemination versions which relates back to the original deposited version".
Documents	Preservation policy; Metadata and Documentation, etc.
Issues	Consider addressing some of these issues: capital equipment inventories; documentation of the acquisition, implementation, update, and retirement of critical archive software and hardware; file retention and disposal schedules and policies, copies of earlier versions of policies and procedures; minutes of meetings; etc.

2.4.3.10. PRESERVATION STRATEGIES AND ACTIONS

Obligation	Recommended
Description	Preservation strategies and actions focus specifically on efforts that can help mitigate the technical challenges of preserving digital materials over time, one of the main challenges being the technological obsolescence of formats. A simple definition of obsolescence is the process of becoming outdated or no longer used. Obsolescence is an issue because all files have their own hardware and software dependencies. Some of the most widely utilised strategies to contain usability and avoid obsolescence is *migration* and/or *emulation*. Format *migration* is different from storage system migration (see policy clause *Redundancy and Diversity*). It involves transferring or transforming (i.e. migrating) data from an ageing/obsolete format to a new format, possibly using new applications systems at each stage to interpret information. Moving from one version of a format standard to a later standard is a version of this method. *Emulation* offers an alternative solution to migration that allows archives to preserve and deliver access to users directly from original files. This technique attempts to preserve the original behaviours and the loo k and feel of applications, as well as informational content. It is based on the view that only the original programme is the authority on the format and this is particularly useful for complex objects with multiple interdependencies (Source: DPC Digital Preservation Handbook, on Preservation action).
Rationale	A core goal of digital preservation actions is to preserve the integrity and authenticity of the material being preserved, despite the generational changes in computing and storage technology.
Related to	Authenticity and Integrity; Files and Formats; Redundancy and Diversity; Metadata and Documentation.
Examples	**FORS: DARIS Preservation Policy** [**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*".
Documents	Preservation Policy; Preservation strategy; etc.
Issues	Are there cases where emulation is more viable than migration?

2.4.3.11. SECURITY

Obligation	Recommended
Description	A policy clause that addresses the implementation of security measures towards the archives' premises and data holdings. Security measures affect both staff and users. The statement should address issues like physical security, (building and perimeter security and security of access (e.g. access by staff, contractors and users to storage areas)).
Rationale	In order to protect the organisation's information assets there should be specific requirements and descriptions of security strategies laid out in an information security policy. It guards against breaches in the organisation's information security.
Related to	Confidentiality and Ethics; Legal and Regulatory Framework.
Examples	UKDA Preservation Policy **[Security]**: "*The UK Data Archive is committed to taking all necessary precautions to ensure the physical safety and security of all data collections that it preserves: fire prevention and protection system; physical intruder prevention and detection systems; [and] environmental control systems. The repository rooms are equipped with multiple key entries and a security-protected swipe-card system linked to an on-site alarm system and to the University Security Office. The swipe-card system is maintained by the Preservation and Systems Manager, and access is restricted to two key members of staff –the Preservation & Systems Manager and the Data Security Manager. The repository rooms are located outside of the secure working area of the Archive. The SSRC building in which the Archive is housed, is locked between 7.30 p.m. and 7.30 a.m. and all weekend, and is regularly patrolled by University of Essex security staff. All machine room computer systems are locked by a logon password system to prevent unauthorised access in the case of a security breach of the room. The Archive's suite of information and premises security are documented in our Statement of Applicability for ISO/IEC 27001. The Archive was recommended for registration to this standard in June 2010 and maintains that registration. These are detailed in the Archive's Information Security Policy, Information Security Management Policy and Premises Security Procedures*".
Documents	Preservation policy; Security policy; Information security policy, etc.
Issues	Consider following the ISO/IEC 27000 family of standards (Information security management systems).

2.4.3.12. TECHNOLOGY WATCH AND MONITORING DESIGNATED COMMUNITY

Obligation	Recommended
Description	The OAIS reference model states that "…the Monitor Technology function is responsible for tracking emerging digital technologies, information standards and computing platforms (i.e., hardware and software) to identify technologies which could cause obsolescence in the Archive's computing environment and prevent access to some of the Archive's current holdings". Further: "…the Monitor Designated Community function interacts with Archive Consumers and Producers to track changes in their service requirements and available product technologies. Such requirements might include data formats, media choices, preferences for software packages, new computing platforms, and mechanisms for communicating with the Archive". Technology Watch deals with how the archive observes/monitors, tracks, filters out and assess potential technologies that may replace and improve upon technologies that are already in use in the archive. The Technology Watch process should be capable of identifying any scientific or technical innovation with potential to create opportunities or avoid threats for the holdings of the archive. Changes might be monitored in the following areas: packaging, storage, formats, tools, environment and access mechanisms. The Technology Watch process should also be able to capture requirements and needs from Designated Communities and other relevant stakeholders (e.g. needs and expectations of users and producers of data, emerging tools for machine to machine access, and formal feedback from users and producer).
Rationale	Monitoring technologies and Designated Communities alert the repository about changes in the external environment and risks that could impact on its ability to preserve and maintain access to the information in its custody, such as innovations in storage and access technologies, or shifts in the scope or expectations of the Designated Community (see Lavoie, 2015). Technology watch represents a safeguard against a constantly evolving user and technology environment. It detects changes or risks that impact the repository's ability to meet its responsibilities, designs strategies for addressing them, and assist in the implementation of these strategies within the archival system (see Digital Preservation Handbook, on Preservation Planning).
Related to	Mission Statement; Designated Community; Preservation Strategies and Actions
Examples	**Parliamentary Archives: A Digital Preservation Policy for Parliament** [**Technology Watch**]: *"Parliament will maintain a technology watch function to monitor technological change within Parliament and the external environment. Significant changes may require risk assessments to be revised. Parliament will ensure that all activities which will lead to technology change take proper account of their preservation impact, and incorporate appropriate mitigation*".
Documents	Preservation Policy; Annual reports, etc.
Issues	The archive should acknowledge and understand that storage technologies, products and services all have a short lifetime; It should use technology watch to assess when migrations might be needed; It should keep an eye on the viability of storage vendors or classes of storage solution; and be proactive in migrating storage before digital material becomes at risk.

2.4.3.13. PERSISTENT IDENTIFICATION

Obligation	Recommended
Description	A persistent identifier is a long-lasting reference to a digital resource. Typically it has two components: a unique identifier; and a service that locates the resource over time even when it's location changes. The first helps to ensure the provenance of a digital resource (that it is what it purports to be); whilst the second will ensure that the identifier resolves to the correct current location.
Rationale	Digital material can easily be copied and altered. Assigning a persistent identifier to the digital object will improve its identification and findability and accessibility. The persistent identifier also adds to the authenticity of the object. A digital archive should use internal identifiers to manage the information objects and their representations and, where applicable, their parts and relationships (part/totality, different variants, versions etc.), especially to ensure unique assignment of the content data to the metadata. The use of externally visible, standardised persistent identifiers ensures reliable tracing of the information objects and their representations, and consequently also access [Source: nestor].
Related to	Authenticity and Integrity; Metadata and Documentation
Examples	FORS Preservation Policy DARIS **[Access]**: "*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*". Harvard Dataverse Preservation Policy **[Preservation of Materials Deposited in the Harvard Dataverse]**: "*…once a dataset is published, the repository guarantees archival and long term access to that dataset with a DOI persistent identifier provided by the California Digital Library's (CDL) EZID service (DataCite member)*". Dryad Terms of Service **[Sustainability]**: "Through registration of DOIs for Data Packages and files, Dryad shows its commitment to online preservation of its Content for many years to come. Registration also allows identifiers to be seamlessly redirected in the event of URL changes. Dryad's participation in the DataONE network ensures that all its data will be available through other institutions if the Dryad organization ever dissolves".
Documents	Access policy; Preservation Policy, etc.
Issues	Does the archive provide persistent identifiers to all versions of data (objects)? How is versioning handled? When is the persistent identifier assigned or minted to the data object?

2.4.3.14. METADATA AND DOCUMENTATION

Obligation	Recommended
Description	Metadata is data about a digital resource that is stored in a structured form suitable for machine processing. It serves many purposes in long-term preservation, providing a record of activities that have been performed upon the digital material and a basis on which future decisions on preservation activities can be made in the future, as well as supporting discovery and use. Documentation is the information (such as software manuals, survey designs, and user guides) provided by a creator and the repository that supplements the metadata and provides enough information to enable the resource's use by others. It is often the only material providing insight into how a digital resource was created, manipulated, managed and used by its creator and it is often the key to others to make informed use of the resource. [Source: Digital Preservation Handbook]
Rationale	The purpose of preservation metadata is to support the goals of long-term digital preservation, which are to maintain the availability, identity, persistence, renderability, understandability, and authenticity of digital objects over long periods of time.
Related to	Relevant for the full lifecycle of data, including issues like technical requirements, change-logs and audit trails, authenticity, rights management, and findability and reuse.
Examples	Metadata may be a recurring theme throughout a policy framework, and/or it can be addresses in a separate segment, like for example in the **ADP Digital Preservation Policy** [**Metadata standards and interoperability**]: "[…] *When defining study descriptions (metadata) the ADP uses the DDI standard (Data Documentation Initiative). The Data Archivist with the help of the data provider checks the submitted documentation for consistency and on this basis prepares the final metadata description according to DDI. The exhaustiveness of metadata differentiates according to the level of study processing and may include, besides the study description, also the full description of variables in cases of the most important studies, including frequency distribution and question texts. The use of standard metadata and interconnectivity of identifiers with other information services enable quality and sustainable directed development of data services. This represents the base of connecting various services of scientific information giving, and at the same time represents the basis for coordination and cooperation in the framework of collective services of the international infrastructural unit of the CESSDA. Interoperability is also one of the exposed focuses of the OAIS standard and the demands of the European Commission in the framework of the open access to research data H2020. Study descriptions according to DDI are interoperable with the catalogs that harvest DDI formats (CESSDA Data Catalogue, DataVerse etc.)* […]".
Documents	Preservation policy; Collection policy
Issues	Does the archive use standard(s) for metadata? Does the archive define a minimum set of required fields in a metadata scheme upon deposit of digital objects? Does the archive distinguish between descriptive metadata, preservation metadata and/or structural metadata? Are the metadata records of the archive open and harvestable?

