



Research and Innovation Action

CESSDA Strengthening and Widening

Project Number: 674939

Start Date of Project: 01/08/2015

Duration: 27 months

Deliverable 3.2 Country report on development potentials

Dissemination Level	PU
Due Date of Deliverable	31/03/2017
Actual Submission Date	6/07/2017
Work Package	WP 3
Task	T3.2
Type	Report
EC Approval Status	16 November 2017
Version	1.3
Number of Pages	p.1 - p. 364 (+ Annexes)
<p>Abstract: The report identifies gaps for further development of data archives services (DAS) in each country by exploring the wider ecosystem of data sharing culture, organisational setting and service operational profile. Where now the research data infrastructure is only emerging, it identifies promising candidate services.</p>	
<p>The information in this document reflects only the author's views and the European Community is not liable for any use that may be made of the information contained therein. The information in this document is provided "as is" without guarantee or warranty of any kind, express or implied, including but not limited to the fitness of the information for a specific purpose. The user thereof uses the information at his/ her sole risk and liability.</p>	

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



History

Version	Date	Reason	Revised by
0.0	12/10/2016	Template for the country reports	Task leader
0.1	11/11/2016	Introduction and country report template	Task leader
0.2	22/11/2016	Introduction and country report template ready for partners' usage	Subtasks lead and support partners
0.3	20/12/2016	Prepare draft country reports	All
0.4	27/01/2017	Editorial review	Subtasks lead partners
0.5	05/02/2017	Country report author address editorial review comments	All
0.6	15/02/2017	Country report author address comments from consultation with the country contact and submit final country report version	All
1.0	29/03/2017	Final integral version	Task leader
1.1	05/05/2017	WP leader and subtasks lead partners' approval	NSD, FORS
1.2	15/05/2017	Review of proto-activities section and document layout	FORS

Author List

Organisation	Name	Contact Information
ADP	Janez Štebe (Task leader)	janez.stebe@fdv.uni-lj.si
ADP	Ana Slavec	ana.slavec@fdv.uni-lj.si
ADP	Irena Vipavc Brvar	irena.vipavc@fdv.uni-lj.si
ADP	Sonja Bezjak	sonja.bezjak@fdv.uni-lj.si
CSDA	Martin Vávra	martin.vavra@soc.cas.cz
CSDA	Tomáš Čížek	tomas.cizek@soc.cas.cz
CSDA	Yana Leontiyeva	yana.leontiyeva@soc.cas.cz
DANS	Ellen Leenarts	ellen.leenarts@dans.knaw.nl
DANS	Marion Wittenberg	marion.wittenberg@dans.knaw.nl
DDA	Anne Sofie Fink	asf@sa.dk
DDA	Bodil Stenvig	bs@sa.dk
DDA	Christian Lindgaard Olsen	clo@sa.dk
DDA	Christina Guldfeldt Madsen	cgm@sa.dk
FFZG	Alen Vodopijevac	alen@irb.hr
FFZG	Irena Kranjec	ikranjec@ffzg.hr
FFZG	Marijana Glavica	mglavica@ffzg.hr

FORS	Brian Kleiner	brian.kleiner@fors.unil.ch
FORS	Christina Bornatici	christina.bornatici@fors.unil.ch
GESIS	Natascha Schumann	Natascha.Schumann@gesis.org
IEN	Aleksandar Zdravković	aleksandar.zdravkovic@ien.bg.ac.rs
IEN	Aleksandra Bradić-Martinović	abmartinovic@ien.bg.ac.rs
LiDA	Algis Krupavicius	algis.krupavicius@ktu.lt
NSD	Trond Kvamme	trond.kvamme@nsd.uib.no
NSD	Vigdis Kvalheim	vigdis.kvalheim@nsd.no
SND	Dimitar Popovski	Dimitar.Popovski@snd.gu.se
SND	Ilze Lace	ilze.lace@gu.se
SOHDA	Johan Surkyn	jrsurkyn@vub.ac.be
SU SAV	Miloslav Bahna	miloslav.bahna@savba.sk
UKDA	Veerle Van Den Eynden	veerle@essex.ac.uk
UniData	Domingo Scisci	domingo.scisci@unimib.it
UniData	Sonia Stefanizzi	sonia.stefanizzi@unimib.it

Time Schedule before Delivery

Next Action	Deadline	Care of
Draft country reports ready	20/12/2016	All T3.2 partners
Final country reports versions submitted	15/02/2017	All T3.2 partners
1.0 Final integral version	29/03/2017	ADP
1.1 Review by the WP leader	05/05/2017	ADP
1.2 Review by the Chair of the Delivery Committee	26/05/2017	CSDA
1.3 Review by the Project Coordinator	04/07/2017	CESSDA AS
Approval and Submission by the Project Coordinator to the European Commission	05/07/2017	CESSDA AS

Executive Summary

The CESSDA SaW (Strengthening and widening) project TASK 3.2 - Audit of the current status of data archive services in each ERA country - is to test and prove the CESSDA Excellence at the level of individual service providers (SP) by examining the barriers to and the potential value and benefits from membership in CESSDA in each European country. It does so by inquiring into the situation in the country regarding support toward organising a sustainable research data service on a country level. The results will serve as a starting point for further internal SP's development planning, to inform external stakeholders on necessary steps to be taken to achieve a national data service sustainability, and to promote CESSDA membership. Currently, CESSDA membership consists of 15 national members and 1 member with observer status. The aim is to achieve full European coverage, to strengthen the network and to ensure sustainability of data services for the widened network.

This document consists of a series of country reports, which will be later on followed up by a synthesis final integrated audit report. The approach that we took was mapping the current state of Data Archive Service (DAS) in each ERA country and reflecting on the conditions that affect their position. An audit of data services was conducted, based on a comprehensive data collection instrument. Multiple sources of information were utilised: via self-assessment, as reported by DAS representatives; from desk research of reports resulting from previous surveys or case studies in the research community projects (CESSDA PPP, SERSCIDA, DASISH etc.); or from country and scientific community representatives' expert interviews.

A DAS does not exist or come into existence in isolation. Therefore, we briefly map the wider ecosystem, which imposes the conditions for the functioning of a DAS, but also the established DAS can have reciprocal impact on the data sharing culture and science system in general. The potential for future activities in existing human resources, technological infrastructures and support services (libraries, research institutes, and research information services) were areas addressed in the DAS proto-activities part of the assessments of the countries with no existing data infrastructure. The second part of the audit was oriented towards the description of the current state of existing social sciences data archives and services in Europe. Even there where the national DAS are formally established, the diversity of situations can be expected and described, nevertheless depending on the specific institutional and organisational setting.

Abbreviations and Acronyms

AAI	Authentication and Authorization Infrastructure
ADP	Arhiv Družboslovnih Podatkov
ADS	Polish social data archive
AuSSDA	The Austrian Social Science Data Archive
CeMI	Centar za monitoring i istraživanje
CIS	Centro de Investigaciones Sociológicas
CNRS	Centre Nationale de la Recherche Scientifique (Progedo)
CPA	Capability Process Area
CPC	Centre for Political Courage
CRA	Capability Requirements Area
ČSDA	Czech Social Science Data Archive
DANS	Data Archiving and Networked Services
DAS	Data Archive Service
DDA	Danish National Archive - Danish Data Archive
DDI	Data Documentation Initiative
DCC	Digital Curation Centre
DMP	Data Management Plan
DOI	Digital Object Identifier
DRAMBORA	Digital Repository Audit Method Based on Risk Assessment
DSA	Data Seal of Approval
EKKE	Ethniko Kentro Koinonikon Erevnon
ESSDA	Estonian Social Science Data Archive
FFZG	Filozofski fakultet Sveučilišta u Zagrebu
FORS	Swiss Foundation for Research in Social Sciences
FSD	Finnish Social Science Data Archive
GERD	Gross Domestic Expenditure on Research and Development
GESIS	Leibniz Institute for the Social Sciences
GDP	Gross Domestic Product
ICS-ULisboa	Instituto de Ciencias Sociais da Universidade de Lisboa
IDM	Institute for democracy and mediation
IEN	Institut Ekonomskih Nauka
IPR	Intellectual Property Right
ISDC	Israel Social Sciences Data Centre

ISPJR	Saints Cyril and Methodius University, Institute for Sociological Political and Juridical Research
JESDA	Joint Economic and Social Data Archive
KIIS	Kiev International Institute of Sociology
LiDA	Lithuanian Data Archive for Humanities and Social Sciences
LISER	Luxembourg Institute of Socio-Economic Research
LSZDA	Latvian Databank of Social Sciences
NSD	Norwegian Centre for Research Data
OAIS	Open Archival Information System
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
PID	Persistent Identifier
Plato	Planets preservation planning component
RDM	Research Data Management
RODA	Asociatia Arhiva Romana de Date Sociale (Romanian Social Data Archive)
So.Da.Net	Greek research infrastructure for the social sciences
SOHDA	Social Sciences and Humanities Data Archive
SU-SAV	Sociologicky Ustav Slovenskej Akademie Vied
TÁRKI	TÁRKI Alapítvány (TARKI Foundation)
UCD ISSDA	University College Dublin - Irish Social Science Data Archive
UGOT-SND	University of Gothenburg - Swedish National Data Service
UKDA	UK Data Archive
UniData	Università degli studi di Milano - Bicocca Data Archive
WOS	Web of Science

Table of Contents

1. INTRODUCTION	9
1.1 ABOUT THE CESSDA SAW PROJECT AND ROLE OF THE COUNTRY REPORTS	10
1.2 PREVIOUS STUDIES	11
1.3 METHODS USED FOR COLLECTING INFORMATION	12
1.4 SCOPE OF COUNTRIES AND SOURCES OF INFORMATION	12
2. STRUCTURE AND CONTENT OF THE COUNTRY REPORTS.....	15
2.1 BROADER ECOSYSTEM OF DAS OPERATION	15
2.2 DAS PROTO-ACTIVITIES	20
2.3 ORGANISATIONAL PROFILE	21
2.4 CAPABILITY REQUIREMENT AREAS OF DAS	21
3. COUNTRY REPORTS.....	24
3.1 ALBANIA	24
3.2 AUSTRIA.....	33
3.3 BELARUS	38
3.4 BELGIUM.....	42
3.5 BOSNIA AND HERZEGOVINA	51
3.6 BULGARIA.....	58
3.7 CROATIA	63
3.8 CYPRUS.....	72
3.9 CZECH REPUBLIC.....	78
3.10 DENMARK	88
3.11 ESTONIA.....	96
3.12 FAROE ISLANDS.....	106
3.13 FINLAND.....	106
3.14 FRANCE.....	116
3.15 GERMANY	125
3.16 GREECE.....	135
3.17 HUNGARY.....	149
3.18 ICELAND.....	157
3.19 IRELAND.....	157
3.20 ISRAEL	167
3.21 ITALY	174
3.22 KOSOVO.....	185
3.23 LATVIA.....	193
3.24 LITHUANIA.....	204
3.25 LUXEMBOURG	216
3.26 FORMER YUGOSLAV REPUBLIC OF MACEDONIA (FYRM)	219
3.27 MALTA	227
3.28 MOLDOVA.....	227
3.29 MONTENEGRO	231
3.30 THE NETHERLANDS.....	238
3.31 NORWAY.....	251
3.32 POLAND	265
3.33 PORTUGAL.....	274
3.34 ROMANIA	283
3.35 RUSSIA.....	291
3.36 SERBIA.....	298

3.37	SLOVAK REPUBLIC	306
3.38	SLOVENIA	312
3.39	SPAIN	323
3.40	SWEDEN.....	326
3.41	SWITZERLAND	333
3.42	TURKEY	345
3.43	UKRAINE	345
3.44	UNITED KINGDOM.....	350
APPENDIX 1: LIST OF TABLES AND FIGURES		365
APPENDIX 2: GUIDELINES AND COMMUNICATION PROTOCOL FOR INTERVIEWERS		368
APPENDIX 3: QUESTIONNAIRE TEXT EXTRACT FROM THE WEB FORM		387
APPENDIX 4: CLASSIFICATION PROTOCOLS FOR ASSIGNING DEVELOPMENT LEVEL.....		411
APPENDIX 5: GLOSSARY		420

1. Introduction

The intention of SaW TASK 3.2 - Audit of current status of data archive services in each ERA country was, firstly to obtain results that could inform external stakeholders on perspectives and necessary steps to be taken in order to achieve certain chosen level of maturity, and secondly to use results for internal planning of existing and future service providers in order to improve the service of the community and to build trust.

First part of the audit was oriented towards data sharing culture and conditions, in which data sharing culture can flourish. Parts of conditions that affect research data infrastructure are financial and institutional statuses of social sciences in each country. In some of the countries, we can find excellence in the development of social science that is supported with robust and multifaceted research data service. What we often encounter at the other end is a syndrome of underdevelopment, where lack of funds affects every other aspect of science system, including the infrastructure. Where social sciences have low budget in general, there are usually also poor conditions for data infrastructure. But in that case, the impact of gradually establishing a robust data infrastructure can have even greater impact on building data sharing culture and improving excellence and efficiency in research in general. If data is shared widely, such immediate effect can be provided by improved quality control and transparency of the research. Individual research projects can be less expensive, if the existing high quality data resources can be findable, assessed, accessed and reused for secondary analysis. Thus, in the following report the individual countries are presented, varied as they appear through the audit methodology developed for this project.

Focus for the countries that do not have national Data Archive Service (DAS) was put on exploring the conditions for establishing a data service that could in the future obtain the role of CESSDA national service provider. These conditions are partly contained in the previous areas addressed, such as financial, structural conditions of social sciences sectors, scientific policy requirements and norms established in science community. These external stakeholders play their role in influencing the existence of DAS and constitute the data sharing culture environment.

There are internal stakeholders that are capable and willing to play a role in establishment of new future services, bring current services to a higher maturity level, or if keep it running on a high excellence if it reaches that level already. For the countries that does not have the running the DAS, it is essential to ground the establishment of its services primarily on internal resources, which means to find the potential for future activities in existing human resources, technological infrastructures and support services (libraries, research institutes, and research information services). These areas were addressed in the DAS proto-activities part of the country assessments, which are the focus of country reports of the countries with no existing data infrastructure.

The second part of the audit was oriented on the description of current state of existing social sciences data archives services in Europe. Even there where the national DAS are formally established, the diversity of situations can be expected and described, nevertheless depending on the specific institutional and organisational setting.

The results can be exploited to:

- Provide base for comparisons across countries - those more mature serving as an example for others,
- Enable informing and support in other CESSDA activities and tasks (staff training, technical support, policy level development, advocacy etc.)
- Use results for promotion of CESSDA membership and SP's sustainability on a national level

Our aim, from a substantive point of view, is to show that there already are quite a lot of communalities in SP's approaches, that CESSDA collectively already shows high levels of excellence, and is continuously striving to improve its services, relevance and impact to the social science community in Europe. Hopefully, this will assure the policy makers that money invested in CESSDA national data archive service providers is well spent, and to raise confidence and awareness among our users about the organisation of sustainable service, oriented toward the fulfilment of the needs of the users' community related to provision of research data.

1.1 About the CESSDA SaW project and role of the country reports

The CESSDA SaW (Strengthening and widening) project examines the barriers to and the potential value and benefits from membership in CESSDA in each European country. Overall aim of SaW TASK 3.2 is to test the CESSDA Excellence both at the level of individual service providers (SP) and on the level of the network. The approaches used by the SaW TASK 3.2 is a bottom-up analysis of current prospects in European Research Area (ERA) countries of becoming full CESSDA members. The findings will serve as a starting point for further development planning, and the results will be used to inform external stakeholders on necessary steps to be taken to achieve a national data service sustainability, and to promote CESSDA membership. Thus, a top-down approach will be developed, which will provide advice and support for initialising and further developing of National Data Archive Service organisations from ERA.

The approach that we took was mapping the current state of Data Archive Service (DAS) in each ERA country and reflecting on the conditions that affect their position. This document consists of a series of country reports, which will be later followed up by a synthesis, a final integrated audit report. In this report, we focus on each of the countries individually, where we stress a specific configuration of factors that affect the development and sustainability of DAS, with an aim to support development planning.

DAS does not exist or comes into existence in isolation. Therefore, we briefly map the wider ecosystem, which imposes the conditions for functioning of DAS, but also the established DAS can have reciprocal impact on data sharing culture and science system in general¹.

The aim of CESSDA is to facilitate and promote the results of social science research, re-use of research data, and thereby to support national and international research and cooperation. In this regard, strengthening and widening of CESSDA are of key importance to achieve this aim. CESSDA achieves its goals collectively through its members facilitating access to important European social science research data resources, regardless of the location of either researcher or data. That is, to enable excellence in social science research and interdisciplinary research based on data resources of high quality, irrespective of national borders.

Currently, CESSDA membership consists of 15 national members and 1 member with observer status. The aim is to achieve full European coverage, to strengthen the network and to ensure sustainability of data services for the widened network.

¹ See reports from Recode project explication what ideally is the role of each of stakeholders. Available at <http://recodeproject.eu/>

1.2 Previous studies

The two key words in the CESSDA SaW project, Strengthening and Widening, refer closely to the two key work package reports of the CESSDA PPP project, which end in 2010:

- WP6: Strengthening the CESSDA RI²
- WP7: Widening CESSDA: Inclusiveness and Comprehensiveness of the upgraded European research infrastructure³

Both topics lay in the heart of past, current and future CESSDA organisation and activities. The widening part built upon the experience that continued since the UNESCO Workshop on Social Science Data Archives in Eastern Europe in 2002. A group of more than 10 countries initiatives to establish the national DAS were engaged in informal cooperation under the EDAN - the East European Data Archives Network and coordinated by the GESIS Leibniz-Institute for Social Sciences. Since the CESSDA PPP, another similar incentive to establish national and regional networks of data archive initiatives came under the SERSCIDA project⁴, and a follow up of this, the SEEDS⁵ project, which both focused on Western Balkan Countries. All the projects and initiatives mentioned provided some overview on the conditions and capacities of different organisations, residing in the national contexts. The results of those projects were heavily used in task 3.2., as first source of reference for contacts, and using the results for desk research information gathering activities. Many-detailed information was here utilized in the new framework and updated with additional evidence from other sources, and reflecting on new developments since first presented.

Similar line of continuous activities follows the Strengthening direction, where in PPP there were already collected some first results that compared some rudimentary capability assessments of different DAS. Similar activities follow into the DASISH project (The Data Service Infrastructure for the Social Sciences and Humanities)⁶, which conducts the survey among at existing and developing DAS in Europe, covering the Humanities and Social Sciences ESFRI infrastructures, and extend it with a selection of interviews. Both provide a baseline for the SaW task 3.2 web form, as related to the organisational capacities evaluation. A comprehensive overview about the national open research data policies was first provided in the IFDO Report from 2014⁷.

We explored other sources of information, e.g. regional, international, specialised reports, such as OECD, EU Commission, etc. Chuck Humphrey's analysis of the profiles and organisational settings that collected information of DAS worldwide⁸, aiming at the proposal how to establish national DAS in Canada, started with the similar assumption as we do in task 3.2: that a comprehensive set of conditions need to be explored in each national setting, that helps to shape further development steps. That is, there is no development model that fits for all.

² Available at https://ppp.cessda.net/doc/WP6_Final_Report.pdf

³ Available at https://ppp.cessda.net/doc/D7.1_Report_and_Recommendations.pdf

⁴ Support for Establishment of National/Regional Social Sciences Data Archives. Reports available at: <http://www.serscida.eu/en/>

⁵ South-Eastern European Data Services, available at: <http://seedsproject.ch/>

⁶ D4.3 - List of Recommended Deposit Services for SSH. Available at DASISH http://dasish.eu/publications/projectreports/DASISH_D4.3_081214-final.pdf

⁷ Vigdis Kvalheim and Trond Kvamme (2014). Policies for Sharing Research Data in Social Sciences and Humanities. A survey about research funders' data policies. Available at http://ifdo.org/wordpress/wp-content/uploads/2015/07/ifdo_survey_report.pdf

⁸ International Models of Data Archiving Services <http://slideplayer.com/slide/5039424/>

1.3 Methods used for collecting information

An audit of data services was conducted, based on a comprehensive data collection instrument, where information about countries *Broader ecosystem influencing DAS operation* and *Capability requirement areas of DAS* organisations, were centralised in a systematic and structured way. Multiple sources of information were utilised: via self-assessment, as reported by SP/DAS representatives; from desk research of reports resulting from previous surveys or case studies in the research community; or from country and scientific community representatives' expert interviews.

For the Capability requirement areas of DAS section data collection was based on self-assessment of activities of current DAS in each country, following the specification of CESSDA SaW Capability Development Model (CESSDA-CDM) that resulted from the SaW Task 3.1. CESSDA-CDM was in the framework of Task 3.2 operationalized through the selection of a limited set of Activities contained in a model. The self-assessment in a form of survey was primarily aimed at SP/DAS representatives as respondents.

Semi-structured expert interviews were used to reach a wider group of stakeholders for assessing the country-specific scientific, cultural, policy, legal and funding setting.

Data from all three data collection methods, as well as references to the relevant resources, were entered in the web form. In countries where no SP/DAS exists, partner responsible for a country assigned scores in the web form, based on information from desk research and semi-structured interviews.

Selection and communication of country contact persons and interviewees were conducted following the common guidance and protocol (see Appendix 1: Guidelines and communication protocol for interviewers).

Limesurvey web form served as a database for data collected via self-assessment, desk research and semi-structured interviews generating scores for the predefined indicators, and including supporting evidence collected through additional references, notes and comments etc. (see Appendix 2: Questionnaire text extract from the Web form).

The central part of the web form was the **self-assessment area** that was accessible for country representatives, having a link (token) specific to the country (sent to the contact person for the country). In addition to self-assessment items, it includes data items/indicators that were collected by desk research and/or semi-structured interviews.

Since data come from various sources (some of which are more representative and objective than others), comparison of the individual scores between indicators, as well as comparison of scores between countries should be interpreted with caution.

1.4 Scope of countries and sources of information

The whole set of ERA Countries were selected for collecting information. The Protocol for selection of interviewees and communication (Appendix 1) was prepared for the task 3.2 partners responsible for the data collection. This was shared activity among all task partners, where each was assigned a set of countries to collect information and finally write Country report. The initial contact (email) includes a brief description of the CESSDA SaW project and CESSDA.

Within CESSDA members, the service providers (SPs) filled in the self-assessment survey. The email was sent to the director of the SP or the contact person for CESSDA SaW, if any was mentioned in the CESSDA SaW contact list. Each contact was documented in the CESSDA SaW contact list.

In countries, where there are already established DAS or grassroots efforts to establish data services, relevant partners of CESSDA PPP, SaW, SERSCIDA and SEEDS projects, representatives of each SP/DAS also fill in the self-assessment survey.

The semi-structured interview was conducted with informants who have a particularly good view of the social science research community in their respective countries. The individual interviewee selection protocol and criteria for selection are described in the Appendix 1. Any particularities during the data collection are noted in the country reports section about the data collection.

Table 1 describes the countries status regarding CESSDA membership, nominated or potential service provider (SP) for CESSDA, task 3.2. partner responsible for data collection and country report, and the range of chapters that were addressed during data collection.

Table 1: Target countries for CESSDA SaW

	Country	CESSDA Member	SP or potential SP	Responsible for data collection	Broader eco-system	DAS proto-activities	Capability requirements	Basic info only	No report
1	Albania	Initiative	IDM	IEN	x	x			
2	Austria	Yes	AuSSDA	GESIS				x	
3	Belarus	No	None	LiDA				x	
4	Belgium	Yes	SOHDA	DANS	x		x		
5	Bosnia and Herzegovina	No	None	FORS	x	x			
6	Bulgaria	No	None	SU SAV	x	x			
7	Croatia	Aspiring	FFZG	ADP	x	x			
8	Cyprus	No	None	CSDA	x	x			
9	Czech Republic	Yes	CSDA	GESIS	x		x		
10	Denmark	Yes	DDA	NSD	x		x		
11	Estonia	Aspiring	ESSDA	SND	x		x		
12	Faroe Islands	No	None	RODA					x
13	Finland	Yes	FSD	DDA	x		x		
14	France	Yes	RQ	GESIS	x		x		
15	Germany	Yes	GESIS	FORS	x		x		
16	Greece	Yes	EKKE	ADP	x		x		
17	Hungary	Yes	TÁRKI	SU SAV	x		x		
18	Iceland	No	None	NSD				x	
19	Ireland	Aspiring	ISSDA	DANS	x		x		
20	Israel	Aspiring	ISDC	UniData	x	x			
21	Italy	Aspiring	UniData	ADP	x		x		
22	Kosovo	Initiative	CPC	FFZG	x	x			

23	Latvia	Aspiring	LSZDA	SND	x	x			
24	Lithuania	Yes	LiDA	SND	x		x		
25	Luxembourg	No	LISER	GESIS	x	x		x	
26	Macedonia	Initiative	ISPJR	IEN	x	x			
27	Malta	No	None	DDA					x
28	Moldova	No	None	LiDA	x	x			
29	Montenegro	Initiative	CeMI	FFZG	x	x			
30	Netherlands	Yes	DANS	DDA	x		x		
31	Norway	Yes	NSD	SND	x		x		
32	Poland	Aspiring	ADS	CSDA	x		x		
33	Portugal	Aspiring	ICS-Ulisboa	DDA	x		x		
34	Romania	Aspiring	RODA	SU SAV	x		x		
35	Russia	Aspiring	JESDA	SND	x		x		
36	Serbia	Aspiring	IEN	ADP	x	x			
37	Slovakia	Observer	SU SAV	CSDA	x		x		
38	Slovenia	Yes	ADP	CSDA	x		x		
39	Spain	No	CIS	GESIS	x	x			
40	Sweden	Yes	SND	DDA	x		x		
41	Switzerland	Yes	FORS	ADP	x		x		
42	Turkey	No	None	SU SAV					x
43	Ukraine	Aspiring	KIIS	LiDA				x	
44	UK	Yes	UKDA	DANS	x		x		

2. Structure and content of the country reports

First part of the audit was oriented towards exploring how different enablers and incentives for sharing data characterize the broader context of DAS activities and development status. It included aspects of general development of social science sector, policy level support and scientific community acceptance that are reflected in data sharing culture. Focus for the countries that do not have national DAS was put on exploring the conditions for establishing a data service that could in the future obtain the role of CESSDA national service provider, labelled as DAS proto-activities.

The main purpose of the second part of country reports was to assess status in each country and identify gaps with regards to minimum CESSDA service provider requirements (SPR) as a starting point for further development planning.

Below is a description of structure of country reports, with a special emphasis on explanation of construction of development indexes for each section, that were used in a standardised visual presentation.

2.1 Broader ecosystem of DAS operation

2.1.1 Development of SSH sector

The first topic in this report is assessment of overall development of SSH and research data production as an important determinant of the quantity and quality of the research data produced and potential of data reuse. Focus is mostly on the issues of financial stability, research capacities and results achieved (funding, human resources and infrastructure conditions, impact and prestige in society).

The assessment involves **three stages of development (see Appendix 3 on Classification procedures)**:

- 0** - Funding of SSH and productivity of the researchers are in the **lowest quantile**; impact on designated community is small or non-existing;
- 1** - Funding of SSH and productivity of the researchers are in the **mid quantile**; impact on designated community is limited and
- 2** - Funding of SSH and productivity of the researchers are in the **highest quantile**; impact on designated community is strong.

The first part of assessment is based on indicators developed by OECD and presented in *Frascati manual: Proposed Standard Practice for Surveys on Research and Experimental Development*. The OECD essentially developed two general measurements of investment (or inputs) into science and technology: the financial resources invested in research and development, and the human resources devoted to these activities. For our country assessments, we used four indicators with focus on SSH sector:

- **I1.1.1 GERD in SSH as percentage of GDP** – shows general intensity of investment in SSH;
- **I1.1.2 GERD in SSH as percentage of GERD** - shows relative intensity of investment in SSH (relative to other scientific disciplines);
- **I1.1.3 Number of researchers (head count) in SSH per 100.000 inhabitants** - shows human resource potential in SSH and

- **I1.1.4 GERD in SSH per researcher in SSH** - shows investment in human resource in SSH.
- Data-gathering method
- Eurostat - <http://ec.europa.eu/eurostat/data/database>, following Science, Technology and Digital Society / Science and Technology (t_scitech) / Research and Development (t_research) / Statistics on Research and Development (t_rd);
- World Social Science Report 2016, ISSC and the Institute of Development Studies (IDS) and UNESCO and
- UNESCO Institute for Statistics Data Centre - <http://www.uis.unesco.org/DataCentre/Pages/BrowseScience.aspx>

We also include **I1.1.5** (see Appendix 2) the ranking of several sources of funding in SSH (Government and higher education sector; Abroad (international and cross-border); Private non-profit sector and Business enterprise sector) as the last indicator on input side. The importance of the source of funding has been recognized in one of the Barcelona targets of the Lisbon agenda where it is said that the appropriate split for R&D is 1/3 financed by public funds and 2/3 by private.⁹

Finally, we include **I1.1.6** Existence of support services or does government or state funding agency in a country support access to commercial databases with scientific papers (i.e. EBSCO, JSTORE) or datasets (i.e. Bankscope) and software purchase.

Development of research data production in SSH

Development of research data production in SSH can be defined as a prevalence of high quality research data with high potential for reuse.

The assessment involves **three stages of development** (again see Appendix 3 on Classification procedures):

- 0** - Rare or no data producing research projects, dispersed and low quality existing data, absence of studies of national importance;
- 1** - Some examples of research excellence, streams of research stand out, either qualitative or quantitative, some examples of international collaborative research and
- 2** - Well established streams of research traditions, national and international, great variety of important types of research data.

The first indicator is **I1.2.1 Existence of international collaborative research or cross-national studies**, as a driver for data production, i.e. involvement in the following international collaborative research or cross-national studies: Comparative Candidate Survey (CCS), European Social Survey (ESS), International Social Survey Program (ISSP), Comparative Study of Electoral Systems (CSES), European Values Study (EVS), World Values Survey (WVS) and Programme for International Student Assessment (PISA). The data has been collected centrally (desktop research) based on web resources.

The second indicator is **I1.2.2 Existence of studies of national importance** as a driver for data production in the country. We are interested in studies that systematically assess conditions and

⁹ See: <http://ec.europa.eu/eurostat/tgm/web/table/description.jsp>

public opinion on the matters of national importance (public opinion survey, election survey, etc.). Service providers (DAS) or key experts provide an answer, based on experience and available published reports.

The third indicator is **I1.2.3 General volume and frequency of data production or average production of research data by the SSH institutions** in a country. Self-assessment instrument and/or key expert can pick one of the following answer: *Rare production* (data are produced ad hoc); *Periodical* (institutions have tradition in producing some type of research data to a certain extent) and *Frequent* (institutions have well established tradition in data production).

2.1.2 RDM Policy setting

The aim of this section was to explore research funders' data sharing policies and underlying strategies about research data management requirements in the country that enable sustainable access sharing practices of publicly financed social science research data. This can range from declared awareness about principles and soft recommendations to explicit requirements, the fulfilment of which is actively supported and rewarded. Data archive services (DAS) can have a recognized and important support role in such an environment¹⁰, which will be highlighted in a report, using comments and qualifications to answers provided.

Clear research data policy in a country is important for paving a way and strengthens the position of existing and emerging DAS. Main points that are addressed in this chapter are: requirement to prepare Data Management Plan (DMP), open access to research data as default principle, recommendation about appropriate place of deposit, selection of data based on quality and reuse potential for long-term curation, importance of legal and ethical guidelines to attain clarity on the legal conditions framing the envisaged re-use of research data.

Most advanced funders consider to or already mandate the requirement that a data producer has to prepare DMP. Some of the enablers in this respect are guidance and adequate support provision planned for a DMP, given by variety of infrastructure activities (see Appendix 2: **I2.1.1 Research data management (RDM) policy requirements: Data management plan**).

Existence of soft and hard infrastructures is also one of the enabling functions that motivate researchers to share data by providing sustainable arrangements for data curation and access and (**I2.1.2 Appropriate place of data deposit defined**).

Sustainability of long-term curation of research data is a recurrent topic in most reports. Awareness raising and incentives and reward systems need to be adapted to support sustainable and managed data curation and access. One of the challenges is to select, which data should be preserved and for how long? For the indicator, we have chosen the question if long-term curation for valuable research data asset, evaluated and selected regarding reuse potential is among funder requirements (**I2.1.3 RDM policy requirements: Sustainability and long-term curation**). Key part for securing cooperation of researchers as data producers are incentives for them to archive and share their data, including cost recovery (**I2.1.4 Incentives for data sharing: Cost for managing the data resources**).

The classification of Research data management (RDM) requirements and incentives fall into the following three high level categories (see the Appendix 3 on Classification procedures):

¹⁰ See reports from Recode project explication what ideally is the role of each of stakeholders. Available at: <http://recodeproject.eu>

0 – Non-existent: Not aware of the need, not seen as a priority.

1 – Emerging: Declared awareness about importance and intentions of formulation of policy principles and strategy supporting data sharing motivation

2 – Developed: Partially or fully operationalized strategy and policy developed and implemented in calls on key aspects enabling data sharing.

To make possible the research data access arrangement that accommodate both legally and ethically sound practices is also one of the responsibilities of funders. The indicator **12.2.1** measures if clarification and support on legal and ethical aspects of data sharing (IPR, data protection) is provided.

Second concept, **A description of ethical and legal framework important for data sharing**, has the following categories (ibid.):

0 – Underdeveloped: Little or no awareness among different stakeholders, legal and ethical concerns are mainly used as an excuse for not sharing data at all.

1 – Developing: Growing awareness about the problems and some general guidance is provided.

2 – Developed: Specific guidance, advice and support services are available to prevent and overcome legal and ethical obstacles to data sharing and to actively support optimal mode of access to data.

Enablers and incentives for sharing data, that come from the top – down, funders **RDM policy requirements and incentives**, are important factors in characterizing the broader context of DAS activities and development, as are important acceptance of those requirements and routine practice of scientific community, that are reflected in data sharing culture described in the following.

2.1.3 Data sharing culture in the scientific community

This section of a country report aims to establish / benchmark the maturity level of the data sharing culture in the social sciences in a country based on indicators measuring data sharing routines, practices and attitudes. This provides (another) perspective for better understanding of the broader ecosystem in which DAS operations are embedded.

As the aim of this section is to provide context for DAS development. The choice of data sources and aggregation methods are closely related to it. Therefore, maturity levels should be seen as indicative, explanatory and contextualizing DAS maturity, rather than as independent characteristics of the maturity of data sharing in a scientific community, moving from informal data sharing to open and transparent data sharing through established procedures and trusted digital repositories.

The maturity of the data sharing culture in a scientific community as part of the broader data sharing ecosystem in a country, can be seen as both a driver for, and be driven by the development of DAS. Data sharing and management practices in research in general, and in the social sciences in particular, influence the availability of data for reuse, for new research, for replication, etc. For this report, we consider four aspects of data sharing: whether researchers provide (share) data to the community and can access the data they need (see Appendix 2: indicators **3.1.1** and **3.1.2**), the common routines or channels via which data are made available (indicator **3.1.3**), and researchers' attitudes to data sharing (indicator **3.1.4**) in the social science research community in the country.

