



---

# ● Digital preservation policy

---

Version 1.0

Date: 28. julij 2017  
Status of the document: Public

## **Arhiv družboslovnih podatkov,**

UL, Fakulteta za družbene vede, Kardeljeva ploščad 5, 1000 Ljubljana

tel: 01 5805 292, e-mail: [arhiv.podatkov@fdv.uni-lj.si](mailto:arhiv.podatkov@fdv.uni-lj.si), <http://www.adp.fdv.uni-lj.si/>, Facebook: [Arhiv Družboslovnih podatko](#), Twitter: [@ArhivPodatkov](#)

## TABLE OF CONTENTS

INTRODUCTION	4
Purpose of the document	4
Structure of the document	4
Control, review and revision of the document	5
1. DESCRIPTION OF THE ARCHIVE	6
1.1 About the archive	6
1.2 Designated Community	7
1.2.1 Usage and impact of archives' holdings	7
1.3 Inclusion in the national and international research infrastructure	8
2. ORGANIZATIONAL INFRASTRUCTURE	11
2.1 Mission of digital preservation and main functions	11
2.2 Legal framework and responsibilities	11
2.2.1 Protection of confidentiality, adherence to disciplinary and ethical norms	13
2.3 Sustainability and funding	13
2.4 Model of digital preservation in ADP	14
2.5 Collection development policy	17
2.6 Principles and strategy of digital preservation	17
2.6.1 Strategy of long-term availability and continuity of access	19
2.7 Roles and responsibilities	20
2.7.1 Working bodies of the ADP	21
2.7.2 Working groups of the ADP	22
2.7.3 Competences and development of the staff	22
2.7.4 Outsource partners and expert guidance	23
3. DIGITAL OBJECT MANAGEMENT	24
3.1 Obtaining holdings and selecting data for ingest	24
3.1.1 Evaluating the quality of studies and adhering to the criteria of ingest	24
3.1.2 Recommended formats	25
3.2 Ingest	26
3.2.1 Submission information package (SIP)	26
3.2.2 Protection of confidentiality and anonymization	28
3.2.3 Metadata standards and interoperability	29
3.3 Archival storage of data	30
3.4 Management of (meta)data	30
3.4.1 Versioning and transformation management	31

3.4.2 Policy of disposal of items from the holdings	32
3.5 Access to data (DIP)	32
3.5.1 Licenses, limitations and types of data access	33
3.5.2 Management of access to microdata	33
3.5.3 Protection of confidentiality and access to data	34
3.5.4 Citation of holdings	35
3.6 Administration	35
4. TECHNOLOGY	38
4.1 Architecture of the information technology	38
4.2 Security and risk management	38
TERMINOLOGY	41
RESOURCES	44
APPENDIX	47
A. List of internal guidelines and instructions	47

# INTRODUCTION

## Purpose of the document

The policy of digital preservation aims to determine the rules, responsibilities, roles and the system of monitoring data management in the Slovenian Social Science Data Archives (Arhiv družboslovnih podatkov, ADP). The policy is developed in line with the core mission of the organization in order to assure permanent accessibility of archives' holdings and thus present the archive as a trustworthy data repository in the eyes of its users, data providers, and funders.

The purpose of the digital preservation policy in the ADP is continuous improvement of its functioning in a way to ensure larger transparency and fulfillment of the core mission of the organization. In this way, the requirements of the membership of the Republic of Slovenia in the international research infrastructure unit CESSDA ERIC are fulfilled, for which the ADP performs the role of the national service provider.

The following document is an attempt to define changes and improvements in the functioning of the ADP in more detail. These changes and improvements were necessary in order to apply for the Core Certificate of Trustworthy Data Repositories (DSA-WDS, see [Core Trustworthy Data Repositories Requirements v01.00](#)), which is one of the main obligations of the membership in [CESSDA ERIC](#). In the following document, a systematic overview of ADP high-level workflow is presented, with references to the internal documentation of detailed daily routine workflow. In some places, suggestions for improvements (following the best practices of similar organizations in the international environment) are presented as well, with some of them being in the process of implementation, whereas others in the phase of preparation.

The ADP has taken into consideration in the development of its policy of digital preservation the characteristics of research environments, established approaches and good practices, such as the [Guidelines for the Creation of an Institutional Policy on Digital Preservation, Authored and Published by the Nestor Working Group Policy](#) (2014), the policies of partner organisations, especially ICPSR (2009) and UK DA ([Preservation Policy](#)) and the scheme within the [SERSCIDA Workpackage D5.1 result](#). Some of the starting points are summarized also in the contribution by [Štebe and Vipavc Brvar \(2011\)](#), where a description of the realization of the general guidelines of Core Certification and the OAIS model in the form of the ADP procedures and workflows is presented. Individual work processes of the ADP are described in detail in internal reference guides and instructions (see [Appendix A](#)).

## Structure of the document

The document consists of four main parts that follow the logic of requirements to obtain the [Core Certificate of Trustworthy Data Repositories](#) (DSA-WDS) and that responds to certain dilemmas in establishing a policy of digital preservation.

In accordance with the structure of the Core certification requirements, the first chapter describes the context of functioning of the archive, which includes the type of the repository, description of the level of curation, its designated communities and the importance of the collection of the ADP in the Slovenian social sciences' environment, including the information on the involvement of the ADP in the national, regional and international research environment.

The second part deals with the organizational infrastructure since it is of vital importance for digital curatorship to have an appropriate organizational structure that reflects the mission and goals of the organization, as well as one that is well supported with human and technical resources and financial viability.

The third part presents the approach of ADP of handling digital objects in more detail and describes the model of digital preservation of the ADP holdings in the catalog of ADP in accordance with the OAIS standard and the requirements of the Core Certificate of Trustworthy Data Repositories.

The fourth part describes technological characteristics of the functioning of the archive. Here the technical and security specifications that are in line with the mission and the goals of the ADP are presented in more detail.

At the end, there is a list of used resources and literature, as well as the terminology used throughout the document. Finally, there is a list of internal guides and instructions that manage the workflow and processes in the ADP.

### **Control, review and revision of the document**

**The following document will be reviewed and, if needed, updated every 2 years.** In the intermediate period, individual elements of the policy and interconnected procedures will be updated internally. The monitoring of the implementation of the digital preservation policy and its updating along individual work areas is the task of the staff of the ADP.

The Head of Digital Preservation is responsible for all the organizational tasks in preparing, updating and implementing the policy. If needed, the Head conveys a working group, consisting of the selected staff of the ADP, to prepare the content of the documents. The Head of the Organization approves the changes of the policy and assures conditions for its implementation.

# 1. DESCRIPTION OF THE ARCHIVE

## 1.1 About the archive

Slovenian Social Science Data Archives (ADP) is a national research infrastructure for social sciences, whose main mission is to manage data and data services in order to support research, education, and general welfare. The ADP performs data-level curation, as it checks received data and documentation in detail, adds detailed metadata to the study, creates new formats, enhances documentation and edits the deposited data for accuracy.

At the essence of the ADP activities is digital curation of high-quality research data that are openly available to researchers and other interested public. This way, financial savings are being made in cases of unnecessary duplication of new research data, as well as higher quality of research findings is assured. This core idea lies in the introduction of the document [Principles and Guidelines for Access to Research Data from Public Funding](#) that commits OECD members to prepare an action plan for open access to research data. By becoming a member of OECD in 2010, the Republic of Slovenia showed continuity in supporting the infrastructure of the ADP, as it committed to the sustainable support of the development of science and its infrastructure.

The ADP is an organizational unit of the [Social Sciences Research Institute](#) (Inštitut za družbene vede - IDV) of the Faculty of Social Sciences, University of Ljubljana (this status is officially recognized in the Rules on Organization and Functioning of the Faculty of Social Sciences of the University of Ljubljana (2017) and in the Rules on Organization and Implementation of the Research Activities of the Faculty of Social Sciences of the University of Ljubljana (2017)). Act on the Establishment of ADP (1997) defines the main activities of the archive as *»collecting, documenting and disseminating original data from social sciences studies, and connecting with similar institutions in national and international merit«*.

The mission of ADP, as defined by the [Rules on Organization and Functioning of the Faculty of Social Sciences of the University of Ljubljana \(2017\)](#) is, *»that it preserves original data from social sciences studies, conducted within the faculty, which all researchers from the faculty are obliged to deposit, including original materials, and all other social sciences studies that are important for the social sciences«*. The functioning of ADP is thus clearly directed towards assuring access to and preserving data in the field of Slovenian social sciences. Within its mission, the ADP establishes itself as a national infrastructure that collects important data sources from a wide range of social sciences, interesting for the analyses of the Slovenian society, deposits, preserves and promotes their further use in scientific, educational and other purposes.

Within its membership in CESSDA the Ministry of Education, Science and Sports of the Republic of Slovenia has appointed ADP as the national data service provider. The long-term national importance of the ADP is clear also in the ongoing support of the ministry, which from the establishment of the ADP onwards ensures funding for its functioning. Since 2004, the funding of operations has been provided within the infrastructure program [Network of research infrastructure centers at the University of Ljubljana](#) (MRIC UL). The current program period of ADP funding is 2015 to 2020.

ADP reports on its activities to the Council of ADP, consisting of honorable social sciences researchers. On an annual basis, it reports also to the [Slovenian Research Agency](#) that funds the infrastructure research program of the University of Ljubljana (MRIC UL).

### Remaining challenges

*The composition of the Council ADP needs to be revised and a new mandate needs to be set. How the council is to be formed needs to be determined, putting into consideration that representatives from different fields are included, which will be able to represent their professional guidelines in order to fulfill the needs of designated users from different fields.*

*Curator of the field: Head of Organization*

## 1.2 Designated Community

Target users are **national and foreign researchers, teachers and students**, who have the knowledge of data handling and are statistically literate for independent understanding and analyzing of materials. Such literacy includes the ability to independently search data for a concrete research problem, to evaluate the usability of data, according to descriptions provided by the ADP about the origins of the data, its methodology, data collection process and detailed descriptions of individual variables, the ability to make analyses of selected data, to interpret results and form conclusions. It is also expected from users of qualitative data that they are familiar with the basics of qualitative methodology and that they understand the context of qualitative research design.

By organizing training for users and by stimulating knowledge exchanges, the ADP actively promotes the secondary use of data amongst its designated communities. The ADP stimulates also the advanced use of data, for example, applied social problem-solving research or various high-quality scientific studies. The ADP aims to promote the secondary use of data even broader and therefore organizes training for non-academic users. **Journalists, policy makers and regular citizens** are considered to be part of the broader public if they express the desire to legitimately use the data and accept the rules and restrictions of working with secondary data.

The ADP adapts access to data, metadata, and other data services to its various target users. Customized views and adjusted services are available that fulfill the needs of different segments of users. The ADP regularly performs quality evaluations of its services (including surveys of users' satisfaction), taking into account internationally comparable criteria of the quality of service ([ADP User Satisfaction Survey 2016](#)). Based on these evaluations it forms and adapts its services to different users' needs and therefore complies with the third and fourth requirement of the OAIS – that is defining its designated communities, considering their levels of knowledge and assuring independent understanding of deposited data by providing enough information on data itself (see [Chapter 3.2.3](#)).

### 1.2.1 Usage and impact of archives' holdings

With the help of a registration form, the ADP follows the usage of its holdings and their importance. In the form, the users define 1) the purpose of the use of the ADP holdings: whether they intend to use data for educational, scientific, public or commercial means and 2) their category: academic user (under-graduate, graduate student, employed in an educational institution (teachers, researchers)), non-profit user (secondary school student, employed in secondary school, employed in non-governmental/governmental organization, non-profit, personal use), commercial user (employed in commercial enterprise or organization) and other user.

The ADP prepares general workshops, intended to train users on the secondary use of data. In order to approach the needs of users as tightly as possible, the ADP asks the users at the point of registration for a workshop to provide information on their educational background,

about their existing knowledge (on data, statistics etc.) and on intended data use. On this basis, the ADP adapts individual workshop to the group of participants (the theme and the degree of complexity). The trainers are always available to participants during and after the workshop to answer specific questions and/or give advice. The ADP also prepares workshops and lectures for specific target groups, such as professors, high school teachers, students of sociology, students of psychology etc. In addition, the ADP conducts special thematic workshops that are of interest to disciplinary researchers and professors. With these special workshops, advanced users are targeted. After the workshops, the ADP invites participants to fill-in a user satisfaction survey. Such feedback is important to form future plans of working with designated communities. Currently, the ADP is preparing a thematic manual on how to use data in teaching together with teachers and professors in order to approach the needs of users even further. By doing so, the ADP on the one side listens to the needs of the teachers, and on the other side tries to promote the use of secondary data amongst students as early as possible in the educational process.

Amongst the most prominent users of data holdings of the ADP are researchers and students, a smaller part of users come from other organizations. Therefore, a conclusion can be made that the data holdings of the ADP are mainly used for scientific and research activities as well as for educational purposes.

Data from the ADP are usually used for seminar work, scientific publication or as an addition to lectures and practical exercises with students – in some cases adjusted modules for teaching are provided, for example, the use of data from Labour Force Survey. The ADP offers study and data related materials to its users, which may be used to develop new studies (for example original questionnaires and collections of indicators for measurement of established theoretical concepts from scientific literature) or to form theoretical grounds of interpreting data (for example titles of publications, based on a certain data source).

With the intention to further stimulate the secondary use of data from the ADP, the special [Klinar Fund Award](#) is offered, for which all on the Faculty of Social Sciences defended undergraduate, graduate works and doctoral dissertations that use data from ADP as primary sources may apply.