2.4.3.15. WITHDRAWAL AND TAKE-DOWN

Obligation	Recommended
Description	A take-down or withdrawal policy clause serves to try and minimise the risk of making inappropriate material available through the archive. This is especially relevant for archives that facilitate (un-curated) self-deposit services. The policy clause should address incidents and occurrences that may commence withdrawals and take-downs. It is important to know in advance what measures should be taken to handle these situations.
Rationale	It is advisable to think through and develop a take-down policy as part of general risk-management. Such a policy puts in place some measure of control if a serious situation does arise and can be seen as a type of insurance. A relevant situation may never happen, but if it does, one should be prepared.
Related to	Legal and regulatory framework; Responsibilities and Rights Management; Security.
Examples	**UK Data Archive** Preservation Policy [**Data collection withdrawal**]: "*The UK Data Archive operates a multifaceted policy towards data collection withdrawal. Like The National Archives, the Archive distinguishes between 'soft deletion' whereby certain references to the withdrawn content are deleted, but not the content itself, and 'hard deletion' whereby the content and all references to it are deleted. In the case of soft deletion the data collection is only accessible to Digital Preservation Systems and Security and Ingest Services staff. The Archive chooses soft deletion as the default method of withdrawal since it is too expensive to remove data collections, and their physical removal would present unacceptable risks to other parts of the collection. The Archive has, in the past, undertaken hard deletion of collections which are archived, preserved and disseminated elsewhere. In cases of the withdrawal of a data collection, the administrative metadata are updated, and the external view of the catalogue record is updated to reflect the change of status of the collection (with information about why the collection had been withdrawn, the dates of its availability, and where appropriate the reasons for withdrawal)*". **ICPSR** Deaccession Policy: "*ICPSR permanently archives deposited files. On an ongoing basis, ICPSR evaluates its data holdings with regard to maintaining access and reserves the right to discontinue the distribution of a data collections when deemed appropriate. When materials are deaccessioned, the data are no longer publicly accessible at ICPSR, although they are still preserved in ICPSR's archival storage. Because digital files are assigned a persistent digital object identifier (DOI), the study description is still available to view, but is not searchable through the ICPSR search index. Web crawlers are instructed to ignore the descriptions (via the robots exclusion protocol)*". **Dryad** Terms of Services [General Miscellaneous Provisions]: "*Each party shall comply with all applicable laws. Dryad reserves the right to take any actions necessary in order to comply with applicable laws or court orders in any jurisdiction. […]*"
Documents	Deaccession policy; Access Policy; Submission policy.
Issues	Under what conditions will the repository choose to remove items? Reasons for withdrawal by repository might include copyright violation, legal requirements and proven violations, (threats of) national security, falsified research, confidentiality concerns, etc. Will items be removed at the request of the depositor? If items are withdrawn, the policy clause should indicate the terms of the withdrawn items, for example: withdrawn items are deleted entirely from the database; withdrawn items are not deleted, but are removed from public view; identifiers/URLs for withdrawn items are retained (indefinitely, transiently, or not at all), etc.

3. CASE STUDIES

3.1. UK DATA ARCHIVE

3.1.1. ORGANISATIONAL INFRASTRUCTURE

The UK Data Archive (UKDA), founded in 1967, is curator of the largest collection of digital data in the social sciences in the United Kingdom[17]. UKDA is primarily funded by the Economic and Social Research Council (ESRC) through its resourcing of the UK Data Service[18], and is a specialist department of the University of Essex[19].

The purpose of the UK Data Archive is explicitly stated in the Preservation Policy:

"The UK Data Archive exists to support high quality research, learning and teaching in the social sciences and humanities by acquiring, developing and managing data and related digital resources, and by promoting and disseminating these resources as widely and effectively as possible. In order to achieve these activities, the primary function of the Archive is to provide long-term preservation activities, and the service-based activities which it carries out on behalf of others, including the provision of access to data which could not occur without the primary preservation activities carried out by the Archive".

The policy statement on purpose is closely connected to the **Strategic Plan** (2010-15) which explicitly states – in broad terms - the mission, strategic challenges, and goals for the archive[20].

The **Preservation Policy**[21] also covers the legal and regulatory framework which the archive operates under and lists the laws, acts, directives and regulations which applies to the preservation of its collection. It also clarifies the relationship between the depositor of a data collection (data set) and the archive, and refers to a specific legally-binding document (the **Licence Agreement**, see below) which confirms the rights and obligations of both parties and offers an opportunity for depositors to specify the conditions under which access may be given to third parties[22].

Further, the policy has a segment which describes the roles and responsibilities of the different sections, or departments, of the archive. It also lays out the roles, responsibilities and accountability of the archive staff.

Example 1 (on roles/responsibilities of sections):

"The Ingest Services section is responsible for the validation of the data and documentation, including checking of consistency, accuracy and suitability for preservation and secondary analysis, and producing and enriching metadata about the data resources in order to populate finding aids.

[17] The Royal Society: Costs of digital repositories: https://royalsociety.org/topics-policy/projects/science-public-enterprise/digital-repositories/
[18] UKDA hosts and manages the UK Data Service, which provides access to data sets (both quantitative and qualitative data) from a wide range of disciplines: http://ukdataservice.ac.uk/get-data/about.aspx
[19] UKDA at University of Essex: http://www.essex.ac.uk/depts/ukda.aspx
[20] UKDA Strategic Plan, 2010-15: http://www.data-archive.ac.uk/media/196518/ukda-strategicplan20102015full.pdf
[21] UKDA Preservation Policy (version 08.00): http://www.data-archive.ac.uk/media/54776/ukda062-dps-preservationpolicy.pdf
[22] UKDA Preservation Policy (version 08.00): http://www.data-archive.ac.uk/media/54776/ukda062-dps-preservationpolicy.pdf

This section also converts the archival copy of data to current distribution formats to meet user needs, and provides other value-added user support".

Example 2 (on staff roles and responsibilities):

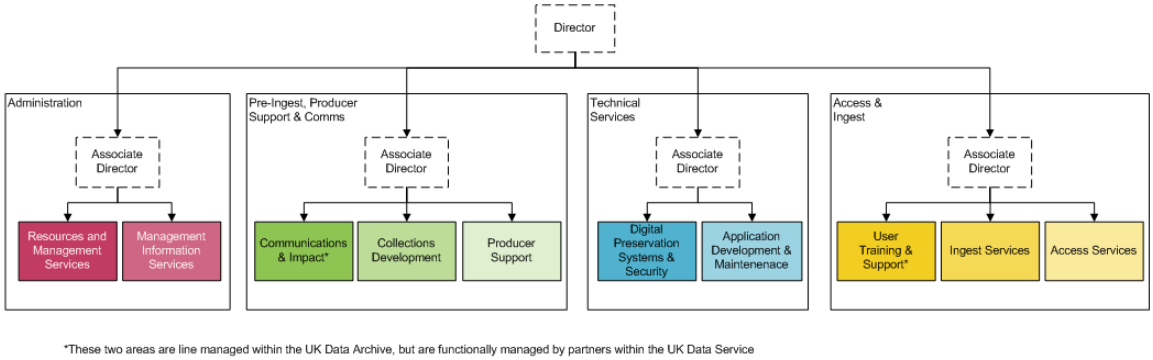
"The **Preservation Planning Manager** monitors the repository environment, providing recommendations and guidance on good preservation management practice and preservation plans to ensure that the information stored remains accessible and understandable over the long term. The role supports our OAIS-compliance and our status as a Trusted Digital Repository by specifying and maintaining a life-cycle, metadata schema-based system for preservation of the data collections and by developing preservation strategies, standards, packaging designs and migration plans across the functions described under the OAIS and associated standards".

In addition to the roles and responsibilities listed in the preservation policy, UKDA provides a separate document that focuses on responsibilities and skills required for operational activities at the archive.

The staffing and management structure of the UKDA is also closely aligned to the OAIS model. That is, four Associate Directors have responsibility for one or more of the OAIS functional areas. The figure below is a representation of the archive's functional and managerial structure. Each coloured box represents a function and associated team, which in turn form parts of groups managed by an associate director[23].

[23] Figure and explanation from UKDA, Staffing and management / Overview: http://www.data-archive.ac.uk/curate/archive-training-manual/staffing-and-management

Figure 3: UKDA's functional and managerial structure



A separate document, **Job Descriptions Summaries for UK Data Archive Roles**[24], focuses on responsibilities and skills required for operational activities at the UK Data Archive (in 2013). The document lists activities and responsibilities within the different archival functions (staffing and management, pre-ingest, ingest, technical service, access, user support, and communications and publicity) and grades the principal duties and key objectives within each functional area. The grades range from 6 (lowest) to 9 (highest). Grade 6 is defined as "likely to be recent graduates; inexperienced in relation to OAIS functions but with excellent general skills and a clear interest in the function of their choice", while grade 9 is defined as "middle level manager with line management responsibilities".

When it comes to general business management UKDA has an internal document, **Business continuity and incident management procedures**, which contains definitions of critical business activities and assets of the archive to ensure that procedures exist for dealing with incidents or events that disrupt such activities[25]. The document contains, among other things, description and examples of the levels of disruption, security events, incidents and weaknesses which may occur and the procedures for logging and reviewing the incident; procedures to monitor, test and backup systems to ensure business continuity along with a succession plan to be implemented in case of loss of funding; a communications plan for informing partners and users of unplanned service problems and the progress of remedial action; and specific procedure for dealing with emergencies, security/access, unplanned service outages and recovery, etc.[26].

For general management of the archive UKDA has a **Records Management Policy** (internal, available on request) which describes the Archive's policy relating to the management of all operational records produced by the archive, including all formal documents and correspondence related to Archive activities. The policy provides an outline of the:

- requirements which must be met for the records of the Archive to be considered a proper record of the activities of the organisation
- archive's commitment to adhering to legislative requirements
- roles and responsibilities of archive staff in general and the specific requirements of individual staff and sections
- requirements for the systems and processes which deal with the records[27]

[24] http://www.data-archive.ac.uk/media/422036/ukda_job_descriptions.pdf
[25] http://www.data-archive.ac.uk/curate/archive-training-manual/staffing-and-management?index=2
[26] Ibid.
[27] http://www.data-archive.ac.uk/curate/archive-training-manual/staffing-and-management?index=3

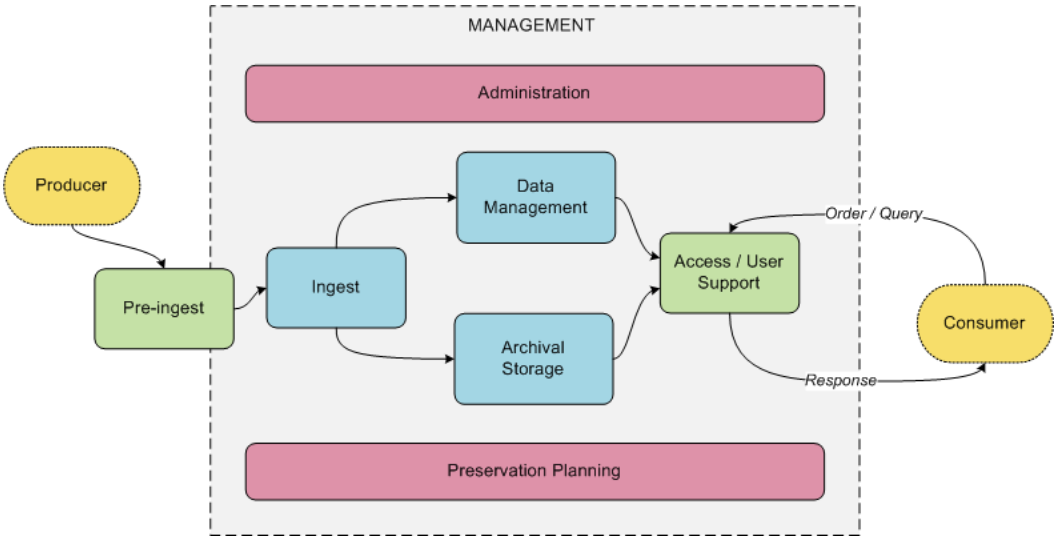
- The Records Management Policy works alongside the **Records Management Strategy** (internal, available on request) which covers the general requirements for records management within the archive. The strategy:
- outlines the types of information that constitute records within the Archive and their properties, and the principles which must govern the development of the records management system
- provides a flexible framework for the records management system
- includes information about the business classification scheme maintained by the Archive (developed from JISC's Higher Education BCS) and the management of legally binding documents held by the Archive
- outlines the general archive approach to meeting the key requirements of making records present, accessible, interpreted, authentic, maintained and documented[28].