The individual indicator scores measuring **Development of Data sharing culture** in the scientific community (indicators **3.1.1 - 3.1.4**) are aggregated to three maturity levels (See Appendix 3):

0 - Underdeveloped: data sharing and reuse are rare or non-existing in the social sciences research community, with no existing routines for sharing, and researchers having negative or indifferent attitudes towards data sharing.

1 - Developing: data sharing and reuse are not that common, with largely informal, non-transparent routines, and indifferent attitudes towards data sharing.

2 - Developed: data sharing and reuse are very common, with formal and transparent routines for data sharing and positive attitudes towards data sharing.

The guiding principle of aggregation leading to the maturity level is that, in order to reach a certain level of maturity, the majority of indicators characterizing the dimension should be on that respective level (i.e. if 3 out of 4 indicators measuring 'data sharing culture' score "0 - underdeveloped", then that is the maturity level of the dimension.). If more than half of the indicators have no meaningful score assigned (due to lack of information or resources to acquire information via alternative sources), the maturity level for a country cannot be estimated.

Enablers for data sharing

Together with data sharing culture, real and perceived barriers to data sharing and enablers and incentives for sharing data in a scientific community are important factors in characterizing the broader context of DAS activities and development.

[Previous research](#) on a local, European or international level allows identifying a list of important factors that form and influence data sharing culture. Enablers that are measured in this section of the report are career rewards (see Appendix 2: **I3.2.2**), the presence of data support services (**I3.2.3**), and existing RDM practices in the social science research community (**I3.2.4**). Other indicators already measured elsewhere in the survey should also be included here as enablers and included in the analysis.

The indicators to combine are:

- data policy at EU, national and institutional level that sets the expectations or requirements for research data sharing, with funding for data management, infrastructure and support development (indicator **2.1.1-2.1.4**);
- the availability (and maturity) of data infrastructure, because the absence of data infrastructure and support services usually means that researchers willing to share their research data with the wider scientific community or publicly, can in practice only share data informally with colleagues and peers, or not share them at all;
- established and supported research ethics as enabler for data sharing (indicator **2.2.1**);
- the availability of support services and tools as enabler for RDM and data sharing (indicator **3.2.3**);
- established and developed RDM practices in the social science research community (indicator **3.2.4**);
- (international) collaborative research as a driver for data sharing (indicator **1.2.1**);
- career progression in academia as a motive for data sharing (indicator **3.2.2**);

As indicated for the section on data sharing culture, data for individual indicator scores are collected from several sources, through self-assessment as perceived by the SP/DAS; from survey

reports or case studies in the research community; or from expert interviews, and again aggregated to three maturity levels (underdeveloped, developing and developed) according to the same principles.

Development of **Enablers for data sharing** (incentives, rewards, skills, support services and tools) has been assigned the following categories (See Appendix 3 for classification):

0 - Underdeveloped: There exist no enablers for data sharing.

1 - Developing: There are some enablers for data sharing.

2 - Developed: There are many enablers for data sharing.

This section of a country report aims to identify meaningful relationships and dependencies with enablers, incentives and barriers for data sharing, for example, data policy requirements, data sharing culture and research data management practices, availability of data infrastructures, funding schemes, etc. This provides context for DAS development. Maturity levels should be seen as indicative, explanatory and providing context for DAS maturity, rather as independent characteristics of the maturity of data sharing in a scientific community.

2.2 DAS proto-activities

Some European countries with no established national data archiving services (DAS) for the social sciences are already far with the activities towards establishment, while others are at the very beginning. In this section, we distinguish for countries that are currently active, out front, and seeking solutions with respect to research data infrastructure from those that are currently less active. We identify the key players involved in DAS-related activities, and describe their current expertise, level of technical infrastructure development, as well as their overall activities and potential regarding open science issues in general.

Thus, this section assesses data archive proto-activities and open access support activities in countries where no formal DAS exists yet. Identifying those activities allows us to detect actors and institutions that could play a key role in the elaboration of new national data services.

Determinant factors of data archive proto-activities include:

- The availability of technical infrastructure (see Appendix 2: **I4.1.1**);
- Organisational first steps (see **I4.1.2** and **I4.1.3**);
- The availability of trainings and capacity development for (future) collaborators in the main DAS activities, which are RDM, data preservation, and data access (see **I4.1.4**).

These factors allow us to assess the overall DAS activities implementation type of a given country. According to the results, we have grouped the countries within three types of DAS activities implementation (see also Appendix 3):

0 - No DAS activities: No DAS exists and no activities – related to technical infrastructure, organisation or capacity building – are provided on a national and/or institutional level;

1 - Basic DAS activities: No DAS exists but basic activities –related to technical infrastructure, organisation or capacity building – are provided on a national and/or institutional level;

2 - Advanced DAS activities: No DAS exists but advanced activities –related to technical infrastructure, organisation or capacity building – are provided on a national and/or institutional level.

Determinant factors of open access support activities include:

- The availability of open access (OA) support for researchers from institutions like university libraries (see I3.2.3);
- Open access projects or initiatives that can be enablers (see I4.2.1).

These factors allow us to assess the OA support activities implementation type of a given country. According to the results, we have also grouped the countries within three types of OA support activities (see also Appendix 3):

0 - No OA support activities: Support activities to encourage and facilitate OA are rare or not existing in the social science research community;

1 - Basic OA support activities: Some support activities to encourage and facilitate OA exist, but not on a regular basis and only in few institutions;

2 - Advanced OA support activities: Support activities to encourage and facilitate OA are well established (common and in most institutions), known and used by the social science research community.

2.3 Organisational profile

The following subheadings were proposed for the Country reports and input based on the questions indicated.

- Organisation (name of the organisation) [input from questions 1.3, 1.4, 1.7]
- Funding [input from question 1.2]
- Core services and activities [input from question 1.1]
- Content current collection [input from questions 1.6 and 1.5]

2.4 Capability requirement areas of DAS

The CESSDA-CDM (Capability Development Model) is a structured collection of elements that identify and describe the characteristics of effective preservation processes and activities. Building on established frameworks for trustworthy data preservation and the CESSDA community's prior experiences, the model provides both a starting point for emerging preservation initiatives and a reference tool for established archives that wants to strengthen their services. It is a model that can be used to appraise and/or improve the capability of a service provider to perform and to provide services. The model focuses on three main subject areas that describe on a high-level, the main objectives and principles of an organisation providing research data preservation services. The subject areas are Organisational Infrastructure, Digital Object Management and Technical Infrastructure. The application of these subject areas in task 3.2 are further explained below. For a more detailed explanation of the full CESSDA-CDM, see the project web page¹¹.

¹¹ CESSDA-CDM. Available at: <http://cessdasaw.eu/content/uploads/2016/06/D3.1.pdf>

Application of CESSDA-CDM in task 3.2

The CESSDA-CDM is inherently hierarchical, where the maturity level of certain sets of activities determines the maturity level of a process area, which in turn determines the 'completeness' of the higher-level requirement areas. The application of the model in the survey in task 3.2 partly ignores this hierarchical structure. Applying it in its full form in combination with the assessment of the "broader ecosystems of the DAS operation" would make the survey (information-gathering tool) too extensive. Considering the short time-frame that in effect was available for the project, it was decided to be pragmatic and to apply a simplified and shortened version of the model.

There is a broad variation between the activities in the model, when it comes to the level of detail. The project group assessed the full model and tried to identify, within in each Capability Requirement Area (CRA), the activities that are broad in scope and that best can be said to cover and represent each specific CRA. These were selected and included in the survey. This cherry-picking and 'flattening' approach (i.e. not considering capability completeness aspect) to the model is far from perfect. However, we still think this approach enables us to capture some valuable information and give us a glimpse into the current capability-level in selected service providers in the European research area. We see this as an opportunity to introduce the CESSDA-CDM to existing and potential CESSDA service providers, capturing some basic information in the process.

2.4.1 Organisational infrastructure

Organisational Infrastructure consists of a set of process areas that support the development and maintenance of a viable preservation organisation. The process goals are as follows: to have a clear definition and delimitation of the mission and scope of the organisation; to keep operations aligned to relevant legal and regulatory frameworks, including the handling of confidentiality issues; and to have adequate funding, valid budget planning, and sufficient numbers of appropriately qualified staff. These processes should be managed through a clear system of governance, where roles and responsibilities are clearly defined, so that the organisation can effectively carry out its mission.

The following elements of Organisational Infrastructure are covered in the survey instrument:

- Mission and scope
- Contracts, licenses and liabilities
- Funding, staff resources
- Confidentiality, ethics and disclosure risk

Self-assessment exercise is limited to the country CESSDA service provider representative, and organisations expected to be in this role for CESSDA membership aspiring countries. Contact person for DAS coordinated which person in the organisation provided information on which section.

2.4.2 Digital object management

Digital Object Management (DOM) consists of the set of processes (e.g. selection, acquisition, ingest, management, preservation) required to maintain and provide access to digital information in an authentic form, for as long as required and across changing technical environments. Digital Object Management is closely related to the term "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. The aim of DOM and digital curation is to mitigate digital obsolescence, keeping the information accessible to users indefinitely.

Listed below are the elements that are selected for the online survey instrument. Each element has been carefully selected from the three process areas in the DOM-part of CESSDA-CDM. The three process areas are Data acquisition and ingest, Data preservation, and Data access / provision.

- Documentation, metadata, provenance
- Data citation mechanisms
- Access conditions
- Legal / contractual arrangements
- Data completeness and correctness (checking of)
- PID handling
- Preservation strategies
- Data / metadata standards
- Authentication and Authorization Infrastructure

2.4.3 Technical infrastructure

Technical infrastructure consists of a set of process areas that provide the technical underpinnings of an organisation so that it can properly fulfil its functions and the provision of services to its designated communities. The technical infrastructure part of the CESSDA-CDM consists of five different process areas, namely Risk assessment, Technical planning and Management, Technical infrastructure, Security, and Disaster planning.

For the survey instrument, we have adjusted two of the process areas (Risk assessment and risk analyses, and Contingency / succession plans) so that they are treated as activities, and applied an adjusted set of maturity level description to each of them. These are combined with one of the broader technical infrastructure activities, namely succession planning.

That leaves us with the following set of process areas that are applied in the survey instrument:

- Risk assessment and risk analyses
- Technical infrastructure resources (suitable software, technical services, and appropriate management plans)
- Succession and contingency plans

3. Country reports

3.1 Albania

The report for Albania was compiled based on two primary sources, CESSDA SaW project survey and SEEDS project survey¹² published as deliverable D3: Report on Evaluation of Research and Legal Conditions in Albania¹³. The SEEDS project (South-Eastern European Data Services) (<http://seedsproject.ch/>) aims to widen efforts to establish new data services in the countries of Albania, Kosovo, Macedonia, Montenegro, and to continue their integration in the national and international landscape for Croatia and Serbia. The project extends the capacities of selected partner institutions, and develops knowledge and tools related to data service infrastructures. It brings partner institutions into the fold of an international movement and network that will provide long-term benefits not only to these institutions, but also, more importantly, to the respective national research communities, including more available data for secondary analyses.

The analysis shows that Albania does not have, at present, an established data service for the social sciences or any data archive that is used by the science research community. Obtaining an objective assessment was connecting with the lack of indicators, such as GERD in SSH, in public reports and databases (Eurostat, World Social Science Report and UNESCO Institute for Statistics Data Centre).¹⁴

The main institution in Albania, which support development of science and technology, established in 2010, is the Agency for Research, Technology and Innovation (ARTI). The ARTI's aim is to build a modern system of science, strengthen of research and technology, as well as their integration inside the higher education system. ARTI as a coordinating and guiding structure, which cooperates with institutions in the field of science and technology for sustainable development of the country, in line with national priorities, development of scientific and technological policies and management of R&D institutes. Its mission is to evaluate, finance, monitor and manage programs and projects in the fields of science, technology and innovation in Albania, and has the tools to provide the research community support in data services and data archiving¹⁵. ARTI is the central

¹² "The survey was conducted on the second half of 2016 and it was shared by email with approximately 400 individuals coming from different backgrounds that potentially produce and store data in Albania, including: social science institutes, universities, non-governmental organizations, the research department at the National Bank of Albania, etc. Aiming at accessing as many researchers operating in Albania as possible, we repeated twice the electronic invitation to participate in the survey. However, out of the total number of people who were contacted, 70 individuals responded to the survey. Even though it is not possible to assess the reasons for the low response rate, we can presume that it reflects the small number of individuals focused on conducting research in Albania as well as a common apathy related to conducting online surveys. Even though the social sciences have shown an increasing trend in the last couple of years, it is important to note that it still fails to attract human capital, whether due to poor infrastructure, lack of financing, or other factors" (SEEDS report, p. 19).

¹³ http://seedsproject.ch/wp-content/uploads/2015/06/SEEDS-Report-Albania_FINAL.pdf

¹⁴ Eurostat does not provide data regarding % of GDP on Research and development intensity. See: [http://ec.europa.eu/eurostat/statistics-explained/index.php/File:Research_and_development_intensity,_2004%E2%80%932014_\(%C2%B9\)_\(%25_of_GDP\).png](http://ec.europa.eu/eurostat/statistics-explained/index.php/File:Research_and_development_intensity,_2004%E2%80%932014_(%C2%B9)_(%25_of_GDP).png).

Additional material stating the lack of data regarding GERD in SSH for Albania can be found in the World Social Science Report. Available at: http://www.assau.org/IMG/pdf/world_social_science_report_2016.pdf (Annex A, pg. 304)

¹⁵ The Agency for Research, Technology and Innovation (ARTI) official web page http://www.akti.gov.al/about_AKTI.html

institution that will lead and give fully support in the process of establishment Albanian data service.¹⁶

Specifics about the data collection in a country

The data collection started 21st September 2016 and was finished 30th September 2016. The information was provided by Institute for Democracy and Mediation (IDM), Tirana, independent non-governmental organization founded in 1999, and potential service provider for Albania. The contact person who filled the survey was Artela Mitrushi, a researcher within IDM, engaged in data archive proto activities within SEEDS project. The data collection process went well without any problems.

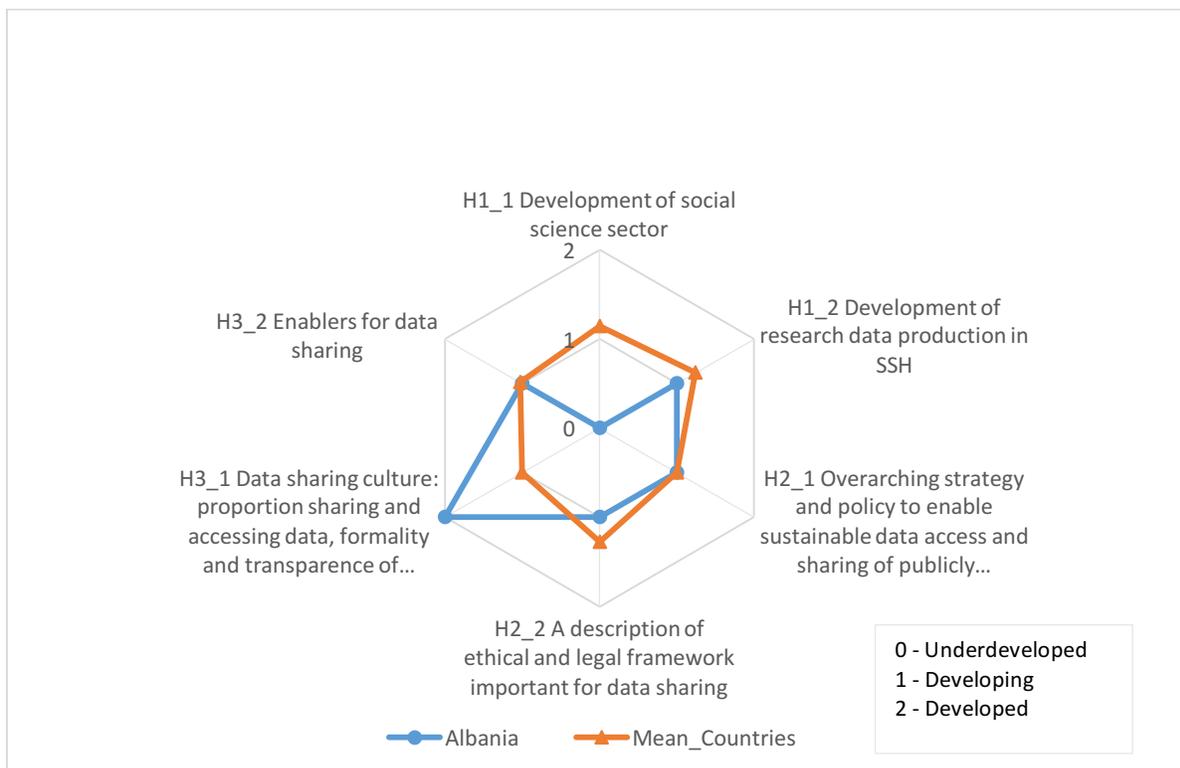
3.1.1 Broader ecosystem of DAS operation

The science system in Albania consists of higher education, scientific research, knowledge and technology (innovation) development institutions. In this respect, it includes not only the public and non-public institutions of higher education and basic research, but also private enterprises that operate in the fields of research, development, and innovation. Research is conducted by a number of institutions, such as Higher education institutions, national research centers, several ministries, and last but not least, private institutions or non-governmental organizations (NGOs).

Capacities to manage both basic and applied research in Albania are limited and generally do not meet EU standards. There are different problems related to scientific infrastructures, and from a cost efficiency point of view, there is a need to align a national scientific infrastructure investment plan with investments being made at a regional (Western Balkans) level, in order to permit sharing of equipment and infrastructure (SEEDS report).

¹⁶ SEEDS Establishment plan for Albania. Available at:
http://seedsproject.ch/wpcontent/uploads/2015/06/Establishment-Plan-ALBANIA_FINAL.pdf

Figure 1: Heading concepts values



Development of social science sector

Overall assessment of SSH development

Figure 3.1.1 shows that Albania is a country with underdeveloped social science sector mainly because level of funding is low.¹⁷ The main funding resources for research in SSH in Albania are the government and higher education sector and international or cross-border funds, with the participation of private non-profit sector and business enterprise sector. In addition, substantial EU funds support Albanian membership in EU framework programs. Albania has followed a centralized model in supporting research. The expenditure on the research system has been and still is almost exclusively financed by the public sector. However, in the past years, there has been a boost on bilateral co-operation with the private sector.

According to Culver, Martin & Maroulis (2013)¹⁸ state budget cuts for 2011–2012 to the national programmes managed by the ART were an indication that overall availability of funding to support R&D in general is reduced and this has affected the development of social science and humanities research projects. Following the same source, in Albania and other non-EU Balkan countries, the allocation of IPA funds for the period 2011–2013 as well as funds for multi-beneficiary programmes and cross-border cooperation represent a big opportunity for organisations involved in SSH, particularly in the fields of law, political science, management, economics and education. The volume of funds allocated for SSH, as well as the number of projects, are very important

¹⁷ The conclusion was made based on the CESSDA SaW survey, SEEDS Report for Albania, and the fact that Albania does not have available indicators, which measures development of SSH, in public reports.

¹⁸ Culver, J., Martin, T., Maroulis, N. (2013): Funding Report - Monitoring Social Sciences and Humanities. METRIS. Available at: http://www.tma.ie/downloads/METRIS_funding_report_2013.pdf.

elements of the environment that can encourage or limit the development of consciousness, needs and infrastructure for sharing research data. The situation in Albania, in terms of infrastructure that provides favourable conditions for Albanian researchers in SSH can be seen through the fact that researchers do not have centralized access, funded by public funds, to commercial bibliographic and full-text databases.

Development of research data production has score 1 which implies some examples of research excellence, and some examples of international collaborative research and national research (explained further in this document). Third component of this indicator is a general volume and frequency of data production. According to the CESSDA SaW survey there are several institutions in SSH that do provide data sets, but not systematically.

Indicator of overarching strategy and policy to enable sustainable data access and sharing of publicly financed social science research data has score 1 which implies emerging stage in this area or declared awareness about importance and intentions of formulation of policy principles and strategy supporting data sharing motivation. Ethical and legal framework also has score 1 which implies that Albania has growing awareness about the problems.

According to CESSDA SaW survey data sharing and reuse in Albania is very common and there are formal and transparent routines for data sharing, together with positive attitudes. On the other hand, SEEDS survey shows that only 9% of researchers are willing to share their data to the broader scientific community.¹⁹ This deviation is caused because the CESSDA SaW survey includes an opinion of few people involved in data sharing initiative, while SEEDS survey includes perception of broader research community. The results of the somewhat more reliable source thus indicate that formal data sharing and reuse is not that common in Albania, and that there are largely informal, non-transparent routines, and indifferent attitudes to sharing culture in this country.

Finally, there are some enablers for data sharing in Albania, which will be explained in the second part of this document.

Development of research data production in SSH

Research data production in Albania can be characterised as only 'periodic' (based on the country informant estimate that 'there is tradition in producing some type of research data to a certain extent'). However, the situation is much better in the participation in main international collaborative research and cross-national studies. Albania was a participant in:

- Comparative Candidate Survey (CCS) in 2013;
- European Social Survey (ESS) in Round 6, 2012;
- Comparative Study of Electoral Systems (CSES) in Module 2 (2001-2006) as Collaborator, 2005, and Module 4 (2011-2016) as Collaborator;
- European Values Study (EVS) in 2008;
- World Values Survey (WVS) in 2001 in W4 PI and
- Programme for International Student Assessment (PISA) in 2000, 2009, 2012 and 2015.

In the field of studies that systematically assess matters of national importance there are sporadic initiatives, but not systematic or harmonized.

¹⁹ Available at: http://seedsproject.ch/wp-content/uploads/2015/06/SEEDS-Report-Albania_FINAL.pdf (p.28)

RDM Policy setting

Funders' data management and sharing strategy and/or policy

Regarding funders' data management and sharing strategy and/or policy, there is a growing recognition and awareness of the need to require Data Management Plans as an integral part of on-going project activity and some institutions have started to pay more attention to open data policies. However, they are still far from the developed maturity level when it comes to Data Management Plans and public research funding organizations do not issue requirements or recommendations about 'archiving quality-assured social science research data with associated metadata'. The funding agencies do not provide financial incentives for sharing research data with associated metadata, but the cost for managing the data and preparing it for access can be implicitly covered by existing budget lines or applicants must include such items/costs in the budget (bottom-up approach). Also, most international donors did not apply this criterion over the data set produced under their funded projects in the past, but since June 2016 IDM (Institute for Democracy and Mediation, Tirana) is a participant of RRPP Data Rescue project funded by Swiss National Science Foundation²⁰. This project has the aim to collect and archive all data sets generated within past RRPP projects in the period 2007-2016. Albania has 11 data sets from those projects to archive and make accessible for secondary use. This project was the first attempt of sharing data according to OAIS standard.

Legal and ethical framework

The legal framework in Albania has faced several changes and improvements in the past years. However, there is still need for 'modernising' the laws related to research, technological development, and innovation. The main laws that are indirectly related to data management are Law on Higher Education and Scientific Research (2015), Law on Archives (2003), Law on Protection of Personal Data (2008) and Law on the Academy of Sciences (2006).

According to SaW survey, development of legal and ethical framework important for data sharing in initial phase but there is a growing awareness about the topic and about the need to provide clarification on legal aspects, scarce or no organised support is given. According to the respondent comment, the main constraint is lack of a cultural ethical and legally sound research. Committee on Higher Education and Scientific Research, "Final Report on Reform of Higher Education and Scientific Research", Tirana, (2014, p. 51). The report argues, among others, that "*The research output is not conceptualized right; in many cases it is justified with publication of textbooks, failing to focus in original and experimental work to bring innovation to science. Cases of plagiarism are quite common.*"

Data sharing culture

The respondent from IDM could not give an estimation about of the proportion of researchers who have shared the research data they produced in the period between 2011 and 2016. The survey conducted within SEEDS project showed that most of the respondents stated that they use international data or data that were produced by themselves in their research, but at the same time it is interesting to see that in order to obtain data they ask their friends and co-researchers personally for help. This implies that the data are either not archived properly, or are not accessible (SEEDS report, p. 21). The condition can be also confirmed by SEEDS report (p. 27) where it states that only 18% of researchers acknowledged that the data they had produced were used for

²⁰ See http://seedsproject.ch/?page_id=93

secondary analysis by other researchers within the past year, and 9% said that their data were used more than a year ago (2015 and earlier). Also, according to respondents, 86% believed that research data sharing within their discipline is very important or somewhat important (13%), while only 1% found data sharing not important (Ibid, p. 27).

Researchers also have routines for data sharing to reflect the preferences of the social science research community in the country between 2011 and 2016, as follows: supplementary data in a journal (alongside paper), via informal contacts (peers and colleagues), via project or personal websites and data archive or repository. However, 57% respondents certainly or probably will provide data to a SSH data archive if the data would be preserved for the long-term in a secure environment and shared only within accredited researcher.

The CEESDA SaW survey covered few questions regarding to researcher's attitudes towards data sharing. Based on the answers we can conclude that Albanian researchers consider that data sharing will bring benefits to all, which is in line with SEEDS survey, and that data sharing will create healthy competition. It is also confirmed that the reuse of existing data can answer new research questions and facilitate advancement of science, with some concern about risk that others may misuse or misinterpret data.

Enablers for data sharing

Enablers for data sharing are underdeveloped in Albania. There are no career rewards related to data sharing in the academic community, nor the data support services available to social science researchers that would facilitate data sharing and/or Open Access to research data. Nevertheless, the situation is changing thanks to the efforts that are being implemented within the framework of SEEDS and RRPP projects²¹.

Despite these shortcomings, some data producers follow data management and data documentation standards and procedures, but only related to some research outcomes conducted in frame of certain projects.

3.1.2 Data archive service (DAS) proto-activities

DAS activities

In Albania, there is no existing technical infrastructure (national and/or institutional infrastructures in the social sciences or other scientific disciplines) that could possibly be used for or applied to a new DAS. However, through SEEDS project, Albania implements basic activities towards establishing a DAS in SSH. The Institute for Democracy and Mediation (IDM) from Tirana, as a partner in this project, delivered in 2016 an Establishment plan for the ADAS - Albanian Data Service in Social Sciences²². According to this document, the Agency for Research, Technology and Innovation (ARTI) will be the leading institution for the establishment of the ADAS. ARTI also want to develop the knowledge and skills of people who might be employed at a data archive at some point.

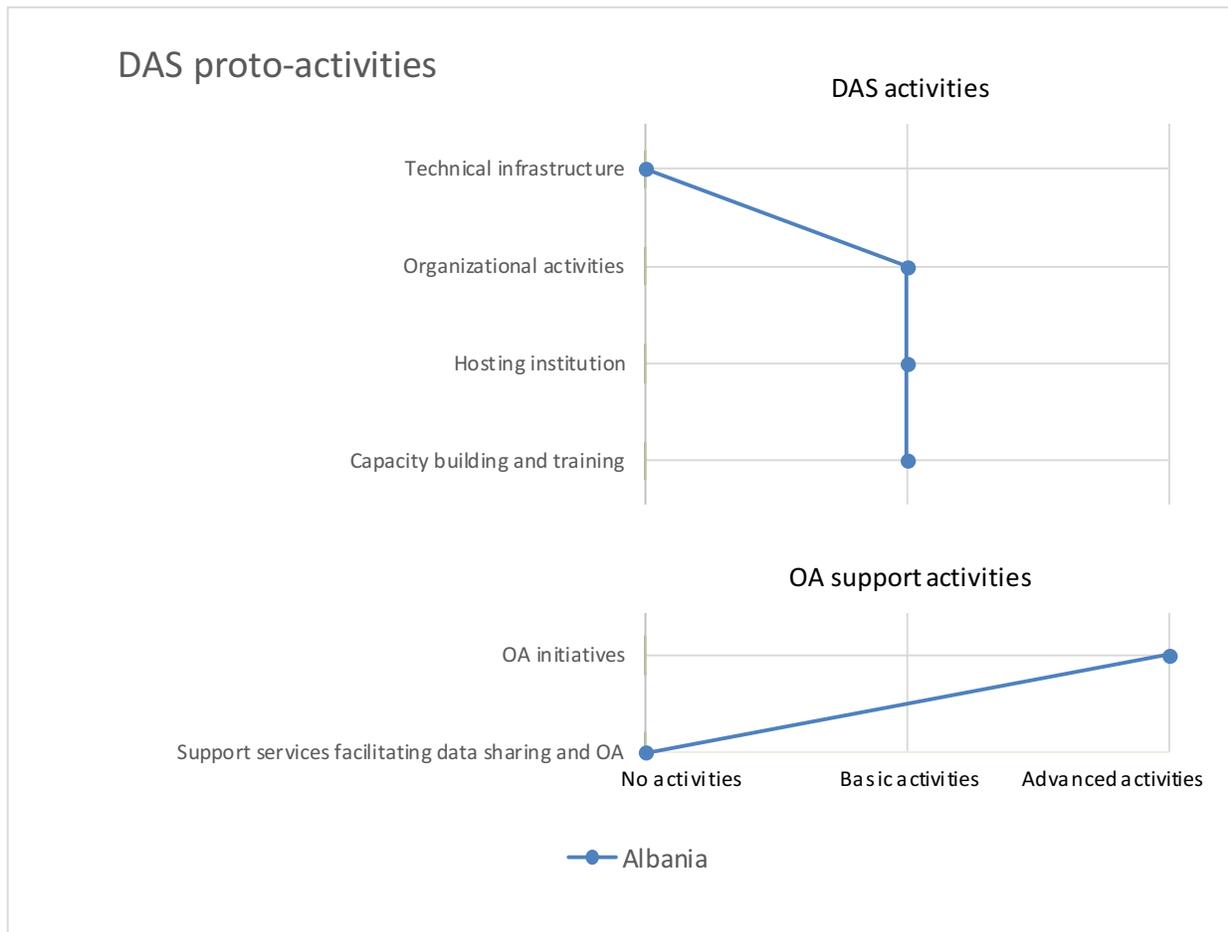
The data archive will offer data services for national and international researchers, students, and teachers. The focus of the archive collection will be on both quantitative and qualitative data in the Social Sciences. The new data service will provide a range of services, including acquisition, preservation and dissemination of data; trainings for its users (i.e. researchers and students); and

²¹ See <http://seedsproject.ch/>.

²² Available at: http://seedsproject.ch/wp-content/uploads/2015/06/Establishment-Plan-ALBANIA_FINAL.pdf

cooperation with the national statistics office as well as other government institutions and stakeholders. Through the SEEDS project, IDM staff gained knowledge, skills and experience, which should be used to influence future developments in the area of social science data management and preservation.

Figure 2: DAS and OA activities implementation type



There are several potential candidates for hosting the new data archive:

- INSTAT, the Albanian statistical institute, which aim is to provide transparent, neutral and timely statistics that help users to judge on the developments of the transformation processes within the country. INSTAT strives to become the major source of statistical information providing decisions makers and researchers, in Albania as well as in the international community, with relevant, reliable and comparable statistical information²³;
- Universities (i.e. University of Tirana the largest university in Albania); and
- Research institutes: Centre for Albanological Studies is a possible choice. The Centre was reorganized in 2008 as a scientific research institute and received the status of

²³ See: <http://www.instat.gov.al/en/home.aspx>

interuniversity organ. It inherited a staff of specialist and a rich archive of a high importance. The Centre does research in various fields, such as history, archaeology, cultural anthropology, art study, linguistics and literature. Part of its work focus is the investigation, evaluation, conservation, and promotion of the material and spiritual inheritance of the Albanian culture, both past and present.²⁴ It also offers master and PhD courses and studies on all the fields that the institute covers.

Open Access (OA) support activities

Albania is involved in the Open Access project named Open Data Albania (ODA)²⁵. This ongoing project is implemented by the Albanian Institute of Science (AIS), which is a non-profit, non-partisan organization established in 2011. The mission of AIS is to promote research activities that offer solutions to socioeconomic problems, increase transparency and strengthen civic engagement. It has accomplished to publish hundreds of datasets in open data formats, deliver more than 300 analyses and studies on socio-economic topics, and launch numerous activities with the goal of creating an Open Data movement. Journalists, students, and civil society activists not only have profited from it, but also became part of the activities establishing in this way a form of eParticipation network. AIS has provided the public with:

- Open data sets, research studies and analyses published in the ODA platform;
- Increased accountability through information about how taxpayers' money are collected and spent via a Tax Calculator²⁶;
- Information about the clients of every state institution through Treasury Monitoring²⁷ including total amounts of expenditures, and transactions;
- Information and increased understanding of electoral spending processes through the Electoral Spending Albania²⁸ project;
- Tools for easier analysis and verification of the amendments to the annual state budget through Money Mapping²⁹.

The platform where the data is published online is the most prominent one in Albania and very popular between the journalists, who are constantly using information from ODA to provide the public with information on different social-economic aspects³⁰. The success of this project is a very good example and incentive for further development of data service in SSH in Albania.

As mentioned earlier, there are no data support services available to social science researchers that would facilitate data sharing and/or Open Access to research data.

²⁴ See: https://en.wikipedia.org/wiki/Centre_of_Albanological_Studies

²⁵ See: <http://www.open.data.al/en>

²⁶ See: <http://spending.data.al/en/openspending/dailybread>

²⁷ See: <http://spending.data.al/en/treasury/index>

²⁸ See: <http://open.data.al/sq/lajme/kat/kid/82/titulli/Shpenzime%20Elektorale%202013>

²⁹ See: <http://spending.data.al/en/openspending/treemap#year:2013/lang:en/budgettype:plan>

³⁰ Hoxha, J., Brahaj, A. (2013) Utilization Cases of Open Data Albania. Available at: <https://www.w3.org/2013/share-psi/wiki/images/6/63/ODA-UtilizationCases-SharePSI-Workshop.pdf>

3.1.3 Conclusions

According to the conducted analysis, we can conclude that:

- Funding of SSH and productivity of the researchers in Albania are at the low level of development and that impact on designated community is small or non-existing.
- Development of research data production in SSH is at the developing phase. There are some examples of research excellence, particular streams of research stand out, either qualitative or quantitative and some examples of international collaborative research.
- Ethical and legal framework important for data sharing is underdeveloped. There is a little or no awareness among different stakeholders, legal and ethical concerns are mainly used as an excuse for not sharing data at all.
- Data sharing and reuse is rare or not existing in the social sciences research community; with no existing routines for sharing, and negative, indifferent attitudes and therefore underdeveloped.
- Albania has basic DAS activities through SEEDS project.

3.2 Austria

As Austria is in a transition phase, there were no reliable statements possible at the time of data collection. After the former archive for social science data was closed in 2014, a new one will be build up beginning in 2017. For this reason, at this moment, no statements are available about the data archive.

Regarding policies in the field of data management and data sharing there are some efforts. Attitudes from researchers towards data sharing and handling of data with regards to storage and documentation have been evaluated. Within the next years, the data archive and CESSDA Service Provider will be building up.

Specifics about the data collection in a country

All information is gathered through desk research.

3.2.1 Broader ecosystem of DAS operation

Austria has started a project called e-infrastructure Austria (<http://www.e-infrastructures.at/en/startpage/>) with the goal to develop and establish a repository infrastructure for digital resources and research data. The project is divided in three subprojects.