### **1.3 Inclusion in the national and international research infrastructure**

Besides its main mission to provide services for its designated communities in the national environment, the ADP is actively involved in a broader national and international research infrastructure. As a data service provider, it actively follows international development initiatives from its discipline (OECD, academic unions, research policies from developed EU countries, USA and Australia) and promotes their implementation into the national environment. The staff of the ADP is involved in different working groups as external experts, preparing drafts or giving advice on the formation of strategic documents and guidelines, such as the strategy of the Ministry of Education, Science and Sports [National Strategy of Open Access to Scientific Publications and Research Data in Slovenia 2015-2010](#) and its Implementation Action Plan, and participation in the working groups for the protocol [Dealing with Research Data at Science Europe](#). Initiatives, guidelines, and demands of national and international organizations and science funders are promoted in the national research community by organizing workshops, roundtables, and seminars, intended for various participants of the scientific research community (researchers, heads of institutions, publishers, libraries etc.).

Due to the interdisciplinary nature of research data holdings of the ADP, the ADP is of importance to users from other scientific fields as well as other infrastructure networks. The content of research data holdings and possibilities of reuse are broad, for example in kinesiology, psychology, geography, history, medicine etc.<sup>1</sup>

On the national level, the ADP connects with other national disciplinary infrastructures, dealing with research data management internationally. ADP works most closely with the national representatives of [Digital Research Infrastructure for the Arts and Humanities – DARIAH](#) (Institute of Contemporary History – SiStory and affiliated research infrastructure ZRC SAZU) and [Common Language Resources and Technology Infrastructure - CLARIN](#) (Institute Jožef Stefan), exchanging experiences, good practices, interdisciplinary exchange of data and discussing the possibilities to develop interconnected infrastructure services on the national level.

The inclusion of the ADP in the activities of the Pan-European research infrastructure – [CESSDA \(Consortium of European Social Science Data Archives\)](#) lies in the forefront of the international cooperation of the ADP. CESSDA was established in 1976 and brings together social science data archives across Europe. In 2013 it reformulated into a European scientific infrastructural unit in accordance with the list of European priorities ESFRI ([European Strategy Forum on Research Infrastructures](#)). The Republic of Slovenia added the membership in the Consortium CESSDA as one of the priorities of the [Plan for the Development of Research Infrastructures 2011–2020](#) and is one of its first members. Under the authority of the Ministry of Education, Science and Sports, the ADP acts as a national service provider of CESSDA. CESSDA is based on mutual expert and organizational support and aims to provide integrated and sustainable data services to the social sciences. Its aim is to spread quality research data, regardless of the geographical location of users, by taking into consideration contemporary international standards, recommendations, and tools.

Membership in CESSDA enables professional exchange and cooperation. Membership in CESSDA also means that national service providers, such as the ADP, need to fulfill several requirements. In the [Work Plans of CESSDA](#), there are several projects that require cooperation such as the use of existing metadata on the level of variables for the creation of EuroQuestion Bank, the inclusion of new types of data and content (for example qualitative data, register data, big data, health data and spatial data etc.). One of the CESSDA activities of the ADP is also working on are shared project applications within the [EU Framework Programme for Research and Innovation Horizon 2020](#).<sup>2</sup>

ADP is particularly cooperating in the following tasks of the working groups within CESSDA:

- (1) harmonized preparation of metadata enabling search in a common CESSDA catalog;
- (2) preparation of documentation and protocol to obtain Core Certificate of Trustworthy Data Repositories ([Core Certification of Data Repositories](#));
- (3) on-going training of staff (training group), including annual expert seminars (CESSDA Expert seminar);
- (4) extended collaboration with national statistics offices and Eurostat;

---

<sup>1</sup> As an example a round table in 2014 can be referenced, where the importance of public opinion data for contemporary historical research was exposed.

<sup>2</sup> An example of such a project is [CESSDA SaW \(2015-2017\)](#) Horizon 2020 project, within which CESSDA aims to expand and integrate the existing research data infrastructure in Europe and broader with an aim to enable researchers from different regions and countries to work together within and between different disciplines. As a project partner, the ADP intensively cooperates in these processes.

- (5) preparation of internationally harmonized thesaurus ELSST ([European Language Social Science Thesaurus](#));
- (6) cooperation in the field of communication of the mission and vision of the CESSDA (Communication group).

In addition to the above, the ADP cooperates in different international initiatives of publishing and citing research data. Amongst the groups, with which the ADP forms professional ties and organizes exchanges of data, are [Interuniversity Consortium for Political and Social Research](#) (ICPSR), [Research Data Alliance](#) (RDA), [Knowledge Exchange](#), [International Federation of Data Organisations](#) (IFDO), [Committee on Data for Science and Technology](#) (CODATA) and others.

## 2. ORGANIZATIONAL INFRASTRUCTURE

### 2.1 Mission of digital preservation and main functions

The mission of digital preservation of the ADP is to ensure and promote sustainable services of ingest, storage and access to quality research data from the field of Slovenian social sciences and broader, with potential for secondary analysis.

Main functions of the disciplinary data service provider ADP are:

- **Acquiring important research data from a wide range of social sciences**, interesting for the study of the Slovenian society.
- **Appraisal of submitted research data and their selection for deposit.** Research data from scientifically important studies that reach theoretical and methodological excellence have precedence, especially longitudinal and internationally comparable data that include data from Slovenia. Investments in dealing with research data must be proportional to their value and benefits of further usage.
- **Ingesting and processing research data and other documentation**, together with **creating metadata** with the goal to prepare a package for long-term digital preservation (AIP) and preparing for access and further use for scientific, educational and other purposes (DIP).
- **Long-term digital preservation.**
- Providing **access to research data** that enables easy and well-informed usage for various purposes.
- **Training researchers** on planning, dealing and preparing data for ingest in open access.
- Actively **promoting the secondary use of research data** through training of users and stimulating exchanges of knowledge amongst users.

### 2.2 Legal framework and responsibilities

Functioning of the ADP in the Slovenian research environment is subjected to certain rules of exchange and use of knowledge and information. The functioning needs to be in line with the relevant national legal framework, defining the area of access and use of digital objects. It is of vital importance that the ADP informs its users on the contents of the relevant legal framework and that it performs monitoring of this fulfillment with the existing rules (see [CESSDA CDM CPA1.2](#)). Within its operation the ADP carefully follows the relevant legislation:

- [The Law on Personal Data Protection Act](#), defining rights, obligations and measures, with which encroachments into privacy and dignity of an individual in dealing with personal data are prevented. The law determines the obligations of the ADP at the point of ingest, storage and access to microdata.
- [Law on Copyright and Related Rights](#), recognizing copyright and related rights, including the rights of authors of research data and other related materials in distribution.

Since the main mission of the ADP lies in the field of providing services supporting research, the ADP is not directly tied to national requirements of the functioning of official archival organizations. The basic guidelines for its functioning are nonetheless extracted from the relevant field of legislation:

- [Law on the Protection of Documents and Archives and Archival Institutions](#), defining the mode, organization, infrastructure and execution of ingest and storage of holdings in physical and digital form.

In collaborating with other organizations, the ADP follows the following legislation:

- [The Law on Access to Public Information](#), defining rules of access and reuse of public data, created within public administration and other public institutions.
- [National Statistics Act](#), defining the functioning of official statistics offices as a professional and independent activity of performing the program of statistical research. The ADP here actively cooperates with the national provider ([Statistical Office of the Republic of Slovenia](#)).

In terms of rules of ingest and rules of accessing materials and research data, the ADP uses fixed standardized forms. Ingest of research data is accompanied by documentation on the fulfillment of [ethical standards](#), legal frameworks, and good practices. Agreement on ingest in the form of a [License Agreement](#), signed by the data depositor and the ADP, defines the relationship between the two parties: it gives rights to the ADP to process the study for the purpose of digital preservation and gives rights of access to research data to users. By signing the agreement, the data depositor in his/her behalf or under researchers authority agrees that the ADP prepares, stores and distributes research data (see [Digital Object Management](#)). With the License Agreement, the ADP gets rights on the basis of which it can (1) prepare materials for storage and distribution and (2) preserve and distribute materials. With such arrangement the ADP is fulfilling the second obligation of OAIS – it is necessary to obtain sufficient rights for dealing with information to be able to provide their long-term storage.

Rights and obligations to prepare and distribute metadata and microdata of register data are subjected to The Agreement of Cooperation with the Statistical Office of the Republic of Slovenia. Contracts, enabling the staff of ADP to access unprotected microdata of SURS, are also subjected to this agreement.

Research data and accompanying documentation are available to users under [Creative Commons 4.0 licenses](#). The ADP uses the following licenses: CC0 – without restrictions, CCBY – attribution alone, CCBYNC – attribution + non-commercial (for more see [section on Access](#)). Final responsibility for thorough use of research data that takes into consideration ethical principles of confidentiality, copyright, and academic honesty lies on the user of the research data. When accessing research data users are warned to carefully respect the principle of ethical reuse of data (see [example](#)). Users can inform themselves on [specific conditions of reuse in the study descriptions](#). By registering at the ADP, the users oblige themselves to respect professional and disciplinary ethical standards (see [General Provisions and Terms of Use](#)), as well as ethical and legislative restrictions on the reuse of data, especially clauses relating to confidentiality. Users can use research data only for the purposes expressed at the point of registration.

#### **Remaining challenges**

*Additions to the process of pre-ingest and ingest are needed.* License agreement needs to be updated with the information on the functioning of the disciplinary data provider as an abstract entity, taking over curatorship of the research data, regardless of its present organizational form, thus enabling possibility of succession and transfer of research data and other related documentation in the system of cooperative long-term preservation.

Rights, connected with the workflows of processing and preparing research data for digital preservation, need to be defined more explicitly in the License Agreement. Good practices of similar organizations need to be taken into consideration (for example [DANS](#)).

The internal procedure of processing received materials (see [Workflow](#)) needs to be updated with a clear pre-step before ingest of examining invulnerability of copyright and taking into consideration possible licenses of third persons.

### ***2.2.1 Protection of confidentiality, adherence to disciplinary and ethical norms***

In the social sciences, it is often the case that researchers deal with personal or other sensitive data and therefore need to be especially attentive to protect the confidentiality of their research subjects. In addition to following the rules set in [The Law on Personal Data Protection Act](#), the ADP takes into consideration relevant codes of ethics of the social sciences research community. Relevant (disciplinary) codes of ethics are:

- [Code of Ethics of the University of Ljubljana](#),
- [European Code of Conduct of Research Integrity](#),
- [Code of Professional Ethics of the Slovene Sociological Society](#),
- [Declaration of Code of Professional Ethics of the Slovene Statistical Society](#).

For every study, deposited in the catalog ADP, the ADP inspects before ingest whether or not it applies to above mentioned ethical norms. This means that the data depositor needs to assure that all research participants are protected against unnecessary damage, that they were notified and they voluntary accepted the cooperation and that the study respects current disciplinary methodological standards.

A *Commission for the Protection of Confidentiality* is an internal body of the ADP. The Commission meets in cases when at the initial inspection of the study a need for further protection of the research data is identified. The Commission has otherwise two main responsibilities: (1) to make decisions on the protection of submitted data before distribution, (2) to deal with applications for researchers to access less protected or unprotected microdata under special conditions.

The protection of research data is primarily the task of the data provider, wherein the ADP may only give support. Researchers may ask the ADP for guidance in the process of managing research data, including the field of protection of confidentiality and anonymisation of research data. The ADP may in extreme cases independently perform the protection of research data, following fixed procedures, however only in cases, where the data provider explicitly demands such intervention. The liability for proper protection of confidentiality is on the side of the data provider and does not transfer to the ADP, regardless of who has made the protection of the research data. The ADP performs the protection of confidentiality in two ways, by anonymizing research data (see [Chapter 3.2.2](#)) and by managing access to various types of data for various types of users under special conditions (see [Chapter 3.5.2](#)).

*Curator of the field: Head of Organization*

## **2.3 Sustainability and funding**

Advanced research policies of open access to scientific results anticipate sustainable functioning of the disciplinary data infrastructure services, such as the ADP. Such research infrastructure is based on the cooperation of different actors, including fulfilling the funder's strategies and policies, the needs represented by the scientific communities and other target users. From the point of resources optimization and utilization, it is essential that the ADP

functions continuously in order to fulfill its mission and functions as a national social science data service provider. To assure transparency, the ADP annually reports to its funders on its activities.

The Republic of Slovenia has accepted the obligation of the membership in the international infrastructural unit [CESSDA ERIC](#) to assure sustainability of research data services for the social sciences and appointed the ADP to be the national service provider, the organization to provide these nationally and internationally integrated services, arising from membership in CESSDA. The Ministry of Education, Science and Sports supports the implementation of these obligations by providing national funding of the infrastructural research program CESSDA, performed by the ADP within the program [MRIC UL](#). By naming the ADP to be the national data service provider, the Ministry explicitly guarantees the sustainability of the ADP's functioning, which is one of the core elements of trustworthy data repositories.

The existence and long-term sustainable functioning of the ADP as a disciplinary data service provider is expanded with obligations of its national membership status in CESSDA and defined in the [Research Infrastructure Roadmap 2011 – 2020. Revision 2016](#) (Government of the Republic of Slovenia 2016). The Roadmap serves as a strategic basis for the long-term financial support by the Slovenian Research Agency within the [Network of Research and Infrastructural Centres MRIC UL](#), in which the ADP is included (University of Ljubljana 2016). The programme is extended every 5 years, and the current program period runs from 2015 to 2020.

The Government of the Republic of Slovenia accepted in 2015 the [National Strategy of Open Access to Scientific Publications and Research Data in Slovenia 2015 – 2020](#) (Government of the Republic of Slovenia 2015) that contains a clause to ensure appropriate sustainable financing of national infrastructure of open access to scientific information in the form of publications and research data, composed of people, organizations, equipment (hardware and software) and content: *»Sustainability of the national open access infrastructure can only be ensured through continuous public funding, which enables the use of international standards for the creation, publication, dissemination, use, processing, preservation and archiving of scientific information, the education of all stakeholders and the notification of the national and international public on the availability and the ways of using the services and information via the national open access infrastructure. Openly accessible scientific information in the form of publications and research data has to be securely preserved to prevent loss, damage, and misuse.«* The existence and sustainable functioning of disciplinary data providers in Slovenia are expected also in the draft [Action Plan to Establish a System of Open Access to Research Data, financed by Public Funds](#) (Štebe et al. 2013).