The preservation policy states that the primary value to the UKDA of the OAIS reference model is "…that it provides a framework on which its activities can be based", and that "…the Archive recognises the benefits of the OAIS model". The archive assessed its conformance with the OAIS model in 2005[29] and found that the main divergence between the model and the practices of the archive was "…the strict separation of Archival Information Packages (AIPs) from Dissemination Information Packages (DIPs), and there were a number of activities within the data management function relating to monitoring and management which were not appropriate to the stated objectives of the Archive"[30].

[28] Ibid.
[29] Assessment of UK Data Archive and The National Archives (TNA) compliance with Open Archival Information System/Metadata Encoding and Transmission Standard: http://www.data-archive.ac.uk/about/projects/oaismets
[30] UKDA Preservation Policy (version 08.00): http://www.data-archive.ac.uk/media/54776/ukda062-dps-preservationpolicy.pdf

The figure below represents the UK Data Archive's OAIS-based functional model. The pink and blue boxes are standard OAIS functions, while the green boxes are either additional (Pre-Ingest) or amended (Access also includes a distinct User Support component). The rounded yellow boxes represent external stakeholders, and arrows the movement of information through the system[31].

Figure 4: the UK Data Archive's OAIS-based functional model



[31] Figure and explanatory text from the UKDA Archival Training Manual: http://www.data-archive.ac.uk/curate/archive-training-manual/introduction

3.1.2. DIGITAL OBJECT MANAGEMENT

3.1.2.1. PRE-INGEST

As the figure above shows the UKDA has specified a pre-ingest function in their functional model, although it is not explicitly specified in the OAIS model. The rationale for including the pre-ingest function is described in the preservation policy: it ensures "quality, comprehensibility and accessibility of all information packages by enforcing quality assurance and minimum standards at the point of 'Producer-Archive interface'". Other benefits that are mentioned are cost-reductions, ease of processing, greater levels of usability through the provision of adequate documentation, decision-making communication with the depositor, and early solving of issues concerning preservation activities (e.g. consent, confidentiality, ethics, legal issues and data formats).

Closely connected to the pre-ingest function is the **Collections Development Policy** (CDP)[32]. The policy provides an overview of the selection and appraisal criteria applied to offers of data, dictating whether or not they are accepted into the collection. This is laid out through a description of a set of 'curation categories' (Data collections selected for *long-term* curation, Data collections selected for *short-term* management, Data collections selected for *delivery only*, and Data collections selected for *discovery only*). The CDP also cover procedures and methods for data acquisition, and selection and appraisal criteria.

The CDP basically outlines the principles by which the archive develops its data holdings to meet the needs of its stakeholder communities. An associate document, the **Collections Development Selection and Appraisal Criteria** (CDSAC)[33], is an implementation of (parts of) the CDP. The CDSAC elaborates on the selection and appraisal priorities and criteria laid out in the CDP by putting the curation categories and appraisal criteria into an applicable 'appraisal grid', and explaining how to use the grid for presenting submissions and to make decisions for ingest and access routes.

The UKDA also makes their licenses and deposit forms publicly available, as part of a description of the deposit process. The aim of the deposit process is to capture information that "…goes on to form the basis of the metadata record for the resulting data collection(s), and should reflect the metadata profile adopted[34]". The deposit process is formalised through a **Data deposit form**[35] and a **License Agreement**[36]. The deposit form is to be filled out by the depositor and describes the data (collections) that are being deposited at the archive (i.e. data and documentation files description, responsible parties, study description, methodology, temporal and geographical coverage, and digitisation). The license agreement provides the archive with the necessary permissions to provide access to data under specified conditions of use, and is a legal agreement that is made between the University of Essex acting by its UK Data Archive and the depositor[37]; it clarifies the rights of both parties, copyright/ownership of the data, and access conditions for the disseminated data.

UKDA also provides a guidance that gives a detailed description of the routes through which data

[32] UK Data Service: Collections Development Policy, version 05.00: https://www.ukdataservice.ac.uk/media/398725/cd227-collectionsdevelopmentpolicy.pdf
[33] UK Data Service: Collections Development Selection and Appraisal Criteria, version 03.00: https://www.ukdataservice.ac.uk/media/455175/cd234-collections-appraisal.pdf
[34] UK Date Archive: How we curate data – Pre-Ingest/Deposit process: http://www.data-archive.ac.uk/curate/archive-training-manual/pre-ingest?index=2
[35] UK Data Service – Data deposit from: http://ukdataservice.ac.uk/media/158016/Data_Deposit_Form.docx
[36] UK Data Archive – License Agreement: http://ukdataservice.ac.uk/media/28102/licenceform.pdf
[37] Ibid.

can be deposited[38], and a best practice / guideline for researchers aimed at increasing the awareness and improvement of data management and data sharing[39].

3.1.2.2. INGEST

The basics of the ingest function in the UK Data Archive is laid out in a reference guide, the **Data Ingest Processing Quick Reference**[40]. The document covers the main stages of data ingest processing (e.g. data and documentation processing, naming of files and checking of study directory structure, cataloguing, indexing and DOI creation, etc.) undertaken on studies/data collections acquired by the UK Data Archive.

More details on the ingest process are provided in the UK Data Archive **Data processing standards**[41] which provide an overall terms of reference for the level of processing applied to a new acquisition. At the UKDA a processing level (A*, A, B or C) is assigned to any new data collection, based on an assessment of its condition, quality, and anticipated potential for secondary use. The standard assigned dictates the level of manual processing that is carried out on a new data collection / data set[42]. Further, the data processing procedures are separated into two distinct documents, one for **quantitative data processing**[43] and one for **qualitative data processing**[44], while there are distinct procedures for **documentation processing**[45]. Processing of sensitive data is treated in an internal document.

In the final processes of the ingest function comes the cataloguing and indexing of the deposited material. Cataloguing is "…the process of adding and enhancing the metadata describing a particular data collection" and the procedures of the UKDA is laid out in the **Cataloguing procedures and guidelines** document[46]. Important factors of the cataloguing procedures include the UKDA metadata profile (i.e. their selected set of metadata elements) - which is based on DDI 2.1 with customisation to fit our user's needs - and the process of assigning a Digital Object Identifier (DOI) to a study.

For indexing UKDA uses the HASSET vocabulary (Humanities and Social Science Electronic Thesaurus). In order to ensure consistent use of keywords across a large number of studies and across a large ingest team, the UK Data Archive maintains a procedural document describing keyword indexing procedures, available on request[47].

[38] UK Data Service – Deposit data: https://www.ukdataservice.ac.uk/deposit-data
[39] UK Data Archive - Managing and sharing data: http://www.data-archive.ac.uk/media/2894/managingsharing.pdf
[40] UK Data Archive - Data Ingest Processing Quick Reference: http://www.data-archive.ac.uk/media/54764/cd080-ingestprocessingquickreference_09_00w.pdf
[41] UK Data Archive – Data Ingest Processing Standards, version 08.00: http://www.data-archive.ac.uk/media/54782/cd079-dataingestprocessingstandards_08_00w.pdf
[42] UK Data Archive – Ingest/Procedures: http://www.data-archive.ac.uk/curate/archive-training-manual/ingest?index=1
[43] http://www.data-archive.ac.uk/media/54770/cd081-quantitativedataingestprocessingprocedures_08_00w.pdf
[44] http://www.data-archive.ac.uk/media/54767/cd093-qualitativedatacollectioningestprocessingprocedures_08_00w.pdf
[45] http://www.data-archive.ac.uk/media/54785/cd078-documentationingestprocessingprocedures_08_00w.pdf
[46] http://data-archive.ac.uk/media/401477/ukda035-cataloguingproceduresandguidelines.pdf

[47] http://www.data-archive.ac.uk/curate/archive-training-manual/ingest?index=2

3.1.2.3. PRESERVATION, ARCHIVAL STORAGE AND DATA MANAGEMENT

The **preservation policy** of the UKDA outlines "…the principles which underpin the main activities of the archive: the active preservation of digital resources for use and reuse within its core user community"[48]. The policy, which is revised biannually, generally conforms to the OAIS Reference Model, with additions and alterations which are specific to the materials held within UKDA (e.g. the inclusion of a pre-ingest function, as discussed above). The policy describes the archival activities connected to each of the functions. Most functions that are presented in the preservation policy are described in more details in other strategies, policies and procedure documents (and are treated separately in this document), so in this segment we will focus in the sections of the preservation policy that focuses on the archival storage function, the data management function, and the preservation planning function (IT architecture and security are dealt with elsewhere).

Important aspects of the archival storage function, as laid out in the preservation policy, are physical data preservation and storage, media monitoring and refreshing strategy, and compression. When it comes to physical data preservation the UKDA follows a policy of multiple copy resilience. That is, a "…main near-line copy and a shadow copy on two separate preservation servers, only accessible using a dedicated preservation account". For housing of magnetic and optical media the archive follows **BS ISO 18925:2002** (*Imaging materials. Optical disk media. Storage practices*) and **BS 5454:2000** (*Recommendations for the storage and exhibition of archival documents*).

Media monitoring and refreshing strategy is briefly explained in the policy: "LTO [Linear Tape-Open] drives that are used for preservation are re-tensioned automatically and retired if worn […] Idle tape media are automatically ejected from the LTO drives and placed in the robotic library at set regular intervals to prevent excessive wear of both tapes and the drive. If any media have either recoverable or non-recoverable errors then they are regenerated from the on-site mirror preservation server".

For compression the Preservation Policy refers to the Information Security Policy which outlines the compression tools and standards that are used by the archive. The Information Security Policy does not seem to be available through the UKDA webpages, but there are references to an internal document called **ISO 27001 Information Security Procedures**[49].

The preservation policy separates the data management function into two sub-functions: Version control/change procedures, and Data collection withdrawal. The most important aspect of the Version control/change procedures sub-function is to ensure "…that any alteration to the preserved version of any part of a data collection is accurately documented" to maintain the authenticity of the data collection at hand.

The most important aspect of the Data collection withdrawal sub-function (as defined in the Preservation Policy) is the distinction "…between 'soft deletion' whereby certain references to the withdrawn content are deleted, but not the content itself, and 'hard deletion' whereby the content and all references to it are deleted".

[48] http://www.data-archive.ac.uk/curate/archive-training-manual/technical-services?index=1
[49] http://www.data-archive.ac.uk/curate/archive-training-manual/technical-services?index=2

3.1.2.4. PRESERVATION PLANNING

The preservation policy states that UKDA's preservation strategy "…is predicated on two basic principles: first, that digital storage media are inherently untrustworthy unless stored appropriately; second, that all file formats and physical storage media will ultimately become obsolete". To achieve this, the preservation decisions at the archive "…must always be made within the context of its [the archive's] Collection Development Policy, balancing the constraints of cost, scholarly and historical value, and user accessibility alongside the requirements of levels of authenticity and legal admissibility".

When it comes to roles and responsibilities of (parts of) the preservation planning, the Digital Preservation Systems and Security (DPSS) section, part of Technical Services of the archive, has the "…responsibility for preserving data and metadata in all forms to ensure they remain usable over time, including monitoring technological changes that will affect preservation and migration decisions. It is also responsible for the creation of preservation metadata". Hence the DPSS is responsible for monitoring developments and advances in the digital preservation area through a technology watch scheme. Based on the technology watch, recommendations are made for migration plans that are then approved by the archive administration and management teams[50].

The preservation policy itself is monitored and reviewed in the light of the technology watch finding and changing technologies on an annual basis.