Within this project, several reports were published. Amongst others, a study was conducted about "Researchers and their data"³¹. The study was directed at the arts, humanities and sciences staff of all 21 public universities and three extramural research institutions in Austria. Some findings were, that with regards to data archiving, that most of the researchers are themselves responsible for archiving the research data. Another result was that there is no common or consistently documentation of the data. Even if the use of external data is considered as useful, there are different approaches to data sharing. The authors have formulated recommendations that include the "*creation of a comprehensive, technological infrastructure in Austria, including specific disciplinary needs, implementation of support services for researchers*" and other measures to develop infrastructures for data management and proper data handling (p. 10).

³¹ »Researchers and their data«. Available at: <https://phaidra.univie.ac.at/o:409318>.

Data sharing culture

The study „Researchers and their data“ has the following results:

Asked to whom they grant access to their data, 13% of social science researchers answered “to no one”, so the rest 87% share their data: 10% with public; 26% - with scientific community; 15% - with members of their institution; 53% - with only selected members of the institution, but 53% - with interested persons by request. It seems that more than one answer was possible (table 61 of the appendix, p. 124). In another question, (table 67, p.130) answers indicate that 24% have said that there is no way others can access their research data, so it could be that the actual level of accessibility is somewhat lower due to lack of technical solutions. More than three quarters answered that their data is reusable for others (table 73, p.135).

Asked, if they have reused data produced by others, 24% of social science, researchers have indicated that they never use such data, so 76% do use data produced by others (table 41, p.110). It does not precisely say how many have got access to data they need, but it is an indicator of access and reuse.

The preferred/ mostly used routines for data sharing is with selected members of own institution and interested persons by request, we could assume that the most preferred channel for sharing data is personal contacts (table 61, p. 124). Sharing with the scientific community might imply use of more formalised and transparent channels like journal supplements and data archives.

The social science researchers were also asked, how others can access their research data: 24% indicate that there is no way to access their data; 9% use data archive, 10% linked supplementary material to journal publications; 16% use websites; the rest of technical solutions seems to be related to data shared via personal contacts - remote server/share drives - 22%, cloud applications - 24%, physical disk/ e-mail - 52% (table 67, p. 130). Thus, we could say that formal and transparent - archive, linked supplementary material - alternatives are not very popular (ranked 3rd /4th), and that personal and project websites and informal contacts are ranked as the first and second.

Concerning the general attitudes towards data sharing - the self-assessment questionnaire asked about trueness/falseness of a set of statements; this survey asks about obstacles to sharing research data (several answers are possible; table 89, p. 150).

As a preferred archive, most social science researchers an International discipline specific data archive was chosen by 44%, followed by a centralized data archive in their own institution (34%) and an International multidisciplinary data archive (32%) (table 95, p. 156).

The self-assessment questionnaire asked about perception of benefits of data sharing; healthy competition vs negative competition and less publication opportunities; reuse related to advancement of science; risk of misinterpretation and misuse of data; effort and resources involved in data sharing.

Attitudes towards data sharing

It seems that the attitudes of social science researchers towards data sharing can be characterised as neutral-negative, even though there is no data at all available for some of the statements.

Table 2: Attitudes towards data sharing in Austria

Data sharing has no benefits at all	No information
Data sharing creates healthy competition	23% of social science researchers consider competition as an obstacle to data sharing, so we could guess that this is not a prevailing obstacle.
Data sharing creates negative competition	See above.
Reuse of existing data can answer new research questions and facilitate advancement of science	No information
Data sharing has as a risk that others may misuse and misinterpret data	43% consider possible misinterpretation of data an obstacle to data sharing 3rd/4th most often named obstacle, and 46% consider misuse possible - 2nd most often named obstacle
Data sharing involves little effort and minimal costs	43% consider effort and/or costs to be an obstacle to data sharing; 3rd/4th most often named obstacle

Source: "Researchers and their data" report (<https://phaidra.univie.ac.at/o:409318>)

Enablers for data sharing

Indicators:

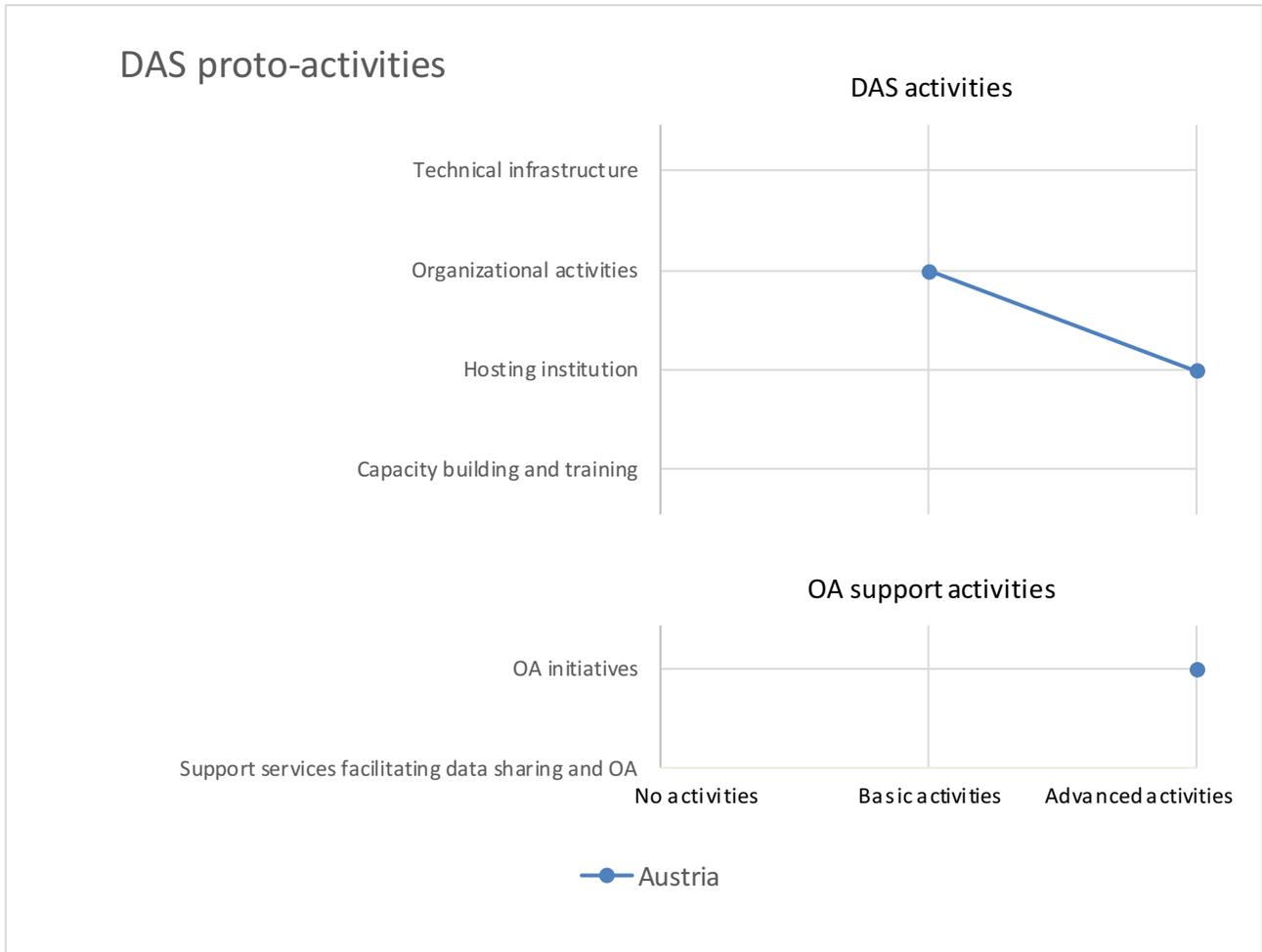
What career rewards are there for data sharing - the report does not allow to say, what rewards are there, but the list of most preferred incentives (table 83 of appendix; p.144) indicate that there is a lot of potential for development, as social science researchers seem to want more of visibility (65%), new contacts (62%), that data sharing is considered relevant for evaluations (52%), etc.

What support is available to social science researchers regarding data sharing - this report does not tell much about the situation regarding support as it is now, but support for making research data available would be an attractive incentive for 30% researchers to start sharing their data (table 83 of appendix; p.144).

How often researchers document their data according to certain standards: 16% of social science researchers document their research data by using suitable standards; 28% individually and consistently, 37% individually and not consistently, 14% do not document their data (table 25, p.97). Thus would allow to say that this indicator scores 2, some researchers document their data.

3.2.2 Data archive service (DAS) proto-activities

Figure 3: DAS and OA activities implementation type in Austria



DAS activities

During the data collection phase, Austria is situated in a transition phase with regard to establish a DAS. The former data archive for the social sciences, WISDOM (Wiener Institut für sozialwissenschaftliche Dokumentation und Methodik), was closed in early 2014. In 2015, a working group started to develop a concept for a new build up data archive. As the data archive is currently in the conception phase it was not possible at this time to involve someone to take part in the survey.

The new data archive will start with its activities in January 2017 and will be funded for 3 years by the Federal Ministry of Science, Research and Economics (BWF). Indeed, this Ministry published an action plan in 2015, in which the establishment of a data archive for the social sciences within the framework of CESSDA is explicitly mentioned. The action plan also stated that existing European ESFRI infrastructures (e.g. SHARE, CESSDA, ESS, CLARIN, and DARIAH) should be more widely used.

The new data archive, the Austrian Social Science Data Archive (AuSSDA), will be located at the University of Vienna with two additional locations in Graz and in Linz. The archive will be built on experiences from the former WISDOM archive and take over its archive holdings, so it will not start from the beginning.

Open Access (OA) support activities

The Austrian Science Fund (FWF) has launched at the beginning of the year 2016 the program “*Open Research Data Pilot*” that focuses on research data as part of an open science strategy. The goal in the medium-term is, that “*all research data and similar materials which are collected and/or analysed using FWF funds must be made openly accessible, whenever legally and ethically possible.*”³² Letters of Interest have been accepted until May 2016 and final funding decisions will be made in March 2017. 41 proposals were submitted, around 19 from SSH.

There is an initiative Open Government Data (<https://www.data.gv.at/>). The Cooperation Open Government Data Austria“, or short „Cooperation OGD Austria“ was founded by the Federal Chancellery, the cities of Vienna, Linz, Salzburg and Graz on the 13th July 2011. Federal Government, States, cities and towns in cooperation with the communities, science, culture and the economy are willing to set the basic agreements for the future of Open Government Data in Austria. By agreeing on common standards, an effective framework for the benefit of all stakeholders should be created. The collaboration between Germany – Austria – Switzerland- and Liechtenstein (D-A-CH-LI) will be intensified. The Cooperation represents the interests of ministries, states, cities and communities – to unite all actors who operate, plan or create an open government data platform. The Federal Chancellery is the link to Competence Center Internet Society (KIG) and the Platform Digital Austria.

3.2.3 Conclusions

To sum up, it seems that social science research in Austria can be characterized as advanced. Even if Austria is currently in a special situation while starting to establish the new data archive for the social sciences it will not start from the beginning. The archive will be built on experiences from the former WISDOM archive and take over its archive holdings. Furthermore, Austria is member of CESSDA and confirms the commitments made in the ministerial action plan.

³² See <https://www.fwf.ac.at/en/news-and-media-relations/news/detail/nid/20160118-2166/>

3.3 Belarus

The analysis shows that Belarusian social sciences sector is underdeveloped. The country has no established data service for the social sciences. Research data management policies and ethical and legal frameworks important for data sharing do not exist in Belarus. The main source of social science research funding is the public sector.

Specifics about the data collection in a country

Information was collected mainly via the webpage of the National Academy of Sciences of Belarus (NASB, <http://nasb.gov.by>). There is no information available from the self-assessment survey provided by the CESSDA SaW project.

3.3.1 Broader ecosystem of DAS operation

Belarus, as a country having emerged from the former Soviet Union, has a tradition of education and research. The National Academy of Sciences of Belarus is the most important scientific organization in Belarus. The NASB is responsible for organization, coordination, and implementation of basic and applied research performed by all public scientific organizations in natural sciences, engineering, social sciences, the humanities and arts. The Academy is accountable directly to the Council of Ministers of the Republic of Belarus. The highly strong status of the NASB is largely a legacy from the Soviet times.

Another important policy-making body in the area of research and development (R&D), carrying out the policy, exercising control in the area of research and innovation is the State Committee on Science and Technology of the Republic of Belarus (SCST). The SCST is ranked as a ministry for science, technology and innovation.

The research system in Belarus includes also higher education organizations that operate in the field of research and development. Major institution here is the Belarusian State University.

On the other hand, there are few independent non-public research institutions as well, for instance, the Independent Institute of Socio-Economic and Political Studies (<http://www.iiseps.org/?lang=en>).

In general, a development of research infrastructures is not a priority in Belarus. Total public funding of research and development in Belarus in 2012- 2014 was less than 1 per cent of GDP or extremely low. An intensity of investment in SSH is not known, but from funding practices of similar countries to Belarus is known that funding for SSH as usual is low.

Development of social science sector

Overall assessment of SSH development

The Academy of Sciences holds most of the country's research capacity. Within the NASB, there are few centers of social research, i.e. Institute of Economics (IE) and Institute of Sociology (IS). The main research areas of the IS are labour and economic sociology, social stratification and demography, sociology of culture and social psychology, social policy. The IS is focused on studies of management of a national economy, political economy, regional economy, market economy and market infrastructure.

Universities have considerable input to SSH development as well. Belarusian State University is a leading higher educational establishment in the country. A Faculty of Philosophy and Social Sciences with three departments, i.e. Sociology, Psychology and Social Communication, is the major player here.

Development of research data production in SSH

The NASB Institute of Sociology is conducting public opinion surveys and producing empirical data. The main research areas of the IS are labour and economic sociology, social stratification and demography, sociology of culture and social psychology, social policy.

The Centre of Sociological and Political Research (CSPR) of the BSU (http://www.cspr.bsu.by/?page_id=63) was conducting empirical studies for the United Nations, UNICEF, UNESCO, EBRD and other international organizations since 1990-s. The CSPR was a participant of several well-known international surveys as European Values Study (EVS), Comparative Study of Electoral Systems (CSES), New Democracies Barometer, LLH: Living conditions, Lifestyles and Health, Eurasia Barometer. In addition, the CSPR has carried out electoral studies in Belarus since 1989.

Among non-public companies in opinion research one of the most important institution is the Independent Institute of Socio-Economic and Political Studies, which is conducting annual polls "Monitoring of a public opinion in Belarus" since 1992. Among research of the IISEPS are "Strengthening pro-European attitudes in Belarus", "Belarus on the threshold of the XXI century: ways of transformations", "Belarus: prospects in the XXI century" and so on.

National Statistical Committee of the Republic of Belarus is the major producer of statistical data. The National Statistical Committee is conducting few surveys as the Household Survey.

RDM Policy setting

Funders' data management and sharing strategy and/or policy

Since sharing of research data is not common in Belarus, requirements about preparing Data Management Plans are non-existent for social science research data.

Legal and ethical framework

There is no reliable data about development of legal and ethical framework for data production and management as well as on data sharing.

Data sharing culture

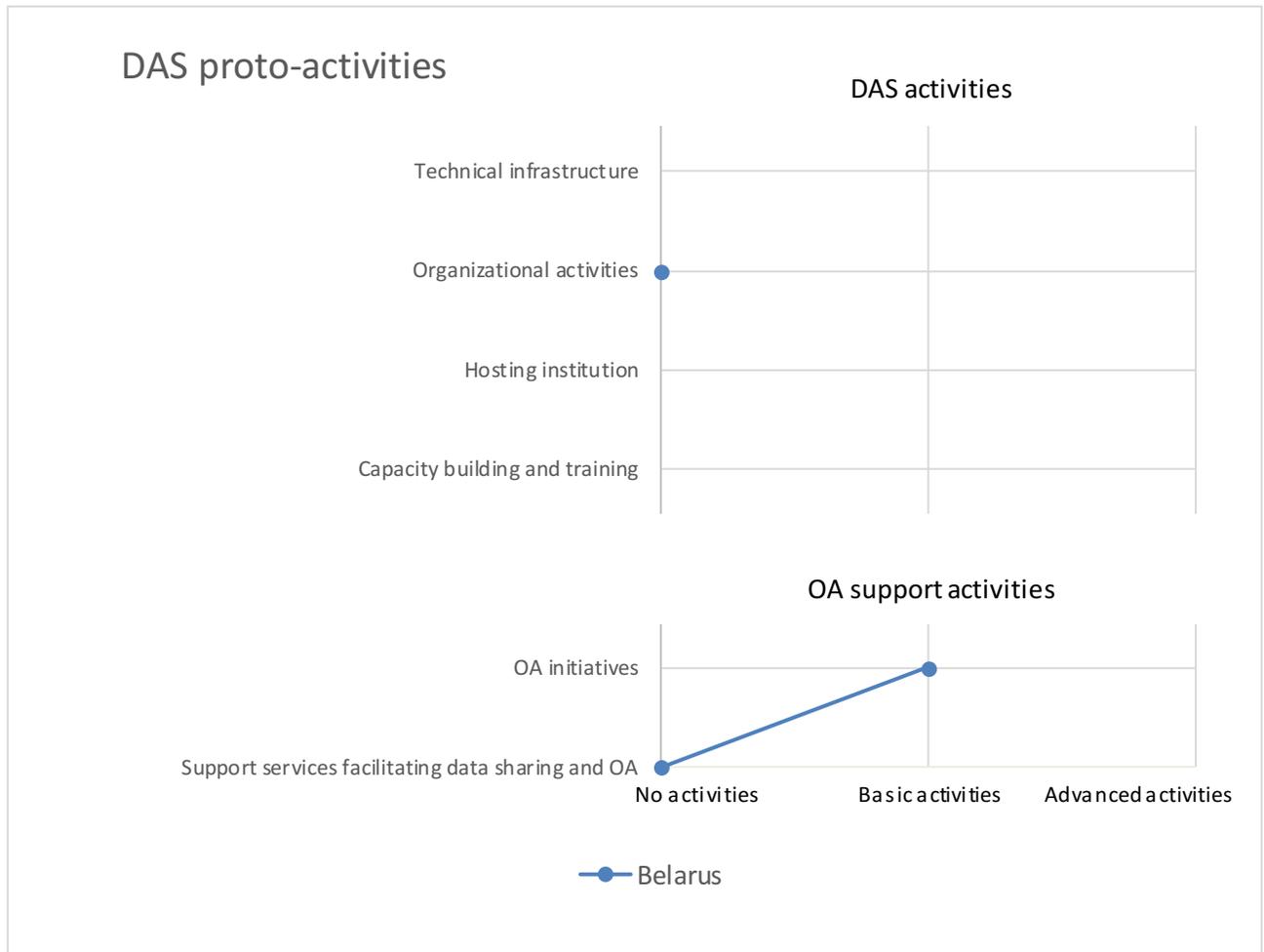
There is no data available on data sharing practices in Belarus. Indirectly it is possible to point out that some data is acquired directly from project and personal websites with individual permissions.

Some Belarusian data is also available via various international projects and data archives. For example, New Democracies Barometer with a data on Belarus is available through the UK Data Service.

3.3.2 Data archive service (DAS) proto-activities

There are no existing national data archiving infrastructures in Belarus. Also, Belarus is one of the late entrants into the Open Access (OA) movement largely due to a lack of widespread information about OA available to researchers. The European Federation for Intercultural Learning (EFIL) supports Belarus in developing OA awareness through Belarus Council of Libraries on Information Cooperation Network. The National Library of Belarus provides methodological supports for OA initiatives.

Figure 4: DAS and OA activities implementation type in Belarus



3.3.3 Conclusions

A development of research data production in SSH is at the initial phase in Belarus. Ethical and legal framework for data sharing is non-existent. Data sharing and reuse is rare in the social science research community. Still the Open Access policies might be a helpful instrument to increase awareness about data management and data sharing in Belarus.

3.4 Belgium

SOHDA is the Belgian preparatory project for sustainable archiving and sharing of research data in the humanities. Because it is a preparatory project not much activity in data archiving is developed yet. The assessment of the CESSDA Saw Project comes for SOHDA too early to draw proper conclusions. It should however be mentioned that a well-equipped Belgian social science data archive (BASS) was operational up until the 1990's. Although no longer operational and requiring substantial efforts for data curation, BASS still constitutes a solid foundation for a revived Belgian social science data archive. These mixed origins (the old BASS and the new but limited scope SOHDA-project) should be understood when reading the Belgian country report.

Specifics about the data collection in a country

The data collection took place in the month October 2016. The information was provided by Johan Surkyn (SOHDA) and Jean-Paul Sanderson (SOHDA/BASS). Before they completed the questionnaire, we had contact via SKYPE with Johan Surkyn to explain the questionnaire and the context of the questions.

3.4.1 Broader ecosystem of DAS operation

Figure 5: Heading concepts values in Belgium

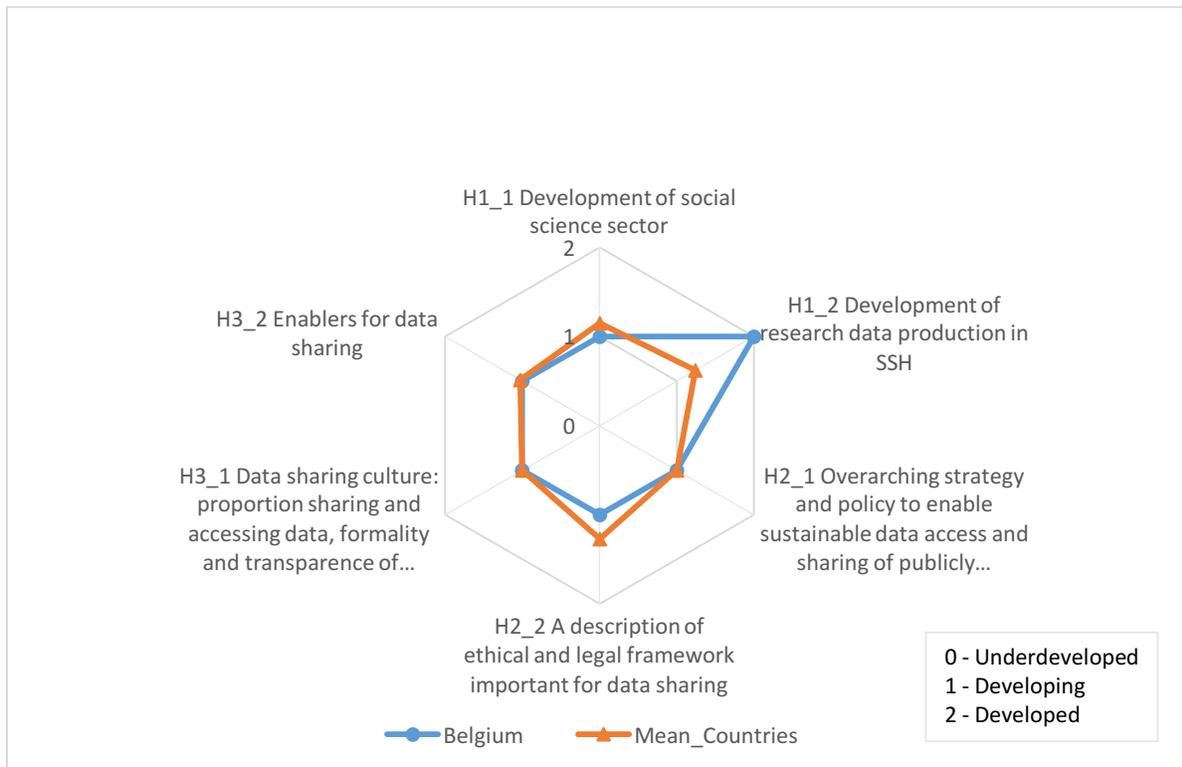


Figure 5 shows that the production of research data in the Social Sciences and the Humanities in Belgium is well developed. Data sharing culture and Research Data Management Policy is, like in most other European countries, starting to develop.

Development of Social Science Sector

Overall assessment of SSH development

The financial stability, research capacities and results achieved in the field of social sciences are on the average. Universities provide access to commercial bibliographic and full-text databases and software licenses.

Development of research data production in SSH

In Belgium, there are well-established streams of research traditions, national and international and there is great variety of important types of research data. Most relevant studies like ESS, SHARE, SILC, GGP, EVS, EES, LFS, Time Use Survey, CCS, ISSP, CSES, PISA and GGP are all conducted. Obviously, the European Census Program is followed. Besides the international surveys and programs, a number of national studies such as post-election studies, migration surveys and divorce studies, are produced.

RDM Policy and Support Setting

The general situation in Belgium with regards to requirements or recommendations about preparing Data Management Plans (DMPs) as an integral part of on-going project activity is initial. There is growing recognition and awareness of need to require Data Management Plans (DMP's). "Initial" describes only the present state. It seems certain that (post-approval) DMP's will become obligatory very shortly at Flemish universities, and the new DMP policy is likely to be officially endorsed before the end of this year. The state of affairs in the French language university system is yet to be confirmed.

Requirements or recommendations set by the Funding Agencies about archiving quality-assured social science research data with associated metadata is developing. There is growing recognition and awareness of the need to have (disciplinary specific) places of deposit and support services. Although the principle of data sharing of outcomes of publicly funded research is stated, it is not reinforced, and standards or provisions for depositing data are lacking.

There is growing recognition and awareness about the value of research data produced and about the need for long-term preservation. This is manifested by a multitude of smaller research data archives kept by research groups at most universities. BASS is probably the only one that went beyond the scope of their own research domain. None of these thematic archives has received public funding for archiving and curation work, at least concerning their own data holdings. However, some projects for data curation have been funded in the past, e.g. in cooperation with the state archives. This situation reflects especially the situation in Sociology, Demography and Political Science.

Cost for managing the data and preparing it for access can be implicitly covered in the overall research project budget. There is no active support in the sense of default expense categories or budget splitting taking into account DMP cost. However, proposals that do devote attention and budgets to data sharing (within the budget rather than as a supplementary cost) are likely to be credited for that. It is very questionable to what degree project proposal evaluations are equipped to assess the quality of proposed data management. In empirical research proposals in general,

there is an unbalance between pressure to revert to existing data as much as possible, and, on the other hand, lack of provisions to make any new data collection re-usable through archiving.

Legal and ethical framework for data sharing

The legal and ethical framework for data sharing is initial. There is growing awareness about the problem and about the need to provide clarification on legal aspects. However, at the moment, scarce or no organised support is given.

The open access movement has helped raising awareness and know-how relating to intellectual property rights and privacy in general, but not always applied to research data. University legal services probably vary in expertise and support. Concerning sharing of statistical data stemming from the statistical agency or from administrative data producers, the "Privacy-commission" is the official legal authority that assesses the balance between the degree of disclosure and the scientific (societal) value of the project. Alongside access-permissions, the privacy commission issues a set of rules and conditions that govern access to sensitive data in terms of research goals, ethics, access termination, etc.

Data sharing culture

It is not possible to estimate the proportion of social science researchers in Belgium that have shared the research data they produced in the period between 2011 and 2016. There doesn't exist a reliable source for assessing this matter. Surely, it varies a lot depending on research groups and their investment in sharing and in publishing their data holdings. Respondents of the self-assessment survey know of a few panel datasets that have been publicly available and shared in socio-demographic research over long periods of time (e.g. Belgian Household Panel survey).

A high proportion (>30%) of social science researchers in Belgium, as estimated in the self-assessment survey, were able to access existing third party data between 2011 and 2016. It is very clear that budgetary limitations have pushed researchers towards using existing data. The social security databank (combining information from many federal state departments) and the population census by the statistical office are today among the most important providers of research data in social science. In the absence of estimations of intra-academic data exchange, respondents of self-assessment survey still think it is safe to say that at least 30% of research activity was based on third party existing data.

The most favourite routines for data sharing is via project or personal websites, followed by informal contacts with peers and colleagues. As mentioned before, administrative data are still growing in importance. In a number of cases, these data providers have developed strong data services with associated user groups. These data services were categorised by the respondents as "*project or personal websites*".

Respondents of the self-assessment survey were not able to estimate what the overall attitude of social science researchers in Belgium is, concerning data sharing; whether Belgian researchers think that data sharing has no benefits at all; that data sharing creates healthy competition in research; that data sharing creates negative competition; that data sharing can answer new research questions; that others may misuse and misinterpret data, nor that the Belgian researcher think that data sharing involves little effort and minimal costs.

Even though some of the objections to sharing are sometimes heard (risk of misinterpretation), it is not possible to generalise. The same for the advantages of sharing, that seem to be generally accepted.

Enablers for data sharing

The data sharing culture in Belgium is comparable to other European countries. There are a lot of indirect career rewards related to data sharing for social science. The respondent stated that in his 30-year professional career he has witnessed a clear link between academic success and an open attitude towards sharing data and expertise for himself and among colleagues. *“This is because data sharing is just one of many forms of academic interaction, all of which benefit from openness and exchange. Also, academic evaluation is essentially done by peers. Turning the question around, the idea that there would be any benefit in keeping research data inaccessible for external users is about as silly as keeping a manuscript of a strong research article from getting published in fear of others stealing the ideas presented in it.”*

According to respondents of the self-assessment survey, there are no domain specific data support services provided to social science researchers in Belgium, that offer methodological support or facilitate data sharing and/or Open Access to research data. Some data producers follow data management and data documentation standards and procedures. Except for international projects with mandatory data management planning, planning and provisions for reuse are probably the exception rather than the rule.

3.4.2 Capability requirement areas of DAS

Organisational profile

Organisation

Social Sciences and Humanities Data Archive (SOHDA) is the Belgian preparatory project for sustainable archiving and sharing of research data in the humanities. The project is funded by Belgian Federal Science Policy. It is a joint enterprise of Interface Demography at Vrije Universiteit Brussel, Institute for the Analysis of Change in Contemporary and Historical Societies (IACCHOS) at the Université Catholique de Louvain (UCL) and the Belgian State Archives.

One of the aims of the SOHDA project is to provide support for reviving the former Belgian Archive for the Social Sciences (BASS)³³ of the Université Catholique de Louvain. However, the scope is more general: it is essentially about developing standards and business models for an operational Belgian SS data archive and doing so in close cooperation with CESSDA.

SOHDA is the acting Belgian Service Provider in the CESSDA ERIC and as such, it can benefit from CESSDA support and shared expertise in preparation of becoming a full service provider. A new operational website is under development, which will be dedicated to archived data of studies carried out at both research centres.

³³ More information about BASS on the website: <http://www.ucl.be/358283.html>

Funding

Belgian Science Policy (BELSPO) finances the pilot project for the Belgian data archive.

At the moment, there is no direct funding for archiving work yet. However, research on BASS contents is financed by internal university research funding, and this work contributes to BASS sustainability.

Core services and activities

Because SOHDA is a preparatory project, there are no real data archiving activities yet. BASS is best described as a "sleeping" archive, which has a very limited level, of mainly internal, use. There is no active data acquisition. Some testing on curation and metadata recovery is currently going on.

Content current collection

BASS has a collection of studies originated from 1972-1998. Major research themes are: censuses; community and regional studies; conflict, aggressions and violence studies; education studies; political sciences and government structures; health care and social services; international relations; European integration; political attitudes; behaviour studies; social indicators; organizations; demographic and family characteristics.

BASS content is physically located at UCL university's central archives. There is limited internal use by researchers, and in rare cases of inter-university exchange for testing purposes.

The linguistic composition of the archive's collections is mainly French and English.

Organizational infrastructure

Figure 6: Organizational infrastructure in Belgium

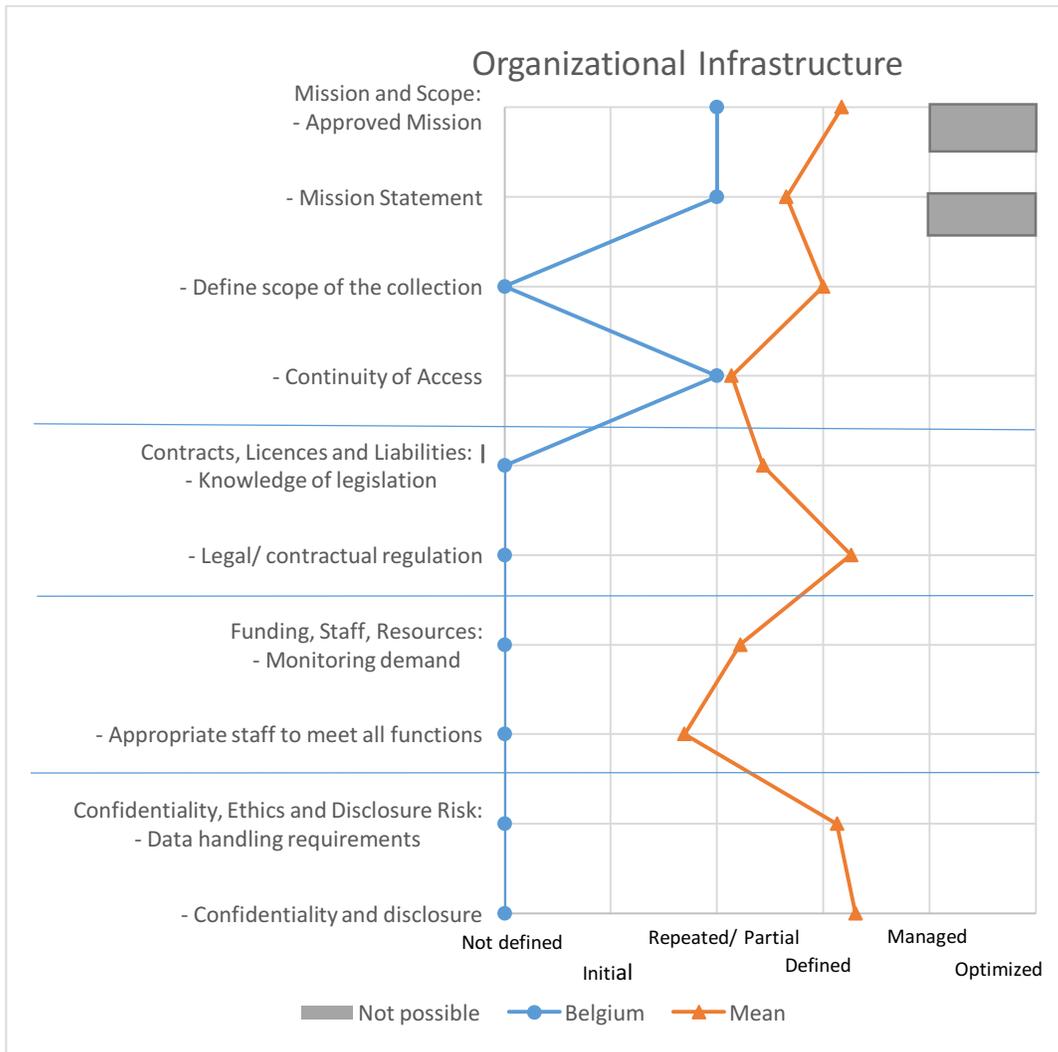


Figure 6 shows that SOHDA is a preparatory project; mission is defined but the real activities are not developed yet.

Mission statement

It must be stressed that the answers in this paragraph describe the BASS archive that ceased to be fully active almost 20 years ago, and did not develop since.

The BASS founded in 1969 and active until 1998, had a dual mission of dynamic and selective archiving on the one hand, and of providing information and services open to users, on the other. Two services were offered: the archiving of statistical datasets and assistance during data processing. The archive provide access to documented datasets (codebook and questionnaire) to faculty researchers.