As a member of international organizations, the Republic of Slovenia bounded to open access to research data from public funding as a public good in accordance with the international recommendation by the [OECD](#) (OECD 2007) and the [European Commission](#) (European Commission Research & Innovation 2016).

## **2.4 Model of digital preservation in ADP**

The ADP follows the OAIS model ([The Open Archival Information System](#); ISO 14721: 2012) and the requirements of the [Core Certificate of Trustworthy Data Repositories](#) in its processes of long-term digital preservation. Image 1 on the next page presents the compliance of ADP's operations with the OAIS model and the Core certificate.

The national document [Guidelines for Ingest, Long-term Preservation, and Access to Cultural Heritage in Digital Form](#) (Ministry of Culture 2013; in continuation Guidelines) recommends compliance with the OAIS standard (ISO 14721: 2012) for archives. The OAIS compliance is

also one of the obligations of national data service providers of CESSDA and one of the requirements of the Core Certificate of Trustworthy Data Repositories (*supporting tasks and functioning of the repository that is in line with internationally accepted archival standards*). By following these obligations and recommendations, the ADP wishes to be recognized as a trustworthy data repository amongst its designated community (users, data depositors, funders) (see [DASISH 2012](#)). The above mentioned Guidelines expose the OAIS model as a basis framework for understanding the concepts needed for the long-term preservation of digital materials and their reuse.<sup>3</sup> The terminology and the OAIS model represent a framework for description of individual phases in the data lifecycle that is on a national level harmonised from the point of view of participating organizations – cultural heritage curators (Guidelines 2013), while at the same time providing grounds for international comparisons and optimizations of procedures of different organizations from the same subject area (see Schumann 2012, Vardigan and Whiteman 2007, Beedham et al. 2007).

The OAIS model is composed of three complex information objects: **submission information package** (SIP), **archival information package** (AIP) and **dissemination information package** (DIP). These three information objects represent the state and transformation of digital content, together with metadata from the point of ingest, through archival storage and finally to giving access to users. Due to the procedure of pre-preparation and active acquisition of materials, a phase of pre-ingest or pre-SIP is added to the model (see [Producer-Archive Interface – Methodology Abstract Standard \(PAIMAS\) - ISO 20652: 2006](#)), where communication with possible data providers is established, and an evaluation of the suitability of the study for the collection of ADP is made. On this basis, a decision is made, which studies to ingest (see [Štebe and Vipavc Brvar 2011](#)).

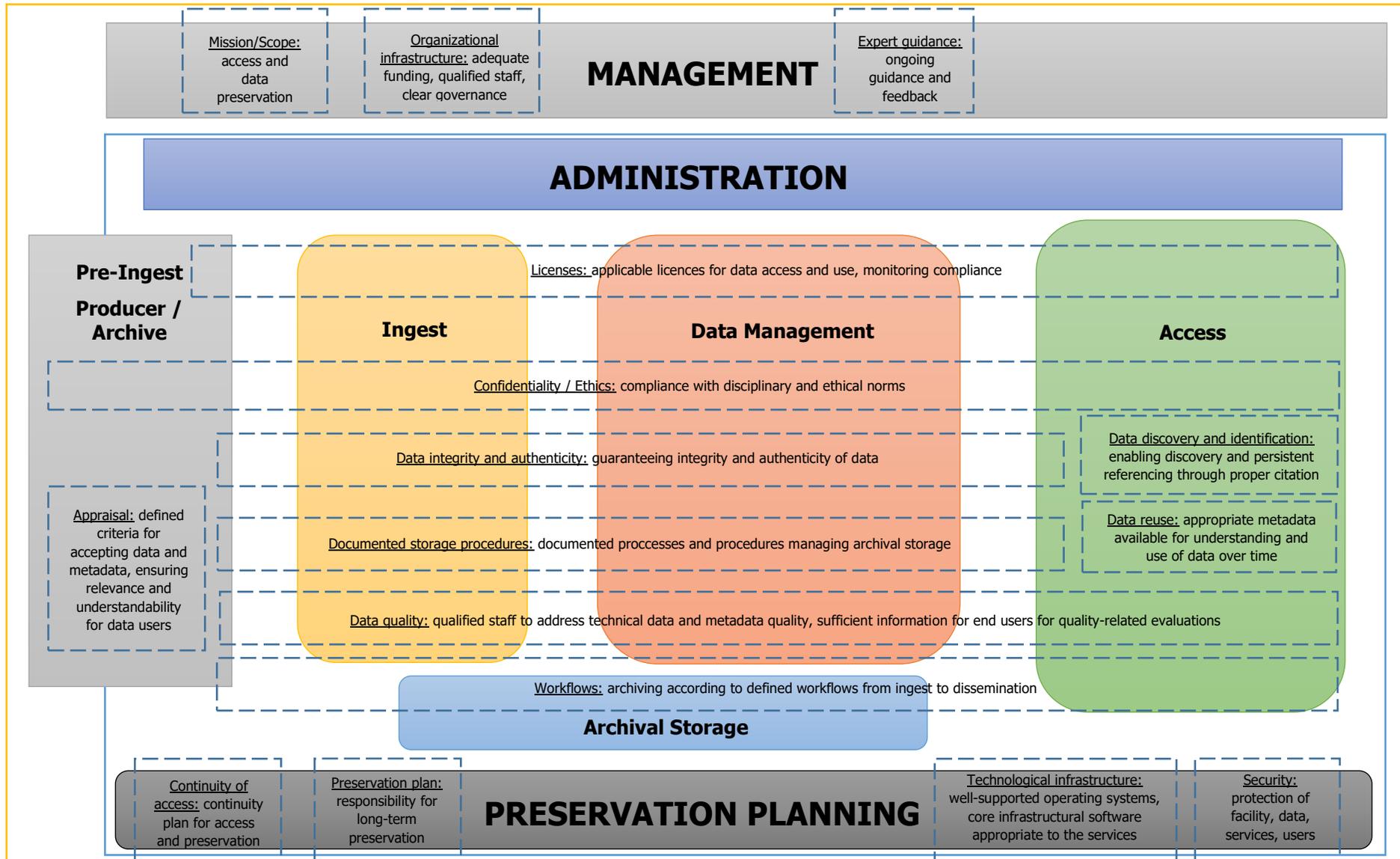
The procedure of archiving in the ADP follows fixed workflow that ranges from ingest to the dissemination of holdings (see [Workflow](#)). Individual workflows of the archive will be presented in the following chapters from the view of following the OAIS standard and taking into consideration the requirements of the Core Certification. More detailed description of the workflow is provided in internal guides and instructions for working in ADP (available in [Appendix A](#) and on the [webpage](#)).

*Curator of the ield: Head of Digital Preservation*

---

<sup>3</sup> By resolution of the expert council SIST from the field of electrical engineering, information technology and telecommunications the standad ISO 14721:2021 is made default as SIST ISO 14721:2013 (Decision 36.3) (Krstulović et al. 2013).

Image 1: Policy of digital preservation:  
 Compliance with the [OAIS](#) and [Core Trustworthy Data Repositories Requirements v01.00](#).



## 2.5 Collection development policy

The ADP holds research data that are of interest for social sciences' research, the emphasis being on problems, connected with the Slovenian society. Studies that reach theoretical, conceptual and methodological excellence have precedence, especially longitudinal and internationally comparable data that include data from Slovenia. Exceptionally, topics, which do not strictly fall into the social sciences' framework, are considered because they either structurally resemble research data from social sciences or lack other possibilities of preservation.

Long-term storage of research data demands additional efforts and costs for their preparation in the format that will enable further use. These costs and efforts are justified by the savings, represented in the continuous reuse of data. The ADP plans to expand its collection of research data also in other fields of science. This will enable preservation of research data of data providers from similar fields, while at the same time promoting interdisciplinary reuse of research data. The ADP connects with other similar disciplines (humanities, medicine etc.) in dealing with certain types, classes, and formats of data, whereas on the other hand it closely follows the development of its collection and stimulates the representation of less represented disciplines (educational sciences, psychology, economics etc.). The ADP strictly follows the current trends in the field of open government data, big data, harmonization and interdisciplinary integration of research data.

### Remaining challenges

The ADP will additionally include the professional public in determining the goals of the archive by stimulating the Council of the Social Sciences at the Slovenian Research Agency to define main problem areas and developmental initiatives to expand the collection of the ADP to new fields.

The ADP's working group for promotion, education, and training plans the development of the collection of the ADP on an annual basis. In the future, requirements of funders will need to be taken into account and thus an update of the criteria for ingest and the policy document will be needed, as well as the development of technical capabilities (for example self-archiving). The services of the ADP also need to follow the policies of scientific journals and demands by funders to enable open access to research data, related to scientific publications (for example H2020).

*Curator of the field: Head of Acquisition and Ingest, Head of Trainings and Promotion*

## 2.6 Principles and strategy of digital preservation

In 2011 the ADP conducted an analysis of gaps in its long-term digital preservation, which was the basis for the upgrade of digital preservation system, including the elaboration of fixed workflows. Internal procedures and instructions now set detailed rules and procedures of processing and managing digital objects (see [Appendix A](#) and our [webpage](#)).

The ADP follows the fourth requirement of OAIS in the requirements of the Core Certification in establishing the strategy for digital preservation. By ensuring that the information is preserved in an independently understandable way, the designated community is able to understand the information without needing the assistance of the experts who produced the information, even if the digital environment, in which they were primarily stored, becomes obsolete. The basic strategy of digital preservation in the ADP is **normalization at ingest**.

For this purpose, the ADP has set rules of appropriate formats of data files that are being accepted from data providers, as well as compatible formats for long-term digital preservation (see [Recommended formats](#)).

The ADP has a mission of long-term digital preservation of its holdings. The main aims of the digital archives are long-term storage of data files and other materials, especially in connection with the following principles:

- Preserving data integrity: ensuring that the archived data are protected against unauthorized alteration.
- Preserving authenticity: ensuring that the digital objects come from a documented originator and that they are what they purport to be.
- Completeness: it is clear that no parts of the digital object are missing.
- Preserving readability: it is possible to show and interpret at least the most recent version of a digital information object at any time. In this context, it is important to know which formats occur for each version.
- Preserving locatability: persistent identifiers (URN) and entries in relevant directories and search engines are planned to be used.
- Preserving confidentiality: taking into consideration relevant legal rules, ethical and disciplinary standards.

By meeting the above principles and creating external transparency of its activities, the ADP considerably contributes to the credibility of the data archive, which is one of the core aims of the present document.

In order to fulfill the sixth requirement of OAIS that is to ensure that data are independently understandable and usable, the ADP follows detailed documentation of all transformations in the processing data by which it assures the preservation of the authenticity of the digital objects. All the changes related to normalization of research data and interventions in data for the purpose of anonymization are carefully documented (see [Chapter 3.4.1](#)). All information mentioned is accessible to final users in the form of metadata of the study. The formats of data files, used by the ADP, are chosen by taking into consideration challenges of long-term preservation, wherein proprietary, outdated and rarely used formats are being avoided. If possible, all textual documentation about the study is saved in PDF/A format, so as to preserve the look and textual status of the documents. Data files are saved in ASCII format, including the DDI record of structure and content of data files, accompanying syntaxes for reading the files are added as well.

The ADP's staff regularly follows the development in the field of preservation of similar organizations, for example amongst the members of the CESSDA, within the professional association IASSIST, and actively cooperates in professional debates. On the national level, the ADP cooperates with the National and University Library, the Archive of the Republic of Slovenia and other similar institutions from the field of the protection of scientific and cultural heritage (conferences in the organization by NUK).<sup>4</sup> The ADP's staff is involved in regular professional training and cooperates in knowledge exchanges between archival and digital humanities in questions concerning digital preservation and efficient development of data services (NUK, UKM, CTK, Archive of the Republic of Slovenia, SURS, and other research libraries).<sup>5</sup> The CESSDA is continuing its tradition of expert seminars ([CESSDA Expert Seminar](#)), where the ADP participates with professional contributions. The ADP also participates in the international association and international conferences of the IASSIST (*International Association for Social Science Information Service and Technology*). The Head of Digital

---

<sup>4</sup> For example, the Congress of Digital Content: the development, storage and access, 5 – 6 June 2014, Ljubljana.

<sup>5</sup> For example [Seminar Practical Prospects of Open Access Publishing](#), 20 May 2015, Ljubljana.

Storage cares for the inclusion of professionally tested approaches and principles of digital preservation in the present strategy.

In the framework of the CESSDA AS the ADP will contribute to the development of educational infrastructure, to the examination of the needs of the designated communities and to the formation of the priority issues/programs for education and training. The Head of Trainings and Promotion keeps the organization of internal and external training in the field of digital preservation.

#### **Remaining challenges**

It is reasonable to search for balance between expenses and the educational and historical importance of the holdings in monitoring and adopting the present strategy. In this process, it is important to take into consideration two core principles: 1) information media need to be stored appropriately in order to avoid damage and 2) data formats and physical information media may become outdated, so a regular monitoring, updating, and maintenance is obligatory.

Rules of monitoring need to be determined – modes of monitoring (for example formats, migrations, media etc.) –, and a strategy needs to be developed for conformance with available means and staff resources, the importance of the collection or its parts (educational, historical), legal regulations, demands of funders etc.

In the future, the ADP will carefully examine individual tools and implement in its workflow the most appropriate tool to convert data files in formats of long-term storage (for example SledgeHammer).

#### ***2.6.1 Strategy of long-term availability and continuity of access***

Concerning the continuity of access, the ADP assures that appropriate local equipment and procedures for long-term digital storage are implemented. In addition, existing tools and services of a backup security duplication in the national environment are used.<sup>6</sup>

The ADP aspires to practice the policy of open access that determines that the service in the framework of open access to research data for final users is free of charge. In extreme cases when funding of ADP would be reduced, ADP could start charging users part of data access services that are in the current system free.

The ADP assures transparency of its operation by responding to the needs of its users. Using diverse ways of promoting its activities, it informs the public on its mission, resources, and services, for examples on workshops, conferences, meetings, round tables. In its annual report, submitted to the Ministry of Education, Science and Sports, the ADP reports on the realization of its plans.