When it comes to security, the above mentioned **ISO 27001 Information Security Procedures** (see section 4) describes the UK Data Archive's information security management policy associated with the ISO 27001:2005 standard [*Information security management*], particularly with regards to data acquisition, preservation and dissemination activities. The policy includes:

- required inputs for the management review which the Archive is required to hold annually
- the role and purpose of the Information Security Management Group
- procedures for authorisation, implementation, monitoring, review, maintenance and improvement of the Information Security Management System
- procedures for analysing and assessing risk levels and the framework for a risk treatment plan
- procedures for corrective and preventative action and incident management[51]"

Regarding other aspects of security, e.g. the physical safety and security of the data collections, the preservation policy lists several measures implemented by the archive, ranging from multiple key entries, security-protected swipe-card system linked to an on-site alarm system, to logon password systems to all machine room computer systems. The preservation policy states that the archive's suite of information and premises security are documented in the Archive's Information Security Policy, Information Security Management Policy and Premises Security Procedures (internal documents).

[50] Information from a report from 2005 (information may be obsolete): *Assessment of UKDA and TNA Compliance with OAIS and METS Standards*: http://data-archive.ac.uk/media/1692/OAISMETS_report.pdf
[51] http://www.data-archive.ac.uk/curate/archive-training-manual/technical-services?index=2

3.1.2.5. ACCESS

UKDA has a **Rights and Access Management Strategy** document (internal, available on request) that "…details the archive's strategy for the management and transfer of rights between parties involved in the production, acquisition, ingest, preservation of and access to digital resources". The areas of coverage include:

- a statement on open access resources
- detailed information on the transfer of rights between depositor and end user, via the depositor agreement (see Pre-Ingest) and the end user licence
- an index of the end user licence types, including tailored licences for teaching data, and for disclosive data
- any other licence agreements between the service and other organisations [52].

A **Data Access Policy** covers the access criteria for the UK Data Service, across all UK Data Service sites and access points[53]. The document describes a generic, three-tier access policy which combines modes of access and conditions of use. (The method of access will depend on the appropriate technical solutions to implement them). The three tiers are:

- *Open data*: data which are not personal and have relatively few restrictions on use. Normally, these data will be available without registration or authentication. The suggested licenses for this tier (as described in the Data Access Policy) are the Open Government Licence (OGL), the Creative Commons Attribution Licence 4.0, or the Open Data Commons Attribution Licence.

- *Safeguarded data*: data which are not personal, but where the data owner considers there to be a risk of disclosure resulting from linkage to other data. Safeguards include knowing who is using the data and for what purpose. Safeguarded data may have additional conditions attached (special conditions, additional special agreements, depositor permissions, limited to non-commercial or academic usage, specific forms of citation, etc.). Data require registration/authentication. Redistribution of data is explicitly denied and restrictions on use for a particular data collection are outlined in an End User Licence (EUL). Some safeguarded data may require an additional Special Licence.

- *Controlled data*: data which may be identifiable and thus disclosive or potentially disclosive. These data are only available to users which have been accredited as Approved/Accredited Researcher (as defined in the official Statistics and Registration Service Act 2007[54] or under an equivalent ESRC (Economic and Social Research Council) scheme). Controlled data require registration/authentication. Multiple secure access agreements are in use to meet the multiple requirements of different data owners. Where access is granted to controlled data, it will be through a physical or virtual secure arrangement, depending on the specific requirements of the data depositor, and users may be required to undertake specific training as part of such arrangements.

[52] http://www.data-archive.ac.uk/curate/archive-training-manual/access?index=1
[53] http://staging.ukdataservice.ac.uk/media/455247/dataaccesspolicypublic_2_00.pdf
[54] http://www.legislation.gov.uk/ukpga/2007/18/pdfs/ukpga_20070018_en.pdf

Table 2: UKDA's three tiers of data access, as defined in the Data Access Policy[55].

	Depositor's Licence	*Registration*	*Authentication*	*Legal Gateway*	*Terms of use*
Open Data	Not required	Not required	Not required	"Not personal"	OGL/CC4.0
Safeguard Data	Required	Required	Required	"Not personal" but with potential residual disclosure risk	End User Licence (additional conditions may apply)
Controlled Data	Required	Required	Required	Defined as personal	Secure Access Arrangement(s)

The **End User Licence** (EUL) specify the conditions under which a user is granted access to a data collection, and how they may use those data. The Licence is an agreement that is made between the user (End User) and the University of Essex and the service funders in order to provide the user with the right to use the collections provided via the UK Data Service and the UK Data Archive, according to a set of terms[56]. The terms and conditions, along with disclaimer statements, are laid out in the licence text and include a total of 16 subjects.

A **Licence Compliance Policy** manages the terms and conditions of use of data services. The policy provides background information concerning the agreements that users of the service enter into and the legal framework that underpins those agreements[57]. Basically it elaborates on the terms and conditions that are laid out in the Licence Agreement and the End User Licence (and/or Special Licences). Among the issues that are treated in the policy are events and incidents connected to 'serious non-compliance' with terms and conditions of the End User Licence / Special Licence and possible penalties; legal framework (mostly the 2007 Statistics and Registration Service Act); commercial use of data; notification of publications, data errors and data enhancements; and possible rights of appeal (for the user who has been accused of disclosure or other non-compliance breaches).

In addition to the licencing framework UKDA offers support and guidance for data access. An **Access Support Procedures** document (internal) supports the activities of the UKDA Access team, and covers issues like:

- The processing of new 'orders' (requests for download access to datasets), particularly any human interaction required to grant an end user access - for example, negotiation of a commercial licence with the original depositor;
- tracking queries from users regarding access to data collection;
- application forms for users to apply for access to restricted resources;

[55] The table is copied from the policy where it is called, 'Matrix of Access'. Open Access applies only when there is no mandatory registration/authentication.
[56] http://data-archive.ac.uk/media/381244/ukda137-enduserlicence.pdf
[57] https://www.ukdataservice.ac.uk/media/311391/CD142-LicenceCompliancePolicy.pdf

- gathering statistics and creating reports on the usage of data;
- a key to the access codes applied to data collections during cataloguing (see section on Ingest)[58].

There is also a web guidance on how to access data[59] and separate documentation on data protection and secure data access (**Microdata handling and security**[60], and **Access Secure Lab**[61]).

[58] http://www.data-archive.ac.uk/curate/archive-training-manual/access?index=2
[59] https://www.ukdataservice.ac.uk/get-data/how-to-access
[60] http://www.data-archive.ac.uk/media/132701/ukda171-ss-microdatahandling.pdf
[61] https://www.ukdataservice.ac.uk/get-data/how-to-access/accesssecurelab

3.2. FSD

3.2.1. ORGANISATIONAL INFRASTRUCTURE

The Finnish Social Science Data Archive (FSD), founded in 1999, provides a single point of access to a wide range of digital research data for learning, teaching and research purposes. The archive is a Finnish national resource centre primarily funded by the Ministry of Education and Culture and the University of Tampere. In addition to archiving and dissemination of data, key services include data-related information services and support for research data management. The archive operates as a separate unit of the University of Tampere.

The purpose of FSD is stated in Section 5 of the Regulations of the University of Tampere:

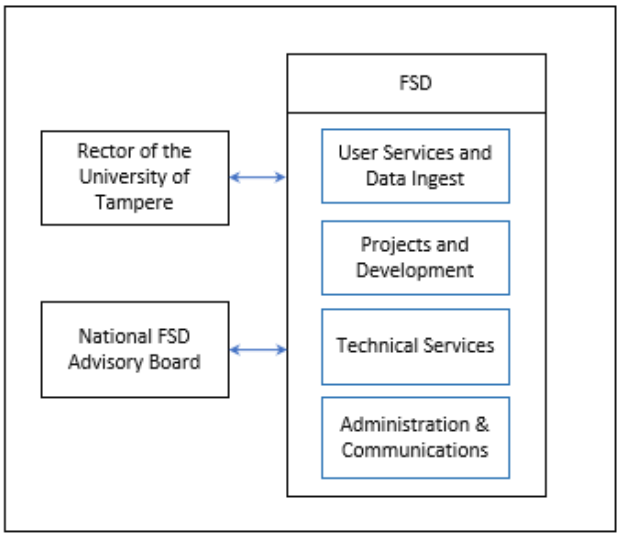
The Finnish Social Science Data Archive (FSD) serves research and teaching on the national level. The FSD archives electronic research data from Finland and abroad and disseminates it for the purpose of research, teaching and studies.[62]

This policy statement on purpose is closely connected to FSD's **Strategic Plan** (2013-2016)[63] which states the mission, operating environment, strategic challenges, and goals for the archive[64].

FSD's **Records Management and Archives Formation Plan**[65], revised annually, is the highest-ranking document concerning FSD's provision of data services. The Plan first describes how current legislation is taken into account in data management and what are the FSD's data acquisition principles. Next, it reviews the FSD's work process and the documents and document series that are processed and produced at different stages of this process. The last part of the Plan focuses on the FSD's data systems, data security and confidentiality measures and practices. The Plan also lays out the key responsibilities of the archive staff.

[62] The Regulations of the University of Tampere:
http://www.uta.fi/english/administration/regulations/TaY_Johtosaanto_EN_voimaan_01012017.pdf
[63] New strategic plan is expected by the end of June 2017.
[64] The Strategic Plan is available in Finnish:
http://www.fsd.uta.fi/fi/hallinto/asiakirjat/strategiasuunnitelmat/strategia1316.pdf
[65] The definitive version of FSD's Records Management and Archives Formation Plan (AMS) is available in Finnish:
http://www.fsd.uta.fi/fi/hallinto/asiakirjat/AMS/ams_index.html. English translation:
http://www.fsd.uta.fi/en/organisation/documents/AMS/index.html

Figure 5: The organisational structure of FSD

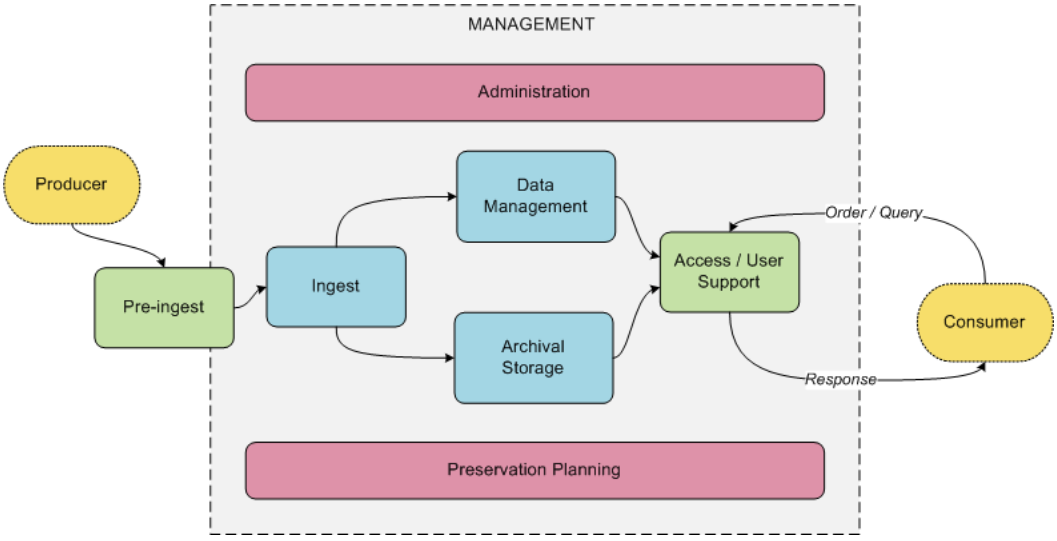


FSD has four functional modules: User Services and Ingest; Projects and Development; Technical Services; and Administration and Communications. Teams that have responsibility of one or more of the OAIS functional areas perform the day-to-day work.