Define scope of the collection

As SODHA is a preparatory project and there are no real data archiving activities yet, the scope of the collection is not yet defined.

Continuity of Access

Having now become part of UCL-university's general archives, the BASS data are subject to the same regime of backup systems as the rest of the university archive. However, this excludes a presently unknown part of datasets that are stored on unmounted tapes and disks.

All the other aspects of the organisational infrastructure, such as knowledge of legislations, legal regulations, appropriate staffing are not applicable to the Belgian situation, because real activities are not developed yet.

Digital object management (data curation) and Technical infrastructure and risk

Figure 7: Digital object management and Technical infrastructure in Belgium

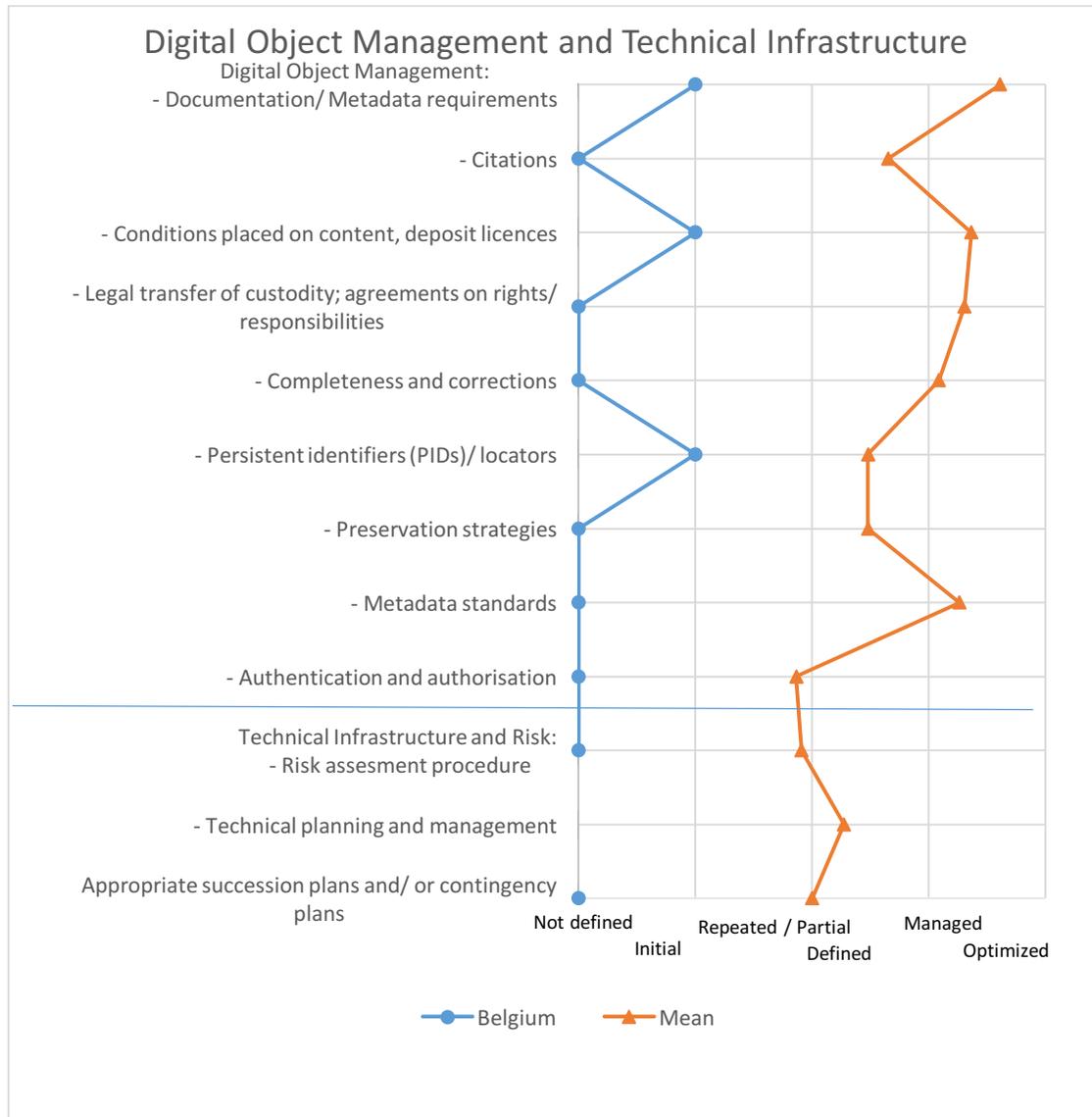


Figure 7 shows that SOHDA is a preparatory project; most of the aspects of digital object management and technical infrastructure are not yet defined and activities are yet not developed.

Documentation and metadata requirements

When it was active, the BASS archive systematically collected metadata. The metadata schema consisted of eleven standardized elements and a content description. The information in the metadata is more or less sufficient to meet DDI requirements. A data sheet accompanies each data file containing the metadata and the degree of accessibility.

Conditions on content and deposit licences

At BASS, access conditions were set up and agreed upon when required by the depositor, but most deposits are done without any conditions on use. The following access conditions were defined: A = no restriction; B = limited to Belgian user community; C = prior notification to the depositor; D = prior authorisation by depositor required

Persistent identifiers

The Dataverse test case used in Belgian SOHDA project made use of DOIs. At BASS, datasets are uniquely identifiable using an ad hoc identification system.

Preservation strategies

At the moment for BASS, no preservation strategies are in place, and an unknown number of tapes and disks might be affected by degradation of storage media, the obsolescence of media drives, and accidental digital corruption of preserved material. Assessing these risks is therefore pressing.

3.4.3 Conclusions

SOHDA is the Belgian preparatory project for sustainable archiving and sharing of research data in the humanities. Because it is a preparatory project, not much activity in data archiving is developed yet. The assessment of the CESSDA Saw Project comes for SOHDA too early to draw proper conclusions.

3.5 Bosnia and Herzegovina

In Bosnia and Herzegovina (BiH), currently only 0.1% of GDP is invested in the social sciences. However, BiH is the country that spends the most in the social sciences relative to other scientific disciplines. Infrastructure conditions, in terms of databases/datasets/software access, are minimal, and overall access to relevant databases is limited. The production of research data by social science institutions (i.e. university institutes, and a few social-science research centres within the NGO sector) is “rare, periodic and ad hoc”. However, the 2012 SERSCIDA survey of social science researchers on “Gathering, preserving and use of research data” showed that 64% of the participants had produced at least one dataset during the last five years. On the other hand, as in many European countries data sharing and reuse are not common in BiH, there are no established formal channels for data sharing, and general attitudes of researchers related to sharing of data are mixed.

Specifics about the data collection in the country

Data were collected through two main channels: the self-assessment survey and desk research. Two experts working in the social sciences in Bosnia and Herzegovina (BiH) answered the online self-assessment survey. Desk research was also used. Information was principally gathered from reports of international capacity-building projects involving Bosnia and Herzegovina, such as SERSCIDA, RRPP, and SEEDS.

3.5.1 Broader ecosystem of DAS operation

Figure 8: Heading concepts values in Bosnia and Herzegovina

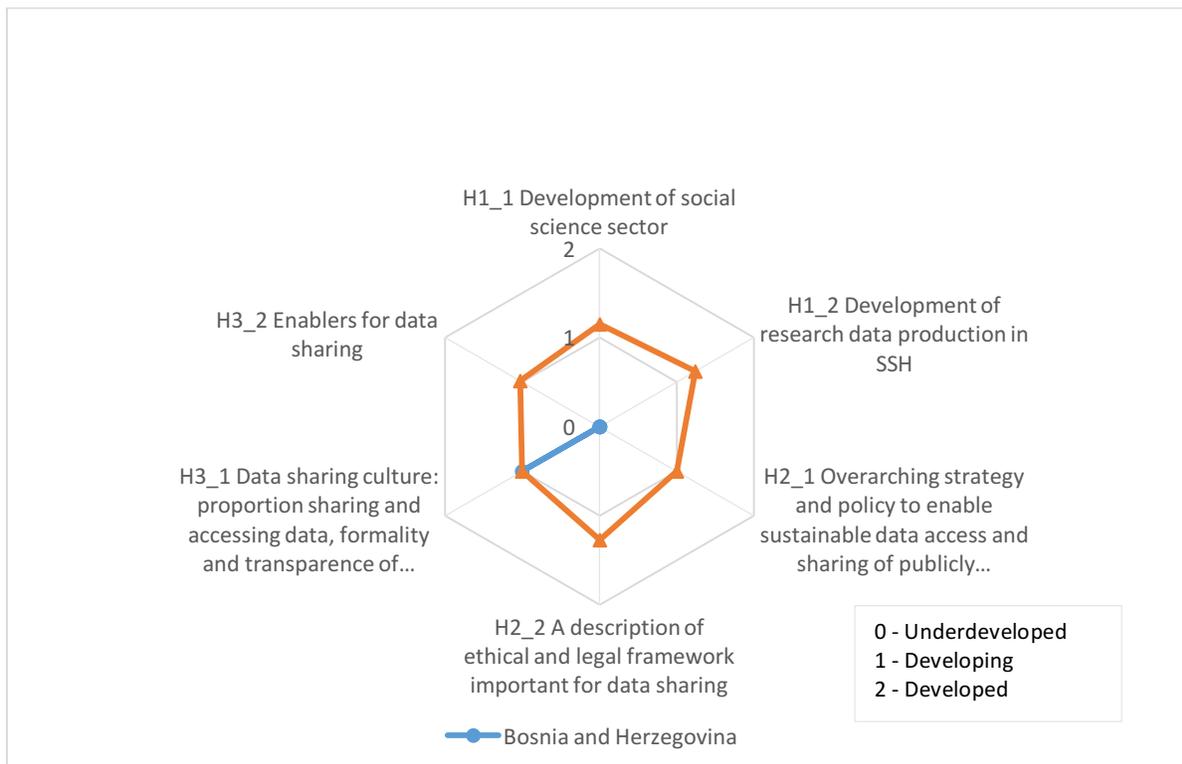


Figure 8 shows that the broader ecosystem of data archive service operation in Bosnia and Herzegovina has been categorized as underdeveloped for all indicators, but for one about data sharing culture which has been categorized as developing.

Development of the social science sector

Overall assessment of SSH development

In Bosnia and Herzegovina, funding in social science research mainly comes from abroad. The government and the higher education sector constitute the second funding sources (sources: self-assessment survey and SERSCIDA report³⁴). An RRPP policy brief³⁵ also mentioned that “the number of social sciences projects financed by the state/entities is too miniscule”.

Currently, 0.1% of GDP is invested in the social sciences in Bosnia and Herzegovina, which is under the mean (0.14)³⁶. This financial support has been qualified as “inadequate” by RRPP37. In comparison with other scientific disciplines, the social sciences receive 37.8 % of the gross domestic expenditure on research and development (GERD). Among the available figures, BiH is the country that spends the most in the social sciences relatively to other scientific disciplines³⁸. The number of researchers in the social sciences per capita (24) is however among the lowest³⁹.

Also, infrastructure conditions, in terms of databases/datasets/software access, are minimal. Indeed, some universities provide access to EBSCO and other full-text databases (i.e. WoS, Scopus, JSTOR, etc.). However, overall access to relevant databases is in principle limited. Access to datasets or to licenses for specific research software (as STATA, SPSS, etc.) is not provided by the government or universities.

Development of research data production in SSH

According to the respondents, “numerous studies” on matters of national importance were produced by social science researchers during last 10 to 15 years in Bosnia and Herzegovina. Nevertheless, the production of research data by social science institutions (i.e. university

³⁴ SERSCIDA, 2012, Analysis of existing potentials for the establishment of a social sciences digitised data service in Bosnia and Herzegovina, http://www.serscida.eu/images/deliverables/SERSCIDA_D_2_1_Mapping_Report_Bosnia_and_Herzegovina.pdf, accessed on 20.12.2016.

³⁵ RRPP, Policy brief, Social Sciences and Research in Bosnia and Herzegovina: (Is There?) Sufficiently Developed Personnel Potential, http://www.rrpp-westernbalkans.net/en/library/Policy-Briefs-and-Policy-Studies/Bosnia-and-Herzegovina/mainColumnParagraphs/02/text_files/file/pol_brief_human_resources_eng.pdf, accessed on 22.12.2016.

³⁶ Social sciences spending in GDP percentages were available for 29 countries (among the countries target within this report). The range is going from 0.03% to 0.061%, the median is 0.12%.

³⁷ RRPP, Policy Briefs and Studies - Bosnia and Herzegovina, <http://www.rrpp-westernbalkans.net/en/library/Policy-Briefs-and-Policy-Studies/Bosnia-and-Herzegovina.html>, accessed on 22.12.2016.

³⁸ Social sciences spending in GERD percentages were available for 31 countries (among the countries target within this report). The minimum is 4.1 %, the mean 13.7% and the median 12.9%.

³⁹ The number of researchers in social sciences per capita was available for 27 countries (among the countries target within this report). The range goes from 13 to 247. The mean is 90.5 and the median 85.

institutes, and a few social-science research centers within the NGO sector) is “rare, periodic and ad hoc”, according to one of the Bosnian experts.

For example, among the eight international surveys selected, Bosnia and Herzegovina only participated in two, namely the European Values Study (EVS) and the World Values Survey (WVS). BiH will take part in PISA 2018 (Table 3.5.1).

However, the SERSCIDA survey⁴⁰ on “Gathering, preserving and use of research data”, conducted in 2012 and that involved 139 students and researchers active in the social sciences, showed that 64% of the participants had produced at least one dataset during the last five years.

Table 3: Participation of Bosnia and Herzegovina in a selection of international surveys

Comparative Candidate Survey (CCS)	No
Comparative Study of Electoral Systems (CSES)	No
European Social Survey (ESS)	No
European Values Study (EVS)	Yes
Generations and gender programme (GGP)	No
International Social Survey Program (ISSP)	No
Programme for International Student Assessment (PISA)	Yes (2018)
World Values Survey (WVS)	Yes

Research Data Management Policy Setting

Currently, in Bosnia and Herzegovina there is no policy, nor strategy to enable sustainable data access and sharing of publicly funded social science research data. Also, no support is provided yet on legal and ethical aspects that facilitate social science data sharing (IPR, data protection, etc.).

The elaboration of a Data Management Plan (DMP) and the deposit of data in an appropriate disciplinary repository are not yet required by public funders of social science research. The public funders do not allocate resources to cover the cost for managing the data and preparing them for access during the research project. There is also no long-term curation of valuable research data assets, evaluated and selected in terms of reuse potential.

Data sharing culture

Data sharing and reuse among social science researchers in Bosnia and Herzegovina is not common, as the self-assessment results indicate, the proportion of researchers sharing data is estimated as low (less than 10%). The proportion of researchers able to access existing third-party data they need was not available. The culture and practice of data sharing seems to be underdeveloped. Data are shared primarily via personal contacts (peers and colleagues) that lack

⁴⁰ SERSCIDA, 2012, Analysis of existing potentials for the establishment of a social sciences digitised data service in Bosnia and Herzegovina, http://www.serscida.eu/images/deliverables/SERSCIDA_D_2_1_Mapping_Report_Bosnia_and_Herzegovina.pdf, accessed on 20.12.2016.

formality and transparency. To obtain data from other researchers is considered “extremely difficult” by the respondent of the self-assessment.

The SERSCIDA 2012 survey indicates that in 33% of the cases, the research data were accessible only by the research team members. However, 38% of the respondents stated that their data were publicly available (20%) or available for the scientific community (18%). Those last results should be put into perspective, as first, the respondents of the survey were self-selected, so a bias might occur with respondents more inclined to data sharing and using of secondary data; and second, these results are inconsistent with the answers given to the question “Where are the data from your last project kept? Indeed, the data are mostly stored on the personal computer of the research team members. Only 6% store their data in an archive or repository or on their institution server. According to the SERSCIDA survey, the respondents are in favour of a wider access, the data should ideally be publicly available for 45% and accessible by the scientific community for 26%. Moreover, 90% of the respondents said that certainly (51%) or probably (31%) would provide their data to a safe and regulated social science data archive. Finally, data sharing is considered as important (24%) and very important (75%) for their respective discipline by 99% of the respondents.

The SERSCIDA 2012 survey showed that most of the respondents have never analysed secondary data, whether it be quantitative data (47%) or qualitative data (65%). The obstacles cited for using secondary data are: few relevant data available (24%), the inaccessibility of the existing data (23%), the poor documentation of accessible data (22%), the lack of knowledge to perform secondary analysis (17%), and the research culture (13%). The secondary data are obtained principally from the websites of national and international projects (26%), from their personal network of colleagues (19%), and from data archives from other countries (15%).

Table 4: Attitudes towards data sharing in Bosnia and Herzegovina

Data sharing has no benefits at all	False
Data sharing creates healthy competition	Probably true
Data sharing creates negative competition	Probably true
Reuse of existing data can answer new research questions and facilitate advancement of science	True
Data sharing has as a risk that others may misuse and misinterpret data	Probably true
Data sharing involves little effort and minimal costs	Neither true, nor false

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five point scale from “5”-True to “1” False!

The attitudes of researchers toward data sharing in the social science community, as the self-assessment shows, can be characterized as ambivalent. In one hand, according to the respondent of the self-assessment, social science researchers in BiH think in general that data sharing has benefits. They consider that reuse of existing data can answer new research questions and facilitate advancement of science. On the other hand, they think that there is a risk of data misuse and misinterpretation, as well as a risk of negative competition (for example, being scooped and therefore reduced publication opportunities) for the researcher.

To summarize the data sharing culture in Bosnia and Herzegovina, data sharing and reuse is not common, there are no established formal channels for data sharing, and general attitudes of

researchers related to sharing of data are mixed. Thus, the data sharing culture can be characterized as underdeveloped (0).

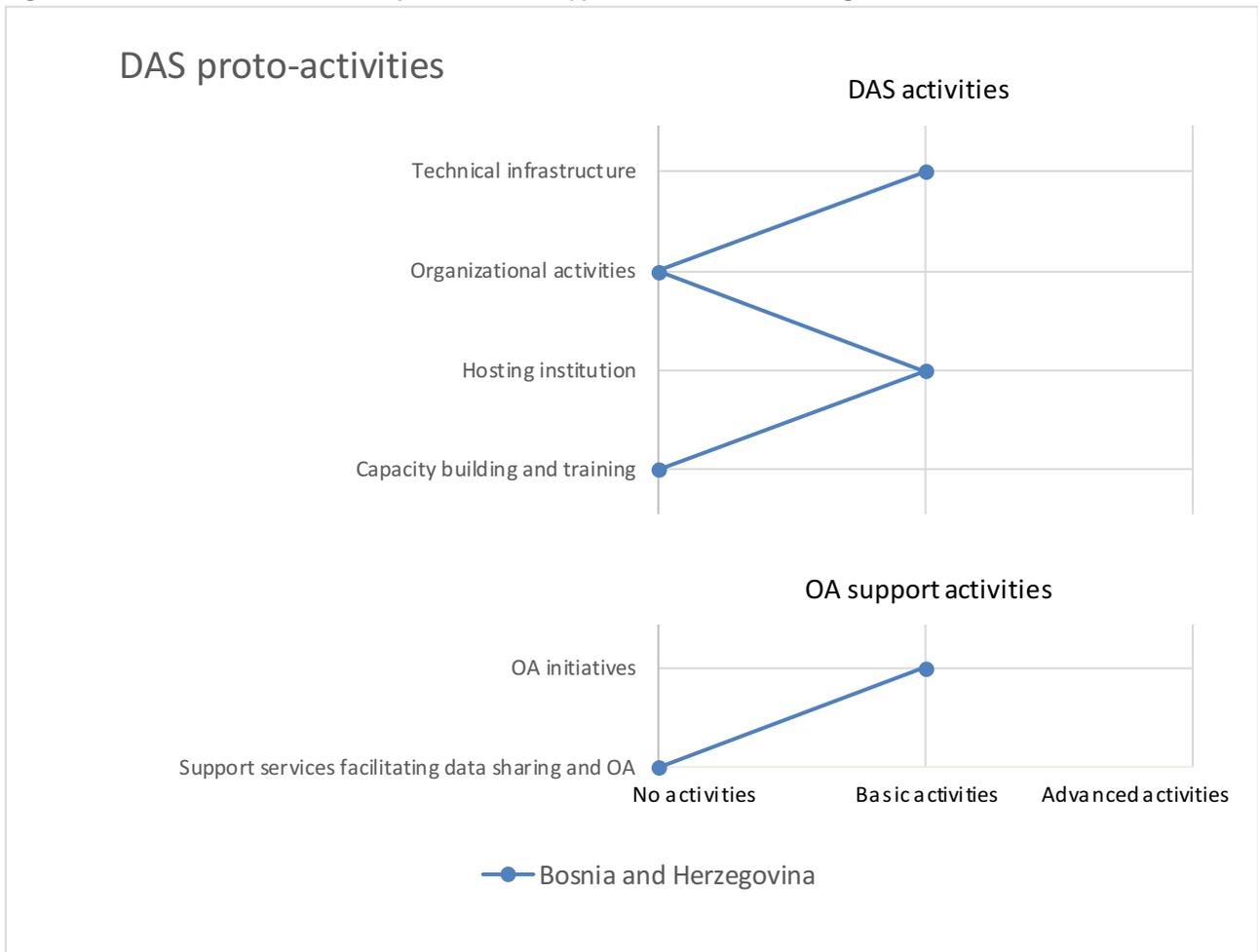
Enablers for data sharing

The incentives and enablers for data sharing within the social sciences research community in Bosnia and Herzegovina are currently non-existent and are therefore benchmarked as underdeveloped (0). Indeed, there is no career reward (e.g. influence on career progression, higher success rate in obtaining research funding, better standing within the research community) related to data sharing within the academic community. Data support services that could facilitate data sharing and/or open access to research data are not yet available to social science researchers.

According to the SERSCIDA 2012 survey, if almost all respondents (97%) stored somehow (personal computer, institutional repository, etc.) their data, only 5% stated that their data were well documented, following institutional or international standards. These results show that currently few social science data producers follow data management and data documentation standards and procedures that facilitate data sharing and reuse.

3.5.2 Data archive service (DAS) proto-activities

Figure 9: DAS and OA activities implementation type in Bosnia and Herzegovina



DAS activities

Bosnia and Herzegovina has currently no established Data Archive Service (DAS), but reflections and activities towards the establishment of a DAS were initiated within the SERSCIDA project⁴¹. Nonetheless, a few services are provided on an institutional level through other projects. Different institutions were cited during the self-assessment survey, which were partially identical to the SERSCIDA 2012 survey results, as technical infrastructures in the social sciences that could possibly be used for or applied to a new DAS:

- The Human Rights Centre of the University of Sarajevo (HRC Sarajevo, <http://hrc.unsa.ba>) was founded in December 1996. Its mission is to contribute to the implementation of international human rights through information dissemination processes and providing documentation and research. Digitization and archiving related projects have been part of the activities of HRC Sarajevo in the past 10 years. Indeed, HRC Sarajevo was the SERSCIDA partner for Bosnia and Herzegovina, and was the former local coordination unit within the international project “Regional Research Promotion Programme” (RRPP, <http://rrpp-westernbalkans.net>).
- Analitika – Center for Social Research (<http://analitika.ba>) is a non-profit NGO established in July 2009 in Sarajevo, whose mission is to enhance the public policy process by conducting socially relevant and high quality research. Analitika is currently involved as the local coordination unit for Bosnia and Herzegovina in the RRPP project, which aims to ensure the long-term preservation of and access to data produced within the project from 2008 to 2016. Through RRPP, Analitika is also in contact with partners of the SEEDS project (South-Eastern European Data Services, <http://seedsproject.ch>), which include FORS and ADP, two CESSDA service providers.
- PHAIDRA (Permanent Hosting, Archiving and Indexing of Digital Resources and Assets) is an institutional repository system for permanent archiving, indexing and use of digital objects that was established within the Tempus project “New Library Services at Western Balkan Universities” (<http://westbulnet.com>). In BiH, the project was carried out by the University library of Sarajevo (<http://phaidra.ba>) and the University library of Tuzla (<http://phaidra.untz.ba>). One of the aims of the project was the implementation of European standards in a university learning environment.
- The National and University Library of Bosnia and Herzegovina (<http://nub.ba>) and the National and University Library of the Republic of Srpska (<http://nub.rs/>) were also cited in the self-assessment survey. A repository for scientific work (<http://registar.nub.ba>) is available through the National and University Library of Bosnia and Herzegovina, including research projects, institutions, registered researchers and bibliographies. This repository is linked with the Slovenian Current Research Information System (SCIRIS, <http://sicris.si/>).
- The survey conducted within SERSCIDA also mentioned the Agency for Statistics of Bosnia and Herzegovina (<http://bhas.ba/>), the Institute for Statistics of the Federation of BiH (<http://fzs.ba>), the Institute for Statistics of the Republic of Srpska (<http://rzs.rs.ba>), the Computer Center of the University of Sarajevo (UTIC, <http://utic.ba>), and the Archives of Bosnia and Herzegovina (<http://arhivbih.gov.ba>).

⁴¹ SERSCIDA - Support for Establishment of National/Regional Social Sciences Data Archives, began in 2012 and ended in 2014, <http://serscida.eu>.

The Universities of Sarajevo and Banja Luka as well as the National and University Libraries of Bosnia and Herzegovina and the Republic of Srpska were also named during the self-assessment survey as institutions that could host a DAS for the social sciences in Bosnia and Herzegovina.

The experts consulted did not know if there are some existing initiatives in BiH to develop the knowledge and skills of people who might be employed at a DAS, such as research and data management (mainly data management plans), data preservation and data access. Moreover, no services that facilitate data sharing are available through their own experience and knowledge.

Open access (OA) support activities

At present, in Bosnia and Herzegovina, support activities or services to encourage and facilitate open access are rare or not existing in the social science research community. According to the self-assessment survey, some NGOs publish and provide open access to publications, but this effort is not coordinated and does not ensue from a general policy on open access.

As mentioned above, data support services that could facilitate data sharing and/or open access to research data are not yet available to social science researchers.

3.5.3 Conclusions

While there is a significant level of social science research being conducted in Bosnia and Herzegovina, existing research infrastructure for the social sciences is underdeveloped, although a few past and ongoing initiatives show the potential for establishing national data archiving services. After considering the past and ongoing projects within Bosnian institutions, it is clear that knowledge about archiving and dissemination processes is already available in various institutions, like university libraries. Also, a research project inventory is already available.

Given the relatively low levels of funding for new data collections, existing data take on even more importance for their re-use potential. Long-term preservation of data and their dissemination is the next step. Some institutions are aware of the importance of the implementation of European standards in research and are active in order to achieve it. They could be interesting partners to establish a data archive in the social sciences in Bosnia and Herzegovina. While building up of the technical infrastructure for data preservation will be a precondition for putting into place a functional DAS, the general culture of data sharing must be improved so that researchers see the value and utility of making their data available to others. Researchers should be provided with the incentives and support needed to make this possible.

3.6 Bulgaria

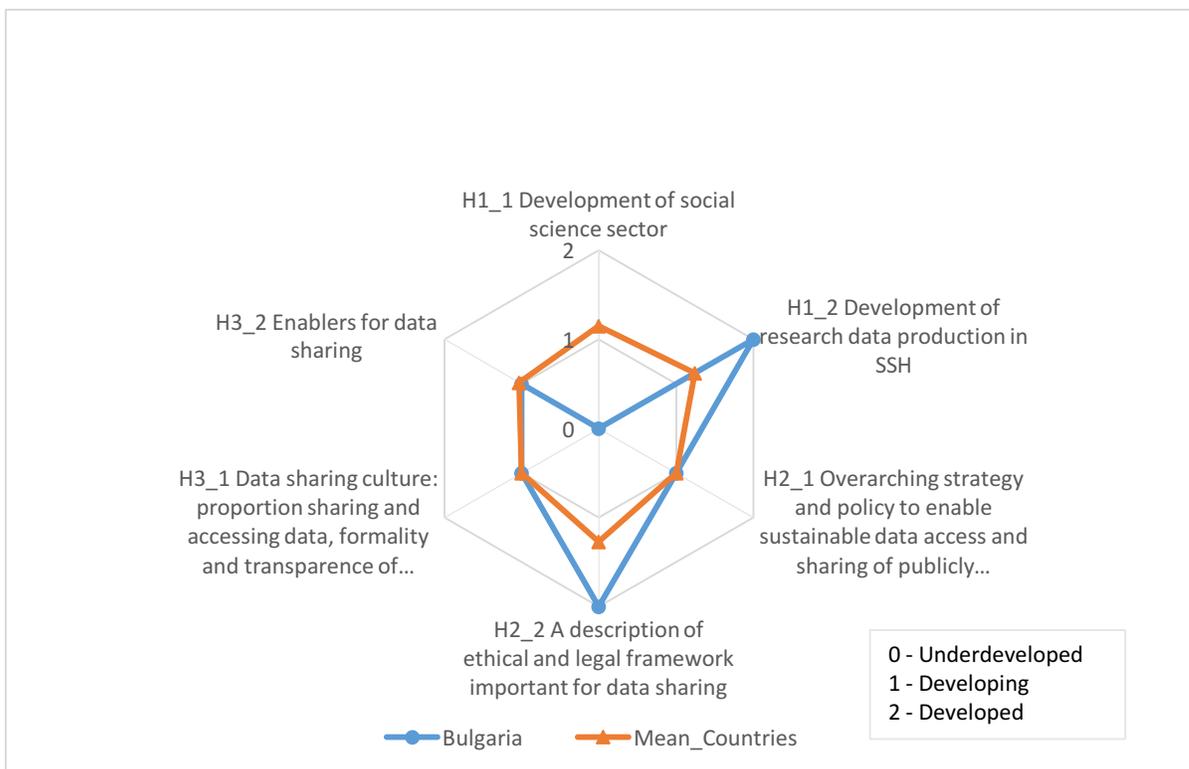
Bulgaria is a country where no DAS activity, and only very limited DAS proto activity has been identified. For data production, funding from EU sources is indispensable.

Specifics about the data collection

To collect data in Bulgaria, an initial attempt to contact the person in the CESSDA SaW contact list (Dona Pickard) was made. After not receiving a response, a contact provided by RODA, who was responsible for help with contacting Bulgaria was used. After sending a first email to the contact person at the Bulgarian Academy of Sciences on October 18, the final answers to the questionnaire were provided by Petia Ilieva Trichkova from the Institute for the Study of Societies and Knowledge of the Bulgarian Academy of Sciences on November 2 2016. Additional, answers related to DAS proto-activities, which were not asked in the first communication, were received on January 5 2017.

3.6.1 Broader ecosystem of DAS operation

Figure 10: Heading concepts values in Bulgaria



Development of social science sector

According to the Bulgarian National Statistical Institute, the main sources of funding of social sciences in Bulgaria are the government and higher education sectors, followed by the business enterprise and the private non-profit sector. Information on the share of funding from abroad is provided by Eurostat (indicator rd_e_fundgerd) and is available only for the whole research and development segment in Bulgaria. In this perspective funding from abroad covers over 50% of the total funding of research and development in Bulgaria, followed by the government and enterprise sectors.

The overall assessment of financial stability, research capacities and results achieved in the social sciences for Bulgaria was characterized as follows:

“As regards to the budget offered to the Bulgarian higher education institutions and the Academy of Sciences (which is very low), the funding from additional sources such as the Horizon 2020 Framework Programme (also FP7 & FP6), Bulgarian-Swiss Research Programme, different operational programmes or other sources (e.g. The Bulgarian Science Fund, but it provides little and not regular support) increase undoubtedly the financial stability, research capacities and results for the social sciences institutions in Bulgaria. Most of the scholars still publish in publications in Bulgarian which make them more accessible for the community. Despite that, the participation in multi-national projects and collecting primary data on various topics (in many cases these data are cross-national) provide good opportunities for collaboration with international teams and for publishing the project results in international volumes, in peer-reviewed journals and in journals with an impact factor. In many of these projects, the stakeholders are well-embedded as a part of these projects. Besides this, from time to time scholars are searched for expert opinions for different Ministries. So, I would say that there is visible impact, although it is still limited.

However, seen in a wider European perspective, the country is with the lowest success rate for applications to Horizon 2020 per EU Member State and overall is among the countries with the lowest share of EU financial contribution to participants in signed grant agreements for both Horizon 2020 and FP7. To the best of my knowledge, so far Bulgaria has not been a coordinating country in these projects in the sphere of social sciences. It has always been involved as a partner-country. In terms of the productivity, the majority of higher education institutions and the Academy of sciences currently have a small number of publications in Scopus and the Web of Science in the case of social sciences. There is only one Bulgarian university in one of the world-class ranking systems—QS World University Rankings: the University of Sofia, which was among the top 700 in 2013 and 2014 and ranked just below the top 700 in 2015. Given this, compared to other countries, the impact on the wider community is very small.”

With regard to systematically assessing matters of national importance, Bulgaria is heavily reliant on its participation in cross-national surveys. As was stated in the survey „National surveys are also conducted on a wide range of topics, but in general they are not so systematic. They depend on the funding available. In the case of electoral surveys: they intensify before elections and they are mainly carried out by private agencies.“

With regard to data production, social sciences were hit by the 2008 crisis and only lately, a revival is felt. Another problem felt by the respondent was that priority was given to natural sciences.

RDM Policy setting

There is initial recognition and awareness of the need to require DMP and of the need to have disciplinary specific place of deposit and support services. The awareness about the value of research data produced and about the need for long-term preservation is also initial.

We were not able to establish any existing incentives for depositing data in an appropriate disciplinary repository, long-term curation for valuable research data assets or covering the cost for managing the data and preparing it for access in Bulgaria.

The situation in the area of Legal and ethical framework is comparatively better as the Bulgarian Sociological Association has its own ethical code which deals with publishing and sharing of the results of sociological surveys.⁴²

Data sharing culture

Data sharing and reuse among social sciences researchers in Bulgaria is low – as self-assessment results indicate, the proportion of researchers sharing data is estimated as low (0-10%), and the proportion of researchers able to access existing third party data they need – is low (0-10%).