The [Rules on Organization and Implementation of the Research Activities of the Faculty of Social Sciences of the University of Ljubljana](#) (2017) determine official obligation of the Faculty to “the continuity of preserving and providing access to research data, preserved by the ADP by creating appropriate organizational and financial conditions for its functioning. This is and will continue to be done by acquiring public funds from the Slovenian Research Agency to support the infrastructural research programme of archiving. The Faculty aspires to preserve these financial flows. In case the ADP will cease to function, the Faculty will, together with the ADP, organize and prepare a plan of transferring data to an external institution that will

---

<sup>6</sup> See the details of transferring backup copies to ARNES and NUK in [Chapter 4.2](#).

responsibly and appropriately assure the archiving of the data, or to another organizational unit, under the condition that it will have assured external funding”.

### **Remaining challenges**

In the preparation of a plan of possible transfer and succession of digital preservation services of research data of the ADP (cases of situational insecurity), there are possibilities of cooperation and usage of infrastructural capacities in the National and University Library of Slovenia. In the document Strategy of Permanent Storage of Digital Resources in National and University Library 2012–2020 the goal 9 says: »*The library will cooperate with other Slovene and international organizations in planning and implementing activities in the field of long-term preservation of digital objects*«.

The national priority areas from the Plan of the Development of Research Infrastructure (Digital national resources, Social Sciences, and Humanities' Research Infrastructure) states that in avoidance of duplication »*participles should aspire to cooperate on the national level in implementing support activities*«. To ensure this possibility, technological demands need to be inspected and harmonized and an agreement about possible succession needs to be established. The cooperation and networking between institutions in the field of scientific information and culture with the Archive of the Republic of Slovenia is already supported in the document Guidelines (Ministry of Culture 2013).

The duplication of preservation and access services may be in the future organized on the level of CEESDA (example Data-Pass in the USA). This solution is mentioned as a recommendation in the project report of DASISH (DASISH 2015). Examination of this possibility for sustainable connection of national data services on the European level is part of CEESDA SaW project.

In the future, the statute of the ADP needs to be updated by incorporating the representatives of key groups and institutions of users, as well as other possible participles in the Strategic council of ADP. This would enable wider support of ADP's operation and provide further directions into the improving the services of the ADP (Measure: revision of the Statute and the expansion of the founders' framework).

*Curator of the field: Head of Digital Preservation*

## **2.7 Roles and responsibilities**

Due to the smallness of the organization, the roles in the ADP are shared between personnel. If needed, some of the responsibilities are assigned to external contractors. The basic structure, roles, tasks and responsibilities in the ADP are:

### **Head of Organization**

Responsible for following the mission of the organization in the field of long-term digital preservation as the basic goal of the organization, assuring financial and organizational viability, preparing the strategic plan, supervising daily activities, making decisions and coordinating all activities of the organization (from international cooperation to establishing relations with designated communities, ingest, preparation and access to research data).

### **Head of Acquisition and Ingest**

Responsible for the recording of new studies and providing for an appropriate responsiveness in communication with data providers as well as key beneficiaries, coordinating the ingest of research data and additional materials and reporting on the state of ingest and the processing of the studies.

### **Head of Digital Preservation**

Responsible for monitoring the management of all of the servers, used in storage and distribution of research data, including the webpage of the ADP. Simultaneously, he/she is responsible for regular implementation of backup copies, updating of the policies and

documentation in connection with long-term storage and acquiring and maintaining certificates (Core Certificate of Trustworthy Data Repositories).

### **Head of Access and Use**

Responsible for recording the publishing of research data and preparation of the statistics of submitted and published studies. He/she is responsible for managing access to research data, including providing guidance to users and informing them through the webpage. He/she registers users of research data and prepares an annual report on the use of data from the archive.

### **Head of Trainings and Promotion**

Responsible for implementing activities in the field of promotion, including the organization of events, informing the users and interested public of important activities of the organization (news, webpage, social media), selecting appropriate internal and external training, organizing preparation and recording the publication of the ADP staff. He/she is especially responsible for the strategic planning and implementing the strategy in the field of acquisition and training of data providers – expanding the data collection to various new fields.

### **System Administrator**

Responsible for managing all of the servers, used in storage and distribution of research data, taking care of regular backup and system recovery.

### **Data Archivist**

Responsible for microdata, accompanying documentation, and metadata processing.

### **Director of Administration**

Responsible for organizing and monitoring of daily administrative tasks, defined by the OAIIS (guidance and control over daily fulfilment of tasks of heads of individual functional units, including the final revision of work results on studies made by the data archivists), and for the preparation of annual reports in cooperation with other departments of the University.

## ***2.7.1 Working bodies of the ADP***

### **Commission for the Acquisition and Evaluation of Studies**

The commission is responsible for making decisions on which studies are suitable for the inclusion in the catalog of the ADP. The role of the Commission is also a review of publicly available studies that could be potentially included in the catalog. After a careful examination of the studies content, the Commission assigns a scientific quality grade for the study that is simultaneously the criteria to obtain scientific points for the researcher/s and is shown in the description of the study in the catalog of ADP.

### **Commission for the Protection of Confidentiality**

The commission is an internal body of ADP that works in the field of data protection and access to microdata. The role of the Commission is to make decisions concerning the need to protect research data, means of protecting data and simultaneously to manage access of users to microdata, intended for secondary use.

### *2.7.2 Working groups of the ADP*

#### **Working Group for the Development of the Webpage**

The role of the working group is to prepare the webpage as the main communication tool of the organization. The preparation includes the graphic design of the webpage, preparation of the website content and logical arrangement of content, taking into consideration the needs of the designated communities.

#### **Working Group for the Development of the Application E-Storage**

The role of the group is to coordinate the preparation of the application for digital storage that will be used for the preparation of study descriptions, storage of data, versioning and distribution of studies. The working groups work in collaboration with external contractors.

#### **Working Group for Official Statistics**

The role of the group is to prepare a working plan for the promotion of the use of microdata from official sources for scientific and educational purposes, in close collaboration with the working group of the Statistical Office of the Republic of Slovenia.

#### **Working Group for Education, Trainings and Promotion**

The role of the working group is to plan, prepare and implement activities of ADP that concern promotion, working with users and training of the staff of the ADP. Its roles include organization of promotional events, the preparation of promotional materials and appropriate communication with designated communities.

#### **Working Group for the Translation of ELSST**

The [European Social Sciences Electronic Thesaurus](#) (ELSST), developed by CESSDA, is a multilingual thesaurus covering data holdings likely to be of particular interest to researchers who work with cross-national social surveys. The role of this group is to translate ELSST from English to Slovenian.

### *2.7.3 Competences and development of the staff*

Despite the relative small size of the organization, the ADP has sufficient number of staff to smoothly and efficiently fulfill its mission and working tasks in the framework of the requirements of the Core certificate for Trustworthy Data Repositories. The staff of the ADP is also included in on-going training and projects on the national and international level (CESSDA), with which appropriate knowledge and professional experiences from the field of digital data storage and data management practices that follow current international standards and practices is assured. Each year the ADP prepares an annual plan of training of its staff, where, according to the needs, a plan of professional education and development of employees is made.

The ADP's staff needs to reach certain requirements concerning their knowledge and qualifications to be able to work in the archive. The defined knowledge and qualifications are the basis for new staff acquisition as well as the basis for the organization of additional training of the staff. The ADP's staff needs to measure up to the following qualifications:

- knowledge of digital preservation in general,
- expertise in specific formats (see [Chapter 3.1.2](#)),
- basic (or advanced) IT and statistical qualifications (depending on the work role),

- communication and organizational skills for cooperation between internal functional entities and between external bodies and individuals from which/whom objects are received, as well as users and external service providers,
- organizational and management skills for overall planning (strategy, resources) and the coordination of the different functional entities.

#### ***2.7.4 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.

*Curator of the filed: Director of Administration*

## 3. DIGITAL OBJECT MANAGEMENT

### 3.1 Obtaining holdings and selecting data for ingest

#### *3.1.1 Evaluating the quality of studies and adhering to the criteria of ingest*

The first function, supported by the OAIS reference model, is ingest. Within this function, the data provider submits research data and other documentation to the archive, where they are, in accordance with fixed procedures, prepared for storage. The function of ingest is an external interface between a data provider and the archive and defines the entire process from ingest to the actual storage of materials.

In the first step, the data providers may [record their study](#) through the webpage of the ADP (where there is also [detailed instruction on ingest](#) available). With this action, the data provider informs the ADP of their willingness to deposit their study in the archive.

The ADP revises the application and evaluates it on the basis of whether or not the submitted study/research data is in accordance with [the quality criteria for ingest](#). If the study is suitable for ingest, the Head of Acquisition and Ingest names a Data Archivist, who will process the study. The Data Archivist calls the data provider to prepare all necessary materials for ingest. The data provider can get professional support from the Data Archivist during this process and can use various guides and [additional recommendations for the preparation of data](#). The data provider can also take part in occasional [free training](#) offered by the ADP in the form of workshops.

In the selection of research data for ingest, the ADP takes into consideration the basic quality criteria and the substantive suitability and attractiveness of the study for secondary analyses. The following criteria are used, when selecting the research data for ingest:

- the richness of the data in terms of relevance of the conceptualization and thematic complementarity to the ADP collection,
- the excellence of the used methodology, integrity and relevance of the research data and additional documentation for secondary analyses,
- the copyright of the data provider over the research data and his/her willingness to deposit data in the archive.

Received materials are firstly subjected to thorough examination, where the focus is put on the completeness of the documentation, substantive suitability of the study, examination of the anonymization of the materials and agreement of the submitted data formats with the recommended formats for ingest (see [Recommended and other forms of formats of individual materials for ingest](#)).

Based on this information, the Commission for Ingest prepares a selection of studies and makes an evaluation of the suitability of research data for ingest. Head of Acquisition and Ingest is in charge of convening the Commission. Besides the Head, the members of the Commission are also the Head of the Archive, the Data Archivist in charge of the study and one other staff members of the ADP.

In the process of the selection, the received studies are [assigned a category](#), according to their relevance:

1. studies outside the geographic or thematic framework of the archive,
2. occasional studies of low quality,

3. non-unique studies of limited theoretical or practical relevance,
4. studies of the limited conceptual extent and methodological sophistication,
5. pilot studies that create or improve the instrument and conceptualisation of a new area,
6. methodological and substantive excellent studies without many implications for a wide range of problems,
7. studies that permit theoretical generalizations or relates to a practical problem, less influential,
8. theoretically or practically important studies, studies that fill the research gap or have many implications for a wide range of practical problems, long-term scholarly value,
9. highest range, comparative or continuous research, influential populations, with methodological excellence.

Submitted studies that fall in the categories 7, 8 or 9 counts as a scientific publication based on the criteria by the Slovenian Research Agency ARRS ([Rule book on the procedure to \(co\)finance, evaluate and monitor the performance of the research activities, Appendix 1](#)).

The Head of Acquisition and Ingest makes minutes on the evaluation of the study and prepares a report that includes the reasoning of the assigned category. The report becomes the basic component of the study documentation.

*Curator of the field: Head of Acquisition and Ingest*

### **3.1.2 Recommended formats**

On the webpage of the ADP, the data provider can find the information on the [recommended formats for quantitative studies](#) that are independent of software and/or computer platforms. At the ingest of research data and other materials the ADP examines the submission formats and if needed, transforms them into formats, appropriate for digital preservation and access. In cases where there is no software available for the transformation of the formats, the data provider is requested to submit materials in other more appropriate formats.

#### **Quantitative data**

For the main group of digital objects, intended for target users, that is for data files of quantitative data, the ADP follows a fixed procedure of ingest of research data in one of the before-determined proprietary formats and its transformation into a non-proprietary format, which is persistent under conditions of long-term storage. The data files are transformed into the CSV format in utf-8, the information on the structure and content of the data file is synchronized in the CSV data file with the DDI record on the level of data and file description. The tool Nesstar Publisher, used for this transformation, is tested to assure the preservation of significant properties with the goal to preserve at least the same possibility of performance of the data analysis as was given by the data provider.

#### **Qualitative data**

In the case of qualitative data, the ADP follows the recommendations on the appropriate formats for long-term storage by [UK DataService](#).

In the future, the ADP aspires to enable access to qualitative studies, therefore it actively follows the area (Štebe et al. 2011) and responds to the developments in the community of users, concerning the creation and the use of various types and formats of qualitative data. In

standardization and integration with other information services, the ADP uses [Guidelines for Ingest, Long-Term Preservation, and Access to Cultural Heritage in a Digital Form](#) (document of the Ministry of Culture 2013) and documented best practices of similar organizations, where standards and typical data formats of multimedia documents, are defined.

The ADP formed detailed rules of the procedure of anonymization of qualitative studies that are thoroughly described in [Chapter 3.2.2](#).

#### **Remained challenges**

The updates of the procedure of ingest need to be made: what is a data unit (an interview, diary, article, story, research day, photograph etc.), on the basis of which data files are defined, as well as metadata and the process of anonymization.

A definition of what a study description should contain has to be done: questionnaire, interviewer instructions, informed consent, documents describing data files etc. – all information that will enable the user to independently understand the research data.

An update of the rules of the procedure of anonymization of qualitative data is needed, where direct and indirect identifiers, specific for this field, need to be determined.

Regulation of access to more sensitive data (where necessary and needed) and a separate regime of safe storage need to be determined.

*Curator of the field: Head of Digital Preservation*

## **3.2 Ingest**

### ***3.2.1 Submission information package (SIP)***

The ADP stimulates cooperation with data providers from the planning of the study until the ingest. It is of vital importance that data creators are familiar with the [basic legal and ethical demands](#) that need to be fulfilled, especially concerning the protection of personal data and copyrights (McGeever et al. 2015). Through the organization of occasional training (see the program of previous training), information on the [webpage of ADP](#) and a special handbook [Preparing research data for open access, Guide for data producers](#),<sup>7</sup> the ADP informs data providers on important steps in the data lifecycle (Štebe et al. 2015). This handbook, together with other guidelines, includes an explanation of data licensing, guidelines for respecting ethical norms and especially recommendations on getting consent from the research subjects during the data collection phase. The ADP promotes a sample of a research consent form that includes an informed and explicit consent to submit and allow access to research data through an authorized data center.