FSD's Records Management and Archives Formation Plan is supplemented by an **Internal Manual** that contains information about all critical business activities. It contains, for example, detailed practical guidelines for archival processes and procedures, customer services, IT services, and communications.

The figure below represents the UK Data Archive's OAIS-based functional model. FSD's model is the same. The pink and blue boxes are standard OAIS functions, while the green boxes are either additional (Pre-Ingest) or amended (Access also includes a distinct User Support component). The rounded yellow boxes represent external stakeholders, and arrows the movement of information through the system[66].

Figure 6: the FSD's OAIS-based functional model



[66] Figure and explanatory text from the UKDA Archival Training Manual: http://www.data-archive.ac.uk/curate/archive-training-manual/introduction

3.2.2. Digital Object Management

3.2.2.1. Pre-ingest

Pre-ingest is not specified in the OAIS model but FSD has data acquisition activities that take place before the Ingest function. FSD increases its collection actively and selectively: the Archive locates and acquires datasets actively but accepts data for archiving selectively. The data to be archived at the FSD must comply with certain qualitative, technical and legislative criteria. The criteria are stated in the Records Management and Archives Formation Plan. The plan also describes the data deposition process. All data deposit forms and licenses are publicly available and the deposit is formalised through a **Deposition Agreement**[67], a contractual document that sets out the conditions for archiving and disseminating the data collection(s).

FSD also provides instructions on how data can be deposited[68], and a **Data Management Guidelines** web resource[69] to advice and help researchers in managing, preserving and sharing their data.

3.2.2.2. Ingest

A general description of FSD's ingest function is part of the Records Management and Archives Formation Plan. FSD's Internal Manual contains very detailed guidance on all aspects of ingest: processing quantitative and qualitative data, creating metadata, anonymising and handling of sensitive data.

FSD creates extensive metadata records for each data collection both in Finnish and in English. FSD's metadata profile is based on DDI Codebook 2.1, with some additional elements to fit FSD's needs. FSD assigns each study a persistent URN identifier. For indexing FSD uses the General Finnish ontology YSO[70], the ELSST Thesaurus[71] and selected DDI Controlled Vocabularies[72].

The study and variable descriptions and related material are stored in FSD's databases, from which online data catalogues are generated.

3.2.2.3. Preservation, archival storage and data management

The FSD has no separate preservation policy document. Instead, FSD's Records Management and Archives Formation Plan outline the preservation, archival storage and data management processes. The Internal Manual contains more detailed descriptions of key processes, version control and change procedures.

In physical data preservation FSD follows a policy of multiple copy resilience. For example, the data in long-term storage on FSD's internal server are copied onto the University of Tampere's IT services server. FSD utilises the National Digital Library's Digital preservation service[73].

---

[67] http://www.fsd.uta.fi/en/forms/deposit.pdf
[68] http://www.fsd.uta.fi/en/data/depositing/
[69] http://www.fsd.uta.fi/aineistonhallinta/en/
[70] https://finto.fi/yso/fi/?clang=en
[71] https://elsst.ukdataservice.ac.uk/
[72] http://www.ddialliance.org/controlled-vocabularies
[73] http://www.kdk.fi/en/digital-preservation

Since hardware, software and storage media become outdated over time, FSD follows closely new developments in the storage and preservation field. When necessary, FSD migrates or converts the records so that they are compatible with other hardware and software environments. The contents of the data are not modified for or during migration. FSD favours formats that are open, compliant with established standards and supported by multiple producers over formats that are closed and tied to specific manufacturers. FSD follows the Finnish national recommendations of acceptable file formats by the National Digital Library[74]. FSD was part of the working group that developed the recommendations.

The FSD uses an operational relational database as an internal registration system for all archival work. Information about the additions, modifications or removals of data collections are recorded in the database.

3.2.2.4. PRESERVATION PLANNING

FSD's preservation plan, strategy and recovery plan are included in the Records Management and Archives Formation Plan that is reviewed annually. FSD's Records Management and Archives Formation Plan team that consists of senior FSD staff members is responsible for keeping the Plan up to date, and for monitoring developments and advances in the digital preservation area. FSD participates actively in national digital preservation projects and initiatives.

Description of the potential (technical and physical) risks, and actions taken to reduce the risks and overcome problems are included in the Records Management and Archives Formation Plan. The Internal Manual contains more detailed information.

3.2.2.5. ACCESS

FSD's Records Management and Archives Formation Plan contains a section about dissemination of data for reuse. Data archived at the FSD can be used for research, study and teaching, according to access conditions set for the data by original data creators and the data archive. Data are disseminated free of charge.

The archived data are disseminated through the Aila data service portal[75]. Aila contains extensive and searchable documentation of all archived data collections in both Finnish and English. In Aila, the data are available free of charge to registered users, according to the conditions set for each dataset. FSD data have four access categories:

(A) freely available to all users

(B) available for research, teaching and study

(C) research only (including e.g. Master's, licentiate and doctoral theses)

(D) access requires permission from the data depositor.

[74] The recommendations in Finnish: http://www.kdk.fi/index.php/fi/pitkaaikaissailytys/maeaerittely-ja-dokumentit/5-suomi/pitkaeaikaissaeilytys/141-kdkn-saeilytys-ja-siirtokelpoiset-tiedostomuodot

[75] https://services.fsd.uta.fi/?lang=en

Aila supports the Haka identity federation[76], which allows students and staff of Finnish universities and polytechnics to register using the credentials provided by their institution. Other users can apply for an Aila username through FSD User Services. Users need to fill in the access application form, providing information on the purpose of data use, their research or study, and information on funders, if any. By downloading data from Aila, users accept and agree to the terms and conditions set out for the use of the downloaded data.

The **General Terms and Conditions for Data Use**[77] specify the terms and conditions applying to the use of any dataset contained in Aila. The **Terms and Conditions for Aila Use**[78] apply to the use of the Aila data service portal.

There is also web guidance on how to access data[79]. FSD requires that data collections and their creators are cited in all publications and presentations for which the data have been used and provides a bibliographic citation for each study.

[76] https://www.csc.fi/-/haka-kayttajatunnistusjarjestel-1
[77] https://services.fsd.uta.fi/docs/terms-of-use?lang=en
[78] https://services.fsd.uta.fi/docs/eula?lang=en
[79] http://www.fsd.uta.fi/en/data/ordering/index.html

3.3. ADP

3.3.1. ORGANISATIONAL INFRASTRUCTURE

3.3.1.1. SCOPE AND OBJECTIVES

Digital preservation policy aims to define rules, responsibilities, roles and systems of control in handling data for assuring long-term access. The purpose of the digital preservation policy in Social Science Data Archives (ADP) is to continuously improve the transparency of its operation.

In developing digital preservation policy, the ADP considers the Guidelines for development of institutional digital preservation policy of the Working Group Nestor, policies of partner organisations, especially ICPSR (2009) and UKDA (Preservation policy), as well as schemes, included in SERSCIDA Work Deliverable D5.1.

The ADP is a Slovenian national disciplinary data infrastructure specialised for social sciences' research, whose mission is to offer data services for supporting research, education and general public benefit. It is involved in international infrastructures and cooperates in developing sustainable policies of open-access.

The goal of digital preservation policy in ADP is to achieve sustainably designed quality assurance and promotion of ingest, storage and access services, supporting dissemination of valuable research data from the field of social sciences in Slovenia and abroad to designated communities.

3.3.1.2. LEGAL AND REGULATORY FRAMEWORK

Sustainable operation of ADP is assured by the following main guidelines:

- Strategy of Open Access in Slovenia 2015-2020: principles and assignment of open access to scientific publication and research data in Slovenia in the period 2015-2020 (Government of the Republic of Slovenia, 2015); it defines sustainable operation of disciplinary research data centres.
- For the purpose of long-term digital preservation, ADP uses existing equipment and preservation services provided by the National and University Library.
- Ingest and access to data are regulated by fixed forms, in which all legal and ethical issues are resolved.
- Data and documentations are available to users under Creative Commons 4.0 licenses.

In its operation, ADP closely follows national legislation from the relevant areas, especially:

- The Law on Personal Data Protection Act, which defines the responsibilities of ADP as micro-data distributor.
- Law on Copyright and Related Rights
- Law on the Protection of Documents and Archives and Archival Institution, that defines the mode, organisation, infrastructure and implementation of ingest and storage of documents in physical and digital form.
- The Law on Access to Public Information: defines access and reuse of public information.

In collaborating with the Statistical Office of the Republic of Slovenia, ADP follows also:

- National Statistics Act.

3.3.1.3. DESIGNATED COMMUNITY

Designated communities of ADP are national and international researchers, teachers and students that are statistically literate for independent use of data. Other target users are journalists, policy makers and regular citizens that express a legitimate purpose of the use of data and accept the rules and limitations of data re-use.

3.3.1.4. ROLES AND RESPONSIBILITIES

Due to the small size of the organisation, each employee can hold several roles within organisation. Some of the work is also outsourced. The basic structure and roles of the ADP are:

Head of organisation: Promulgates the mission of the organisation in the field of long-term digital preservation as the primary goal of the organisation.

Head of ingest and documentation acquisition: Registers new studies and communicates with data producers, controls the ingest of data and other related documentation and reports on ingest and processing stage of the new studies.

Head of digital preservation: Oversights the management of all servers used in preservation and distribution of data, oversights regular backup procedures and is responsible for periodic update of policy and documentation regarding long-term preservation and corresponding certification (DSA).

Head of access and use: Controls data publishing and prepares statistics of received and published studies.

Head of trainings and promotion: Responsible for performing activities in the field of promotion, including organisation of events.

System administrator: Controls the management of all servers used in preservation and distribution of data.

Data archivist: Manages the processing of microdata, metadata and related documentation.

Director of administration: Organizes and controls implementation of daily administrative assignments and prepares reports in cooperation with other departments of the University.

3.3.2. DIGITAL OBJECT MANAGEMENT

The ADP follows the OAIS model (The Open Archival Information System; ISO 14721: 2012) that is composed of complex information objects: submission information package (SIP), archival information package (AIP) and dissemination information package (DIP), which represent the current state and versioning of digital content, together with metadata from the point of ingest, through archival storage to final access of users. Due to the elaborate process of pre-ingest and data acquisition, the ADP adds to the model the function of pre-ingest or pre-SIP (see Producer-Archive Interface – Methodology Abstract Standard (PAIMAS) – ISO 20652: 2006), where the ADP establishes communication with possible data depositors and makes an evaluation of the appropriateness of new data. The OAIS model defines also six functional units, developed to implement the tasks of storage and to assure access to information. Together with ingest, archival storage and access, those are data management, preservation planning and administration.

The OAIS standard defines a list of mandatory responsibilities in defining tasks of the archive. In the field of access to data, designated communities need to be defined, as well as their needs and knowledge, and this is then used when preparing data and metadata for access, so as to assure independent understanding of given information. Individual mandatory responsibilities in connection to the OAIS standard will be shown in appropriate places in the following text defining our individual policy segments.

3.3.2.1. PRE-INGEST FUNCTION AND DATA ACQUISITION

The ADP curates research data, interesting for social sciences research, focusing especially on problems, connected with the Slovenian society. Priority is given to research data that contain important topics relevant for research and that are methodologically well-prepared, especially longitudinal data and international comparative data, which include Slovenia. ADP accepts also data that do not fit strictly the field of social sciences, but are structure-wise similar and/or are lacking other appropriate curator. Long-term data preservation demands extra efforts and finances for preparation of data in a form appropriate for further use. These expenses are justifiable considering savings made by reuse of data.