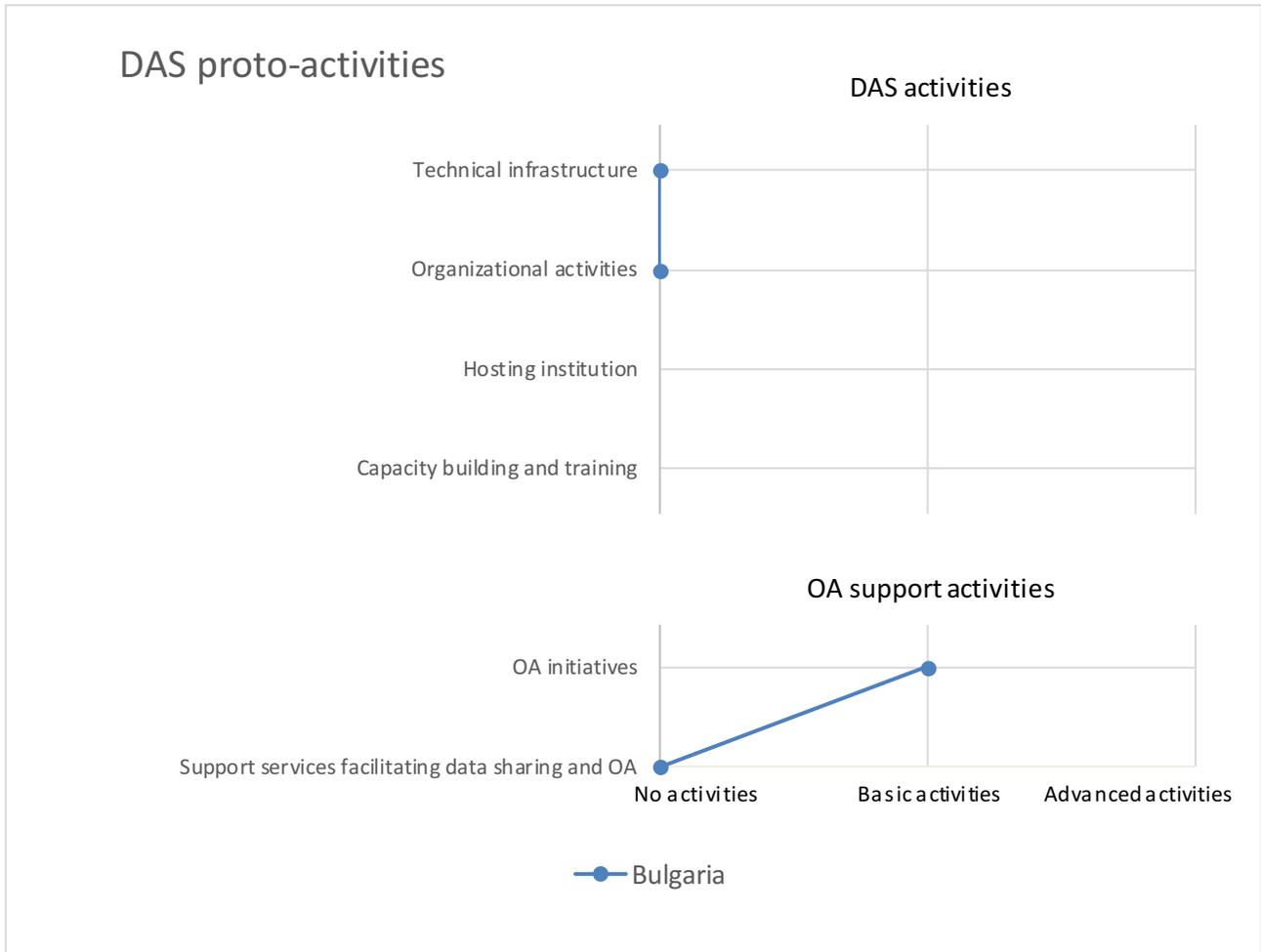
The most common channel of data sharing is via informal contacts, data archive or repository and project or personal websites are used less often.

No direct career rewards related to data sharing were mentioned in the interview. Still, according to our contact some data producers follow data management and data documentation standards and procedures.

⁴² Available only in Bulgarian: http://www.bsa-bg.org/documents/ethical_code.pdf

3.6.2 Data archive service (DAS) proto-activities

Figure 11: DAS and OA activities implementation type in Bulgaria



DAS activities

Bulgaria has currently no established DAS. There are also no existing technical infrastructures that could possibly be used for or applied to a new DAS.

Previous DAS proto-activities in Bulgaria included the project Establishment of Social Science Data Archive in Bulgaria at the Regional and Global Development Research Centre (REGLO) in Sofia and are documented in the CESSDA PPP FP7 project.⁴³ These activities have been discontinued due to lack of funding.

Currently, one proto-activity towards establishing a data archive for the social sciences involving the Institute for the study of societies and knowledge of the Bulgarian Academy of Sciences has been identified. The institute is in the process of preparing a project proposal with representatives

⁴³ See https://ppp.cessda.net/doc/D7.1_Report_and_Recommendations.pdf

from the European Academies Internet Gateway (AGATE), project coordinated by the Union of the German Academies of Sciences and Humanities and funded by the Federal Ministry of Education and Research (BMBF) between October 2015 and March 2017. The project proposal aims at establishing an Internet platform – archive for projects in the area of social science and humanities. However, since the aim is to create an archive of projects in the SSH area, this activity has only a loose connection with establishing a data archive in Bulgaria.

Open access (OA) support activities

With regards to open access, no supporting activities in Bulgaria were identified. However, Bulgaria participated in the expanded OpenAIREplus project in 2011 - 2014. Also, several institutional open access repositories exist in other fields- e.g. at the Institute of Mathematics and Informatics at Bulgarian Academy of Sciences, Sofia University "St. Kliment Ohridski" and New Bulgarian University.⁴⁴

3.6.3 Conclusions

No DAS activity and only very limited DAS proto-activity has been identified in Bulgaria. The social science ecosystem is rather undeveloped with regards to RDM policies or data sharing culture.

⁴⁴ See http://www.math.bas.bg/infres/IS-publ/IS-2011-Simeonov_Stanchev-DiPP.pdf

3.7 Croatia

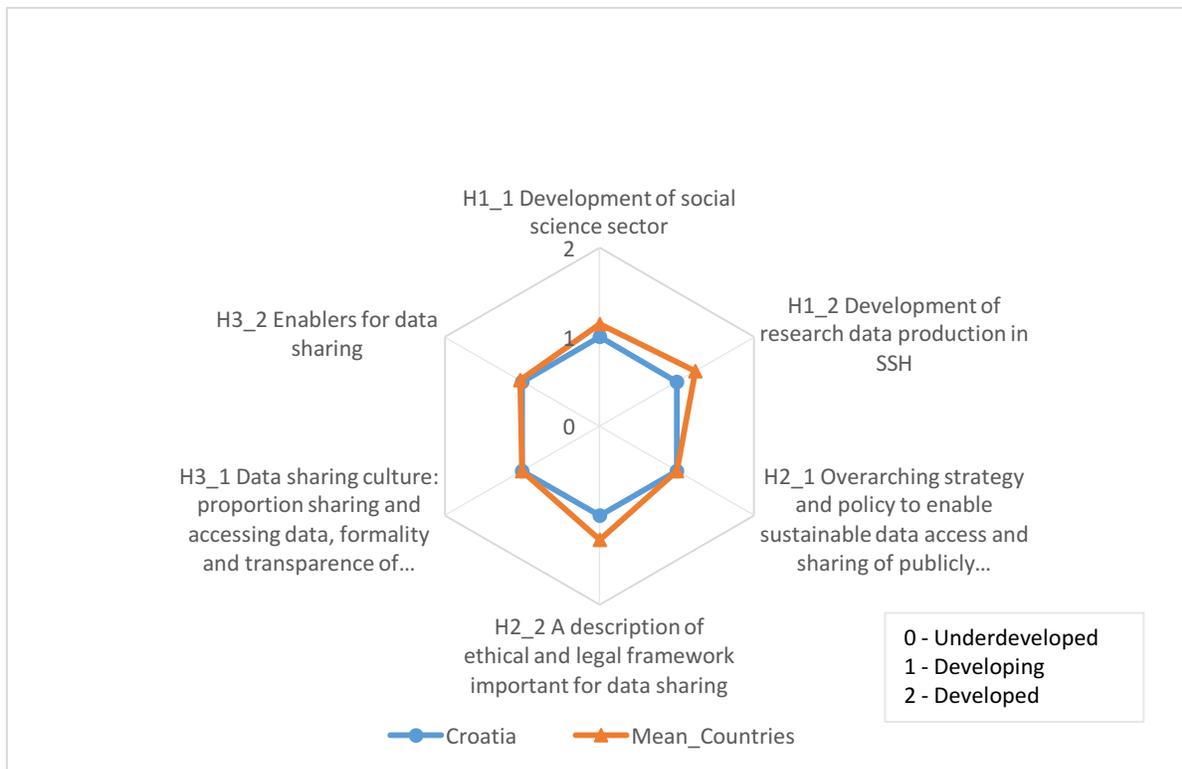
The analysis shows that the social sciences sector is still developing in Croatia, as well as the research data production, research data management policies and ethical and legal frameworks important for data sharing. Data sharing and reuse is not very common, informal channels have precedence over formal ones and researchers consider data sharing to involve lots of effort. Incentives and enablers for data sharing are also still developing. However, Croatia has basic DAS prototype activities established through the SERSCIDA and SEEDS project. It is expected that the support environment and the culture of data sharing will gradually get off the ground.

Specifics of data collection in country

The information was entered by Marijana Glavica (librarian at Faculty of Humanities and Social Sciences, University of Zagreb), Irena Kranjec (librarian at Faculty of Humanities and Social Sciences, University of Zagreb) and Alen Vodopijevac (head of IT section at the Centre for Scientific Information at the Ruđer Bošković Institute). The main method used was self-assessment survey. Self-assessment scores were followed with extended explanations of answers in the notes and with references to written resources. The results from the SERSCIDA survey were used as appropriate.

3.7.1 Broader ecosystem of DAS operation

Figure 12: Heading concepts values in Croatia



Development of social science sector

Overall Assessment of SSH development

In Croatia the social sciences and humanities sector is developing. Funding of the SSH and productivity of the researchers in Croatia are in the mid quantile; impact on designated community is limited.

The general intensity of investment in SSH is 0.11 of the GDP. Relative to other disciplines, the intensity of investment in SSH is 13.5% and the number of researchers in SSH per capita (pop. 100,000) is 63. The investment in human resources in SSH is 17.883 (per researcher). The most important source of research funding in SSH in Croatia is the government and higher education sector, followed by foreign funding (international or cross-border), the business enterprise sector and last, the private non-profit sector.

The government or university provides access to commercial bibliographic and full-text databases and to software licences but not to datasets. *“Access on the national level is provided by the Office for coordination of e-resources acquisitions via the Portal of e-resources for Croatian academic and scientific community (<http://baze.nsk.hr/>). Some faculties and institutes acquire additional resources for their own needs, including additional full-text databases, datasets and software licences. On a national level, only Microsoft products licences are acquired.”*

Development of research data production in SSH

In Croatia the research data production in social sciences is developing. There are some examples of research excellence, particular streams of research stand out, either qualitative or quantitative, and some examples of international collaborative research.

Croatia is involved in some international collaborative research and cross-national studies, i.e. ESS (2008, 2010), ISSP, CSES(2006-2011, 2011-2016), EVS (1999, 2008), WVS (1996), PISA (2006, 2009, 2012, 2015), while it has never participated in the CCS and GGP studies. Apart from the international studies, there is also a regular post-election study conducted by the Faculty of political science. Moreover, there are some studies conducted by agencies but this is not a regular activity. The average production of research data is periodical production, i.e. *“institutions have tradition in producing some type of research data to a certain extent”*.

Moreover, *“results of the survey conducted in 2012, during the SERSCIDA project, shows that 83% of the researchers have produced some sets of data in the past five years.”* ([SERSCIDA report](#)). Even if this estimate is based on a self-selected sample of researchers, together with the evidence of international collaboration and some continuous national studies of importance, this shows established tradition of high quality research data production. This could be taken as incentive for establishing systematic research data management service nationally, in particular in the field of political science and some others social science fields.

RDM Policy setting

Research data management policies in Croatia are developing. Currently Croatia's National Science Foundation does not require RDM but there is growing awareness by public research funding organizations to consider offering incentives for sharing research data with associated metadata:

- There are initial recommendations and requirements for DMPs, i.e. a growing recognition and awareness of need to require DMP. *“There is still a lack of knowledge about DPMs among public funders. Researchers are becoming familiar with it through international projects. Several workshops have been conducted by the Rudjer Boskovic Institute in Zagreb, supported by OpenAIRE and FOSTER projects, about open science in general and about data management in particular. DMP issues have been presented on several conferences in the area of information science. During the SERSCIDA project, funders were informed about the importance of data archiving and implementation of data management plan was suggested as an integral part of project application.”*
- In the initial phase is also offering or depositing data in an appropriate disciplinary repository, i.e. there is growing recognition and awareness of the need to have disciplinary specific place of deposit and support services. *“There is still a lack of knowledge about depositing data in appropriate disciplinary repository among public funders. As with DMP, researchers are becoming familiar with it through international projects.”* The grass root initiatives are arising which shows a need to coordinate National Science Foundation policies and requirements among different scientific disciplines: *“DMP plan was also recently presented at Open Access Week aimed at researchers in the area of natural sciences.”*
- Long-term curation for valuable research data assets, evaluated and selected in terms of reuse potential is also only in the initial phase, i.e. there is growing recognition and awareness about the value of research data produced and about the need for long-term preservation; scarce or no investment and support for long-term curation provided. *“Together with activities mentioned in previous answers, we cannot say that the need for long-term curation of datasets is fully recognized among funders, but there are some activities going on in the country which contributes to rising awareness about data management and archiving for long-term use.”* Wider national initiatives could help in raising awareness about needs for long-term curation of national heritage in general: *“...there is also growing interest about long-term preservation among some infrastructure or cultural heritage institutions, particularly traditional archives.”*
- Initial is also the involvement in the costs for managing the data and preparing it for access, i.e. the cost for managing the data and preparing it for access can be implicitly covered in the overall research project budget.

Data sharing culture

The data sharing culture in Croatia is developing. Data sharing and reuse among social sciences researchers is not common – as self-assessment results indicate, the proportion of researchers sharing data is estimated as low (0-10%)⁴⁵, and the proportion of researchers able to access

⁴⁵ “The results the from SERSCIDA survey conducted in 2012 for the period 2008-2012 can be used to give estimates, because not much has probably changed in the last five years. When interpreting the results of this survey, it has to be kept in mind that the respondents were self-selected, so not representative for the whole population of social scientists. We can assume that the respondents were more inclined to data sharing and secondary use practices because they showed the interest for the survey. Only 10% of respondents stated that the data are available to the academic community or the public. It is indicative that this practice was more frequent among the internationally financed projects. In 20% of the cases only a team leader had a privilege to use data for secondary analysis. However, the dominant

existing third party data they need – as medium (10-30%)⁴⁶. There is growing awareness about importance of providing support on legal and ethical aspects of social science data protection but no organised support is given.

There are established data sharing channels and routines. The most popular data sharing channels are project or personal websites, which are ranked first. Data is shared also via informal personal contacts (ranked second) and data archives or repositories (third). *“Data sharing is not yet a common practice among researchers in Croatia, so there are also no routines for data sharing among researchers.”* The source of the data is the SERSCIDA survey (2008-2012) referenced in the previous paragraph.

The observations that the use of existing data sharing channels is already common and the perception that the competent service of archiving and data sharing could be best performed outside the research institutions with limited resources, both forms a strong incentive for the establishment of a national data service.

Attitudes towards data sharing

The attitudes of researchers toward data sharing based on self-assessment can be characterized as neutral. The country representatives in Croatia consider false the statement “Data sharing involves little effort and minimal costs” (5) and mostly false the statement that “data sharing has no benefits at all” (4). On the other hand, the statements “data sharing creates negative competition” (2), “reuse of existing data can answer new research questions and facilitate advancement of science” (2) and “data sharing has a risk that others may misuse and misinterpret data” (2) mostly true, while no evaluation was given regarding the statement about healthy competition.

scope of access, stated by somewhat more than half of the respondents, is that data access is limited to the research team that ran the project. It must be born in mind that this, is usually a very small group comprised of 2- 4 researchers, meaning that the scope and potential for data sharing (and capacity for secondary analysis) is very limited. Institutions to which researchers belong are seldom perceived as appropriate not only for archiving but for the data sharing as well. Only in 10% of the cases did other researchers in the institution have exclusive access to data.” See: http://www.serscida.eu/images/deliverables/SERSCIDA_D_2_2_Mapping_Report_Croatia.pdf (pp.21).

⁴⁶“Again, from the SERSCIDA survey, for the period 2008-2012: Currently, data for secondary analysis is obtained in various ways. Project websites or data repositories abroad were mentioned quite frequently as a source, and 33% of respondents mentioned using Croatian Bureau of Statistics to obtain data for secondary analysis, as this is an institution where a significant amount of research is conducted, resulting in macro and micro-data relevant for social science. Therefore, despite lack of national data archive, access to secondary data through formal channels is already common. The data is seldom obtained through colleagues working on the same project as researcher, or from the institution in which the respondents works; the most common way of obtaining the data for secondary analysis is to rely on one’s own professional network outside of institutions.” See: http://www.serscida.eu/images/deliverables/SERSCIDA_D_2_2_Mapping_Report_Croatia.pdf (pp.24-25).

Table 5: Attitudes towards data sharing in Croatia

Data sharing has no benefits at all	Probably false
Data sharing creates healthy competition	Estimate not available for 2011-2016
Data sharing creates negative competition	Probably true
Reuse of existing data can answer new research questions and facilitate advancement of science	Probably true
Data sharing has as a risk that others may misuse and misinterpret data	Probably true
Data sharing involves little effort and minimal costs	False

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from "5"-True to "1" False!

To summarize data sharing culture in Croatia, data sharing and reuse is not very common, informal channels have precedence before formal and transparent channels. Researchers consider data sharing to involve lots of effort and high costs. Although they see the benefits and potential for advancement in science, they are sceptical because of negative competition and the risk of misuse and misinterpretation. In summary data sharing culture can be characterized as developing.

Enablers for data sharing

Enablers for data sharing in Croatia are developing. There are no career rewards related to data sharing in the academic community and there are no data support services available to social science researchers that would facilitate data sharing and/or Open Access to research data, except for some informal support provided by library staff. However, a "prototype data services website was built during the SERSCIDA project (<http://www.hr.serscida.eu/>), which provides basic information on data depositing, including the Guidelines for data preparation. Through FOSTER workshops, a series of lectures about Open science, Open access to publications and Open access to research data were held in several Croatian cities, and presentations from the workshops are available online: <https://www.lib.irb.hr/foster/>."

There are also some data producers that follow data management and data documentation standards and procedures that facilitate data reuse. In the SERSCIDA survey, the "existence of raw data was acknowledged by 87% of respondents, "clean" and anonymized data sets in 75% of cases, and use of transformed data was reported by 68% of data producers. The existence of well documented datasets, described by metadata was mentioned by only 12% of participants. Out of those, most could not name the standards used for documenting and there was only one person that mentioned one international standard, while some them mentioned using in-house or institutional standards for data documentation."⁴⁷

⁴⁷ See: http://www.serscida.eu/images/deliverables/SERSCIDA_D_2_2_Mapping_Report_Croatia.pdf (pp.19-20).

In summary, the incentives and enablers for data sharing within social science research community in Croatia can be benchmarked as developing. There is a potential for emerging data service to continue with the education and training activities along the lines of open science and research data management in general.

3.7.2 Data archive service (DAS) proto-activities

DAS activities

In Croatia, no Data Archive Service (DAS) exists yet, but advanced activities related to technical infrastructure, organisation and capacity building are provided on a national and institutional level. There are also activities towards the establishment of a DAS for the social sciences.

DAS proto-activities started in Croatia during the SERSCIDA project (<http://www.serscida.eu/>) from 2012 to 2017 at the Faculty of Humanities and Social Sciences (FFZG) at the University of Zagreb. A website for data services was built - <http://www.hr.serscida.eu/> where the CRO-DAS (Croatian Data Archive for the Social Sciences) datasets are available for academic use. FFZG is also the Croatian partner currently involved in follow up international projects that aim to assist the establishment of a DAS for the social sciences, namely the SEEDS and CESSDA SaW projects. The Swiss funded SEEDS project (<http://seedsproject.ch>) helped FFZG to publish 10 datasets and to improve technical tools and develop human resources through workshops about data management organized through the SEEDS project. Within the CESSDA SaW project, the aims are to further plan for the establishment of the DAS and to strengthen the relationship between the future DAS and the policy makers and funders.

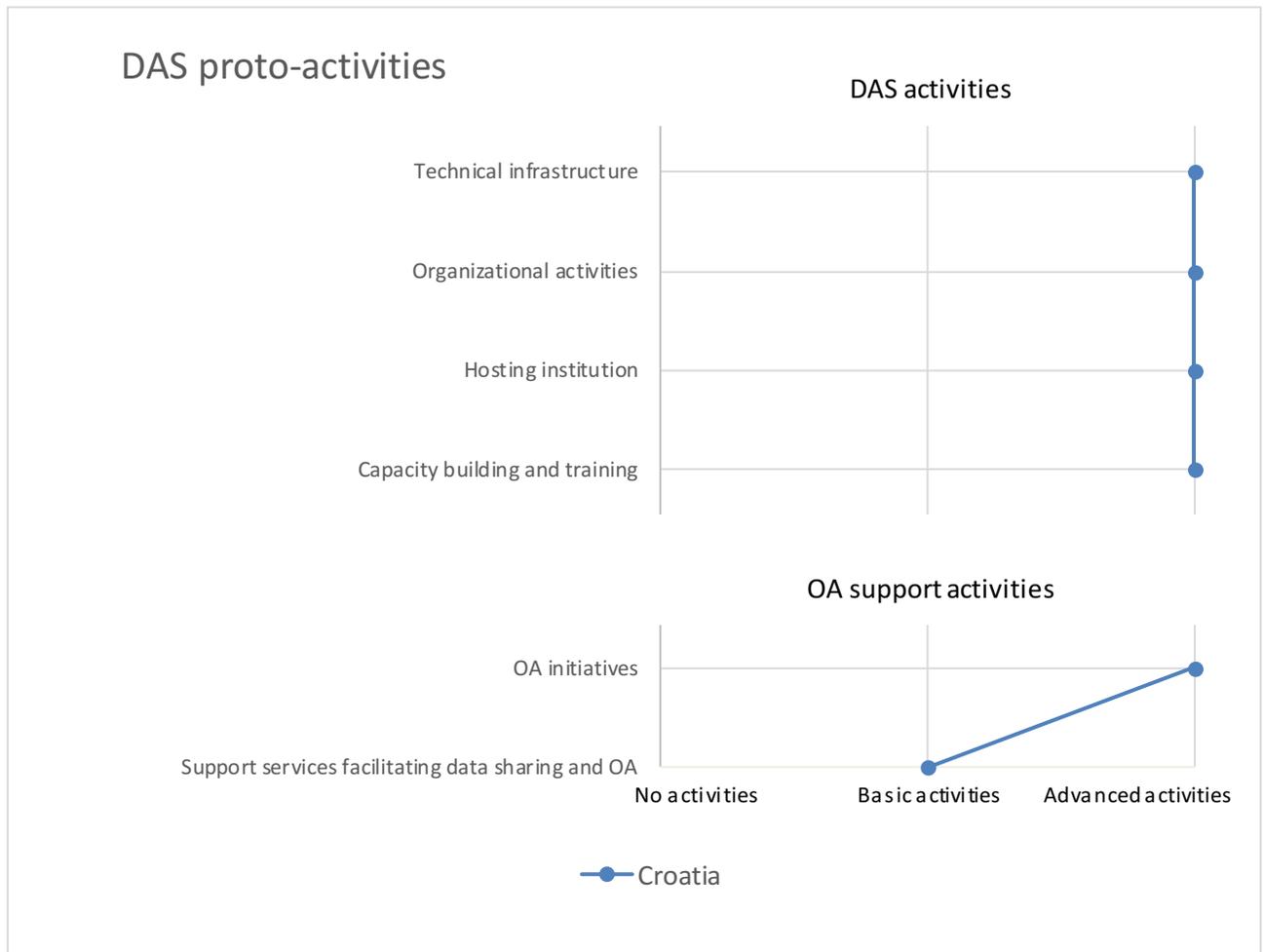
Furthermore, in 2015, DABAR (Digital Academic Archives and Repositories) was established as the key component of the Croatian e-infrastructure's data layer. It provides technological solutions that facilitate maintenance of higher education and science institutions' digital assets, i.e., various digital objects (i.e. journal articles, diploma theses and dissertations) produced by the institutions and their employees. DABAR also facilitates establishment and maintenance of a large number of reliable and interoperable institutional and thematic digital repositories and archives, without any costs to the institutions, i.e., DABAR users. By establishing a repository within DABAR, an institution is provided with a reliable, flexible and ready-to-use environment that can be used to collect, store and disseminate various digital objects, as well as maintain those digital collections. Cooperation between FFZG and DABAR is established and different models of using DABAR or similar infrastructure for DAS operations are currently explored.

The Faculty of Humanities and Social Sciences at the University of Zagreb is the institution that has the knowledge and infrastructure for hosting the DAS in Croatia, and is already involved in international projects and collaborates with other institutions in Croatia. The University Computing Centre at the University of Zagreb could be used for technical facilities for storing digital objects

In Croatia, there are already specific expertise on research and data management (mainly data management plans), data preservation and data access and initiatives to develop such knowledge and skills. The development of the knowledge and skills started during the SERSCIDA project, which is continued through the SEEDS, and CESSDA SaW projects. The issues about research data management have already been incorporated in some courses at the postgraduate program of the Faculty of organization and informatics in Varaždin, University of Zagreb. The main topics of the courses were:

- Data and information; Data life cycle; Exchange problems; Linked open data; Access and charging policy; Connection and integration policy; Security and integrity; Personal data and intellectual rights protection; Access to public information; Implication of these policies on the development of society⁴⁸.
- Research process documentation; Research project documentation; Primary and secondary data; Data organization and documentation; Metadata standards; Archiving and enabling access to data⁴⁹.

Figure 13: DAS and OA activities implementation type in Croatia



⁴⁸http://nastava.foi.hr/public/course?study=33958&course=45241&academicYear=2016%2F2017#nastavni_plan_i_program).

⁴⁹http://nastava.foi.hr/public/course?study=33958&course=80639&academicYear=2016%2F2017#nastavni_plan_i_program

Open access (OA) support activities

Croatia's supports activities to encourage and facilitate OA are mostly related to access to publications which are known and used by the social science research community. Most of the current scientific journals in Croatia, funded by public funds, are freely available via Hrčak - Portal of Croatian Scientific Journals (<http://hrcak.srce.hr/>).

In 2012, Croatian Open Access Declaration was published and signed by more than 650 individuals (<http://www.fer.unizg.hr/oa2012/declaration>). According to the Law on Science and Higher Education from 2015, open access publishing is mandatory for graduate thesis and dissertations. DABAR - Digital Academic Archives and Repositories (<http://dabar.srce.hr>) currently accepts diploma theses and dissertations and half of its content is in open access. Through FOSTER (Facilitate Open Science Training for European Research) workshops, series of lectures about Open science, Open access to publications and Open access to research data were held in several Croatian cities, and presentations from the workshops are available online: <https://www.lib.irb.hr/foster/>.

Croatia is also involved in OpenAIRE project, where 50 partners, from all EU countries, and beyond, collaborate to work on this large-scale initiative that aims to promote open scholarship and substantially improve the discoverability and reusability of research publications and data.

As mentioned above, there are no data support services available to social science researchers facilitating data sharing and/or Open Access to research data, except for some informal support provided by library staff and basic information available at some websites.

3.7.3 Conclusions

According to the conducted analysis, it can be concluded that:

- Funding of SSH and productivity of the researchers in Croatia are at the mid-level of development and that impact on designated community is limited.
- Development of research data production in SSH is developing. There are some examples of research excellence and international collaborative research.
- Research data management policy is not required by the national institutions and data management plans are still developing.
- Ethical and legal framework important for data sharing is developing but no organised support is given.
- Data sharing and reuse is not very common in the social sciences research community; informal channels have precedence over formal channels and researchers consider data sharing to involve lots of effort and high costs but the area is still developing.
- Incentives and enablers for data sharing within the social science research community in Croatia are currently not existent, except for some informal support. Thus, it can be benchmarked as developing.

Croatia has basic DAS activities through the SERSCIDA and SEEDS project that have created capacities in terms of technical infrastructure, organisation and human resources. Individual

people were trained to work, basic protocols and technical solutions are prepared and pilot operations of data services will get off with the promised data from 10 studies. Gradually it is expected that the support environment and the culture of data sharing will get off the ground. Activities for training and informing the scientific community about open access can be seen.

It is expected that also the support environment on the level of liabilities and principles of open access policies will follow. The research sector development is in the mid stage, which is a possible obstacle for the organisationally sustainable functioning of data services. This could be overcome with connecting different scientific information services on the national level.

3.8 Cyprus

SSH sector in Cyprus is well established, at least when it comes to the relative amount of money allocated. Cyprus does participate in some important international research series, but when it comes to the research data production in general, it is quite limited and irregular.

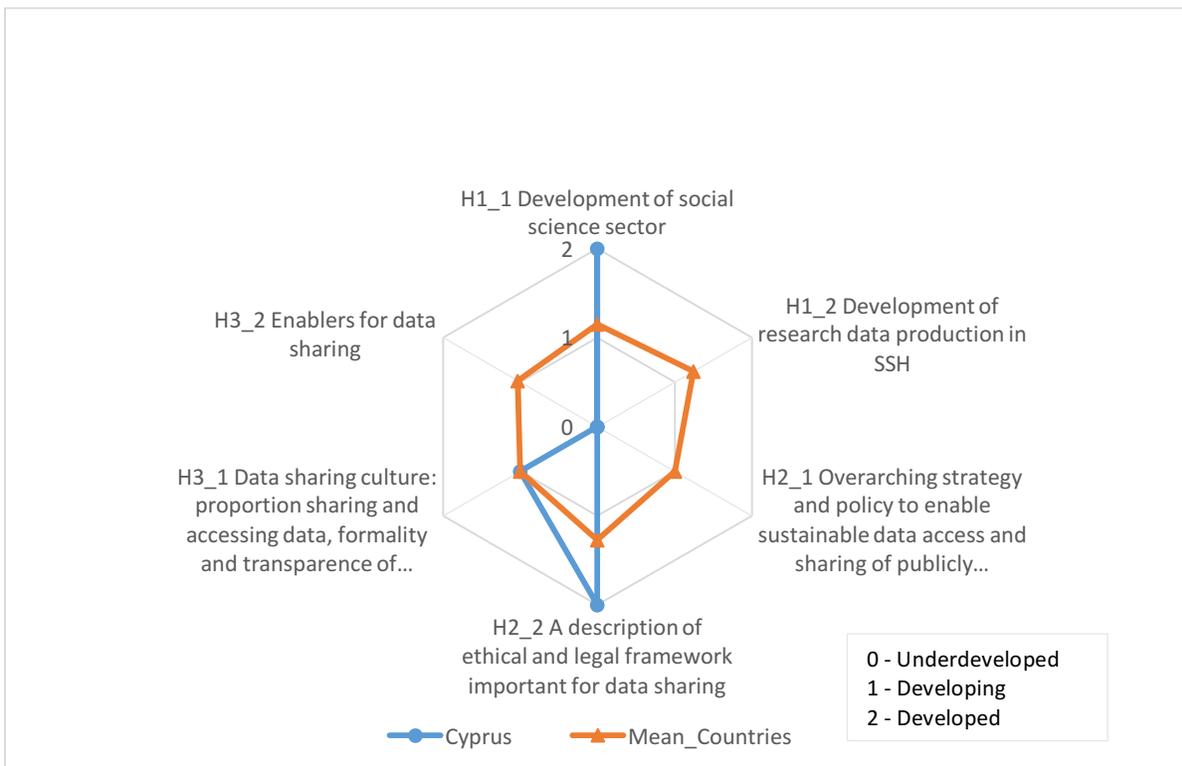
The data sharing culture in the country is underdeveloped since only small part of the research community in SSH shares produced data and comparatively small part of the researchers is able to get to the needed data for the secondary analysis. According to the report of Directorate General for Research and Innovation, at the moment, there is a lack of open access policies, institutional strategies or subject-based initiatives for research data; and no steps towards development in this area is foreseen for the nearest future. Important reason behind this is probably the fact that in Cyprus, the service for the social sciences data is not yet established and the preparation of such a system is not planned for the moment.

Specifics about the data collection in the country

Data for this report were collected mostly through email communication with Maryos Vryonides (Associate Professor of Sociology and Research Methods at the European University Cyprus, active in international research collaboration, e. g. leading the ESS survey in Cyprus) and through the desk research. Two more experts were contacted in September and October of 2016 (Christos Aspris, delegate for CESSDA SaW in Cypriot public administration and Constantinos Phellas, professor of sociology at University of Nicosia) but no feedback was received.

3.8.1 Broader ecosystem of DAS operation

Figure 14: Heading concepts values in Cyprus



Development of the social science sector

Overall assessment of SSH development

According to Vryonides the rank of the sources (sectors) of research funding in social science in Cyprus by the amount they provide is following: 1st abroad (international or cross-border), 2nd government and higher education sector, 3rd private non-profit sector, 4th business enterprise sector. This ranking is in accordance (where comparable) with the official Eurostat data on SSH – „Total R&D personnel and researchers by sectors of performance, sex and fields of science" [rd_p_perisci] and "Total intramural R&D expenditure (GERD) by sectors of performance and fields of science" [rd_e_gerdsc].

Table 6: Assessment of SSH development in Cyprus

Indicator	Values for Cyprus	Rating of Cyprus (0 low level, 1 medium level, 2 high level)
GERD in SSH as % GDP	0.12	1
GERD in SSH as % of GERD	25.1	2
Number of researchers in SSH per capita	108	2
GERD in SSH per researcher in SSH	23556	2
Access to databases	-	2
WoS publications	missing	
Average	-	1.8
Overall	-	2 (high level)

Development of research data production in SSH

The production of research data can be characterized as “periodical”. Cyprus participates in some international research projects like for example ESS (waves 3, 4, 5, 6), ISSP (however last participation dates to 2009 - Social Inequality IV), EVS, and PISA. The country also participates in Eurobarometer as this opinion survey is organized on the European level. However, the studies that systematically assess matters of national importance seem to be lacking (based on Vryonides estimates).

Table 7: Development of research data production in SSH in Cyprus

Indicator	Rating of Cyprus (0 low level, 1 medium level, 2 high level)
International collaborative research	1
Studies of national importance	0
Average production of data	1
Sum	2
Overall	0 (underdeveloped)

Research Data Management Policy Setting

Cyprus does not seem to have any requirements or recommendations about preparing Data Management Plans; there are no explicit principles on budgeting costs for managing research data and preparing it for the access (based on Maryos Vryonides expertise). Public research funding organizations operating in Cyprus (the most important of them is the Research Promotion Foundation) haven't issued any systematic requirements or recommendations on open access to or long term curation of the research data. There are only some recommendations and guidance provided on how to respect the legal requirements while sharing data.

In the document called *Background Note in Open Access to Scientific Publications and Open Research Data* published by the Directorate-General for Research and Innovation (RTD) in April 2016 it is mentioned that the open access policies, institutional strategies or subject-based initiatives for research data are missing at the moment and that in the nearest future the development towards more effective policies in this respect could not be expected in Cyprus.

Data sharing culture

Data sharing and reuse among social science researchers in Cyprus is not common. Proportion of social science researchers that have shared the research produced research data during the past five years is estimated as low (up to 10%). Similar situation is with the access to the existing third party data; the share of social science researchers who were able to access the data they needed during the mentioned period is estimated as very low.