In the phase of ingest of research data, the ADP resolves with the data providers all possible questions, connected with the study, such as ethical dilemmas, confidentiality and anonymization of research data, copyrights. This includes rights of the archive and its successor to ingest data, to subject them to digital curation, to preserve and allow access to them under specified condition of access – for example possible embargo, special treatment of sensitive data and/or special restrictions of commercial use of research data. On this basis, the archive and data provider sign an agreement in the form of a [License agreement](#). All information on rights and conditions of access is written in the administrative metadata, which regulates the further processes of preservation and access regimes.

By signing the agreement, the data providers confirm to have rights to dispose the research data, assure that care was made to enable confidentiality of personal data, and define licenses, under which the research data and additional documentation can be distributed to users. The

---

<sup>7</sup> Available in Slovenian and English (digital and print versions).

data provider can also define possible exceptions concerning access. If not, the research data and other documentation are accessible to users under Creative Commons 4.0 licenses (see [Chapter 3.5.1](#) and the section on [Access](#)).

In exceptional cases, for example, when we are dealing with older data without accessible information on the holder of the copyright, the ADP assess the risks of unresolved copyright in case of distributing data and if it resolves that there is a negligible risk, it may decide to provide access to such data under a CCBY license. Access to such data is accompanied by information that the ADP may withdraw this study in case it receives a warning from the side of the holder of the copyright (see an [example](#)).

The ADP allows the data provider to safely and swiftly submit materials through the ADP Cloud (see [Instructions for ingest](#)). In accordance with the information model of an OAIS-type archive, the research data and other additional documentation, submitted by the data provider, compose the submission information package (SIP). This package consists of:

- filled-in [Study Description form](#),
- appropriately edited and documented data file (in accordance with the [Recommendations on how to edit data files](#)),
- original questionnaire (if available),
- [other materials](#) that help understand the research data, such as list of codes, codebook, frequencies, instructions for interviewers, data on the execution of the survey, copies of publications and other supporting documents that were part of the data collection or are relevant to their understanding,
- filled-in and signed [License Agreement](#) together with a list of submitted materials (2 copies).

It is recommended (but not obligatory) that the data provider adds to the submission package also research reports and other accompanying publications that could be beneficial in the secondary use of research data.

The submitted materials are reviewed by the Data Archivist and all the possible questions concerning content and formats are discussed. If needed, the Data Archivist calls upon the data provider to complete the submitted materials in order to assure the completeness of the study. The Data Archivist prepares all necessary metadata to register individual units of materials, connected in the ingest package. He/she also imports descriptive metadata for study description, according to DDI, and structural metadata that enables understandability and maximal usability for future users. The materials are then transformed to recommended formats for long-term storage and access. The distributed materials, intended for final users, are saved separately from archival materials.

The ADP stores also all physically submitted printed materials and materials on different media.

#### **Remaining challenges**

Automatic antivirus and format checking will need to be introduced. The ADP will need to define the activities of following the needs of new formats (that occur considering new types of data) and will need to find their equivalent long-term formats. By doing so, the possibility of completing the procedures on the basis of the Guidelines for Ingest, Long-Term Storage, and Access to Cultural Heritage in a Digital Form (Krstulović et al. 2013) considering various types of materials, will need to be studied.

It is expected that in the future years the number of requests to archive qualitative studies will rise. One of the future developmental needs is thus training of working with qualitative studies and the standardization of the procedure to archive qualitative research data.

*Curator of the field: Head of Aquisition and Ingest*

### ***3.2.2 Protection of confidentiality and anonymization***

The *Commission for the Protection of Confidentiality* of the ADP deals with the field of the protection of confidentiality. The Commission meets only in cases when in the process of ingest an additional need to protect research data is identified. The Commission has two main tasks: (1) to make decisions on the protection of submitted data before their distribution, (2) to deal with requests from researchers for access to less protected or unprotected microdata.

The protection of data is primarily the role of the data provider, where the ADP can only offer support. The ADP can advise data providers on research data management planning, including the protection of confidentiality and restrictions of access to research data. In extreme circumstances, if the data provider asks for such an intervention, the ADP can independently perform the confidentiality protection of research data under its fixed procedures. The liability for the appropriate protection of confidentiality lies on the side of the data provider and does not transfer to the ADP, regardless of who made the initial protection of the data.

#### **Methods of implementing protection of confidentiality**

The ADP carries out the protection of confidentiality in two ways, firstly with the protection of data and secondly by managing access to various types of data for various types of users (see [Chapter 3.5](#)). The data provider prepares all the different versions of microdata, which have different levels of protection for different levels of access with the support of ADP.

#### **Protection of data**

The ADP advises data providers on the protection of data. Simultaneously two types of protection are used: anonymization of microdata (quantitative and qualitative) and anonymization of descriptive statistics (frequencies of categorical variables and statistics for numerical variables).

##### **a) Anonymization of microdata**

Anonymization is the basis for the protection of confidentiality. The ADP uses the following principles of anonymization of microdata, wherein a distinction is made concerning the type of microdata:

#### **Quantitative microdata**

There may be no direct identifiers (for example names, telephone numbers, and addresses) in distributed data. In cases of public access to microdata, indirect identifiers need to be protected, if needed, in order to prevent identification of a unit in the data. Different methods and techniques may be used: aggregation of the values of variables, deletion of variables, deletion of values in all units and deletion of values in certain units. The fact that the user may identify a unit by using publicly accessible databases (for example administrative sources on the Internet) are taken into consideration. A special emphasis is put on the protection of sensitive microdata, for example, microdata that has companies or organizations as units of analysis, medical microdata, microdata with extremely sensitive research topics (mobing, sexual harassment etc.). In cases, where such data are submitted to the archive, the Commission for the Protection of Confidentiality needs to convene and decide whether or not to include such study in the catalog of the ADP.

## **Answers to open survey questions**

Answers to open survey questions need to be appropriately coded so that they can be distributed within a unified microdata file. The data provider must normally make the coding, and the ADP may only offer support. In cases where the coding was not done correctly, the answers to open ended survey questions are deleted from the distributed data file or they are separately distributed in an aggregate form.

## **Qualitative microdata**

There may be no direct identifiers in the distributed data. Direct and indirect identifiers need to be replaced by pseudonyms or surrogate terms that do not enable identification of individuals who were included in the study (for example in-depth interviews, focus groups). The data in the qualitative form (for example transcripts) must be edited by the data provider, and the ADP may only offer support. The main principle that is followed in protecting the qualitative microdata is that it is meaningful to preserve the highest usability of research data, therefore it is advised to avoid overprotecting or deletion of information that may be important for further analyses.

### **b) Anonymization of descriptive statistics**

The ADP uses an internal program for the protection of descriptive statistic, which, if necessary, protects the univariate aggregated data in the description of the study. Because these statistics are freely available on the webpage of ADP, without the need for users to register in order to use them, they need to be protected in certain cases. The tool is used mostly in protecting aggregated data of register data, and enables simple automatic protection of data with the use of different methods of data protection, such as recode, protection of the lowest and the highest values, deletion of variables (numeric information), protection of descriptive statistics of numerical variables and protection of frequencies of variables, where the rule of the lowest allowed shown frequency is applied.

### ***3.2.3 Metadata standards and interoperability***

The ADP follows the OAIS model and the requirements of the Core Certificate of Trustworthy Data Repositories by enabling its designated communities sufficient amount of information on the study (metadata) that enable final users to independently understand individual parts of the data file or additional materials. By doing so, the ADP fulfills one of the mandatory responsibilities of OAIS since it regularly updates and improves the metadata of studies, according to the identified needs of users.

Metadata are the main components of all versions of information packages in the system and are, together with preservation metadata, the basics for the appropriate semantic and structural use of data and additional documentation. 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 ([PaSC](#), [DataVerse](#) etc.).

The ADP's staff cooperates in the projects of harmonization of the use of metadata in the dedicated project group in CESSDA, which includes debating on the fixed use of terminology and developing methods to adjust the metadata to the technological advancements.

#### **Remaining challenges**

The ADP regularly follows the trends in the field of new services and identifiers. On the list of monitoring initiatives are the following: CONOR, COBISS, PREMIS. Currently, the ADP is developing a program for repository control of the holdings description, which will contain FOXML and accompanied identifiers, including the information on checksum (MD5), traceability, accessibility of different versions etc.

In the next period the ADP anticipates the assurance of conformance with requirements of various external services of merging and displaying information, such as [OpenAire](#), [DataCite](#), [Data Catalogue Vocabulary Application Profile](#) (DCAT - AP), [Digital Library of Slovenia - dLib](#), [National Open Science Portal](#), and integration with the existing and developing activities of other networks of scientific information and open access.

*Curator of the field: Head of Digital Preservation*

### **3.3 Archival storage of data**

The main mission of the second OAIS function is to assure long-term preservation, which means to save data in appropriate formats for long-term preservation and on appropriate locations. To assure this requirement, regular monitoring and updating of the procedure of digital preservation are needed (updating of media, migration of formats etc.).

The ADP monitors its holdings, and, in cases of changes in formats for long-term storage, adapts them to its needs. The ADP makes sure as well that all materials and metadata are machines readable on the long-run. The function of archival storage includes many security mechanisms, such as checking for errors in the information package, evaluation of the preparation of materials for long-term storage, as well as a policy of dealing with disposal of items from holdings.

Archival information package (AIP) consists of all submitted and accessible digital objects, which were transformed into formats that are appropriate for long-term storage. Together with the accompanying metadata, they are stored in a separate location from the dissemination information package (DIP). Backup copies of all of the packages are made on a regular basis, by which greater protection before the loss of data is assured (see [Chapter 4.2](#)).

*Curator of the field: Head of Digital Preservation*

### **3.4 Management of (meta)data**

The third function of the OAIS model is Data Management function. The Data Management function maintains databases of metadata identifying and describing the archival holding. The primary functions of data management also include performing queries on these databases and generating reports in response to requests from other functional entities within the OAIS (for example from ingest, administration, access) and performing updates to the databases as new information arrives, or existing information is modified or deleted.

The ADP stores administrative metadata in different databases that are necessary for reporting on new materials, users, data providers etc. There is:

- **a database of recorded studies**, where all the basic metadata of every study are recorded (for example address, author, submitted materials etc.);
- **a database of users**, in which all data on users and their assigned access levels are provided;
- **a database of license agreements**, where all the basic information of the agreements is given: ID of the study, number of the agreement, the name of the data file, URN, date of signature, data provider;
- **a database of access regimes to studies**, in which a list of all accessible studies is given, together with the assigned level of access (ACU).

With a regular updating of above-mentioned databases of administrative data and by assuring appropriately updated metadata to final users, the ADP enables long-term independent understanding and usability of research data/deposited studies. The ADP's staff is [adequately trained](#) to manage and add all technical changes to the holdings and to reach the quality of metadata, in accordance with the international standards and best practices, which in turn enables sufficient information in the form of quality metadata, available to final users. With such functioning, the ADP reaches the requirements of the Core Certificate of Trustworthy Data Repositories.

### ***3.4.1 Versioning and transformation management***

The ADP developed a system of monitoring changes that enable traceability of the transformation of digital objects. Two types of changes are possible that is: smaller changes, which are run as *revisions*, and larger transformations of objects, which are run as *versions*. The revision and version of the document are visible in the document name.

All materials, regardless of their revision or version, are stored in a common directory, their security copies are held at various locations (see [Chapter 4.2](#)).

#### **a) Versioning and noting down changes in studies and accompanying documentation**

Changes are most carefully noted down at the level of data files so that in each step it is possible to trace them and get to the original version of the data file. The changes in the study descriptions, data files, and additional materials are noted down with versioning in naming individual files. The explication of the changes is written down in Jira.<sup>8</sup> For data files, all changes are noted also in the SPSS syntax.

The system of naming the documents is the following:

[IDstudy\\_type of material\\_ language\\_ version\\_ revision.ending](#)  
(for example *SJM13\_vp1\_sl\_v1\_r0.pdf*)

---

<sup>8</sup> Software used to note down the state and changes of individual studies.

## **b) Versioning and noting down changes in internal documents**

Changes are noted down also in internal documents. This is done in a way that the name of the documents is given a version and a revision. When smaller changes are made, a new revision of the document emerges, when a bigger change is made, a new version of the document is made.

The system of naming the documents is the following:

[name\\_version\\_revision.ending](#)  
(for example *navodila\_v1\_r2.pdf*)

## **c) Versioning and noting down changes on the webpage of ADP**

A Subversion SVN client is used to track changes on the webpage of the ADP. All the changes on the webpage of the ADP are noted down and at the same time, there is a possibility to revive at every step the previous version of the webpage.

### ***3.4.2 Policy of disposal of items from the holdings***

For now, the ADP has not come across a case, where one of the materials would need to be disposed of. In the future, it may occur that some documents, whose lifespan will end, will need to be disposed. This destruction needs to be carried out through a decision by a Commission, which means that the commission will review the outdated item and order destruction. The revision track of destruction, as well as the existence of the document, need to be noted down.

#### **Remaining challenges**

The current system of access does not allow users to access different versions of studies. By developing the new software Fedora, this will be enabled.

In the future, a policy of conduct in cases of disposal of items from holdings (from any reasons) needs to be established.

*Curator of the field: Data Archivist*

## **3.5 Access to data (DIP)**

Assuring access to holdings is the fifth function in the OAIS model. The function deals with user applications (designated communities) and allows access to contents based on the applications and rights of individual users. The function is responsible also for technological surveillance over managing access, according to the user rights.

Metadata and other materials, connected with the study, are available to all users of the [webpage ADP](#) without registration. Access to microdata is subjected to prior [registration](#). By registering, users obtain the possibility to make online analyses on [Nesstar browser](#) and/or the possibility to download data files in the selected format on their personal computers.