Firstly, data producers preliminary register the study candidate for ingest. Detailed instructions on how to submit data are published on the website of the ADP. By respecting the conditions that data producers need to fulfil in order to submit the study to the archive, the first responsibility of an OAIS archive is fulfilled: “Negotiate for and accept appropriate information from information producers.”

In the phase of pre-ingest, the ADP carefully inspects the application for study submission. In evaluating the submission, the ADP uses quality criteria and content appropriateness of the study for further analysis. The criteria for selection are the following:

- Richness of data in terms of appropriateness of conceptualization and thematic completeness of ADP studies’ collection.
- Soundness of methodology, completeness and appropriateness of data and other documentation for further analyses.
- Data producer holds the IPR of the data and is willing to give the data to the archive for dissemination.

Selection and evaluation of appropriateness of data for ingest is made by the ingest commission.

The submitted documentation is firstly carefully reviewed, putting emphasis on the completeness of the documentation, appropriateness of data content, state of data anonymization and the type

of data formats. If needed, the data producer is contacted to provide further clarification.

There is a list of recommended file formats on the ADP website that are independent of software and computer platform. When data is given to ADP these formats are inspected and, if needed, converted to formats appropriate for long-term data preservation and access.

Studies that are accepted by the commission are assigned a category of scientific importance. On the basis of judgment, whether or not a study is theoretically or practically important, it can be assigned a status of scientific data publication, considering the ARRS (Slovenian Research Agency) scientific evaluation standards (Rule book on the procedures of (co)financing, evaluating and monitoring of implementing research activities, Appendix 2, Point 2.H.).

Head of ingest and documentation acquisition notes down the evaluation of the study and prepares a report, which contains the argumentation behind the review score. The report becomes an integral part of the documentation on ingest.

3.3.2.2. INGEST FUNCTION

The first function, supported by the OAIS model, is ingest. Within this function the data producer deposits the study to the archive, and, according to the procedures, the archive prepares the study for preservation. At this point, it needs to be ascertained whether or not the given documentation is appropriate and complete. The ingest function works as an external interface with the data producer and defines the entire process from data acquisition to data storage.

In the pre-ingest phase the archive resolves all possible questions, regarding the study, with the data producer, (for example ethical questions, confidentiality, data anonymization, authorship and rights of the archive or its successor to take over the data and to assure the process of digital preservation). All the rules on data access are also established. On this basis, an agreement is signed between the archive and the data producer (License Agreement between Archive and Data Producer). The archive assures for safe and swift data submission. In accordance with the information model of an OAIS-type archive, the research data and other accompanying documentation from the data producer, given to ADP, compose the submission information package (SIP). The package of quantitative data is composed of:

- Filled-in Study Description Form
- Appropriately managed and documented data file (in accordance with Recommendations for editing the data file)
- Original questionnaire (print and digital copy)
- Other documentation that helps to clarify the content of the data, such as: codebook, frequency list, instructions for interviewers, reports on study fieldwork, copies of publication and other accompanying materials, which was part of data collection or are important for their understanding
- Filled-in and signed form License Agreement between Archive and Data Producer, together with the list of submitted materials (2 copies)

Data producer can add reports and other related publication that would help in secondary data analysis to the submission package.

The data archivist prepares all necessary metadata for registration of individual parts of the submission information package. All the descriptive metadata of the study in accordance with DDI, as well as structural metadata, enabling clarity and usability of data to future users, are provided. The package is then transformed into recommended formats for long-term preservation and access. The distribution information package, intended for data users, is kept separately from the

archival information package.

ADP keeps the following data and metadata:

- Data with corresponding metadata accessible through ADP: data is archived and accessible through the ADP.
- Collections of metadata with data accessible through other organisations, together with links to the data access: metadata of data, accessible through other sources, that are of considerable interest content-wise for ADP's designated communities. In these cases, only metadata is kept and disseminated (for example some data from international surveys and official statistics).

The OAIS system of data preservation demands adequate quantity of contextual information (metadata) that enable target users to independently understand data files or related materials. Descriptive, administrative and structural metadata are the basic elements of all versions of information packages in the system. Together with a subset of preservation metadata, they represent the basis for semantic and procedural adequate use of data and accompanying documentation.

The use of standard metadata and interconnectivity of the used identifiers with other information services, assures for quality and sustainable development of data services and is the basis for coordination and collaboration in the field of collective services within the international infrastructure CESSDA. Interoperability is one of the exposed themes of the OAIS standard and a requirement of the European Commission in the framework of Open Access Data Policy H2020.

Descriptions of studies, according to DDI, are interoperable with catalogues that use DDI records (CESSDA common catalogue, DataVerse etc.). ADP also collaborates in projects aimed at coordinating the use of metadata on the level of CESSDA.

ADP regularly follows the development of new services and identifiers.

3.3.2.3. PRESERVATION, ARCHIVAL STORAGE AND DATA MANAGEMENT

Main role of the second OAIS function is to assure for long-term preservation. This means that all the documentation needs to be archived in appropriate formats for long-term preservation and kept on appropriate locations. Archival information package consists of all given and available documents, which are, together with all metadata, kept on a separate location. Additionally, backup copies are made on a regular basis.

All documentation, regardless of their version or subversion, is being kept on a shared directory, whereas their backup copies are being kept on different locations.

Currently, the ADP does not have a withdrawal policy (for any possible reason). It would be advisable to develop such a policy in the future.

Within the (meta)data management function, the metadata databases are refreshed and queries within these bases are made. On the basis of other OAIS functions (ingest, system management, access) reports are made. The ADP keeps such information in different databases. There is a database for study recording, database of data users, database of license agreements and database of data access.

3.3.2.4. DATA ACCESS

Providing access to data is the fifth function in the OAIS model. This function deals with data users' claims and assures access to content based on these claims and rights of individual data users. The function is responsible also for the technological control over data access, based on rights of users.

Metadata and other documentation, connected with the study, are accessible without registration to all visitors of the ADP webpage. However, to access the microdata, registration is required. Registered users have the possibility to make online analyses with Nesstar and also the possibility to download the data files in selected formats to their own computers.

Anyone can register to access the data. To register, a user needs to fill in a form Registration to access data, in which he/she states his/her personal information, status (student, researcher, public official, other etc.) and states the purpose of the use of data (educational, scientific, public, commercial). In order to register, the data user needs to conform to the terms of use of the data, which amongst others, demand respect for professional codes of ethics and obligation of full citation of the author/s and the archive.

Possible restrictions of access and exceptions, made by the data provider, are described in the study description in the catalogue ADP. Information on appropriate data citation is provided to data users for every documented study.

Registered users have the possibility to access most of the microdata from the ADP catalogue. Sensitive or undisclosed data are accessible in two ways: through portable media or through a secure room (office of ADP), for which a procedure of approval is foreseen.

By registering, data users confirm to cite the used data. The ADP is committed to develop ways, how to follow the citation of used data sources (for example DataCite). The ADP asks researchers to report on publications that use data from the catalogue of ADP.

3.3.2.5. PRESERVATION PLANNING AND STRATEGY

The ADP provides researchers the services of ingest, scientific validation of data, selection, preparation for long-term preservation and access to data. Its basic mission is long-term data preservation.

To assure long-term data preservation, the ADP regularly follows and maintains its system of digital storage. The administrative aspect includes the following tasks:

1) Regulation of outsourcing contracts (National and University Library, ARNES, keeper of servers and IT support, programmer).
2) Regular follow-up of backup management and documentation of digital document processing in Jira task management system (audit trails and documentation).

The basic preservation strategy of the ADP is normalization of data in the ingest phase. This is achieved by following a list of appropriate formats of data for long-term preservation (following best practices of similar organisations and international standards).

A procedure of data submission in one of the pre-assigned formats is made for the lead group of archived digital objects (quantitative data files). The ADP uses Nesstar publisher as a tool for transformation of data files. Nesstar publisher is tested to assure the preservation of significant properties of data files with the goal to keep at least the same possible performance of data, as the original data producer had.

In the framework of fulfilling the sixth obligation of OAIS – *to keep the authenticity of data by documenting changes in the processing of data* – an administrative system of changes (audit trail), in connection with data normalization, anonymization and versioning of data is established.

Head of digital preservation controls the inclusion of professional approaches and principles of digital preservation in the present strategy. Other employees of ADP are actively professionally trained and collaborate in the knowledge-sharing community with other libraries, archives and digital humanistic professionals on national and international level (CESSDA Expert Seminar, conferences of IASSIST etc.).

3.3.2.6. IT ARCHITECTURE, HARDWARE AND SOFTWARE UPGRADES

Technological support is based on the repository software, with customization taking into consideration data specificities and characteristics of documentation and metadata from the field of social sciences. The ADP invests in development of services that will enable interconnectivity between different organisations, use of versioning and persistent identifiers, overview over copies of materials in different formats and regimes of access. The ADP will continue the collaboration with the National and University Library and ARNES on the field of interconnectivity by assuring safe digital preservation on different locations.

One of the continuous activities of the ADP is cooperation in developing, testing and implementation of tools in the international environment. The ADP is planning to enable technological support for researchers to assure for a controlled management of research data during the duration of the project, which will also enable easier delivery of data to the ADP at project end. This will be accomplished with a combination of open access tools, such as DataVerse, and by adapting and opening developed repository tools to manage processes in the archive.

Keeping of databases, distribution of data and overall infrastructure of the archive is based on an adapted IT infrastructure. Only the employees and registered users have access to the system. A firewall is put in place, as well as limited access to hardware to assure higher security of access.

To enable long-term preservation, security copies of the entire system are being made on different locations. Security copies on the location of the archive enable synchronised security copies in case of a defect of one of the servers. These security copies provide also copies in cases of more serious events in the server room (destructive fire). In case of a more serious event (a destructive fire, earthquake, bomb etc.) on the area of the Faculty of Social Sciences, security copies are being kept on locations outside of the faculty. Software and hardware are regularly checked and updated. Storage capacities are also adequately managed regularly.

3.3.2.7. SECURITY AND RISK MANAGEMENT

There are three rooms in the ADP, where computer equipment and materials are stored. All of them are part of the Faculty of Social Sciences of the University of Ljubljana, which under its protocols, cares for physical safety. This includes permanent presence of a security guard, fire safety procedures and physical security services.

To assure safety, all rooms, containing computer equipment and materials, need to be locked if none of the employees of ADP are present. Physical access to servers is given only to employees of the Computer Centre of the Faculty of Social Sciences, external partners with a valid contract and employees of ADP, accompanied by a representative of the Computer Centre of the Faculty. Only the employees of Arnes have access to the servers in Arnes.

3.3.2.8. COLLABORATION WITH OTHER INSTITUTIONS

On an international level, the ADP is involved in activities of the pan-European research infrastructure CESSDA. One of the priorities of Slovenia in the Plan of development of research infrastructures in the period 2011-2020 is to get involved in CESSDA. As a result, Slovenia became one of the first members of CESSDA.

Nationally, the ADP is involved in a wider national research infrastructure. Employees of ADP cooperate as external experts in different workgroups of Slovenian ministries (Ministry of Education, Science and Sport; Ministry of Culture), where they prepare initiatives and guidelines and promote the requirements of national and international organisations and science funders in their national research community. The ADP cooperates also with libraries of research institutions in raising awareness of researchers about open access to research data and accompanying materials.

On the national level the ADP cooperates with other national, disciplinary research data infrastructures that are also included in international organisations, such as DARIAH and CLARIN.