Attitudes towards data sharing

Table 8: Attitudes towards data sharing in Cyprus

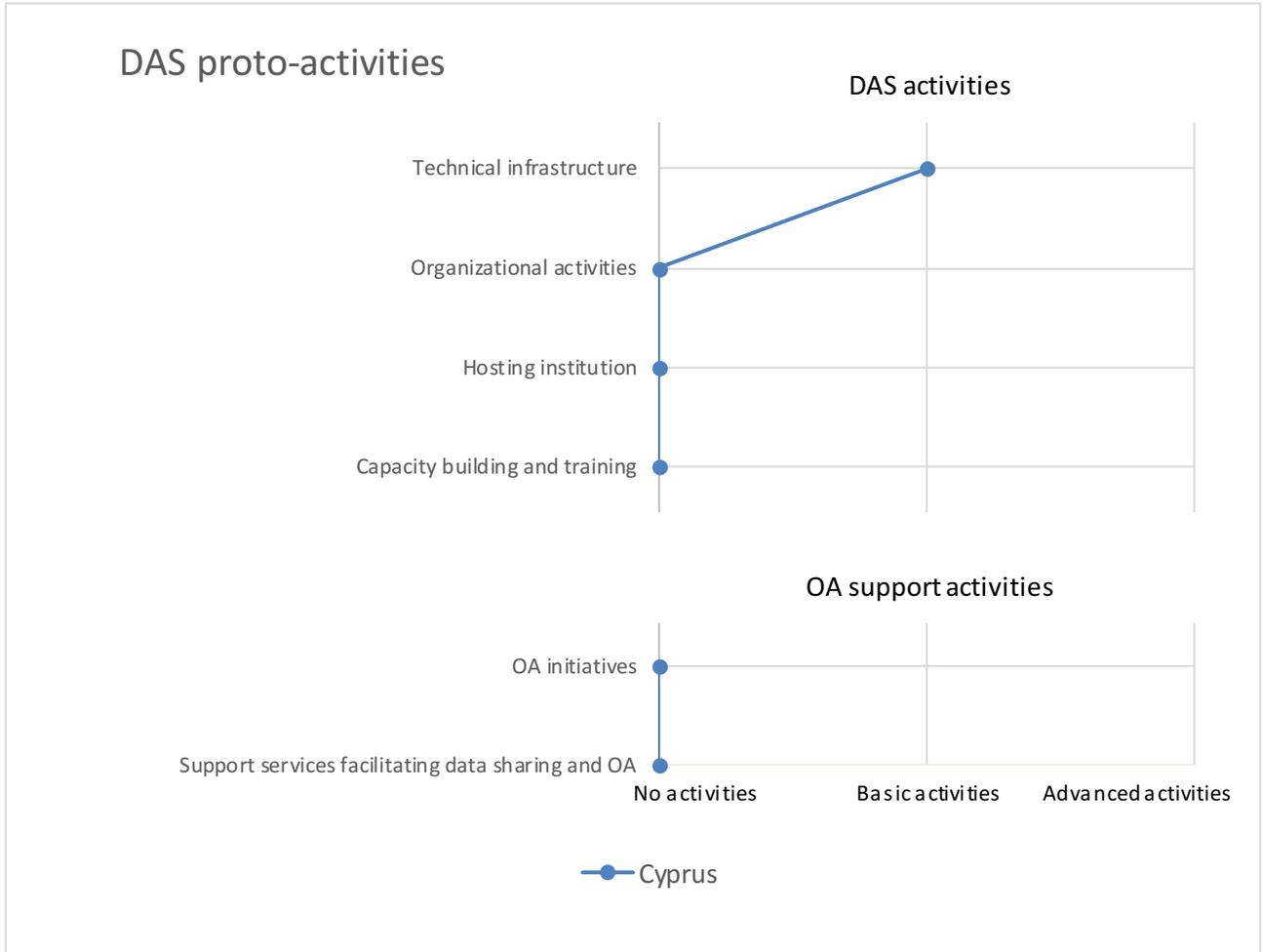
Data sharing has no benefits at all	False
Data sharing creates healthy competition	Probably true
Data sharing creates negative competition	Neither true, nor false
Reuse of existing data can answer new research questions and facilitate advancement of science	True
Data sharing has as a risk that others may misuse and misinterpret data	True
Data sharing involves little effort and minimal costs	Probably true

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from "5"-True to "1" False!

Social science data producers in Cyprus do not follow data management and data documentation standards and procedures that facilitate the reuse of the data. The most important ways for data sharing in the social science in Cyprus seem to be informal contacts among colleagues and (less frequently) supplementary data in journals. This situation is probably the result of the absence of institutional facilitators for data sharing, as well as the lack of the career rewards related to data sharing and the dearth of data support services for social science researchers in Cyprus.

3.8.2 Data archive service (DAS) proto-activities

Figure 15: DAS and OA activities implementation type in Cyprus



DAS activities

Currently, there is no institutional and technical infrastructure for data deposit in SSH in Cyprus; and we have identified neither official steps, nor public initiatives towards establishing a DAS for the social sciences in this country. It is a task for relevant institutions in Cyprus to do steps towards establishing needed infrastructure.

In other disciplines (e.g. natural and technical sciences), there are some initiatives for data depositing, for example DARECLIMED data repository for data on climate and natural resources. Maryos Vryonides mentioned that no existing technical infrastructures could possibly be used for or applied to a new DAS. However, during the desk research we identified the Cyprus Institute and its Science and Technology in Archaeology Research Center (STARC), which could be a potential partner for future DAS in Cyprus. According to available information, this institution develops tools for Digital Humanities (mainly archaeology) and as part of the Cyprus Institute has developed data infrastructure.

Open Access (OA) support activities

According to the self-assessment survey results, there are no support services provided to social science researchers that facilitate data sharing and/or Open Access to research data. The respondent was not aware of OA initiatives in Cyprus.

3.8.3 Conclusions

In general social science sector in Cyprus seem to be well established when it comes to the institutional and financial background. Cyprus participates in some important international research series as ESS or ISSP. However, research data production is generally quite limited and irregular. Sharing data is not widespread among social researchers in the country and a large part of them fails to access needed data for the secondary analysis. The nonexistence of institutional repository for social sciences, insufficient institutional support and low level of data sharing culture were identified as the main obstacles for any development in this area at the moment.

3.9 Czech Republic

On a policy level, there is a growing awareness towards research data management and a culture of data sharing. Currently there are no requirements for data management plans from funders. In the context of Open Access, a recently published national policy recommends amongst others the implementations of data archives. Thus, there are activities in this field to some extent but these could be improved. It seems that requirements from the EC for research funding like Horizon 2020 encourage the developing and establishing of policies and infrastructure in this field.

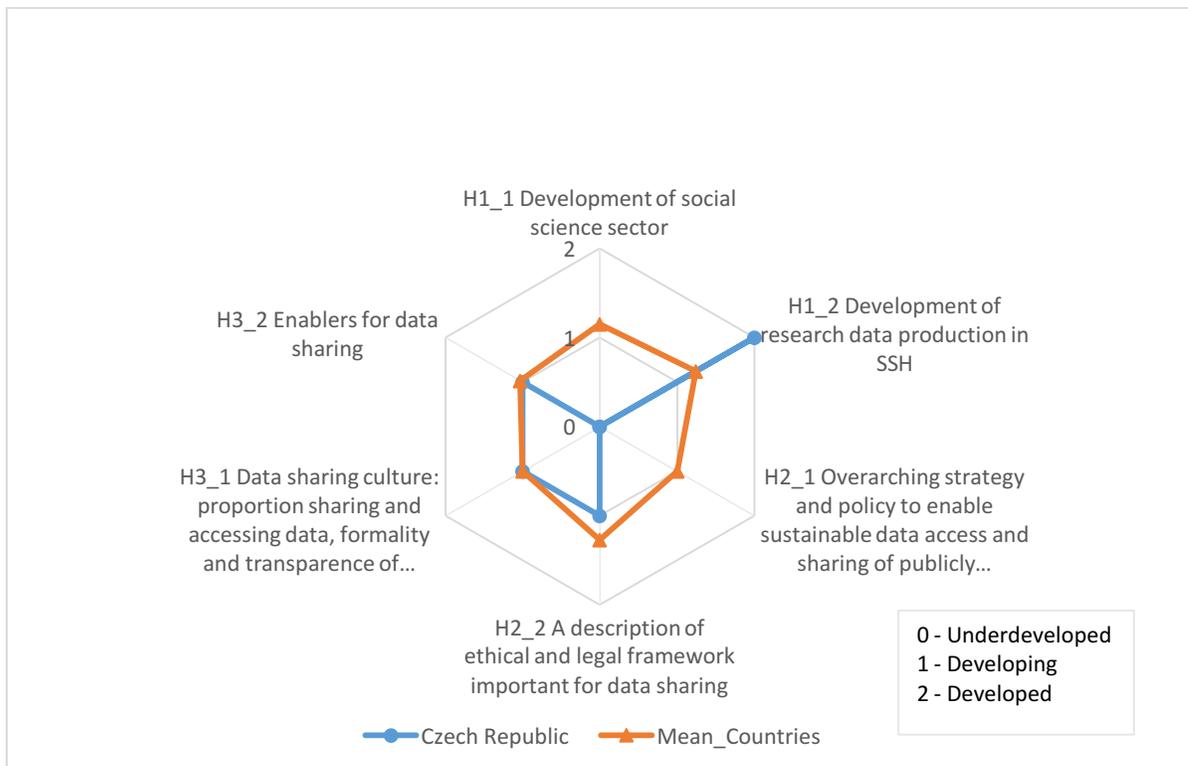
The Czech Social Science Data Archive CSDA is the national Service Provider for CESSDA and is, taken as a whole, well positioned and well established. The organizational infrastructure as well as handling of the data is managed and documented. Sufficient information can be found at the website.

Specifics about the data collection in a country

CSDA staff provided almost all information. Many questions from the survey were answered and additional information was given. Communication was very easy and agreeable. Just a little desktop research was added.

3.9.1 Broader ecosystem of DAS operation

Figure 16: Heading concepts values in Czech Republic



Development of social science sector

In the field of research data production in the social sciences and humanities the Czech Republic is very engaged. The land is participating in almost all mentioned cross-national surveys. For studies of national importance is mentioned:

„Public opinion research center (department of Institute of sociology) produces ten times a year opinion poll *Our society* which is targeted at important social and political issues (<http://cvvm.soc.cas.cz/en/>). The project Dynamics of change in Czech society⁵⁰ is long-term (2014-2018) panel survey targeted at political, social and economic issues. “[Comment to question I1.2.2, see appendix 2]

There is established research in the field of empirical social sciences and institutions with an established tradition of data production.

RDM Policy setting

The awareness about RDM is not very high in the Czech Republic. Funders do not require any Digital Management Plans. Information about RDM and DMP for the social sciences is provided by CSDA as well as recommendations and training to these issues⁵¹.

With regard to quality assurance of social science research data and metadata the situation is similar. There are neither requirements made by funding agencies nor any recommendations about it. There are also differences between the research fields and communities. In the self-assessment survey is stated: *“In sociology the view to open access issues is more positive than e.g. in economics, psychology, etc.”* (Comment to question I2.1.2)

Indeed, at the level of evaluation committees and among peer reviewer, there could be observed a growing awareness of quality standards of research data and metadata enhancement.

In September 2016, a national open access policy was published. According to this policy, it is recommended that research institutions should *„support an implementation of institutional repositories to archive research outputs and open data“*.⁵²

One central motivation in promoting open access activities seems to be European research activities and specially the Horizon 2020 program. Supporting open access activities on a national level, including research institutions as well as research funders is just starting and the emphasis seems to be on publications and not on research data:

“At present there are discussions about policy of open access to scientific information in academic field, but mostly related to open access to publications.” (Comment to question I2.2.1)

Clarification and support provided on legal and ethical issues is not yet established in depth. There is information available on a very general level like “researchers take care about data”. Detailed descriptions or standards related to data sharing or open access in the social science are still lacking.

⁵⁰ Available at: <http://www.soc.cas.cz/en/project/dynamics-change-czech-society>

⁵¹ See <http://archiv.soc.cas.cz/en/social-science-data-management>

⁵² See <http://openaccess.cz/en/open-access-in-the-czech-republic>

As the national open access policy is quite new and has more recommendations rather than requirements, funding for data sharing or DRM is not yet common. In the self-assessment survey is mentioned: “Cost for data management is usually eligible for funding. However, the open access is not required and higher cost make project less competitive. So, it depends on the cost – if it is low, there is no problem. If higher, it is better rather to skip it in the project application.” (Comment to question I2.1.4)

With regard to long-term preservation of research data there is growing awareness but no sustainable investment or support for it.

Data sharing culture

Until now, there is no representative survey of attitudes among researchers about sharing their data. Statements are made according the experiences from CSDA.

The attitude towards data sharing differs between institutions. For the Institute of sociology, departments of sociology or political sciences at public universities it is estimated “according to our experience” (Comment to question I3.1.1) that more than the half are sharing their data. However, even within these communities, the attitude towards data sharing varies.

For other disciplines, the situation seems to be worse, especially for demography, psychology, social geography, but maybe this estimation is due to the fact that in these disciplines data sharing may be more informal as no data archive was available.

Attitudes towards data sharing

Table 9: Attitudes towards data sharing in Czech Republic

Data sharing has no benefits at all	Probably false
Data sharing creates healthy competition	Probably true
Data sharing creates negative competition	Neither true, nor false
Reuse of existing data can answer new research questions and facilitate advancement of science	Probably true
Data sharing has as a risk that others may misuse and misinterpret data	Probably false
Data sharing involves little effort and minimal costs	Probably false

Source: self-assessment survey. Question asked: *Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from “5”-True to “1” False!*

Enablers for data sharing

Data producers in Czech Republic at the moment follow no standards or data management plans. No career rewards for data sharing.

"The only support service represents Czech social science data archive with information provided on web pages (<http://archiv.soc.cas.cz/en/data-deposition>) and with personal information service for (potential) depositors." [1.3.2.3]

3.9.2 Capability requirement areas of DAS

Organisational profile

Organisation

The Czech Social Science Data Archive (ČSDA) accesses, processes, documents and stores data files from social science research projects and promotes their dissemination to make them widely available for secondary use in academic research and for educational purposes. It is part of the Institute of Sociology of the Czech Academy of Sciences and located in Prague (<http://archiv.soc.cas.cz/en>) and was founded in 1998 as Sociological Data Archive (SDA) and renamed in 2011. The Institute of Sociology is a non-profit public research institution within the structure of the Academy. The data archive CSDA is housed in the premises of IS and it is using services of the service departments of IS, e.g. the Economic Department (including administrative and IT services) and the Press and Publications Department. CSDA also co-operates with IS research departments in different projects. CSDA is a department within the structure of the Institute of Sociology CAS (IS). There are no units of lower level within CSDA. There are 8 employees (5.8 FTE).

In September 2016, the CSDA has received the Data Seal of Approval.

Funding

The CSDA receives 100% public funding from the following institutions:

1. Ministry of Education, Youth and Sports (MEYS),
2. Institute of Sociology (IS) of the Czech Academy of Sciences,
3. Different projects (e.g., Horizon 2020, Czech National Foundation, CESSDA).

In the self-assessment survey is mentioned, that "the main funds come from the "targeted support" provided by MEYS to research infrastructures listed in the Roadmap of Large Infrastructures for Research, Experimental Development and Innovation of the Czech Republic. It is based on short-term projects (e.g., 2010 - 2015, 2016 - 2019) and regular evaluations of the infrastructures. However, the funding rules and procedures consider continuity and requirement of long-term sustainability of research infrastructures. The funding from the operational budget of IS comes from the "institutional support" for operation and development of the research institutes of the Czech Academy of Sciences. This funding of IS is long-term and based on regular evaluations of the research programme and the performance of IS."

Core services and activities

The main tasks of CSDA are the following:

- Acquiring, archiving and preserving datasets from Czech social science research projects and making them publicly available for secondary analysis
- Providing technical and organisational support for large-scale survey research programmes, e.g. Czech participation in the International Social Survey Programme (ISSP) or the Czech Household Panel Survey (CHPS)
- Promoting secondary data analysis in general, i.e. mapping and analysing available sources of data and facilitating their utilization, linking different Czech and international data resources, conducting research in the field of data quality, data harmonisation, disseminating information and providing training in data management and survey methodology, etc.

ČSDA conducts professional data processing, connects data with relevant research information and the context of other databases and materials and produces and publishes overviews of available data and data sources. It also carries out independent research activities in the field of methodology including analysis of data quality and collaborates in many research projects at the Institute of Sociology ASCR. The Archive is also a source of research instruments and procedures which have been tested in prior research, thus creating a basis for the implementation of new survey strategies⁵³. Acquisition of social science research data is pursued both, passively and actively. ČSDA also offers training for researcher in the field of data management and data analysis.

Content current collection

The majority of the collection is questionnaire-based survey data. The rest are qualitative data and historical data. Data in CSDA collection are related to society (and territory) Czech republic or Czechoslovakia (before 1993). Data sets covers period from 1946 until present. For 1946–1972, no microdata were preserved, but there are aggregated data from opinion polls in form of Nesstar cubes.

With help of CESSDA topics classification, we can say that following topics cover most of the content of CSDA data collection:

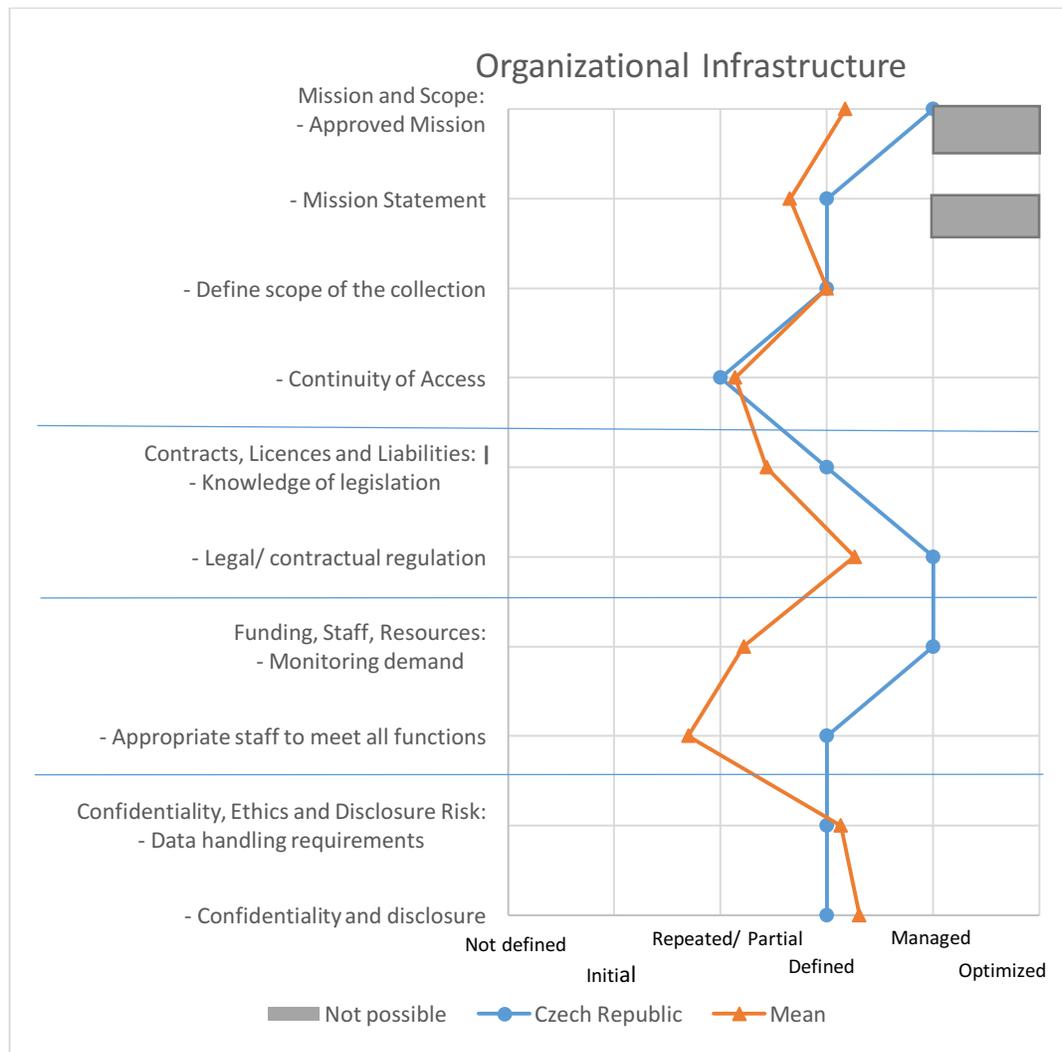
- society and culture
- social stratification and groupings
- politics
- labour and employment
- health
- education
- economics
- history
- social welfare policy and systems

⁵³ See <http://archiv.soc.cas.cz/en/about-czech-social-science-data-archive>

At present there are 774 archived studies – 36 of them is (both data and metadata) in English (4,7% of all), the rest is in Czech.

Organisational infrastructure

Figure 17: Organizational infrastructure in Czech Republic



The mission of CSDA is summarized on the archives website:

The Czech Social Science Data Archive (ČSDA) at the Institute of Sociology of the Academy of Sciences of the Czech Republic accesses, processes, documents and stores data files from social science research projects and promotes their dissemination to make them widely available for secondary use in academic research and for educational purposes. ČSDA promotes the objectives of the OECD Principles and Guidelines for Access to Research Data from Public Funding (OECD 2007) in the area of social science research in the Czech Republic.

Core services and activities

ČSDA is the Service Provider for the Consortium of European Social Science Data Archives (CESSDA).

The main activities of ČSDA are the acquisition, archiving and dissemination datasets from Czech social science research. It also provides technical and organisational support for large-scale survey research programmes, like the International Social Survey Programme (ISSP) and the European Social Survey (ESS) or the newly established Czech Household Panel Survey (CHPS).

It also is promoting secondary data analysis in general, i.e. mapping and analysing available sources of data and facilitating their utilization, linking different Czech and international data resources, conducting research in the field of data quality, disseminating information and providing training in data management and survey methodology, etc. In addition, the ČSDA pursues independent research activities in the field of methodology including analysis of data quality and collaborates in many research projects at the Institute of Sociology ASCR. The Archive is also a source of research instruments and procedures which have been tested in prior research, thus creating a basis for the implementation of new survey strategies.

Contracts, licences, liabilities

CSDA follows all relevant legislation, which is in force at the territory of Czechia - mainly "Personal data protection act" and "Copyright act".

Data archiving and data usage is regulated by agreements and licenses. These are in accordance with legal regulations and consider data protection and copyrights.

CSDA uses a license agreement which is a model contract with data depositors/producers⁵⁴. The License Agreement is in accordance with the Czech law (e.g. Personal data protection act, Copyright act, etc.). Enforceability of License agreement conditions has not been tested.

Data users are required to agree to the conditions for data usage. These are in accordance with the Czech legislation (namely The Personal Data Protection Act and The Copyright Act). They are publicly available on the website⁵⁵ and data users should agree by registration (<http://archivreg.soc.cas.cz/register>).

The Institute of Sociology of the Academy of Sciences of the Czech Republic is registered with the Czech Office for Personal Data Protection and its research activities are governed by Act No. 101/2000 Coll. on personal data protection. In line with the Act the CSDA currently deposits and disseminates data of non-personal character.

To sum up, the CSDA in all categories is higher ranked than the mean of all archives concerning the organizational infrastructure. All items are either defined or managed except the aspect of continuity, which is not yet solved satisfying for many digital archives.

⁵⁴ See http://archiv.soc.cas.cz/sites/default/files/agreement_on_data_deposition_en.doc

⁵⁵ See <http://archiv.soc.cas.cz/en/obligatory-rules-data-access-datasets-csda-and-information-about-their-use>

Digital object management (data curation) and Technical infrastructure and risk

Information that needs to be associated:

CSDA has published a preservation policy. It contains information about the archiving procedures from the pre-ingest phase to access. The different stages are analogue to OAIS functions. The policy also describes the managing and handling of data, list data formats suitable for preservation and contains all activities about strategy, risk management and preservation actions. It also contains all requirements for data deposition. The policy is available via the website⁵⁶.

Bibliographic citation for each data file is included in the metadata section of CSDA Nesstar catalogue. In case of a publication based on data from the CSDA the user is required to refer to the depositor as well as to the CSDA.

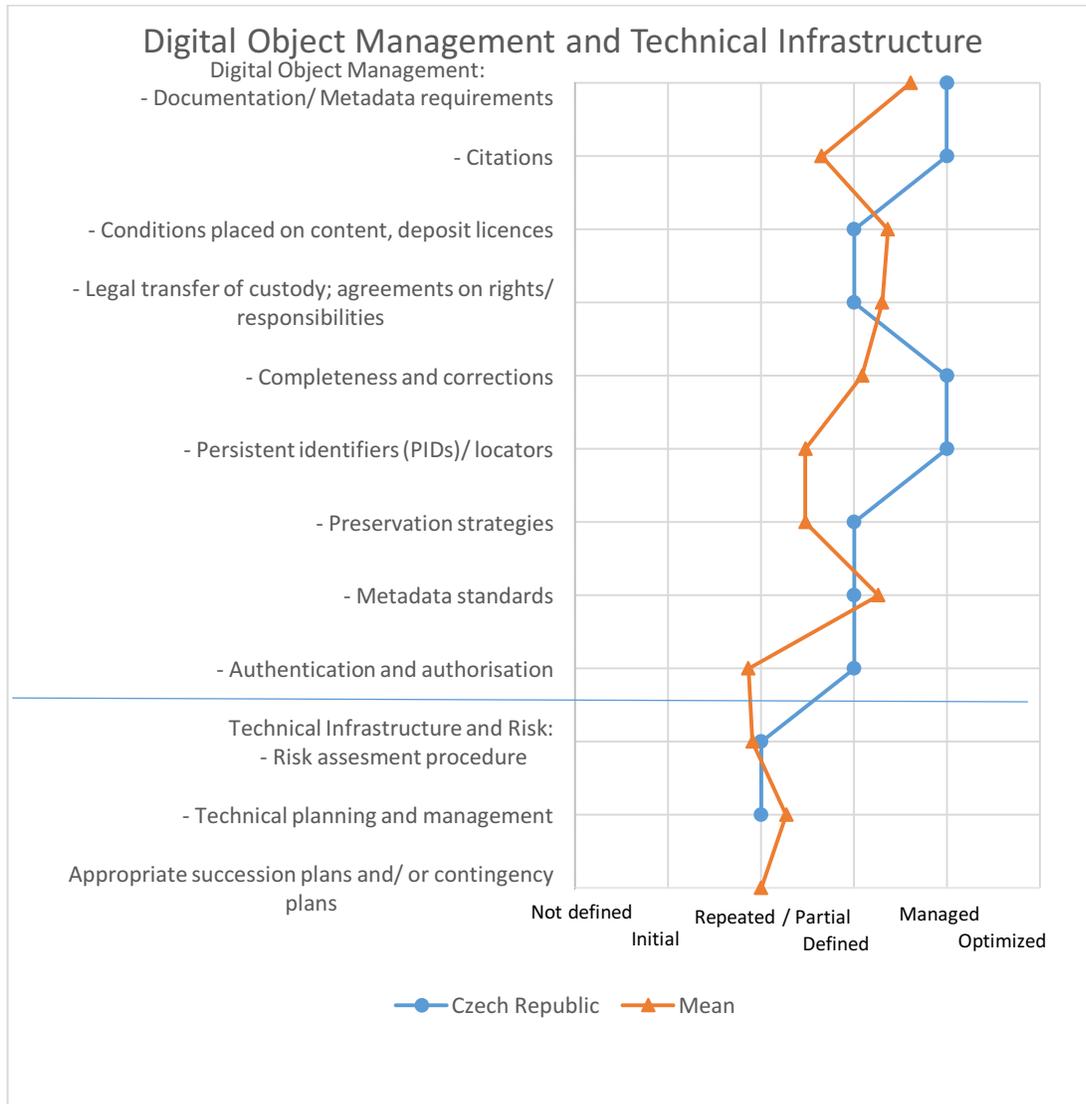
The CSDA uses DOI registration of the data files through da-ra registration agency (<http://www.da-ra.de>)

The CSDA provides different access conditions. Users have to register and agree to conditions of use before getting access. The range is from accessible via internet to a written permission from the data depositor. Data depositors may choose special access conditions in the license agreement⁵⁷.

⁵⁶ See http://archiv.soc.cas.cz/sites/default/files/csda_preservation_policy_0.pdf

⁵⁷ Available at: http://archiv.soc.cas.cz/sites/default/files/agreement_on_data_deposition_en.doc

Figure 18: Digital object management and Technical infrastructure in Czech Republic



Standards and designated community:

CSDA is fully compliant with the DDI metadata scheme (see <http://www.ddialliance.org/>). At the moment, Nesstar is used as software platform even though some other solutions tested.

A succession plan is in accordance to the general internal regulations of the Czech Academy of Sciences in place. In case of liquidation of the Institute of Sociology CAS archived materials will be deposited in the Masaryk Institute and Archive of the CAS. The original succession plan specific for the CSDA is under preparation.

To summarize, digital object management and the technical infrastructure are well organized. Information is available about procedures and processes according to the OAIS reference model. Persistent identifiers are mandatory for all studies and agreements with the depositors have to be signed. The users have to accept usage conditions. Succession regulations are under preparation, but meanwhile in a case of closing of the archive data will be hosted by the Academy of Science.

3.9.3 Conclusions

On a national level, there is a growing awareness of the importance of data management activities and data sharing. For open access, short steps have already been taken but in general there should be more support and more requirements from governmental and funders side. This could include information for researchers and infrastructure institutions about data management, data sharing and digital preservation. Furthermore, infrastructure on a national level should be promoted and policies should require data management plans and regulations concerning publication of data.

The social science infrastructure services seem to be well positioned through the Czech Social Science Data Archive CSDA. Next to its mission of providing and preservation of social science research data it also offers information for researchers about data handling and data management on a more general level. Not the whole funding is on a sustainable basis.

The standards given e.g. by Horizon2020 programs seems to push developments in the area of (social science) research. In addition, CESSDA, with its requirements for Service Providers on the one hand, and advantage through collaboration on the other hand, are motors that strengthen services and scientific research.

3.10 Denmark

In Denmark, the broader ecosystem of DAS operations seem to be well-developed. Development of the social science sector and the development of the SSH research data production is on a high maturity level, and above average. It is close to average on indicators measuring data sharing culture.

In this report the name Danish Data Archive (DDA) is used to summarize DAS provided by the Danish National Archive (Rigsarkivet). Since May 2014 the DDA, established 1973, has been organized as dedicated functional sections for acquisition, preservation and dissemination & marketing for both public records and research data.

The Danish Data Archive (DDA) is on a comparatively high maturity level. The archive scores generally high on most of the capability requirement indicators (organisational infrastructure, digital object management and technical infrastructure). However, DDA has recently been integrated into the Danish National Archives - "Rigsarkivet" - and is still in a transformation phase. The research data archive is about to expand the coverage of research disciplines substantially. This will be an expansion from social science and health science at the moment to all scientific disciplines in the future. This may have affected the score on some of the capability requirement indicators: within the organisational infrastructure there are some uncertainties regarding mission and scope. But on other indicators within this capability requirement area DDA is above or on average.

When it comes to digital object management DDA is on a comparatively high maturity level. Most indicators are either defined, managed or optimised. There is still some potential for improvement when it comes to technical infrastructure and risk assessment. Again, this may be due to the transitional phase that DDA is currently undergoing.

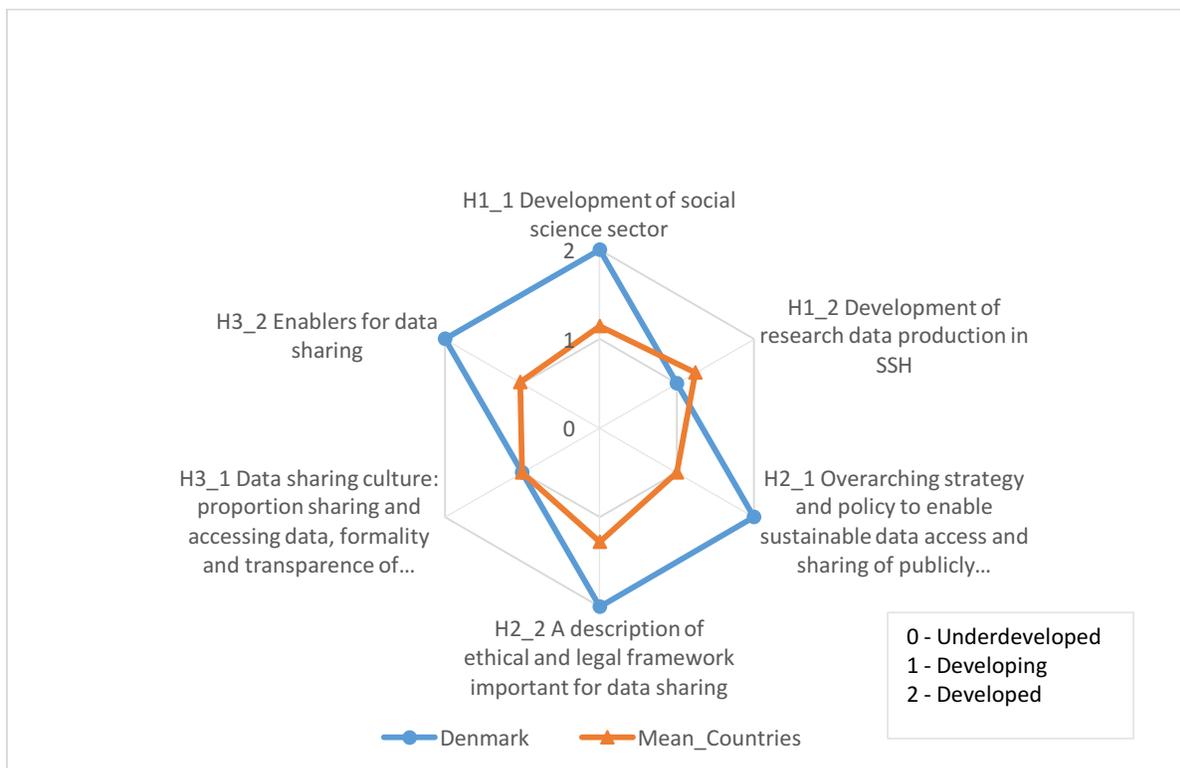
Specifics about the data collection in a country

All the information in the survey was provided by the Danish National Archive by Anne Sofie Fink.

3.10.1 Broader ecosystem of DAS operation

Figure 17 shows that Denmark is on the highest quantile for the general development of the social science sector. There is also a top-level score for the development of research data production in SSH. Denmark is ranked as 'developing' within the categories 'A description of ethical and legal framework important for data sharing', 'Strategy and policy to enable sustainable data access' and 'Enablers for data sharing'.

Figure 19: Heading concepts values in Denmark



Development of social science sector

The financial stability, research capacities and results achieved in the field of social sciences are on the developed level and above the average. When ranking the sources (sectors) of research funding in social science in Denmark (by the amount provided), the business enterprise sector is ranked highest, while the government and higher education sector is ranked second. Ranked third and fourth is funding from abroad (international or cross-border funding) and the private non-profit sector, respectively.

Development of research data production in SSH

Figure 19 shows that Denmark is well developed when it comes to research data production in SSH. In Denmark, the government and/or universities provide access to bibliographic and full-text databases, datasets and software licenses of full-text databases like WoS, Scopus, EBSCO, JSTOR, etc. Denmark is also part of international collaborative research and cross-national studies, like CCS, ESS, ISSP, CSES, EVS, PISA. They also have studies that systematically and frequently assess matters of national importance (public opinion survey, election survey, etc.).