Users may access the following study metadata on the webpage of ADP:

### **a) Metadata of studies with microdata available at ADP**

The ADP holds metadata together with microdata that is accessible through ADP. The user needs to agree with the rules of the ADP in order to use such microdata.

**b) Metadata of studies with microdata available in other organizations**

The ADP prepares also metadata of studies that have microdata available in other archives or organizations, such as some internationally comparable data and unprotected microdata from official statistics. The ADP saves and disseminates their metadata and links to the data access landing page. The access to data files is subjected to the rules of the organization that is responsible for the preservation and dissemination of the study.

### ***3.5.1 Licenses, limitations and types of data access***

At the point of submission of the materials to the ADP, the data providers, by signing the License Agreement, confirm to have rights to deal with the research data, they assure confidentiality of research subjects and define licences under which research data and other documentation will be distributed to users, including possible limitations of access. Research data and documentation is available to users under [Creative Commons 4.0 license](#). ADP uses the following licenses:

**CC0** – without restriction

**CCBY** – authorization only

**CCBYNC** – authorization + non-commercial

Three types of users are distinguished by ADP: **registered researchers**, who may access also less protected microdata, public users such as **students**, who may access most of the microdata in the catalog of the ADP, and **commercial users**, who may access only a limited amount of microdata that is not distributed under the CCBYNC license. The user is informed about possible restrictions or exceptions concerning access, given by the data provider at the point of ingest, through the study description on the webpage of the ADP (more on access and terms of use of individual studies can be read on the [webpage under Access](#)).

### ***3.5.2 Management of access to microdata***

Access to most microdata in the catalog of ADP is openly available to users, however, a [registration](#) is needed on the webpage of ADP:

Anyone can register to access microdata, regardless of their purpose of use or employment status. At registration on the webpage the user needs to provide his/her name, status (student, researchers, employed in various institutions, personal use etc.), define the purpose of the use of data (educational, scientific, public, commercial) and agree with the conditions of usage, which among others demand respect of [professional disciplinary and ethical codes of conduct](#) together with [full citation of the author of the research data and the archive](#). The information on the status is demanded in order to keep track of the nature of ADP's users, so as to form future data services (preparation of different versions of microdata, the organization of training etc.).

Access to microdata is regulated by the form [Registration to access data](#) on the webpage ADP.

In the case when **microdata is sensitive or unprotected**, they are accessible only to a **limited** range of users. [Possible types of access](#) are two: access of portable media or access in a secure environment (office of ADP).

The process to obtain access to less protected microdata is the following:

1. the user registers on the webpage of ADP,
2. the user's fill in the form Access to microdata,
3. the *Commission for the Protection of Confidentiality* conveys and inspects the application,
4. on the basis of the application, the commission evaluates the appropriateness of the user/research group to access the microdata,
5. main criteria for decision are the status of the user, previous research results, the appropriate reasoning of the reasons to access unprotected/sensitive microdata and appropriate definition of publication of research results,
6. the user is given the decision on access and a type and condition of access is determined,
7. the user and the ADP sign a contract, in which the user agrees to respect all determined conditions of access and ethical and legal obligations concerning the access to microdata.

### ***3.5.3 Protection of confidentiality and access to data***

Protection of confidentiality can be managed with different restrictions of access to microdata, depending on the level of protection of microdata and the type of the user. If needed, the data provider, with the support of the ADP, prepares different versions of the microdata with different levels of protection for different levels of access. The types of microdata differentiate in details of the microdata itself, which in practice means greater or lesser sets of variables (for examples detailed geographical information), sets of values of variables (for example levels of aggregation) and availability of values of variables (for example subsamples, hidden values).

There are four types of accessing microdata in ADP: access to microdata in Nesstar, accessing protected microdata (PUF – Public Use Files), access to less protected microdata (SUF – Scientific Use Files) and access to unprotected sensitive microdata (ScUF – Secure Use Files).

#### **Access to microdata on Nesstar browser**

Access to microdata on Nesstar browser is intended for registered users, where dealing with data does not demand the transfer of individual microdata files on local computers. These users make statistical analyses on [Nesstar browser](#), where they can use only aggregate data as a result of desired statistical analyses, and not microdata.

#### **Access to PUF microdata**

Access to microdata is intended for registered users, who conduct analyses of microdata on local computers, using selected statistical software. PUF microdata does not contain detailed data, as they have been protected or they primarily do not contain (combination of) variables that needed to be protected.

#### **Access to SUF microdata**

Access to such microdata is available only for registered users that need access to microdata that is less protected than PUF versions of microdata. Access is possible only in cases when the user makes a special application for access that is evaluated by the ADP's *Commission for the Protection of Confidentiality*. SUF microdata file is available only for specific studies, for example in cases where a protected PUF file was prepared for general distribution.

## Access to ScUF microdata

Access is possible only for registered users on the basis of approved application for access. Access to such sensitive unprotected data is possible only in a safe environment in the offices of the ADP, where there is no Internet access and impossible to transfer data on portable media. It is forbidden to use widgets for making screenshots or videos. Each application is individually dealt with by the *Commission for the Protection of Confidentiality*, which carefully examines the application and determines the possibilities of access.

### 3.5.4 Citation of holdings

The ADP provides a standardized format to cite research data and additional materials from the catalog, which includes a persistent link to access landing page of research data and materials. The goal is to instruct the users about the requirement to include appropriate and permanent citation format that enables traceability. Each study in the catalog of ADP is accompanied with the information on proper citation (see [How to cite this study?](#)). By registering, the users oblige to properly cite the used materials (see [General Provisions and Terms of Use](#)). The ADP calls researchers to inform the ADP of all possible publication on the basis of the data used from the catalog of ADP.

#### Remaining challenges

The sanctioning of improper use of data has not been necessary so far. In the future, ADP will, if needed, develop a system for monitoring and a procedure in case of sanctioning the improper use of data files, in accordance with ethical codes of conduct and active consultation with ethical commissions.

In the future, the ADP will develop new ways to monitor citations in the environment of bibliographic information services (for example in the framework of DataCite model).

*Curator of the field: Head of Access and Use*

## 3.6 Administration

The Administration function of the OAIS model covers managing daily tasks of the OAIS model and the coordination of its individual parts. The ADP performs administrative tasks in the following areas:

### a) Ingest

At the level of ingest, administrative work is tied mostly to the communication between the data provider and the curator of the study and managing the documentation for ingest. The following tasks are included:

- Newly reported study gets recorded in Jira and a caretaker of the study is assigned, who takes over the entire processing of the study.
- A list of potential studies that serves as a basis for acquiring new studies is made.
- All necessary information for ingest are sent to the data provider (the procedure, what and how to submit materials).
- A *Commission for Ingest and Evaluation of the Study* is convened that determines whether or not to deposit the study in the archive. Minutes of the meeting are being made that are saved in Jira and are accessible to the data provider on demand.
- Curator of the study makes sure that the submitted package is complete and that all the materials are appropriately edited.

- Curator of the study makes sure that the License Agreement is filled-in and signed by both parties. The curator scans the agreement and notes it in the database and in Jira. The physical copy of the agreement is saved in the folder in the office of the ADP. It needs to be assured that all possible specific demands from the data provider are noted down in Jira.
- The curator makes sure that the entire communication with the data provider is documented in Jira.
- Information on the authorization of the study description by the data provider is noted down in Jira, including all possible subsequent changes in the study description.

### **b) Archiving**

It is important that all the materials are in formats that are appropriate for long-term storage and that there is a regular notation of all the changes on individual materials. Administrative tasks at archiving are:

- For every study archival versions of submitted and distributed materials are made. The log is made in the database of Microsoft Access and in Jira.
- All possible subsequent changes of archival materials are noted in the database of Microsoft Access and in Jira.

### **c) Managing data (processing)**

It is important to regularly monitor and note all changes in dealing with data. To assure the traceability of transformations the entire process of changes is noted in Jira, where there is information on:

- the entire process of dealing with the study,
- the list to review studies, research data, and additional materials.

Internal instructions of dealing with data are stored in the local ADP directory.

### **d) Storage planning**

In order to assure long-term preservation, it is necessary to regularly monitor and update the system of archiving. Administration of this area involves the following tasks:

- caring for the agreements with external service providers (NUK, ARNES, server maintenance and IT support, programmer): a caretaker of the agreement is assigned, who monitors the content of the agreement (validity, payments, regular reporting),
- appropriate and safe storage of backup copies (one copy on the local drive, one physical copy in the folder in the office of ADP),
- regular monitoring of making the backup copies and noting all of the activities in Jira.

### **e) Access**

The phase of access includes the following administrative tasks:

- helping users at registration,
- processing and storing of registration forms (printed versions stored in folders in the offices of ADP, digital versions stored on the local server),
- managing lists of access to individual studies,
- granting access to users, in accordance with filled-in registration forms.

## **Additionally, administrative tasks are carried out in the following fields as well:**

### **f) Interaction with data providers**

Interaction with data providers in the administrative respect involves organization of training and direct support (e-mail, telephone, personal contact, handbooks, webinars). The data providers need to be motivated and requested for ingest, a negotiation needs to be made with them concerning the accessibility of data and additional training need to be offered, where they will be informed about the importance and the process of ingest.

Meetings with the representatives of institutions are being organized in order to expand the support environment that will enable prompt ingest of quality research data in the frameworks of the policies of institutions and scientific journals. The plan of working with data providers for the purpose of thematic and quality expansion of the collection of ADP is included into the document of promotion of ADP and is prepared at the beginning of each calendar year for the current year.

### **g) Interaction with users**

Head of Access and Use offers support to users through e-mail, telephone or personally. In addition, several group and individual trainings are organized, within which the users are trained to search and use data from the ADP's data catalog. The plan of working with users is included in the document of the promotional activities of ADP in the field of education, training, and promotion that the Head of Trainings and Promotion prepares at the beginning of the calendar year for the current year and is negotiated and confirmed within the working group.

In this plan the following activities for users are defined:

- workshops on the use of data from the catalog of ADP,
- lectures for users on the use of data from the catalog of ADP on their institutions (for example on faculties),
- workshops on the use of official statistics.

Users are regularly informed about new studies in the ADP's data catalog. Promotion takes place through different digital channels:

- column »News« on the webpage of ADP (<http://www.adp.fdv.uni-lj.si/novice/>),
- [ADP's monthly newsletter](http://www.adp.fdv.uni-lj.si/enovice/) that is sent to subscribed users (<http://www.adp.fdv.uni-lj.si/enovice/>),
- Facebook page of ADP (<http://sl-si.facebook.com/Arhiv.Druzboslovnih.Podatkov>),
- Twitter account of ADP (<https://twitter.com/ArhivPodatkov>),
- webpage of the Faculty of Social Sciences (FDV), where the ADP has its own subpage (<http://www.fdv.uni-lj.si/raziskovanje/publikacije/arhiv-druzboslovnih-podatkov>),
- internal news of FDV that are weekly sent to all employees of FDV,
- Facebook page of FDV (<https://www.facebook.com/fdv.si>),
- Twitter account of FDV (<https://twitter.com/FDVLjubljana>).

### **h) Management/Maintenance**

The ADP takes care that its functioning takes place according to accepted standards and policies. The administrative tasks in this area are to monitor the functioning of the entire ADP system and to determine appropriate updates of procedures.

*Curator of the field: Director of Administration*

## 4. TECHNOLOGY

### 4.1 Architecture of the information technology

The ADP is functioning on a well supported operating system and other infrastructural program foundations that are suitable for the data services that the ADP offers to its designated communities (see the internal document [Scheme of the ADP system](#)). For its activities, the ADP uses the network of the [Academic and Research Network of Slovenia ARNES](#) that assures a stable, secure and efficient functioning of the information-communication infrastructure (see [General provisions of the use of ARNES network services](#)). The Arnes assures that it manages the e-infrastructure, networking resources and services carefully and in accordance with the best practices and technological standards and thus with the best effort attends to the needs of its users.

In forming its information technology, the ADP takes into consideration the recommendations and requirements of CESSDA (for example recommendations regarding PID and AAI). It follows current trends in the development of hardware and software, which are regularly monitored and updated. Efficient capacities of storage of the entire infrastructure are provided for. The preservation of data, its distribution and the entire infrastructure of the archive is based on an adapted IT infrastructure, where only employees and registered users have access.

The ADP is currently developing technological upgrade to support internal processes and enable their automatization that is based on a repository software with certain adaptations that take into consideration the specificities of data, documentation, and metadata from the field of social sciences. The ADP regularly invests in upgrades and developments of services that will enable interconnectivity between different organizations and will enable the use of versioning and persistent identifiers, an overview of the copies in different formats and regimes of access. The ADP will also continue cooperation with the NUK and the ARNES regarding the work on safe digital storage on different locations.

One of the ADP's continuous activities is its cooperation in developing tools in the international environment, their testing, and implementation. In addition to Nesstar and online environments with contemporary documentary systems (Django, Wordpress, Jira) that the ADP is regularly maintaining, the ADP is planning to introduce technological support for researchers that will enable controlled handling of research data throughout the duration of the project and will facilitate the transfer of data, gathered within the project, to digital preservation in ADP after the end of the project.

Stated goals of long-term digital storage will be reached with a combination of introducing open access tools, such as [DataVerse](#), and by adapting and the opening of internally developed repositories tools to manage processes in the archive (application based on Fedora). These tools will serve as external support in preparing data and documentation and will enable the rule based machine processing of ingest.

### 4.2 Security and risk management

The ADP is committed to the security of its physical spaces, equipment, research data and other materials, its services and users. There are three physical rooms, where all the computer equipment and materials are stored. The rooms are part of the Faculty of Social Sciences of the University of Ljubljana, which under its own protocols attends for the security of its

premises, such as fire safety and physical security of premises, including the permanent presence of a security guard.