3.3.2.9. FUNDING AND RESOURCE PLANNING

Long-term national importance of the ADP is shown in consistent financial support of the ministry, responsible for scientific-research work. From 2004 onwards, funding is enabled through the infrastructure programme "Network of Research and Infrastructural Centres". The current program period of financing is from 2015 to 2020.

Advanced research policies of open access to scientific results anticipate the sustainability of functioning of infrastructure services of area centres, such as the ADP. Such research infrastructure is based on cooperation between different parties, both in the field of fulfilling funders' policies and in the field of fulfilling the needs of the scientific community and other designated communities. From the point of view of the economy and utilization, continuity of functioning of ADP is the basis of fulfilling the role of national disciplinary research data centre.

The Ministry named the ADP as the service provider of the national social science database infrastructure within the Slovenian membership at CESSDA. The sustainability of services of the national data centre of social sciences is an obligation, made by the Republic of Slovenia, in the framework of membership in the CESSDA ERIC.

3.3.2.10. MONITORING, REVIEW AND FEEDBACK OF THE POLICY

Policy of digital preservation will be revised and updated every 2 years. In the intermediate period, individual elements of the policy and connected procedures will be internally updated. The employees of ADP have the responsibility to control the implementation and updating of the policy.

3.4. FORS / DARIS

3.4.1. Organisational Infrastructure

FORS, the Swiss Centre of Expertise in the Social Sciences, is a research infrastructure, serving researchers from Switzerland and abroad. FORS is a legal entity in the form of a national foundation, established on 1.1.2008 in the terms of the Swiss Civil Code Article 80 ff.[80] The Foundation Board is the supreme governing body of the foundation. The Organizational Chart describes the structure of the foundation.

DARIS is the Data and research information services department of FORS. Integrated within DARIS is the COMPASS service (Communication Portal for Accessing Social Statistics). COMPASS serves the scientific community with the intention of promoting access to microdata from public statistics.

The purpose and mandate of DARIS and FORS are outlined in a series of interrelated documents: FORS' Deed of Foundation stipulates that FORS shall gather, archive and distribute all types of data for scientific use. The DARIS Mission Statement outlines a twofold mission:

„First, we aim to acquire, document, preserve and disseminate high-quality quantitative and qualitative data and research information in conformity with national and international standards; Second, we aim to make these data and services known and to promote a research culture of data sharing and secondary analysis for the social sciences in Switzerland".

The FORS Preservation Policy defines the purpose and function of the institutional setting: "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."

The Preservation Policy also lays out the legal and regulatory framework within which DARIS operates:

- Federal Act on Data Protection (FADP) of 19 June 1992 (Status as of 1 January 2014)
- Federal Act on the Promotion of research and Innovation, 14 December 2012 (Status as of 1 January 2015)
- DARIS deposit licence
- DARIS end user licence

FORS is funded primarily by the State Secretariat for Education, Research and Innovation (SERI), by contributions from the Swiss National Science Foundation (SNSF) as well as by the University of Lausanne that provides a host institution for FORS. Third party projects represent additional funding sources.[81] The Service Level Agreement with the SERI is renewed for a three-year-period at the beginning of 2017, the SNF funding for the FORS surveys 2017-2020 is guaranteed. Though funding is stable and reliable FORS is evaluating different options of succession and continuity of access planning.

[80] Swiss Civil Code of 10 December 1907 (Status as of 1 January 2016).
[81] http://forscenter.ch/en/about-us-2/foundation/funding/

3.4.2. Digital Object Management

3.4.2.1. Pre-Ingest

The content and activities of the pre-Ingest function which DARIS has added to the OAIS Reference Model in order to ensure the quality of the data and determine problematic issues such as confidentiality, etc. prior to the official deposit, is described in the Preservation Policy: It includes solicitation of data, guidance and technical support for data producers wishing to deposit data. Data management guidelines provide data depositors with clear instructions on how to properly prepare, document and submit their data, also for qualitative data. The preferred and accepted formats are communicated transparently in the FORSbase list of recommended formats and the Collections Policy as well as in the Policy on Archiving Qualitative Data. The actual procedures of the deposit of data in our information system FORSbase are described in the FORSbase guidelines.

The principles and criteria by which the archive develops its data collection and defines the designated community are indicated in the Collections Policy. It details the purpose and scope of the collection, the selection and appraisal criteria, formats, eligible depositors and acquisition strategies. Solicitation of data is done mainly by way of the Research Inventory, which is maintained by DARIS. This database currently contains over 10'000 descriptions of research projects conducted in the social sciences.

The deposit licence, an agreement between FORS and the data owner, specifies the rights and responsibilities of the repository and the depositing institution: The depositor maintains the copyright and is entitled to be regularly informed about the reuse of the data by third parties.

3.4.2.2. Ingest

The legal transfer of the data is managed through our archiving platform FORSbase: Once the depositor uploads the data in the system, he/she formally accepts the electronic deposit as well as the end user licence. Only subsequently can the data be formally deposited for archiving. Only a registered user whose identity has been authenticated by name and institutional email address can be validated and in consequence deposit and download data.

The citation is generated automatically by FORSbase based upon the metadata entered by the depositor, yet it can be adapted by the depositor. It is recorded in the end user licence, the catalogue, and the metadata report that we add to the DIP. The end user licence commits users to using the citation contained therein.

Following the introduction of our information system FORSbase we have not yet fully developed our new workflow and have not fully determined and drawn up the standards and procedures in processing data, documentation and metadata. The following procedures will be described in more detail in a future document on Ingest processing procedures in FORSbase.

FORSbase allocates a unique study and dataset number to each project and dataset while depositors create an entry. The unique study and dataset numbers appear in the catalogue and the metadata report that we create and add to the DIP.

As FORSbase is geared towards self-archiving we do not perform systematic plausibility tests. We do however perform some quality assurance routine checks for completeness, integrity and validation of the data files: Variable names, variable labels, value labels and missings, and control for adequate anonymization of the data. These checks are usually performed in SPSS, the most common ingest format. We have SPSS-syntaxes for performing those checks and plan to draft more

comprehensive syntax matrices for data processing in SPSS, STATA and R soon.

The treated data and documentation files are subsequently renamed according to our naming scheme.[82]

The documentation is checked for completeness and if necessary converted to PDF. We are currently testing the conversion to our preferred preservation format PDF/A, but still experience some problems with the recording of the metadata of the files.

Data depositors create an extensive metadata record in FORSbase while establishing an entry (title, authors, principal investigator, data type, time method, media, data collector, available documentation, access conditions, embargo, keywords, disciplines, period concerned, geographical space, abstract, results, methods, instruments, financing, grant number, missings, weighting, etc.). Our metadata standard is DDI Codebook compliant. In the course of cataloguing and indexing (including ELSST topic classification) we enhance the metadata if necessary.

3.4.2.3. PRESERVATION, ARCHIVAL STORAGE AND DATA MANAGEMENT

The Preservation Policy that is revised biannually 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.

The Preservation Policy is the most comprehensive policy describing the digital preservation strategies and principles of DARIS, the scope and mission, legal aspects, the preservation strategy, as well as the responsibilities and procedures involved in ensuring adequate preservation of and access to the data held within the archive. It adheres to the terminology and preservation practices outlined by the Open Archival Information System (OAIS) Reference Model, with the addition of the pre-Ingest function (see pre-Ingest above), and reports how the functions of the reference model have been implemented in FORSbase.

The Archival Storage function focuses on data preservation, monitors the technical fitness of its archive, does regular risk assessments of the stored digital objects. The data management function occurs independent of primary data within FORSbase. Our archival system is Fedora, which enables long-term access to digital resources. The AIP and DIP from the Ingest function are saved in Fedora to a permanent storage facility. Our preservation formats are .csv for quantitative tabular data and PDF/A for documentation files.

The Data Management function occurs independently from the primary data in FORSbase, meaning the descriptive and administrative metadata of the SIP are treated separately. Changes to the data or documentation automatically trigger a new version in FORSbase, as do changes to the metadata with very few exceptions, which are recorded in version comment metadata. Changes to the data require the creation of a new AIP and SIP and the modification of the metadata report provided in the DIP.

We intend to introduce into FORSbase a system of persistent identifiers and are currently at the beginning stages of the implementation of DOIs provided by da¦ra. The versioning, handled so far

[82] Containing the dataset number, the acronym of the title of the study, file type, cohorts and wave where necessary, language and version (e.g. 726_Selects2015_PES_Doc_Questionnaire_IT_v.1).

by FORSbase would partly adapt to the good practices of da¦ra.

The functional entity Preservation Planning encompasses tasks such as development of preservation strategies and standards, migration plans, and monitoring of technology (innovations in storage and access technologies) and the designated community (shifts in scope or expectations). The Preservation Strategy of DARIS is outlined in the Preservation Policy: It addresses how the preservation is carried out at DARIS and is based on the principles of open preservation file formats and data migration. Format obsolescence is prevented by migration into new formats.

The preservation of our holdings relies on the IT technology. Approved storage technology is provided and guaranteed by the IT department of the University of Lausanne based upon a Service Level Agreement for each server that FORS uses. Data recovery provisions, back-up strategies and risk management are carried out by the IT department of the University of Lausanne based on measures for physical security, redundant server infrastructure on different locations on campus.

Data integrity is guaranteed by checksums, which are applied not only at Ingest (SIP), but also during the archival storage (AIP). Currently we are dealing with the question of how to correctly apply checksums for compressed uploads, and are examining how to guarantee the integrity of compressed objects.

3.4.2.4. ACCESS

DARIS is currently lacking a written Access Policy, although access conditions and eligibility criteria are detailed in a series of other interrelated documents and policies. Eligible users are defined in the DARIS Mission Statement, the Policy on Archiving Qualitative Data, and on our website as researchers and students affiliated with or enroled at a research institution or an institution of higher education. Our Preservation Policy states that data are accessed by users under controlled conditions in accordance with agreed standards and guidelines. The DARIS Vision 2020 commits itself to the further removal of barriers to accessing data. The FORS Policy Statement on Open Data is also dedicated to free and easy discovery of and access to data.

The end user licence restricts the use of the data to scientific research and teaching within an academic framework. In accordance with the author, some data are restricted to research purposes (excluding teaching purposes), and/or their dissemination subject to prior approval. These restrictions imposed on the reuse of data and the access conditions are indicated in the catalogue.

Unauthorized access to data is prevented by an authentication process: Registrations in FORSbase are vetted and authenticated by name and institutional email address, and only validated users can download and deposit data. The guide Get data explains how to access and download data in FORSbase.

Data can be discovered and accessed through two channels: FORSbase and NESSTAR. Data discovery in the FORSbase catalogue is possible by a simple and advanced search of the metadata by title, abstract, methods, funding institution, study type, research domains or keywords, research disciplines and topic class (CESSDA topic classification ELLST), etc. The guide Explore the FORSbase catalogue provides some basic information about the search function.

3.4.3. FUTURE PLANS

The DARIS policy framework is revised biannually. The next and first revision of the existing policies, most notably the Mission Statement, the Preservation Policy, as well as the Collections Policy, is due in spring 2017. This revision will take account of the change of workflow with the introduction of FORSbase and the implementation of the DOIs. It would also need to define more clearly the designated community of depositors, users, and other stakeholders.

The policy framework also needs to be completed by releasing among other things an Access Policy, as well as a document for Ingest processing procedures in FORSbase which will detail our data and documentation processing activities.

FORS is currently evaluating different options with respect to ensuring continuity of services and succession planning.