RDM Policy setting

Figure 19 shows that Denmark is on the average level when it comes to strategy and policy to enable sustainable data access and sharing (DMP).

Regarding the general situation for requirements or recommendations about preparing Data Management Plans among public funders of SSH research, there is a growing recognition and awareness of the need to have DMP requirements in place (level: Initial). Regarding the public funders willingness to issue requirements or recommendations about depositing data in an appropriate disciplinary repository, Denmark is 'partially' developed; there is the 'expectation or recommendation to offer or deposit data in an appropriate disciplinary repository or equivalent data archive service'. Same goes for the requirements or recommendations on long-term curation for valuable research data assets, evaluated and selected in terms of reuse potential; Denmark is on the 'initial' level.

When it comes to the cost for managing the data and preparing it for access, the impression is that the costs are in general not adequately resourced during the research project lifetime.

Legal and ethical framework

Figure 19 shows that Denmark is a little below average on the level of importance of ethical and legal issues when it comes to data sharing. There is a growing awareness (level: initial) with regards to the clarification and support provided on legal and ethical issues that better can facilitate social science data sharing (IPR, data protection, etc.). There are among other things initiatives to implement a national Code of Conduct.

In 2017, the Danish National Archives will start to utilize the Archival Act as legal basis for acquiring research data across scientific disciplines. A legal framework for preservation and dissemination is additionally found in the Archival Act.

Data sharing culture

Figure 19 shows that Denmark is on the average when it comes to data sharing culture.

The DDA contact estimates that approximately 10-30 % of social science researchers in Denmark have shared the research data they produced in the period between 2011 and 2016. Social Science Survey data is to a large extent archived and shared by nationally provided DAS'. Other kinds of data are shared within formal and informal research networks.

Only an estimated 0-10% of social science researchers have been able to access existing third party data they need in the period between 2011 and 2016.

When ranking the routines for data sharing to reflect the preferences of the social science research community in Denmark (for the period 2011-2016), informal contacts (peers and colleagues) is ranked highest. Data archives and repositories are ranked as number two, while project or personal websites are ranked as number three. Supplementary data in a journal (alongside paper) is ranked as number four.

The DDA contacts says that *"...the reason why informal contacts are estimated ranked higher than sharing through data archives is pure convenience. The fact that the DDA does not provide free download of data is expected to produce significant negative behavioural consequence. Why will the researcher engage in a formal request to archive if she can ask a colleague to email a file?"*

When it comes to the overall attitudes of social science researchers, results from the survey are provided below.

Attitudes towards data sharing

Table 10: Attitudes towards data sharing in Denmark

Data sharing has no benefits at all	Probably true
Data sharing creates healthy competition	Estimate not available for 2011-2016
Data sharing creates negative competition	Estimate not available for 2011-2016
Reuse of existing data can answer new research questions and facilitate advancement of science	Probably true
Data sharing has as a risk that others may misuse and misinterpret data	Probably false
Data sharing involves little effort and minimal costs	Probably false

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from "5" - True to "1" False!

Enablers for data sharing

The DDA contact states that there are in general few or no career rewards related to data sharing in Denmark. However, there are data support services provided to social science researchers, that facilitate data sharing and/or Open Access to research data (regarding for example, data management plans, data preservation, and data access). Support is offered by university libraries and the data archives (through training, workshops, presentations, personal guidance, etc.). However, only parts of the social science research community is covered, primarily quantitative survey-based research.

Some data producers follow data management and data documentation standards and procedures (no numerical estimate provided).

3.10.2 Capability requirement areas of DAS

Organisational profile

The Danish Data Archive (DDA)⁵⁸ is integrated into the national archive “Rigsarkivet” (Danish National Archives). The departments of DDA consist of Acquisition, Processing curation and preservation, and Dissemination and marketing (for both public records and research data). Primary users of the archive are researchers and students from the social sciences and health sciences.

DDA is publicly funded. Core services and activities include acquisition, curation and dissemination (including data referencing (DOI)). Content of the current collection are primarily surveys (including cohorts, time series, case-control and epidemiology studies).

Organisational infrastructure

Figure 3.10.2 shows that DDA scores below average on all three indicators that deal with ‘Mission and scope’, while they are somewhat more mature and above average for on indicators for ‘Contracts, Licences and Liabilities’, ‘Funding, Staff, Resources (except on the monitoring of demand) and ‘Confidentiality, Ethics and Disclosure Risk’.

DDA is in a transformation phase, which affects the development level on several of the indicators, especially concerning mission and scope. Approved mission is in place, but when it comes to medium- and long-term planning, there is still room for further improvement.

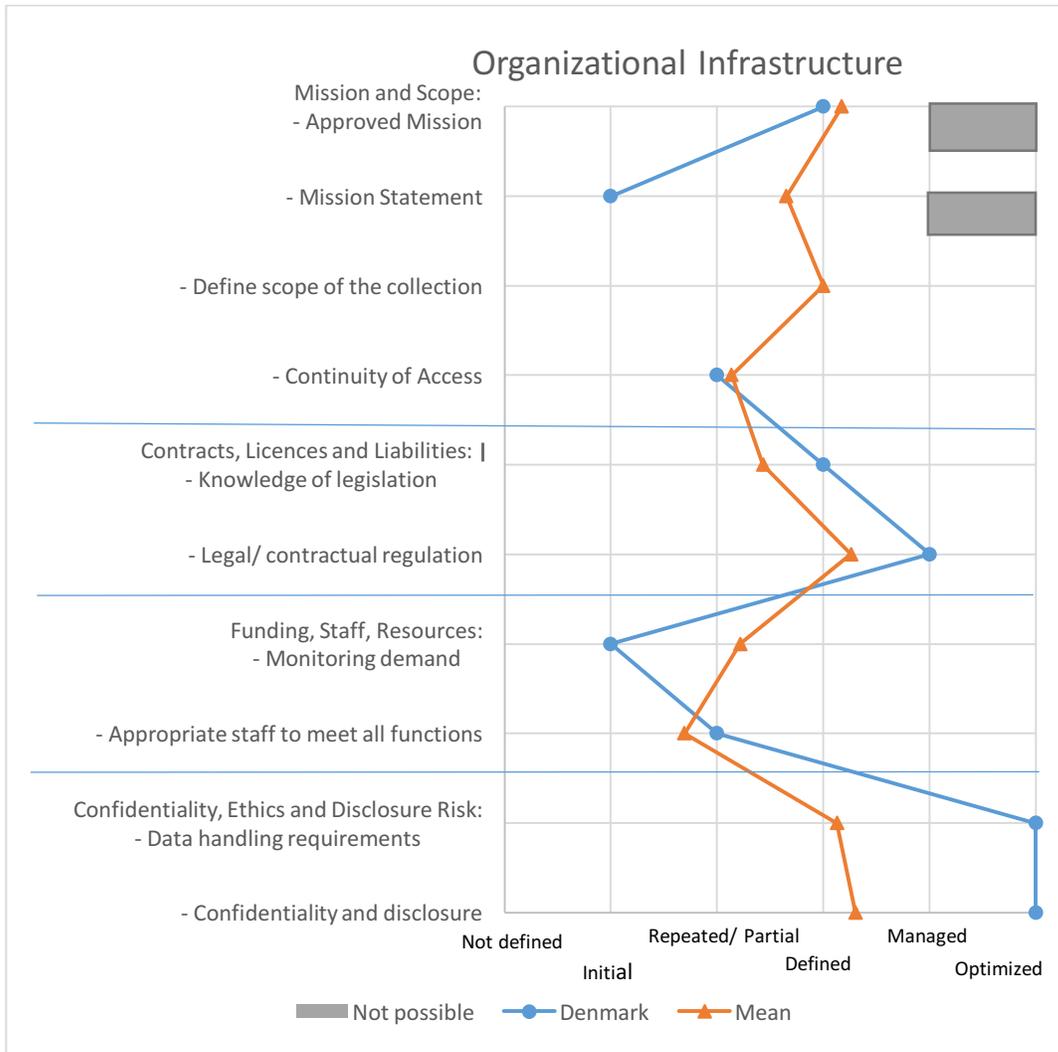
DDA scores better than the average on ‘knowledge of legislation’ and ‘legal and contractual regulation’.

DDA is in the ‘initial’ phase of monitoring demand for its repository services, growth and funding, but there seem to be an appropriate number of staff to support all repository functions and services (level: partial).

DDA scores very high (optimised) on compliance with applicable research discipline norms and legal data handling requirements, and on data handling requirements towards data depositors.

⁵⁸ See <https://www.sa.dk/en/services/danish-data-archive>

Figure 20: Organizational infrastructure in Denmark



Digital object management (data curation) and Technical infrastructure and risk

Figure 21 shows that DDA scores mostly above average on indicators for Digital Object Management, and mostly below average on Technical Infrastructure and Risk.

For indicators on Digital Object Management (See section II.2, Appendix 2), DDA scores high (optimised) on metadata requirements [2.1.1.2], checks of completeness and correctness [2.1.3.1], and preservation strategies [2.2.3.3]. Two indicators are ranked as ‘managed’, namely legal of custody (agreements on rights and responsibilities) [2.1.2.6], and the provision and use of metadata standards [2.3.1.5]. When it comes to mechanisms for proper data citations [2.1.2.4], for setting access conditions [2.1.2.5] and for generating PIDs [2.2.1.2], these are partially implemented.

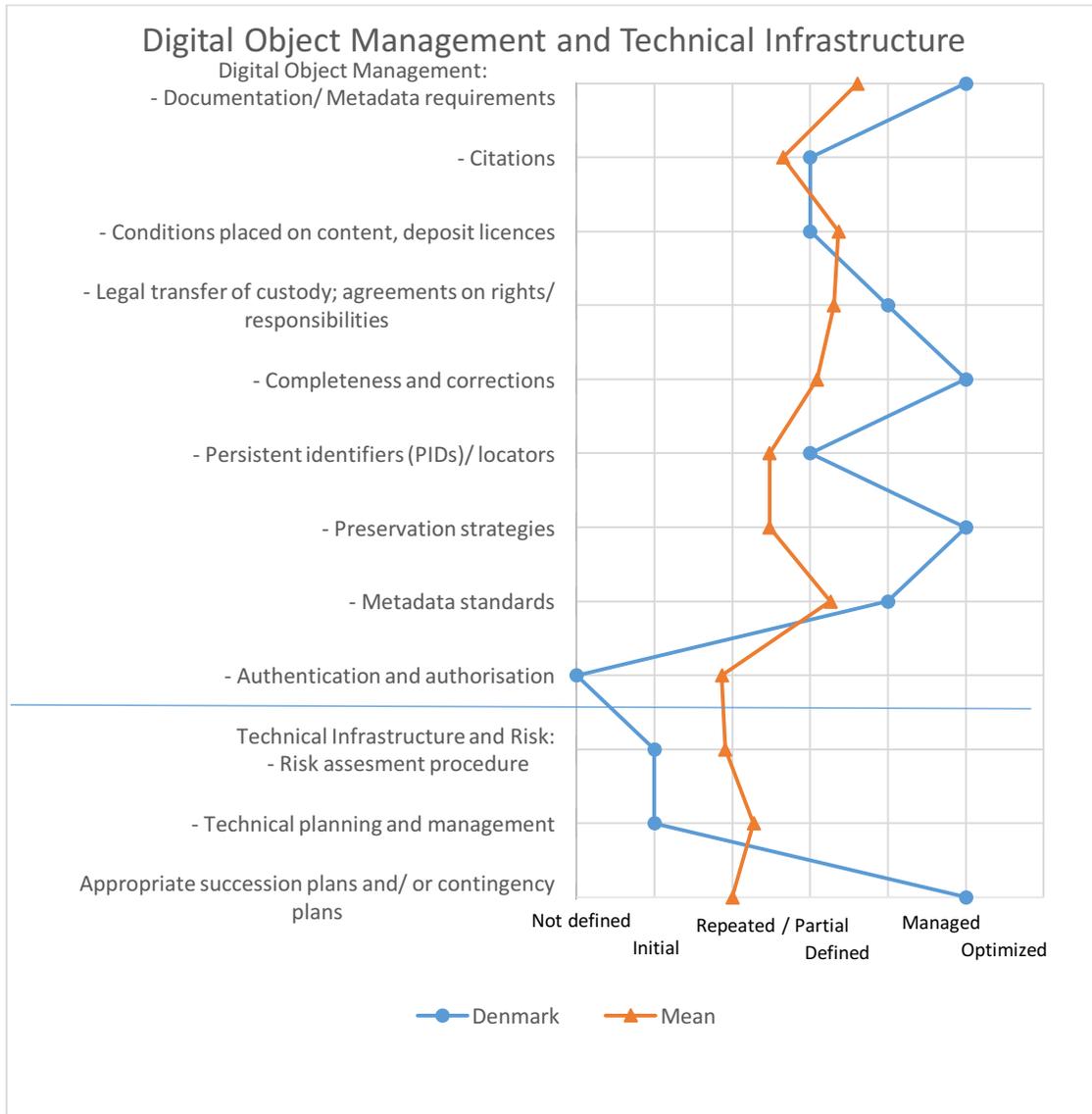
Measures for authentication and authorisation are not defined at DDA.

When it comes to technical infrastructure and risk, DDA is in the initial phase regarding risk assessment procedures [3.1] and technical infrastructure resources [3.2]. This may be due to the fact that DDA is currently in a transformation phase, which may cause some uncertainties

regarding the general organisation and resource availability of the current and future technical infrastructure.

However, implementation of successions plans / contingency plans [3.5.1] are considered optimised at DDA.

Figure 21: Digital object management and Technical infrastructure in Denmark



3.10.3 Conclusions

In Denmark, the broader ecosystem of DAS operations seems to be well-developed. Development of the social science sector and the development of the SSH research data production is on a high maturity level, but scores somewhat lower on indicators measuring data sharing culture (but still around the average).

The Danish Data Archive (DDA) in the Danish National Archive is on a comparatively high maturity level. The archive was established in 1973 alongside similar social science data archives in Europe and North America, and have a long tradition of handling preservation and dissemination issues. Since May 2014 the DDA has been organized as dedicated functional sections for acquisition, preservation and dissemination and marketing for both public records and research data.

The archive scores generally high on most of the capability requirement indicators. However, DDA has recently been integrated into the Danish National Archives - "Rigsarkivet" - and is still in a transformation phase. This may have affected the score on some of the capability requirement indicators. An important factor at this point in time is that the research data archive is about to expand the number of research disciplines targeted substantially. This will be an expansion from the current situation, with a majority of data within social science and health science, to include all scientific disciplines in the future.

Where potential gaps can be identified - mission statement, scope, technical infrastructure and risk assessments - work is in progress.

3.11 Estonia

Broader ecosystem of DAS operations in Estonia is characterized by relatively highly developed social science sector and data production in social sciences, even though social sciences sector has in recent years been under pressure due to lower level of finance than other disciplines. General ethical and legal framework for data sharing is in place, and even though it could be more elaborated, it can be considered as developed. Overarching policy and strategies for data sharing are developing. Data sharing culture is slowly developing and characterized by use of mainly informal and not transparent data sharing channels.

Estonian Social Sciences Data Archive was founded in 1996. However, from 2000-ties, lack of sustainability in funding has severely limited development possibilities of ESSDA. ESSDA was member of CESSDA 1997-2013, when it was forced to drop out due to lack of commitment of the country to national membership in CESSDA. Maturity level of capability requirement areas – organizational infrastructure, digital object management and technical infrastructure - are consistently below the average.

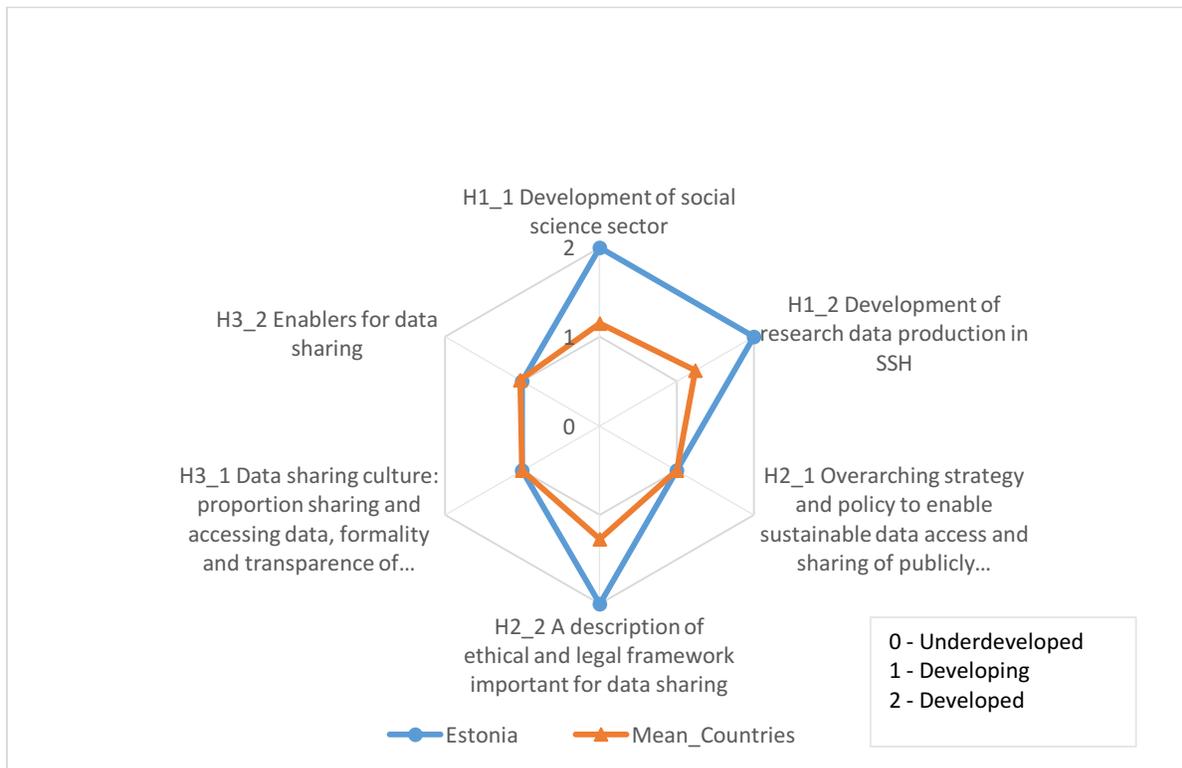
Specifics about the data collection in a country

Data was collected mainly via self-assessment performed by ESSDA representative Rein Murakas and, where possible supported by documents and information from ESSDA's webpage.

3.11.1 Broader ecosystem of DAS operation

In Estonia three of broader DAS ecosystem dimensions – development of social science sector, development of research data production and ethical and legal framework important for data sharing are developed on the highest level (and above average in the audited countries). Three other dimensions – RDM policy settings dimension, data sharing culture, and enablers for data sharing - are developing, being on the same level as the average level in the audited countries.

Figure 22: Heading concepts values in Estonia



Development of social science sector

In this section, basic profile regarding scientific development in Estonia is described thus giving insight on general conditions influencing the DAS operation.

Overall assessment of SSH development

General intensity of investment in SSH-sector in Estonia is above medial value, if compared to other countries audited – 0.17% (GERD in SSH as percentage of GDP), ranking 8th from top among all audited countries having provided data. Intensity of investment in SSH (relative to other scientific disciplines) in Estonia is slightly below median value – 11.9%, ranking 14th from bottom (GERD in SSH as % of GERD). Human resource potential in SSH in Estonia can be estimated as high – with 159 researchers per 100.000 inhabitants it ranks 5th from the top among the audited countries (number of researchers (head count) in SSH per 100.000 inhabitants). Investment in human resources in SSH is relatively somewhat lower rather low – 15518, slightly above median (but below the average) value among audited countries with data available. At the same time, the expert indicates that there are actual problems connected with underfinancing of social sciences compared with natural and life sciences, and even humanities in last years.

Most important sources of funding in social sciences and humanities, based on self-assessment by a DAS expert, are ranked as follows: (1) government and higher education sector; (2) international and cross-border funding; (3) private non-profit sector, and, finally, (4) business enterprise sector.

Researchers in the academic sector have publicly financed access to bibliographic and full-text databases, datasets and software licences.

Altogether, social science sector in Estonia can be considered as developed.

Development of research data production in SSH

There are well-established traditions of international collaborative research or cross-national studies in Estonia. The country has participated in Comparative Candidate Survey (CCS) (2011); European Social Survey (ESS) rounds 2 to 7 (2004-2014); International Social Survey Program (ISSP); Comparative Study of Electoral Systems (CSES) Module 3 (2006-2011) and Module 4 (2011-2016); European Values Study (EVS) waves 1990, 1999 and 2008; World Values Survey (Wave 6 2011); Programme for International Student Assessment (PISA) –2015, 2012, 2009, 2006; and Generations and gender programme (GGP) Wave 1. Estonia has at some point participated or is still participating in all eight international studies enlisted.

In addition to participation in above-mentioned international studies, Estonia is involved in international long-term research projects like SHARE⁵⁹, ESPAD⁶⁰, ISCWeB⁶¹, Youth in Europe⁶², EU Kids Online⁶³. Social sciences in Estonia have a long-time tradition in longitudinal studies, like Paths of Generation⁶⁴ started in 1983.

Besides, social scientists in Estonia produce also systematic studies of national importance. According to self-assessment, the average production of research data in the SSH institutions in Estonia can be characterized as frequent - institutions have well established tradition in data production.

So overall, research data production in Estonia is developing and as such, this dimension lies slightly below the average in all audited countries.

⁵⁹ Survey of Health, Ageing and Retirement in Europe, available at: <http://www.share-project.org/contact-organisation/share-eric.html> (last accessed 2017-02-22). Estonia has participated in waves 4 (2010/2011) and 5 (2013).

⁶⁰ European School Survey Project on Alcohol and Other Drugs, available at: <http://www.espad.org> (last accessed 2017-02-22). Estonia has participated in all six waves, 1995-2015.

⁶¹ Children's Worlds, the International Survey of Children's Well-Being (ISCWeB), available at: <http://www.isciweb.org/> (last accessed 2017-02-22). Estonia has participated in the last two waves.

⁶² Youth in Europe is a large-scale monitoring/survey on the level of local communities in different European countries about risk behavior of youth started in 2006, available at: <http://youthineurope.org/> (last accessed 2017-02-22).

⁶³ EU Kids Online is an international research network, which aims to coordinate and stimulate investigation into the way children use new media, with a particular focus on evidence about the conditions that shape online risk and safety, and has until now had four phases of work, available at: <http://www.lse.ac.uk/media@lse/research/EUKidsOnline/About-the-project.aspx> (last accessed 2017-02-02). Estonia has participated in all of them.

⁶⁴ "Paths of a generation" is a longitudinal research project of secondary school graduates of 1983 in Belarus, Estonia, Kazakhstan, Latvia, Lithuania, Moldova, Russia, Tajikistan, Ukraine; started in 1983 with follow-up waves in 1989, 1993 and mid 2000ties. For description see, for example, Titma, M. People in Post-Soviet Transitional Societies. International Journal of Sociology, vol. 34, no. 2, Summer 2004, pp. 3-10. ISSN 0020- 7659/2004.

RDM Policy setting

Overall incentives and high-level policy requirements in Estonia have not reached a level that provides appropriate level of sustainable conditions for data service provision. Of the four policy setting requirements audited, all are on 'Initial' level.

There are policy guidelines aiming at wider accessibility of research data and thus implying indirectly need for RDM activities. General principles are covered in "Public Information Act"⁶⁵, and "Implementation Plan 2016-2019 for achieving the objectives of the Estonian Research and Development and Innovation Strategy 2014-2020 "Knowledge-based Estonia"" priority no. 5 "Optimal circulation, accessibility and transfer of scientific knowledge", activity 2⁶⁶ has as the goal to promote open access to research publications and to encourage open access to public-financed research results and research data. As a result, regarding open data from public sector, an Open Data Portal of Estonia⁶⁷ has been established, but the volume of available data is currently rather low there and mostly from municipal institutions. At the same time, further developments to the other sectors (including state financed scientific data) may be possible also.

There is no explicit and clear request for a DMP, but there is some recognition and awareness of need to require DMP. There are different ideas, but no general regulations, only suggestions, the only rules in written form being requirements for Horizon 2020 projects. Following these developments, universities are developing support DMP related support services, for example, University of Tartu Library has a developed web-based resource about RDM, including sub-page about DMP⁶⁸. Recommendations for depositing data in an appropriate disciplinary repository are on 'Initial' level as well, as there is some recognition and awareness of the need to have disciplinary specific place of deposit and support services. There is some awareness about the need for research data to be carefully evaluated and selected, in terms of reuse potential. Managing of the data and preparing it for access can be implicitly covered in the overall research project budget.

Therefore, overall RDM policy setting in Estonia can be characterized as developing, as there can be there is some awareness and initial developments towards establishing the key aspects enabling data sharing.

Ethical and legal framework important for data sharing in SSH in Estonia can be characterized as developed. According to self-assessment, clarification and support provided on legal and ethical aspects that facilitate social science data sharing (IPR, data protection etc.) are on 'Partial' level – there are recommendations and guidance provided on how to respect the legal requirements while sharing data, for example, well-developed regulations about personal data protection as well as university Research ethics committees evaluating the ethical aspects of research involving humans⁶⁹. However, as commented by the expert, concrete and detailed regulations supporting data sharing on local level are often missing.

⁶⁵ Public Information Act, available at: <https://www.riigiteataja.ee/en/eli/514112013001/consolide> (last accessed 2017-02-02).

⁶⁶ Implementation Plan 2016 - 2019 for achieving the objectives of the Estonian Research and Development and Innovation Strategy 2014 -2020 "Knowledge-based Estonia", pp 16-17; available at: https://www.hm.ee/sites/default/files/tai_rakendusplaan_2016-2020-ennewtextpdf.pdf (last accessed 2017-02-02)

⁶⁷ Open Data Portal of Estonia, available at: <https://opendata.riik.ee/en/> (last accessed 2017-02-22).

⁶⁸ Scientific data management [Teadusandmete haldamine], in Estonian only,, available at: <https://sisu.ut.ee/teadusandmed/teadusandmete-haldamine> (last accessed 2017-02-22).

⁶⁹ The Statutes of the Ethics Review Committee on Human Research of the University of

Data sharing culture

Data sharing among social sciences researchers in Estonia is estimated in self-assessment as being on medium level (10-30%). As explained by expert, data sharing works better via personal contacts, while extensive, systematic and transparent data sharing is connected only with international research projects (like, European Social Survey) or data from Statistics Estonia, but not on research data from local projects. The proportion of social science researchers able to access existing third party data they need could not be estimated for 2011-2016.

A survey about production, documentation and sharing of research data, conducted by Estonian Research Council in 2014⁷⁰, indicates that 75% of the respondents - principal investigators in research area of culture and society - collect or create research data⁷¹. Data is most often shared only with colleagues and on request (40%). Additional 14% indicate that on request data is shared with everybody, and about one fourth, 24% have answered that their data is freely available, but 12% of principal investigators surveyed indicate that they do not share data⁷².

There are, as self-assessment shows, established channels and routines for sharing data in Estonia, though more data is sharing via informal and not transparent channels, with informal contacts (peers and colleagues) ranked 1st, and project or personal websites ranked 3rd. The formal and transparent channels - data archive or repository and supplementary data in a journal (alongside paper) are ranked 2nd and 4th respectively.

The attitudes of researchers towards data sharing, based self-assessment, can be characterized as mainly neutral. The social science community in Estonia is perceived as having concerns about misinterpretation of data or time and effort to be invested in preparing data for sharing, but moderately positive with regards to data sharing facilitating advancement of science. The social science research community in Estonia is perceived in self-assessment as neutral regarding benefits and competition in data sharing.

Tartu (In English), available at:

https://www.ut.ee/sites/default/files/ut_files/Tartu%20%C3%9Clikooli%20inimuuringute%20eetika%20komitee%20statuut_eng_Dima_puhas.pdf (last accessed 2017-02-02).

⁷⁰ Muuli, V. 2014. Research Data in Estonia: collecting, storing, availability. Some findings from questionnaire, available at: <http://hdl.handle.net/10062/44052>. Web-survey of project leaders of European Social Fund (ESF) grants, targeted financed projects and personal research grants, total 471 persons, with 174 completed answers, 37% response rate. 52 respondents (30%) represented research area of culture and society, making it 17% of surveyed principal investigators in social sciences.

⁷¹ Muuli 2014:7

⁷² Muuli 2014:10. It might, however, be that the respondents represent the 17% of surveyed principal investigators in social sciences that filled in the questionnaire, and therefore might be more knowledgeable and positive towards data sharing than the general population of social science researchers.

Attitudes towards data sharing

Table 11: Attitudes towards data sharing in Estonia

Data sharing has as a risk that others may misuse and misinterpret data	Probably true
Data sharing involves little effort and minimal costs	Probably false
Reuse of existing data can answer new research questions and facilitate advancement of science.	Probably true
Data sharing has no benefits at all	Neither true, nor false
Data sharing creates healthy competition in research	Neither true, nor false
Data sharing creates negative competition (for example, being scooped and therefore reduced publication opportunities) for the researcher	Neither true, nor false

Source: self-assessment survey. Question asked: *Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from "5"-True to "1" False!*

The survey of principal investigators by Estonian Research Council in 2014 indicate that in general, overwhelming majority (93%) of principal investigators of social sciences research projects that answered the question about the attitudes towards data sharing (56%), had positive attitude towards data sharing⁷³. As for the costs, one third, 33% of principal investigators in social sciences, have indicated that they have no idea, what are the costs associated with data sharing, and 44% answered that they think there are no costs; only 6% had actual knowledge of the subject⁷⁴.

To summarize data sharing culture in Estonia, data sharing is rather common, there are established both formal and informal channels (the informal channels are prevailing) for data sharing, and general attitudes of researchers related to sharing of data are neutral or positive, so in summary data sharing culture can be characterized as developing.

Enablers for data sharing

There are no career rewards related to data sharing the social sciences academic community in Estonia, as self-assessment indicates.

There are data support services available to social science researchers that facilitate data sharing and/or Open Access to research data, provided by Statistics Estonia, Estonian Social Sciences Data Archive (ESSDA), as well as University of Tartu Library. For example, University of Tartu Library has a webpage about RDM⁷⁵. In collaboration with GESIS, CESSDA workshop "Introduction to Research Data Management for Social Scientists," was held there in 2015⁷⁶ as well.

⁷³ Muuli 2014:12

⁷⁴ Muuli 2014:11

⁷⁵ Scientific data management [Teadusandmete haldamine], in Estonian only, available at: <https://sisu.ut.ee/teadusandmed/teadusandmete-haldamine> (last accessed 2017-02-22).

⁷⁶ Course syllabus (in English) available at: https://dspace.ut.ee/bitstream/handle/10062/50138/Syllabus_RDM_Tartu_Oct2015-1-1.pdf?sequence=2&isAllowed=y (last accessed 2017-02-22).

Based on the initiative from ESSDA, data sharing problems were discussed on annual social science conferences. A special section dealing with data sharing and RDM problems is planned on the 10th Annual Conference of Estonian Social Sciences in March 2017.

Thus, there are some data producers that follow data management and data documentation standards and procedures.

In summary, the incentives and enablers for data sharing within social science research community in the Estonia can be benchmarked as developing.

3.11.2 Capability requirement areas of DAS

Organisational profile

Organisation

The Estonian Social Science Data Archive (ESSDA) was established at Tartu University in March 1996, but actual work began earlier. In the fall of 1993, a group of sociologists, psychologists, political scientists, and human geographers of Tartu University formed an initiative group for creating a data bank on social sciences. In 1994, a project for creating a data bank was presented to the Higher Education Support Programme (HESP) at Open Estonia Foundation and a grant was awarded for the years 1994-1996. The formation of the data bank started by transferring the data deposited in the computing centres of Tartu University and Radio Estonia in order to ensure the maintenance of the information.

For continuing the activities after the expiration of the HESP grant, preparations for creating the Estonian Social Science Data Archives begun. In early 1996 ESSDA was founded as an interdisciplinary centre at the Department of Social Sciences and functions as a national social science data bank. Currently ESSDA is functioning as a consortium inside University of Tartu.

From 1997 ESSDA was a member of CESSDA, dropping out in 2013 when CESSDA started require national membership.

Primary users: researchers, students (different levels). Secondary users: public institutions, private institutions (a small number: newspapers, etc.).

Funding

Lack of funding seriously restricted the development of ESSDA in the late 1990s (Murakas and Rämmer 2002). The situation was critical and the main goal for ESSDA in that period was survival. Several applications for funding were rejected, and the only regular funding source was from the Faculty of Social Sciences, University of Tartu. From 2005 until 2008, ESSDA's activities were partially financed from state-level program for developing of scientific collections. The following attempts to get support from similar financing instruments, the last one made in 2016, were unsuccessful. Also, financial support from structural units inside University of Tartu has been quite irregular in the past years. Currently the data archive has no regular staff. Data archive services are performed with one-time projects and voluntary work, and the possibility to use the university's general infrastructure.

Core services and activities

- Consultation about data archiving and analysis
- Depositing and storing of data
- Data checking and quality control
- Cataloguing of data
- Data and metadata publishing using NESSTAR
- Data dissemination
- Digitizing of non-digital study materials
- Data support for teaching processes, teaching of archiving principles

Content current collection

A unique collection of research data from Soviet and transition times (from 70s to 90s) Research data from different newer local research projects (non-systematic coverage) Digitized social research materials (reports, questionnaires, etc.) Non-digitized social research materials (including fulfilled questionnaires and answering lists) Main topics: media research, stratification, life paths (longitudinal data).

Material is available mostly in Estonian, in few cases - English, Russian, Estonian/Russian, Estonian/English (by digitized research materials also Finnish).