The ADP follows security instructions and rules of the Faculty of Social Sciences in cases of different natural disasters:

- *Rules on the Measures to Protect the Information-Communication Systems at the Faculty of Social Sciences (2010),*
- *House Rules of the Faculty of Social Sciences,*
- *Fire Regulations of the Faculty of Social Sciences.*

To assure security, all rooms with computers and materials need to be locked if there is no staff member present. Two laptops are available in ADP, for the safety of which their users are responsible. The ADP assures secure access to its servers that are available only to the employees of ADP and their external service providers. Physical access to servers are enabled only to the staff of the Computer Center of the Faculty of Social Sciences (RC FDV), external service providers with a valid contract and the staff of ADP, but only in the presence of a representative of the RC FDV. Access to servers at ARNES is available only for the staff of ARNES.

Access to the system and materials is enabled only to users and registered users. A firewall is in place for the greater security of access. To assure security, there is a limited physical access to the hardware. The transfer of research data between data providers and the ADP is possible through a safe cloud transfer. All of these measures assure appropriate information security of the ADP.

The network of ARNES that the ADP uses in its functioning is under the management of ARNES. ARNES has in place established mechanisms of automatic control and a control center of technical assistance, whereas, outside the office hours, issues in the working of the network are eliminated by the emergency services. With these measures, ARNES assures the reliability, quality and security of their services. ARNES is trying to establish redundancy in all of its network services, which diminishes the possibilities of fallouts. In the case of unpredictable events, ARNES assures to eliminate issues in the shortest reasonable time.

For assuring physical storage, the ADP developed a strategy and thorough instructions of the process of making backup copies and the recovery of the system in cases of fallout (see field [Technology and Security in Appendix A](#)). Software solutions enable automatic synchronization of backup copies on a predetermined backup server, where data files are saved in encrypted versions. Encrypted data files are then stored in three different locations: NAS server, external iSCI disc ([Arnes Storage](#) – Technological park of Ljubljana) that copies through SSH, and on an external hard disk. This way backup copies are made in the case of severe events in the server room (for example destructive fire). Backup copy outside the location of the archive on the other side enables backup copies in cases of greater disruptive events at the Faculty of Social Sciences (destructive fire, bomb, earthquake etc.).

#### **Remaining challenges**

More thorough procedures of long-term storage and handling of materials are in the process of preparation and will be fully implemented in 2017. For these purposes, the ADP is developing a new Fedora based software environment in cooperation with NUK, which will enable safer and more automatized method of digital preservation of materials. In the new system, one copy of data and additional materials will be saved on the server of the ADP, and one copy will be saved at NUK.

Currently, the ADP is introducing new open access software to follow some of the administrative tasks and working processes. After its implementation, some of the procedures in the policy will need to be updated.

The ADP currently does not have a defined standard to follow in formulating its technological infrastructure. The ADP follows requirements, defined in [ISO 16363/TDR](#), and tries to, as possible, put its functioning within the framework of these requirements. In the future, it would be sensible to examine this standard more carefully and adapt the technological infrastructure accordingly. The ADP as well does not have a systematic analysis of possible security threats or a risk management plan. In the future, a policy of acknowledging and managing of risks should be prepared, following the standards requirements.

*Curator of the field: System Administrator*

## TERMINOLOGY

**ADP:** Slovenian Social Science Data Archive (<http://www.adp.fdv.uni-lj.si/>)

**AIP, Archival Information Package:** the state of digital objects in the archive after storage in the preservation system.

**Authenticity:** a digital object is authentic if it can be said to rely on three significant provable properties: that the object is what it purports to be, that it was created by whomsoever is purported to have been created by; and that it was created at the time when it is purported to have been created.

**CESSDA AS:** Consortium of European Social Science Data Archives (<http://cessda.net/>)

**CONOR:** is a [normative data file](#) of personal and corporate names in the system COBISS.SI. The normative control is based on the connectivity between the bibliographic and normative records.

**DANS:** Data Archiving and Networked Services (<https://dans.knaw.nl/en/>)

**DARIAH:** Digital Research Infrastructure for the Arts and Humanities (<https://www.dariah.eu/>)

**Data database:** a relational database with basic data on originally received data files and their variations, including the identification number, type of data file and basic thematic characteristics (data file or other connected materials).

**Data description:** the description of the data file with data in accordance with DDI (position in the data file and the markings of variables and units), together with a detailed description of the content of variables (conceptual meaning of the variable, set of values with codes) that define the thematic and conceptual level of data file that is under digital preservation and part of DDI codebook file.

**Data Documentation Initiative (DDI):** is a project of the social science community to establish an international standard and methodology for describing the content, presentation, transport, and preservation of metadata about datasets in the social and behavioral sciences. (<http://www.ddialliance.org/>).

**Data producer:** is an individual who is named on a deposit agreement as having sufficient responsibility to grant particular rights to the Archive on behalf of a data collection. The depositor may be the investigator, creator or the copyright owner of a data collection, but does not have to be. In some terminology, the term depositor is used in a similar meaning.

**Data users:** those persons, or client systems, that interact with OAIS services to find and acquire preserved information of interest. A special class of Consumers is the Designated Community. The Designated Community is the set of Consumers who should be able to understand the preserved information.

**Data:** constitute primary sources that underpin scientific research and enable derivation of theoretical or applied findings. More precise definitions of data vary according to context. Quantitative data may refer to just the matrices of numbers or words that comprise a data file but may also cover other information (metadata) held within a statistical package data file, such as variable labels, code labels, and missing value definitions. Qualitative data might include interview transcripts as well as audio and video recordings (analog or digital).

**DataCite:** DataCite is a leading global non-profit organization that provides persistent identifiers (DOIs) for research data (<https://www.datacite.org/>).

**Data file description:** data file description of quantitative data includes at least information on the number of variables and units in the datafile and their corresponding data file format: part of the DDI codebook file.

**Datafile:** a digital object that is the core unit of preservation and distribution in the ADP. The data are symbolic representations of reality, acquired with the data collection process.

**DataVerse:** is an open source web application to share, preserve, cite, explore, and analyze research data (<http://dataverse.org/>).

**DDI codebook file:** digital object, formatted according to the DDI2.1 DTD standard in the XML language that includes all descriptive data, data on provenance and other metadata, which are important for understanding and reproducing data files. It includes the marking of identifiers and metadata of digital objects that form the collection of data materials (<http://www.ddialliance.org/specification>).

**Designated Community:** an identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities. A Designated Community is defined by the Archive and this definition may change over time.

**Digital data curation:** Data curation is the selection, preservation, maintenance, and archiving of digital assets and it establishes, maintains and adds value to data for present and future use.

**Digital Object:** An object composed of a set of bit sequences.

**Digital preservation:** is a series of managed activities necessary to ensure enduring access to authentic versions of the content of digital materials for as long as necessary.

**DIP, Dissemination Information Package:** the state of digital objects as available to final users in different periods.

**DSA:** Data Seal of Approval (<http://datasealofapproval.org/en/>)

**Fedora:** is the flexible, modular, open source repository platform with native linked data support.

**Independently Understandable:** A characteristic of information that is sufficiently complete to allow it to be interpreted, understood and used by the Designated Community without having to resort to special resources not widely available, including named individuals.

**Information Object:** A Data Object together with its Representation Information.

**Information Package:** A logical container composed of optional Content Information and optional associated Preservation Description Information. Associated with this Information Package is Packaging Information used to delimit and identify the Content Information and Package Description information used to facilitate searches for the Content Information.

**Integrity:** refers to its completeness and to a continued state of un-alteration of a digital object.

**Jira:** software used to note down the state and changes of individual studies as well as other administrative tasks.

**Long-Term Preservation:** The act of maintaining information, Independently Understandable by a Designated Community, and with evidence supporting its Authenticity, over the Long Term.

**Metadata:** information that describes significant aspects (e.g. content, context and structure of information) of a resource.

**Nesstar:** Nesstar WebView is a web-based system for the dissemination of data (<http://www.nesstar.com/>).

**Normalisation at ingest:** file format conversion by the repository when data are submitted.

**OAIS:** The Open Archival Information System (OAIS) Reference Model is a conceptual framework for an archival system dedicated to preserving and maintaining access to digital information. It addresses a full range of archival preservation functions including ingest, archival storage, data management, access, and dissemination. It is not a metadata standard but rather it outlines a taxonomy that defines the information types deemed necessary for the understanding of digital content over an indefinite period of time (<https://public.ccsds.org/pubs/650x0m2.pdf>).

**OECD:** Organisation for Economic Co-operation and Development (<http://www.oecd.org/>)

**OpenAire:** The FP7 project OpenAIRE aimed to support the implementation of the EC and ERC Open Access policies. Its successor OpenAIREplus was aimed at linking the aggregated research publications to the accompanying research and project information, datasets and author information. Open access to scientific peer reviewed publications has evolved from a pilot project with limited scope in FP7 to an underlying principle in the Horizon 2020 funding scheme, obligatory for all H2020 funded projects. The

goal is to make as much European funded research output as possible available to all, via the OpenAIRE portal (<https://www.openaire.eu/>).

**Other materials, connected with the research data:** data files of textual and other materials, which are connected with the research data, such as questionnaire, external code etc.

**Other materials:** textual and other data file documents that are important for the additional understanding of the content and context of the study, such as preliminary reports, publications etc. They are usually not part of ADP's digital preservation but are accessible through a link on the primary point of access.

**Preservation strategy:** a digital preservation strategy is a particular technical approach to the preservation of digital materials. This document contains the strategy and policy of the Archive.

**Reliability:** relies on having trusted and dependable contents of a digital object.

**Semantic Information:** The Representation Information that further describes the meaning beyond that provided by the Structure Information.

**Significant properties:** are those elements of a digital object which need to be preserved in order for it to be used by the designated user community. They will almost always include information content and a level of functionality, but might also include formatting and look and feel.

**SIP, Submission Information Package:** the state of digital objects in the archive after ingest.

**Structure Information:** The Representation Information that imparts meaning about how other information is organized. For example, it maps bit streams to common computer types such as characters, numbers, and pixels and aggregations of those types such as character strings and arrays.

**Study database:** a relational database with basic metadata on the level of study that includes the unique study code, title of the study, data provider identification and the date of signing the license agreement.

**Study:** is the Archive's basic content-tracking concept used for a data collection during the stages of acquisition, ingest, preservation and dissemination. There is normally a one-to-one relationship between a study and a data collection. Similarly, there is often a one-to-one relationship between a deposit and a study, but a deposit may, depending on circumstances, be divided into more than one study. Equally, more than one deposit from the same depositor may be combined into a single study, similar to accumulating archival fonds.

**Succession Plan:** The plan of how and when the management, ownership and/or control of the OAIS holdings will be transferred to a subsequent OAIS in order to ensure the continued effective preservation of those holdings.

**Transformation:** A Digital Migration in which there is an alteration to the Content Information or PDI of an Archival Information Package. For example, changing ASCII codes to UNICODE in a text document being preserved is a Transformation.

**UK DA:** UK Data Service (<http://www.data-archive.ac.uk/>)

**Usability:** relies on the ability of a digital object to be located, retrieved, presented and interpreted.

## RESOURCES

- Arhiv družboslovnih podatkov. 2017. *3. Prepare Data*. Accessed: <http://www.adp.fdv.uni-lj.si/eng/deli/postopek/priprava/> (03 June 2017).
- Arhiv družboslovnih podatkov. 2017. *Deposit Procedure*. Accessed: <http://adp.fdv.uni-lj.si/eng/deli/postopek/> (03 June 2017).
- Arhiv družboslovnih podatkov. 2017. *2. Sign a License Agreement*. Accessed: <http://adp.fdv.uni-lj.si/eng/deli/postopek/izjava/> (03 June 2017).
- Arhiv družboslovnih podatkov. 2017. *4. Prepare the Study Description*. Accessed: [http://adp.fdv.uni-lj.si/eng/deli/postopek/opis\\_raziskave/](http://adp.fdv.uni-lj.si/eng/deli/postopek/opis_raziskave/) (03 June 2017).
- ARRS. 2011. *PRILOGA 2: Bibliografska merila znanstvene in strokovne uspešnosti*. Accessed: <http://www.arrs.gov.si/sl/akti/inc/prav-raz-dej-jan11-priloga2-1.pdf> (03 June 2017).
- Beedham, Hilary, Julie Missen, Matt Palmer, in Ruusalepp Raivo. 2007. *Assessment of UKDA and TNA compliance with OAIS and mets standards*. Accessed: <http://www.webarchive.org.uk/wayback/archive/20140615012529/http://www.jisc.ac.uk/media/documents/programmes/preservation/oaismets.pdf> (16 March 2016).
- CESSDA AS. 2016. *The Open Archival Information System (OAIS) Reference Model*. Accessed: [http://cessda.net/content/download/496/4465/file/CESSDA%20User%20Guide%20for%20digital%20preservation\\_2\\_OAIS.pdf](http://cessda.net/content/download/496/4465/file/CESSDA%20User%20Guide%20for%20digital%20preservation_2_OAIS.pdf) (17 March 2016).
- CESSDA. 2014. *Statutes for CESSDA Annexes: Consortium of European Social Science Data Archives*. Accessed: [https://www.google.si/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwjRgpWd-4LAhXGPxQKHYYHyCmIQFggbMAA&url=http%3A%2F%2Fcessda.net%2Fcontent%2Fdownload%2F316%2F2908%2Ffile%2FAnnexes-to-Statutes-for-CESSDA-210213-Final-Version-brand.pdf&usq=AFOjCNEgpbI7A\\_S9-13vebFbPpWjDZSI9g&sig2=hCjwzpNxNBjHtG252zHm7Q](https://www.google.si/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwjRgpWd-4LAhXGPxQKHYYHyCmIQFggbMAA&url=http%3A%2F%2Fcessda.net%2Fcontent%2Fdownload%2F316%2F2908%2Ffile%2FAnnexes-to-Statutes-for-CESSDA-210213-Final-Version-brand.pdf&usq=AFOjCNEgpbI7A_S9-13vebFbPpWjDZSI9g&sig2=hCjwzpNxNBjHtG252zHm7Q) (24 February 2016).
- DASISH. 2015. *Data Service Infrastructure for the Social Sciences and Humanities*. Accessed: [http://dasish.eu/publications/projectreports/DASISH\\_D4.3\\_081214-final.pdf](http://dasish.eu/publications/projectreports/DASISH_D4.3_081214-final.pdf) (25 February 2016).
- European Commission Research & Innovation. 2016. *Policy Initiatives*. Accessed: <http://ec.europa.eu/research/openscience/index.cfm?pg=openaccess> (18 April 2016).
- *Fakulteta za družbene vede*. Accessed: <http://www.fdv.uni-lj.si/> (3 March 2016).
- Interuniversity Consortium for Political and Social Research (ICPSR). 2009. *Principles and Good Practice for Preserving Data. IHSN Working Paper No 003. December 2009*. Accessed: <http://www.ihsn.org/home/sites/default/files/resources/IHSN-WP003.pdf> (24 February 2016).
- Jones, Sarah. 2010a. *DCC preservation policy template*. Accessed: <http://www.dcc.ac.uk/sites/default/files/documents/Preservation%20policy%20template.pdf> (4 March 2016).
- Jones, Sarah. 2010b. *Preservation policy template for repositories*. Accessed: <http://www.dcc.ac.uk/sites/default/files/documents/Preservation%20policy%20template.pdf> (4 March 2016).
- Kavčič-Čolić, Alenka, Melita Ambrožič, Janko Klasinc, Eva Kodrič-Dačić, Matjaž Kragelj, Zoran Krstulović, Aljoša Nikl, Irena Sešek in Renata Šolar. 2012. *Strategija trajnega ohranjanja digitalnih virov v Narodni in univerzitetni knjižnici 2012-2020*. 2012. Ljubljana: Narodna in univerzitetna knjižnica. [URN:NBN:SI:DOC-HPEWEXEN](http://www.nuk-uz.si/URN:NBN:SI:DOC-HPEWEXEN)
- Krstulović, Zoran, Tatjana Hajtnik in Mitja Doma. 2013. *Smernice za zajem, dolgotrajno ohranjanje in dostop do kulturne dediščine v digitalni obliki*. Ljubljana: Ministrstvo za kulturo.