4. CONCLUSION

By analysing relevant literature and assessing existing policy recommendations, guidelines and templates, combined with analyses and assessments of the content of state-of-the-art, real-world data management policies and strategies (case-studies), we have created a comprehensive CESSDA Policy Framework Template.

It should be noted that the model presented in this report may be refined through future iterations; a comprehensive policy framework template must be able to adapt and absorb ongoing changes and developments in the field in which it operates.

The policy elements included in the CESSDA Policy Framework Template builds on the OAIS model and corresponding criteria catalogues for trustworthy preservation of digital objects. Future versions should also be coherent with funder requirements and in line with relevant data management plans, which will further contribute to the accessibility and integrity of preserved research output.

However, the policies should not be dictated by funder requirements or pre-customised data management plan templates. There should rather be a dialogue between the different stakeholders in the further refinement of a full-scale policy model framework.

REFERENCES

LIST OF TRUST FRAMEWORKS

- CCSDS 652.0-M-1: Audit and Certification of Trustworthy Digital Repositories (ISO 16363): https://public.ccsds.org/pubs/652x0m1.pdf
- CCSDS 650.0-M-2: Reference Model for an Open Archival Information System (OAIS): https://public.ccsds.org/pubs/650x0m2.pdf
- CCSDS 651.0-M-1: Producer-Archive Interface Methodology Abstract Standard (PAIMAS): https://public.ccsds.org/Pubs/651x0m1.pdf
- DSA–WDS Partnership Working Group - Catalogue of Common Requirements v2.2, 2016: https://docs.google.com/document/d/1_DPwSA5P8LpK9Q34BhxJmX8So2GKL7eSLa-G-z5JvVg/edit#heading=h.xw6ahc91jrrs
- nestor Certification Working Group - nestor-materials 17: Explanatory notes on the nestor Seal for Trustworthy Digital Archives, July 2013: http://files.dnb.de/nestor/materialien/nestor_mat_17_eng.pdf

LIST OF GUIDELINES AND SOURCES

- PREMIS Data Dictionary for Preservation Metadata, v3.0, June 2015: http://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf
- Digital Preservation Coalition: *Digital Preservation Handbook*: http://www.dpconline.org/handbook/
- ICPSR/Nancy Y. McGovern, 2007: *Version 2.0 Digital Preservation Policy Framework: Outline*: https://www.icpsr.umich.edu/files/ICPSR/curation/preservation/policies/dp-policy-outline.pdf
- ERPANET, 2003: *Digital Preservation Policy Tool*: http://www.erpanet.org/guidance/docs/ERPANETPolicyTool.pdf
- Meta Archive Cooperative, 2010: *Preservation Policy Template*: https://metaarchive.org/public/resources/pres_comm/policy_planning/Digital_Preservation_Policy_Template.pdf
- The National Archives, 2011: *Digital Preservation Policies: Guidance for archives*: https://www.nationalarchives.gov.uk/documents/information-management/digital-preservation-policies-guidance-draft-v4.2.pdf
- British Library, Preservation Advisory Centre, 2013: *Building a Preservation Policy*: https://www.bl.uk/aboutus/stratpolprog/collectioncare/publications/booklets/building_a_preservation_policy.pdf
- Digital Curation Centre / Sarah Jones, 2010: *Preservation policy template for repositories*: http://www.dcc.ac.uk/sites/default/files/documents/Preservation%20policy%20template.pdf
- DISC-UK / DataShare Project, 2009: *Policy-making for Research: Data in Repositories: A Guide*: http://www.disc-uk.org/docs/guide.pdf
- nestor, 2014: nestor-materials 18: *Guidelines for the creation of an institutional policy on digital preservation*: http://files.dnb.de/nestor/materialien/nestor_mat_18-eng.pdf
- Directory of Open Access Repositories (OpenDOAR), 2014: *Policies Tool*: http://www.opendoar.org/tools/en/policies.php
- Repositories Support Project, 2013: *Policies and legal issues*: http://www.rsp.ac.uk/start/policies-and-legal-issues/
- DASISH, 2014: Deliverable 4.4: *Comprehensive Policy-Rules for Data Management in SSH*:

http://dasish.eu/publications/projectreports/DASISH_D_4.4-desember2014.pdf

- Charles Beagrie Limited (Neil Beagrie, Najla Semple, Peter Williams, and Richard Wright), 2008: Digital Preservation Policies Study: https://www.webarchive.org.uk/wayback/archive/20140615022334/http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_p1finalreport.pdf
- SCAPE, 2014: Catalogue of Preservation Policy Elements: http://wiki.opf-labs.org/display/SP/Catalogue+of+Preservation+Policy+Elements

List of Policies

CESSDA Archives

- FORS Preservation Policy v3.0, April 2017: http://forscenter.ch/wp-content/uploads/2015/05/Preservation-Policy_E_v3.pdf
- ADP Digital Preservation Policy, Working Version, 13. January 2017: http://www.adp.fdv.uni-lj.si/media/publikacije/ADP_policy_WorkingVersion.pdf
- DANS Preservation Policy, v1.1, 20. January, 2015: https://dans.knaw.nl/nl/deponeren/toelichting-data-deponeren/DANSpreservationpolicyUK.pdf
- UKDA Preservation Policy, v09.00, 15. June, 2016: http://www.data-archive.ac.uk/media/514523/cd062-preservationpolicy.pdf
- UKDA Reports & Publications (policies, processes, licences, etc.): http://www.data-archive.ac.uk/about/publications
- GESIS Digital Preservation Policy - Principles of digital preservation at the Data Archive for the Social Sciences, v.1.4.8 (English version), 7. April, 2015: http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/DAS_Preservation_Policy_eng_1.4.8.pdf
- ICPSR Digital Preservation Policy Framework, November 2016 (revision): https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/preservation/policies/dpp-framework.html
- Czech Social Science Data Archive Preservation Policy, Ver. 1.3 2016: http://archiv.soc.cas.cz/sites/default/files/csda_preservation_policy_0.pdf

Other Policies

- Parliamentary Archives: A Digital Preservation Policy for Parliament, 1st edition March 2009: http://www.parliament.uk/documents/upload/digitalpreservationpolicy1.0.pdf
- Harvard Dataverse Preservation Policy: http://best-practices.dataverse.org/harvard-policies/harvard-preservation-policy.html
- Dryad, Terms of Services: http://datadryad.org/pages/policies
- Figshare, Preservation Policies: https://support.figshare.com/support/solutions/articles/6000079077-preservation-policies

LIST OF FIGURES

- Figure 1: The PAS 197 Collections Management Framework
- Figure 2: Data Management Policy Framework for CESSDA Service Providers
- Figure 3: UKDA's functional and managerial structure
- Figure 4: UK Data Archive's OAIS-based functional model
- Figure 5: The organisational structure of FSD
- Figure 6: FSD's OAIS-based functional model

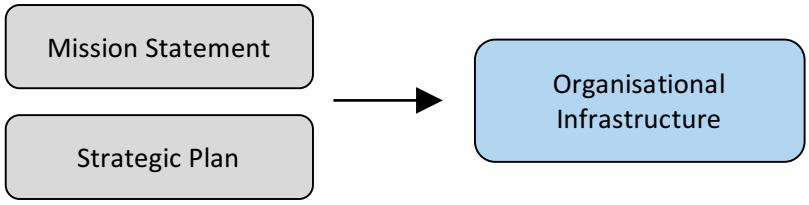
LIST OF TABLES

- Table 1: Policy description template
- Table 2: The three tiers of UKDA data access

APPENDIX 1: HOW TO PREPARE A SERVICE PROVIDER POLICY

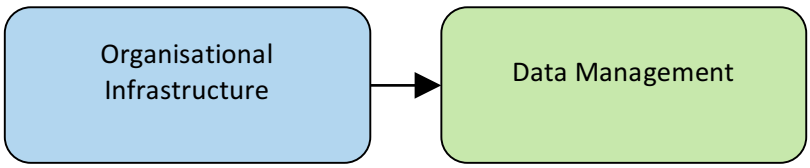
STEP 1: Start by making the core mission of the service provider explicit through a written and publicly available ***Mission Statement***. The statement describes the commitment of the organization for the management of digital objects in its custody.

STEP 2: Develop a ***Strategic Plan*** that operationalise the mission into specific goals and objectives for achieving the mission. The Strategic Plans can consist of long-term and/or short-term plans. Short-term planning looks at the characteristics of the organisation in the present and sets out strategies for improving them. Long-term planning addresses the situation of the organisation in its social, economic and political environment and develops strategies for adapting and influencing its position to achieve long-term goals.



STEP 3: Based on the Mission Statement and the Strategic Plan, articulate a set of ***Organisational Infrastructure policy clauses***. These clauses should be statements on a high organisational level which applies to all parts of the organisation. They should aim to capture the general business drivers, i.e. the conditions, resources and processes that are vital for the existence and continuation of the organisation. Key aspects of the Organisational Infrastructure policy clauses are *Scope of the Archive*; defining a *Designated Community*; clarify *Funding, Staff and Resource Planning*; define *Roles and Responsibilities*; and identification and clarification of the *Legal and Regulatory Framework* which the archive operates under.

STEP 4: Based on the Organisational Infrastructure policy clauses, articulate a set **Data Management Policy Clauses**. These are the *implementations* of the organisational clauses and they describe the approaches taken to fulfil the organisational clauses. These implementations can apply to specific parts of the organisation, to specific processes, and, in the archival context, to specific parts of the data holdings or data collections.

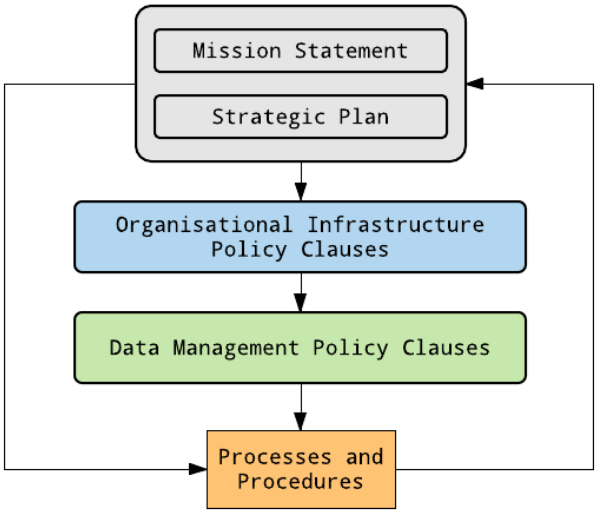


STEP 5: Decide on the Policy Document Model. When the policies are articulated they may be linked together in, either in one master document with an overarching Preservation or Data Management policy statement that covers information common to all aspects of the organisation, or simply by cross-referencing the policies to each other.

The CESSDA Data Management Policy Framework suggests following the approach taken by the Audit and Certification of Trustworthy Digital Repositories, namely to arrange the different policy clauses into three sets of policies: Collection policy, Preservation policy and Access policy.

It is important however, to underline that the policy framework and the distinction between Organisational Infrastructure and Data Management with a sub-set of policy clauses, is an *abstract* model. It can result in different policy documents, a different distribution of subjects between policy documents, different document names, etc.; it can be concentrated or 'compounded' into one or a few documents or it can be modulated and distributed among several different and specific policy documents and policy elements. Hence, the clauses are in the subsequent segments presented independent of the suggested arrangement of sub-set policies

STEP 6: Regular audit and review of policies and procedures feeds back into Mission and Strategy which in turn refines and reiterates the organisational and operational policy clauses that leads to a cycle of improvement.



The CESSDA Data Management Policy Framework - full model