- Lavoie, Brian. 2014. *The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition) DPC Technology Watch Report 14-02 October 2014*. Great Britain: Charles Beagrie.Ltd. DOI: <http://dx.doi.org/10.7207/twr14-02>
- Lavoie, Brian. 2014. *The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition)*. Accessed: <http://dx.doi.org/10.7207/twr14-02> (24 February 2016).
- McGeever, Mags, Angus Whyte in Laura Molloy. 2015. *Five Things You Need to Know About RDM and the Law: DCC Checklist on Legal Aspects of RDM*. Accessed: <http://www.dcc.ac.uk/resources/how-guides/rdm-law#sthash.IMdJCLdo.ygq01jsc.dpuf> (3 March 2016).
- Nestor. 2014. *Guideline for Preservation Planning: Procedural Model and Implementation*. Accessed: <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:0008-2014102709> (3 March 2016).
- OECD. 2007. *OECD Principles and Guidelines for Access to Research Data from Public Funding*. Accessed: <http://www.oecd.org/sti/sci-tech/oecdprinciplesandguidelinesforaccesstoresearchdatafrompublicfunding.htm> (25 February 2016).
- OpenAIRE. 2016. Dostopno preko: <https://www.openaire.eu/> (4 March 2016).
- OpenAIRE. 2016. *OpenAIRE Guidelines for Data Archives*. Accessed: <https://guidelines.openaire.eu/en/latest/data/index.html> (4 March 2016).
- PIS. 2016. *Print preview Zakon o državni statistiki*. Accessed: <http://www.pisrs.si/Pis.web/pregledPredpisa?id=ZAKO424> (4 March 2016).
- PIS. 2016. *Zakon o avtorski in sorodnih pravicah (ZASP)*. Accessed: <http://www.pisrs.si/Pis.web/pregledPredpisa?id=ZAKO403> (4 March 2016).
- PIS. 2016. *Zakon o Informacijskem pooblaščenju*. Accessed: <http://www.pisrs.si/Pis.web/pregledPredpisa?id=ZAKO4498> (4 March 2016).
- PIS. 2016. *Zakon o varstvu osebnih podatkov (ZVOP)*. Accessed: <http://pisrs.si/Pis.web/pregledPredpisa?id=ZAKO1213> (4 March 2016).
- Schumann, Natascha and Astrid Recker. 2012. *De-mystifying OAIS compliance: Benefits and challenges of mapping the OAIS reference model to the GESIS Data Archive*. Accessed: <http://www.iassistdata.org/iq/de-mystifying-oais-compliance-benefits-and-challenges-mapping-oais-reference-model-gegis-data-arc> (24 February 2016).
- Slovenian Research Agency. 2017. *Slovenian Research Agency (ARRS)*. Accessed: <https://www.arrs.gov.si/en/index.asp> (3 March 2016).
- Slovenian Research Agency. *Zakoni, predpisi, normativni akti*. Accessed: <http://www.arrs.gov.si/sl/akti/prav-sof-ocen-sprem-razisk-dej-nov-13-spr.asp> (3 March 2016).
- Štebe, Janez, Jože Hudales and Boris Kragelj. 2011. *Archiving and Re-using Qualitative and Quantitative Longitudinal Data in Slovenia*. Accessed: <http://www.iassistdata.org/iq/archiving-and-re-using-qualitative-and-quantitative-longitudinal-data-slovenia> (15 April 2016).
- Štebe, Janez, Sonja Bezjak in Irena Vipavc Brvar. 2015. *Priprava raziskovalnih podatkov za odprti dostop: priročnik za raziskovalce*. Ljubljana: Fakulteta za družbene vede, Založba FDV. [URN:NBN:SI:DOC-06SLBVXX](http://nbn-resolving.org/urn:nbn:si:doc-06slbvxx)
- Štebe, Janez, Sonja Bezjak in Sanja Lužar. 2013. *Akcijski načrt za vzpostavitev sistema odprtega dostopa do raziskovalnih podatkov financiranih z javnimi sredstvi*. Accessed: [http://www.adp.fdv.uni-lj.si/o\\_arhivu/publikacije/odpp10\\_akcijski\\_nactr/](http://www.adp.fdv.uni-lj.si/o_arhivu/publikacije/odpp10_akcijski_nactr/) (24 February 2016).
- Univerza v Ljubljani. 2017. *Network of Research and Infrastructural Centres*. Accessed: [https://www.uni-lj.si/research\\_and\\_development/research\\_infrastructure/network\\_of\\_research\\_and\\_infrastructural\\_centres/](https://www.uni-lj.si/research_and_development/research_infrastructure/network_of_research_and_infrastructural_centres/) (03 June 2017).

- Vardigan, Mary in Cole Whiteman. 2007. *ICPSR Meets OAIS: Applying the OAIS Reference Model to the Social Science Archive Context*. Accessed: <http://hdl.handle.net/2027.42/60440> (24 February 2016).
- Vlada republike Slovenije. 2015. *Nacionalna strategija odprtega dostopa do znanstvenih objav in raziskovalnih podatkov v Sloveniji 2015-2020*. Accessed: [http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Zakonodaja/Strategije/Nacionalna\\_strategija\\_odprtega\\_dostopa.pdf](http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Zakonodaja/Strategije/Nacionalna_strategija_odprtega_dostopa.pdf) (24 February 2016).
- Vlada republike Slovenije. 2016. *Načrt razvoja raziskovalnih infrastruktur 2011–2020. Revizija 2016*. Accessed: [http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Strategije/NRRI\\_2016\\_SLO.pdf](http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Strategije/NRRI_2016_SLO.pdf) (23 December 2016).
- Zakon o dostopu do informacij javnega značaja. Accessed: <http://www.uradni-list.si/1/objava.jsp?urlid=200651&stevilka=2180> (3 March 2016).
- Zakon o varstvu dokumentarnega in arhivskega gradiva ter arhivih. Accessed: <https://www.uradni-list.si/1/content?id=72425> (3 March 2016).

# APPENDIX

## A. List of internal guidelines and instructions

### Workflow of handling a new study

- Procedure of working with study (workflow\_v1.pdf)

### Obtaining and selecting data for ingest

- Instructions for submitting data for ADP (Navodila za predajo podatkov-za ADP\_V1.3.docx)
- Instructions for submitting data for data producers (Navodila za predajo podatkov-za dajalca\_V1\_2.docx)
- Workflow for the protection of data (Workflow za zaščito podatkov.pdf)
- Functioning of the Commission for the protection of confidentiality (Komisija za zaščito zaupnosti – opredelitev\_v2.docx)
- Main criteria for evaluating studies and marking of scientific successfulness (Ocene raziskav - definicije\_v3.docx)
- Instructions for checking and storing of holdings (Navodila za pregledovanje in spravilo gradiv\_V5\_R6.docx)
- Internal instructions for the protection of microdata (Interna navodila za zaščito mikropodatkov.docx)

### Ingest

- Instructions on how to add new studies in Jira (Uvoz raziskave v Jira.docx)
- The procedure for entering a study (Navodila\_raziskave\_JIRA\_v3.docx)
- Additional recommendations for editing the data file (DodatnaPriporocilaSPSS-V2.docx)
- Instructions for the program of correcting the Access database (Navodila za program popravkov Accessove baze.docx)
- Instructions for the creation of CDB (navodila\_za\_izdelavo\_cdb\_14\_5\_02.doc)
- Instructions for database export (Navodila za izvoz baze raziskav.docx)
- Instructions for database users' export (Navodila za izvoz baze uporabnikov.docx)
- Instructions for publishing on Nesstar (v 4.0.9) (Navodila za objavo na Nesstarju.docx)
- Instructions for publishing on web (Navodila za objavo na Web-u\_v2.docx)
- Instructions for the use of programme sintaxe.exe to build SPSS, SAS, STATA and R syntax (Navodila za sestavljanje sintakse iz xml in ascii.docx)
- Instructions on how to prepare the final data files (NavodilaSPSS.docx)
- The procedure for preparation of data file and related documentation (Postopek za pripravo podatkovne datoteke\_V2\_R2.docx)
- Instructions on how to process study descriptions in xml format in accordance with DDI DTD standard (PRAVILA ADP1.doc)
- Instructions on how to process study descriptions (pravilaADP.doc)

## **Archival information package**

- Instructions for the programme for the creation of XML for NESSTAR (Navodila za sestavljanje xml za NESSTAR.docx)
- Instructions for the programme for the creation of XML for web (Navodila za sestavljanje xml za WEB.docx)
- Instructions for the use of programme to generate ACU for NESSTAR (Navodila za generiranje ACU.docx)
- Instructions for publishing a study on NESSTAR (Navodila za objavo na Nesstarju\_v2.docx)
- Instructions for publishing a study on webpage ADP (Navodila za objavo na Web-u\_v3.docx)
- Instructions for using the application e-Trajna hramba [eng. e-Long-term curation] (Fedora\_navodila.docx; Navodila za pregledovanje in spravilo gradiv\_V5\_R7.docx)

## **Access**

- Instructions for assigning passwords for access to NESSTAR (Navodilo za dodelitev gesel1.doc)
- Instructions on how to use Nesstar (Nesstar guide.docx)
- Instructions on how to use Nesstar in Slovenian (nesstar navodila\_sl.doc)

## **Administration**

- Instructions on how to use the Admin module for users (Admin modul za uporabnike.docx)
- Instructions on how to use export user database (Izvoz uporabnikov.docx)
- JIRA-instructions for use (JIRA navodila za uporabo.docx)
- Administration of NESSTAR and Jira server (Navodila za administracijo Nesstar in Jira strežnika.docx)
- Instructions for using the programme Nesstar\_logi (Navodila za generiranje poročil o obiskanosti Nesstarja.docx)
- Instructions to generate reports from Jira database (Navodila za generiranje JIRA porocil.docx)
- Instructions to generate reports from web database (Navodila za generiranje porocil iz WEB baze.docx)
- Instructions on how to use import a CSV file to Excel (NAVODILA ZA UVOZ CSV DATOTEKE V EXCEL.docx)
- Diagram of directories on »Linux« server (Shema direktorijev na Linux strezniku.docx)
- Instructions for maintaining the website of ADP (Navodila za urejanje spletne strani\_v2.docx)
- Instructions for adding a new webpage outside of Django (NAVODILA ZA DODAJANJE NOVE SPLETNE STRANI - NOVI STREZNIK.docx)
- Instructions on how to translate website in English (Prevodi na spletni strani.doc)
- Instructions for publishing on the webpage of Faculty of Social Sciences (Objava spletna stran FDV.docx)
- Instructions for administration in Wordpress (Wordpress admin.docx)
- Instructions on how to be present on the social media (Prisotnost na socialnih omrežjih v1\_6.docx)
- Instructions on how to prepare and send e-News (Priprava in posiljanje e-novic.docx)

- Instructions for saving e-News (Navodila za shranjevanje novic v PDF.docx)
- Instructions for using Google Analytics and preparation of monthly reports (Navodila\_google\_analytics.docx)
- Instructions on how to prepare a travel order (Navodila za pripravo POTNEGA NALOGA.doc)
- Instructions on how to set up rights on server directories (Streznik pravice.docx)
- Instructions on how to deal with frequent problems of users (Tezave uporabnikov.docx)
- Prof. Klinar Prize - Instructions (NAGRADE prof Klinarja - navodila.doc)

### **Technology and Security**

- Diagram of the ADP system (skica\_ADP\_v10.pdf)
- Instructions on how to migrate the Nesstar server (Migracija Nesstarja.docx)
- Instructions on how to migrate Nesstar (Nesstar migracija.docx)
- Instructions on how to deal with a non-functioning webpage, Nesstar or Jira (kaj\_narediti\_ko\_spletna\_stran\_ne\_deluje.docx)
- Instructions on how to protect website content with password (Zascita strani z geslom.docx)
- Instructions on how to provide backup (Dokumentacija\_varnostnih\_kopij.docx)
- Instructions on how to inspect backup copies (pregled\_varnostnih\_kopij.docx)