Organisational infrastructure

Figure 23: Organizational infrastructure in Estonia

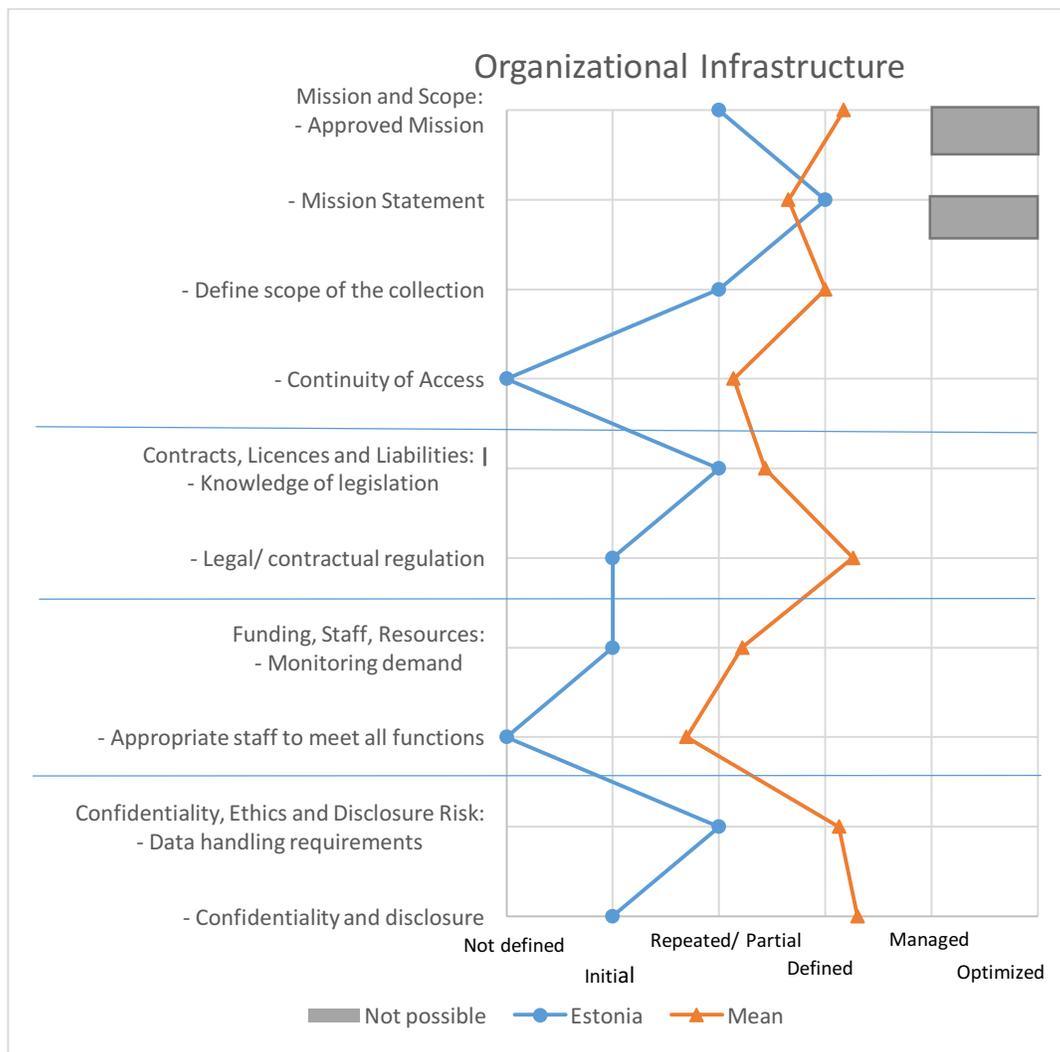


Figure 23 shows that the level of the organizational infrastructure is placed below, and mostly much below, the average in all aspects except for the mission statement.

Digital object management (data curation) and Technical infrastructure and risk

Figure 24: Digital object management and Technical infrastructure in Estonia

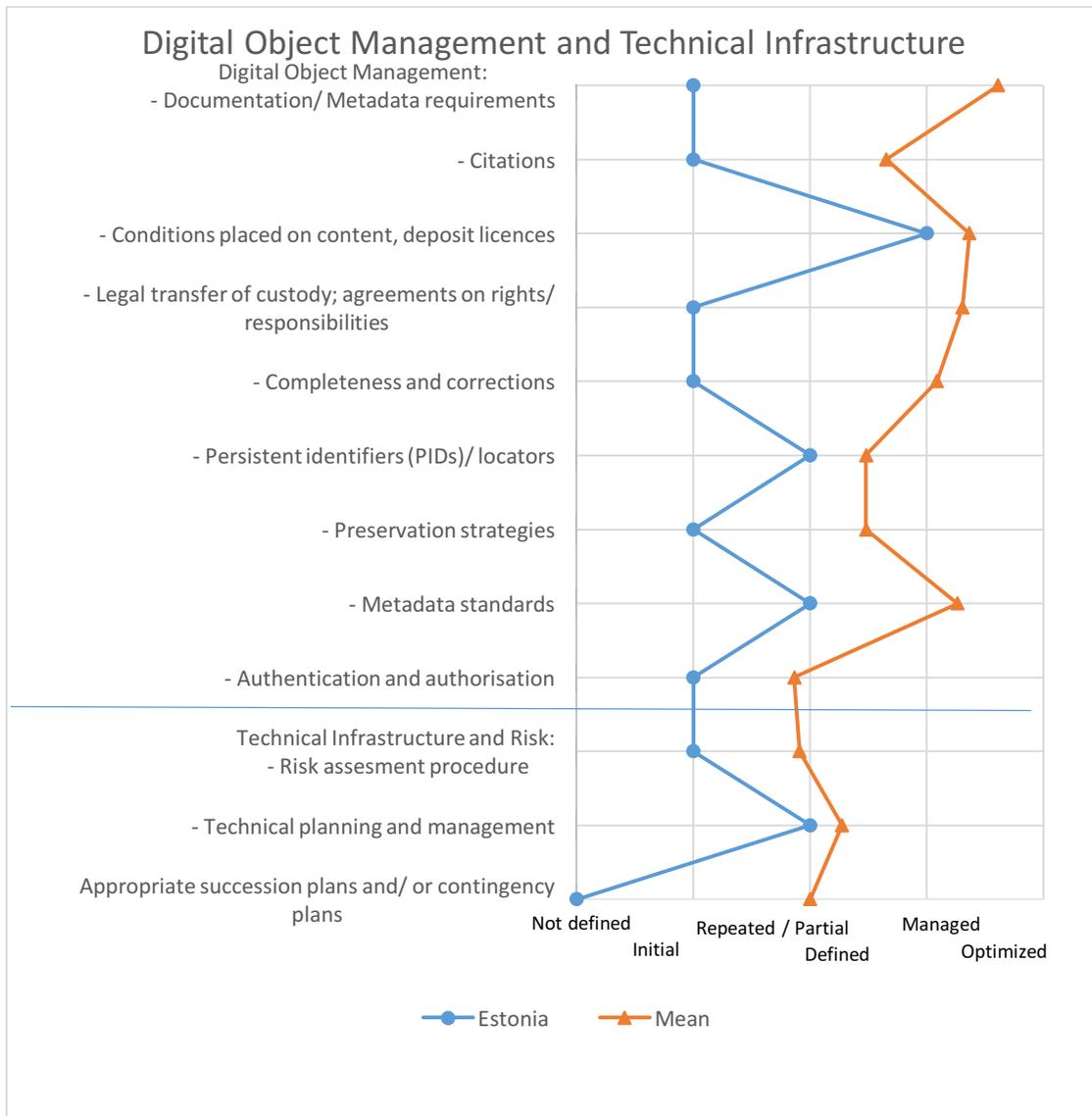


Figure 24 shows that Digital object management and Technical infrastructure in most aspects are on the initial or partial level. In all aspects, it is placed mostly much below the average.

3.11.3 Conclusions

Broader ecosystem of DAS operations is characterized by a developed social sciences sector and data production as well as general ethical and legal framework important for data sharing ensured, among others, by Research Ethical committees in universities. Overarching strategies and policy to enable data sharing, data sharing culture and enablers for data sharing have not yet developed to the level to provide sustainable framework for DAS operations.

After 20 years of operation with insufficient level of funding Estonian Social Science Data Archive (ESSDA) has potential to develop, but sustainable funding and policies enabling data sharing are essential.

Capability requirement areas of DAS operation lie below the average maturity level in audited countries. This reflects the unsustainable funding scheme, as there have never been enough resources for development.

3.12 Faroe Islands

No report. No contact was achieved with representatives of the Faroe Island research community.

3.13 Finland

All in all, Finland has a fine, medium-high maturity level with potential for development.

Finland scores high in relation to the broader eco-system of DAS-operation and medium-high to the organizational profile and digital object management (data curation) in relation to the capability requirement areas of DAS. The last area, technical infrastructure risk, can be further developed.

This fine maturity level reveals a lot of strengths in relation to the broader eco-system, such as participation in a lot of (inter)national surveys and having requirements/recommendations for DMP's. There is potential for developing the way of depositing data in an appropriate disciplinary repository. The data sharing culture is developing rapidly.

The maturity level in relation to the capability requirement areas shows room for developing a broader and more detailed plan for continuity of access, handling the shortage of IT-personnel, and formulating formal statements/definitions of appropriate levels in relation to technical planning of management. Furthermore, there is a need for developing exit plans/scenarios concerning succession/ contingency plans and need for a formal methodology in relation to risk assessment procedure.

The conclusion is that Finland is doing fine and will continue to develop in the future.

Specifics about the data collection in a country

The data collection started 21st September 2016 and was finished 30th September 2016. The information was provided by Finnish Social Science Data Archive (FSD). FSD is a national resource centre and a separate unit of the University of Tampere. The data collection process went well without any problems.

3.13.1 Broader ecosystem of DAS operation

Figure 25: Heading concepts values in Finland

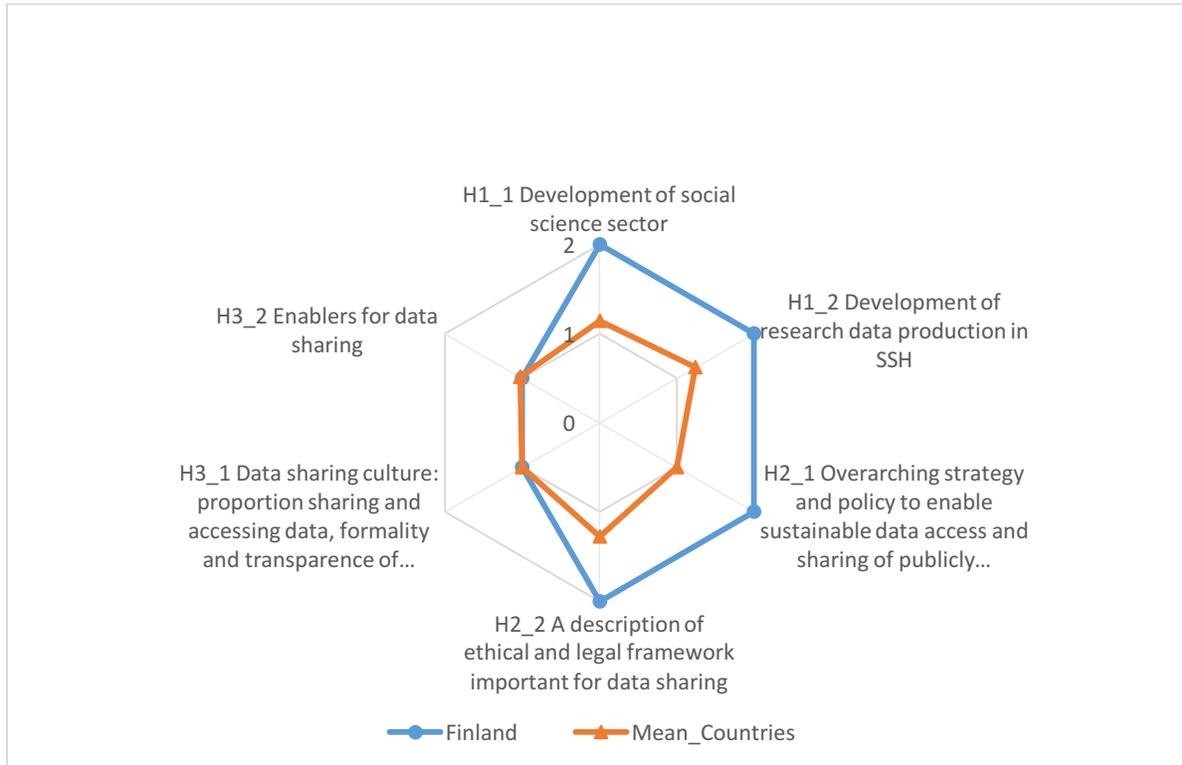


Figure 25 shows that Finland is on the developed level at all areas with exception of ‘Data sharing culture’ and ‘Enablers for data sharing’. Here is the country on the developing level.

Development of social science sector

Overall assessment of SSH development

The financial stability, research capacities and results achieved in the field of social sciences are in the highest quantile (on the developed level) and above the average. FSD (and several other SSH RIs) are mentioned in the Finland’s Strategy and Roadmap for Research Infrastructures 2014-2020⁷⁷. Furthermore, the national Etsin catalogue is a research data finder that contains descriptive information – that is, metadata – on research datasets. It is funded by the Ministry of Education and Culture (<https://etsin.avointiede.fi/en>).

In relation to research funding in social science, government and higher education sector is ranked first. The research is also funded abroad (international or cross-border) (ranked second), business enterprise sector (ranked third) and private non-profit sector (ranked fourth).

⁷⁷ http://www.aka.fi/globalassets/awanhat/documents/firi/tutkimusinfrastruktuurien_strategia_ja_tiekartta_2014_en.pdf

The FinELib consortium centrally acquires electronic materials for its member organisations. The consortium members are Finnish universities, universities of applied sciences, research institutes and public libraries. The consortium's service unit, the FinELib office, is located at the National Library. The FinELib office negotiates the licensing of e-materials on behalf of the consortium members⁷⁸.

Development of research data production in SSH

The research data production in SSH is on the developed level and above the average.

There are surveys from national election studies, ISSP, European Social Survey, and European Values Study. FSD is involved in all these surveys (for reports, see FSD's data documentation and information elements on publications).

The average production of research data by the social science institutions is frequent; the institutions have well established tradition in data production.

RDM Policy setting

Funders' data management and sharing strategy and/or policy

The overarching strategy and policy to enable sustainable data access and sharing of publicly financed social science research data is on the developed level and above the average.

The requirements or recommendations for DMPs in most cases are on the defined level; formal requirement. The major research funder, the Academy of Finland, has this autumn enforced a new data policy, which says: *"We require that principal investigators of Academy-funded research projects see to that the projects data are stored and made available through major national or international archives or storage services that are important in the fields concerned. Data may for justified reasons, however, come in varying degrees of openness, ranging from fully open to strictly confidential. The research project concerned and the publisher of the data must ensure that publishing the data will not be in breach of the Finnish Act on the Openness of Government Activities, the Finnish Personal Data Act or the Finnish Copyright Act. When making data openly available, the parties involved must also consider licensing issues"*.

The way of depositing data in an appropriate disciplinary repository is on the defined level, while the long-term curation for valuable research data assets is partial. The public research funding organizations in Finland provide partial incentives for sharing research data with associated metadata. The Academy of Finland's practices are described above. In addition, several other funders require data management plans and recommend archiving or sharing data. Examples are:

⁷⁸ See <https://www.kiwi.fi/display/finelib/In+English>.

- TEKES: "There should be a materials management plan...how and what kind of research material/data is collected and/or re-utilised"⁷⁹.
- KONE Foundation: "We recommend that data collected during research funded by Kone Foundation be archived for possible future use"⁸⁰.

Legal and ethical framework

The ethical and legal framework important for data sharing is on the developed level and above the average.

There are recommendations and guidance provided on how to respect the legal requirements while sharing data. The guidance of data protection officials is restricted due to its limited resources. The FSD has resources and gives information services for researchers concerning informing research participants⁸¹, anonymization of data files⁸², as well as copyright and agreements⁸³.

Data sharing culture

The data sharing culture in Finland is – like the average – on the developing level.

Data sharing and reuse among social sciences researchers in Finland is not so common - as self-assessment results indicate, the proportion of researchers sharing data is estimated as low (0-10%). It can be explained by the fact that until this autumn the main research funder has not required but just recommended research data to be archived for further use. The situation is expected to change in the following years. There is an ongoing Open Science and Research Initiative (ATT) where objective is to make Finland the leading country for openness in science and research by 2017 (<http://openscience.fi/>). The project is funded by the Ministry of Education and Culture. Also, many universities have published data policies in the last couple of years, and generally they encourage sharing data.

Concerning the proportion of researchers able to access existing third party data they need, it is medium (10-30%). It is because at least the FSD's user statistics from 2015 show that the use of FSD data collection continues to grow. This trend has been clearly visible since the launch of the data service portal Aila in mid-2014. The portal allows users to download data online. Compared to 2014, the number of datasets downloaded grew by 60%, reaching 2,200 by the end of the year 2015. It is still burdensome to get access to official administrative records and registries for research purposes. There is an introduction of a new law aiming to help the possibilities of using health and social service records and databases for research purposes.

⁷⁹ See section 15.1 of TEKES report. Available at:

https://www.tekes.fi/globalassets/global/asiointi/ehdot/en_julkisen_tutkimuksen_rahoituksen_vleiset_ehdot.pdf

⁸⁰ See <http://www.koneensaatio.fi/en/tuemme/forgrantrecipients/>

⁸¹ See <http://www.fsd.uta.fi/aineistonhallinta/en/informing-research-participants.html>

⁸² See <http://www.fsd.uta.fi/aineistonhallinta/en/anonymisation-and-identifiers.html>

⁸³ See <http://www.fsd.uta.fi/aineistonhallinta/en/copyright-and-agreements.html>

There are established data sharing channels and routines, as self-assessment indicates. Most popular data sharing channels includes informal contacts (peers and colleagues) (ranked first) and then data archive (ranked second).

According to self-assessment, attitudes towards data sharing in the period 2011-2016 are following:

Table 12: Attitudes towards data sharing in Finland

Data sharing has no benefits at all	False
Data sharing creates healthy competition	Neither true, nor false
Data sharing creates negative competition	Probably true
Reuse of existing data can answer new research questions and facilitate advancement of science	Probably true
Data sharing has as a risk that others may misuse and misinterpret data	Probably true
Data sharing involves little effort and minimal costs	False

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from "5"-True to "1" False!

The scores indicate that data sharing in Finland is beneficial but resource demanding. It probably creates negative competition and has a risk. At the same time, reuse can secure future development in research.

Enablers for data sharing

Enablers for data sharing in Finland are – like the average – on the developing level.

Finnish Social Science Data Archive (FSD) is a data support service, which is available to social science researchers.

Concerning career rewards, there are no career rewards related to data sharing.

Finally, concerning data management and data documentation standards and procedures, some data producers follow these.

3.13.2 Capability requirement areas of DAS

Organisational profile

FSD is a national resource center. It operates as a separate unit of the University of Tampere⁸⁴.

FSD has 15 permanent full-time staff members. About 10 staff members have temporary contracts, some of them part-time. There are four operational Modules: Administration and Communications; User Services and Data Ingest; Projects and Development; and Technical Services. FSD is led by the Director, assisted by the FSD's Management Team (Deputy Director, Module Managers and PA to the Director). Strategic and management direction is provided by the National Advisory Board.

Data archived at the FSD are mainly available for research, teaching and learning purposes. Through FSD's Aila data service portal the archived datasets are available free of charge to registered users, according to the conditions set for each dataset. Terms and conditions vary from dataset to dataset. There are a number of datasets that are (A) freely available to all users. Most datasets are (B) available for research, teaching and study, while some are for (C) research only (including e.g. Master's, licentiate and doctoral theses). In some cases access requires (D) permission from the data depositor. FSD serves both national and international research community.

Funding

FSD is publicly funded;

- Basic funding from the Ministry of Education and Culture and the University of Tampere
- Project funding (about 40% of budget) mainly from Academy of Finland
- Project funders include also CESSDA and EU

Core services and activities

The Finnish Social Science Data Archive (FSD) provides a single point of access to a wide range of digital research data for learning, teaching and research purposes. The archive is a national resource centre funded by the Ministry of Education and Culture. In addition to archiving and dissemination of data, key services include data-related information services and support for research data management. FSD promotes open access to research data as well as transparency, accumulation and efficient reuse of scientific research. Core services include depositing data, information services, Aila data portal, Data Management Guidelines, Research Methods Web Resource MOTV.

⁸⁴ See: <http://www.fsd.uta.fi/en/organisation/index.html>

Content current collection

FSD's data collection includes 1300+ studies, of which about 1100 are quantitative studies (mainly survey data) and about 200 qualitative studies (mainly textual data). The data holdings cover a wide variety of topics, including sociology, social policy, elections, voting, religion, welfare, local government studies and media studies.

Most of the datasets held by FSD are in Finnish but there are also data in English and Swedish. Study-level metadata are always provided in both Finnish and English. Variable-level metadata is provided in the language of the data file, most cases Finnish, but data is also translated into English on request.

Organisational infrastructure

Figure 26 Organizational infrastructure in Finland

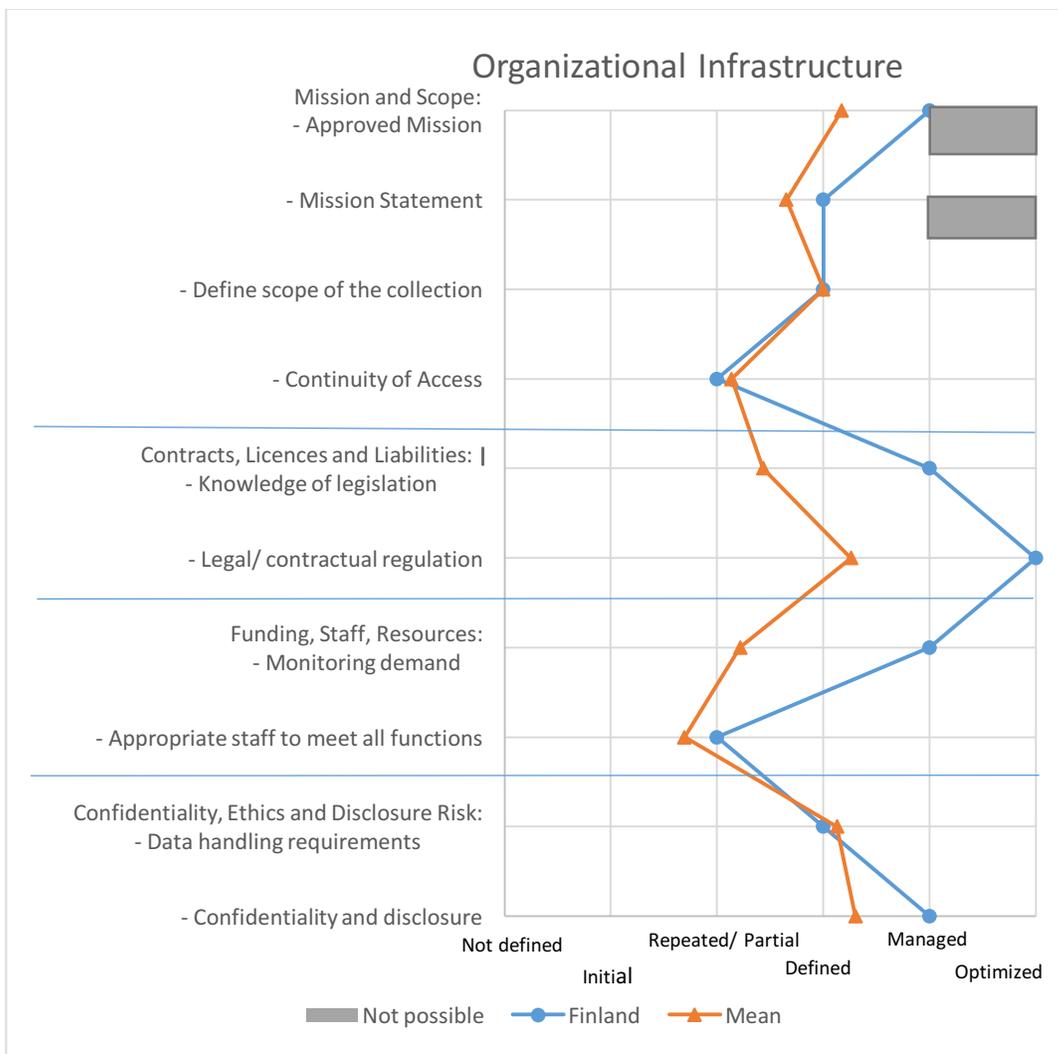


Figure 26 shows that the organizational infrastructure on some areas is on the repeated/partial level, on other areas on the defined level and in other areas again on the managed level. Finland scores on some areas like the average and in general above the average.

Mission statement – unofficial translation

The Finnish Social Science Data Archive (FSD) serves research and teaching on the national level. The FSD archives digital research data from Finland and abroad and disseminates it for research, teaching and learning purposes.⁸⁵

Continuity of access is on the repeated/partial level. A partial plan is in place but limited in detail and scope⁸⁶.

Appropriate staff to meet all functions is on repeated/partial level. The staff number is evaluated at regular intervals as the University requires and there have been recognised a shortage of IT personnel. In January 2017, the Academy of Finland granted FSD five-year infrastructure funding. The new funding will strengthen FSD's IT competence and capacity. It allows FSD to focus on meeting the requirements set by CESSDA as well as actively taking part in development work within the research infrastructure⁸⁷.

Digital object management (data curation) and Technical infrastructure and risk

Digital object management (data curation)

Preservation strategies are on the managed level⁸⁸.

Technical infrastructure and risk

Technical planning and management is shown on the repeated/partial level in Figure 3. Finland is somewhere between partial and defined. Internal handbook, asset management systems and documentation provide evidence that FSD likely to be able to support all current functions and services. However, there are no formal statements/definitions of what appropriate levels are. Also, monitoring (managed) is partially implemented. Long term funding for sufficient resources is lacking (including sufficient IT staff).

Appropriate succession plans and/or contingency plans are shown on the repeated/partial level in Figure 3.13.3. Finland is somewhere between partial and defined. Data holdings and their storage is well enough documented for successfully managed exit scenario. However, actual exit plans or scenarios do not exist.

Risk assessment procedure is on the initial level. Risks are assessed whenever there are changes in the tech infrastructure. There is however no formal methodology. Documentation of assessments is lacking.

In January 2017, the Academy of Finland granted FSD five-year infrastructure funding. The new funding will strengthen FSD's IT competence and capacity.

⁸⁵ See: <http://www.fsd.uta.fi/en/organisation/documents/AMS/index.html>.

⁸⁶ See FSD's Records Management and Archives Formation Plan. Available at: <http://www.fsd.uta.fi/en/organisation/documents/AMS/>

⁸⁷ News item can be found at http://www.fsd.uta.fi/en/news/index.html#v17_4

⁸⁸ See FSD's Records Management and Archives Formation Plan. Available at: <http://www.fsd.uta.fi/en/organisation/documents/AMS/>

Figure 27: Digital object management and Technical infrastructure in Finland

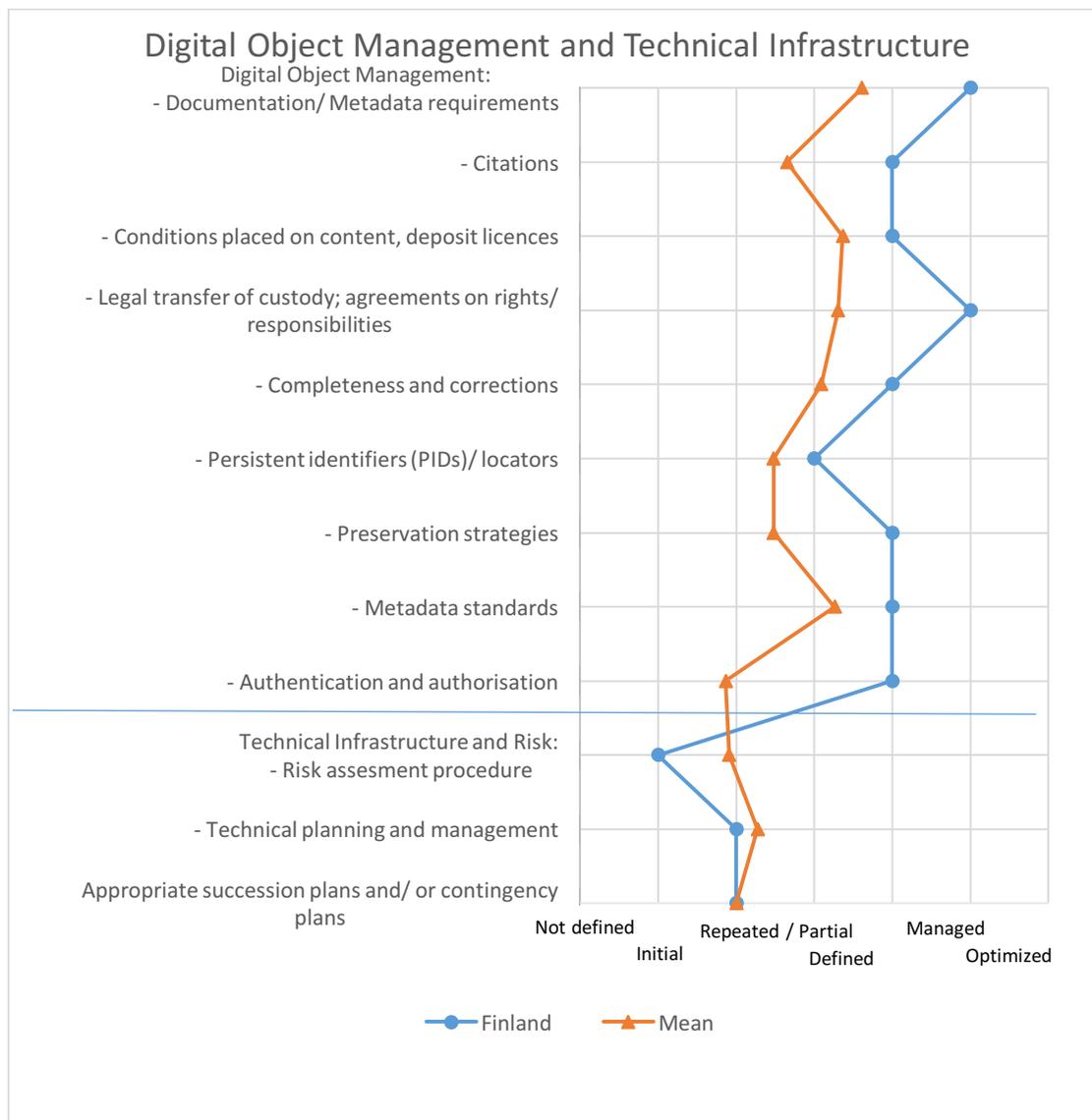


Figure 27 shows that digital management (data curation) in general is on the managed or optimized level and above the average. The technical infrastructure and risk is on the initial and repeated/partial level and below or like the average.

3.13.3 Conclusions

The self-assessment was conducted by Finnish Social Science Data Archive (FSD) and the data collection process went very well. The results show that Finland in general has a fine, medium-high maturity level with potential for development.

In relation to the broader eco-system of DAS operation, Finland is on the developed level in the development of social science sector. FSD is involved in a lot of (inter)national surveys and the social science institutions have well established tradition in data production. In relation to the

RDM Policy setting, Finland is also on the developed level. There are formal requirements for DMPs in most cases and a new policy was enforced by the major research funder, the Academy of Finland, autumn 2016. There are recommendations and guidance provided in relation to legal requirements while sharing data. There are potential for developing the way of depositing data in an appropriate disciplinary repository. The data sharing culture is developing rapidly and there is an ongoing Open Science and Research Initiative (ATT) where objective is to make Finland the leading country for openness in science and research by 2017. Concerning enablers for data sharing, FSD is a data support service available to social science researchers.

In relation to the capability requirement areas of DAS, the organizational profile shows that FSD in 2016 had 15 permanent full-time staff members meanwhile about 10 staff members had temporary contracts and some of them part-time. Data archived at the FSD are mainly available for research, teaching and learning purposes. The organizational infrastructure is both assessed at the repeated/partial level, on the defined level, and on the managed level. In relation to continuity of access, a partial plan is in place but limited in detail and scope. It is possible to imagine that it can be developed further in the coming years. The appropriate staff to meet all functions is on the same level because there have been recognized a shortage of IT-personnel. The digital object management (data curation) is in general on the managed or optimized level and above the average. The technical infrastructure and risk is on the initial and repeated/partial level and below or the same as the average. In relation to technical planning and management, FSD is likely to be able to support functions and services but there are missing formal statements/definitions of appropriate levels. There is also a need for developing exit plans or scenarios (in regards to succession/contingency plans) and a formal methodology for risk assessment procedure. These are all issues that FSD will address during the next five years in an infrastructure project funded by the Academy of Finland.

3.14 France

On a policy level, there is a growing awareness towards data management and data sharing.

There are also several activities considering data management policies. Information cannot always be found easily for people not familiar with the structure of research funding in France. Efforts could be extended to more comprehensive policies and guidelines on a national level.

France is participating in cross-national studies and there is funding of government resource. PROGEDO is also funded publically.

On the organizational level, there are some strengths and gaps. There is a good support for example in the field of data confidentiality, handling of the data. Preservation strategies and single measures should be developed and implemented.

Specifics about the data collection in a country

Survey was filled in by the head of PROGEDO, Bénédicte André, as PROGEDO will be the CESSDA Service Provider.

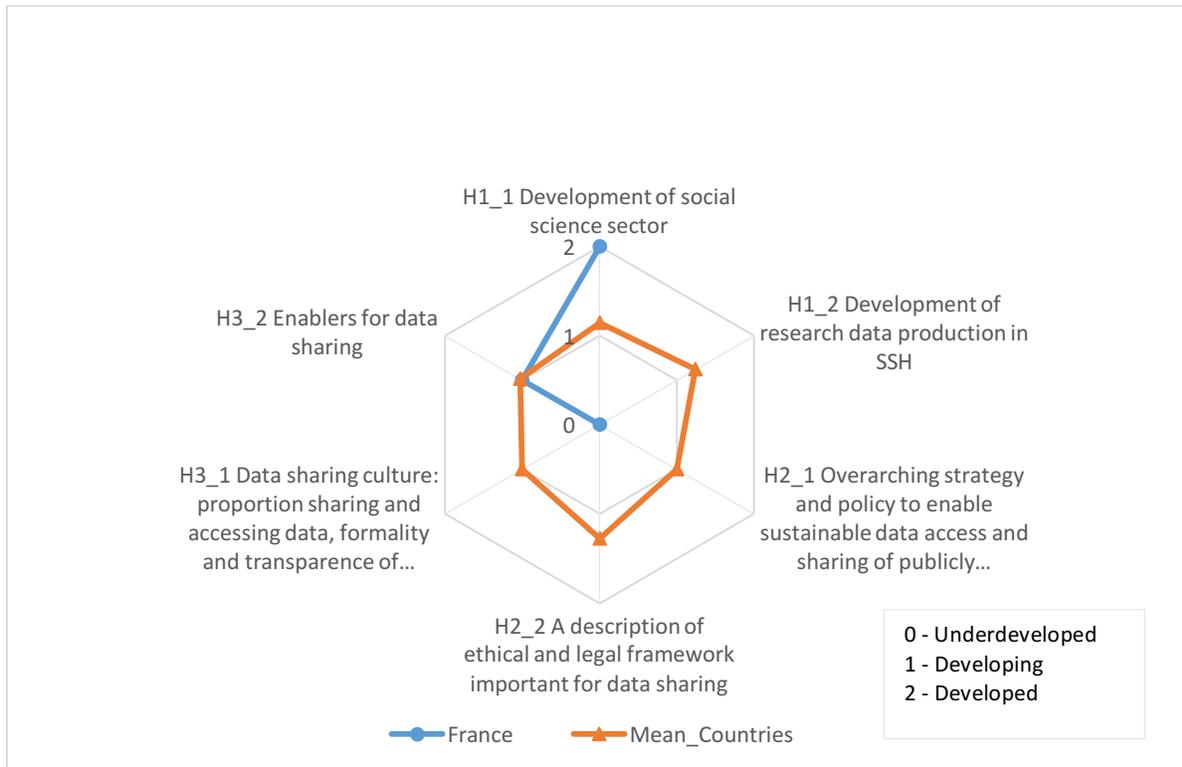
The situation in France is a little bit confusing as there are several institutions involved in activities and there were many changes of responsibilities. There are also a lot of abbreviations used at the website that are not so easy to assign if one is not familiar with the structures in France.

Initially it was not clear who should be contacted, but Nathalie Patón (also involved in CESSDA SaW) recommended contacting Bénédicte André.

Additional to the data collection through the survey, a desktop research was conducted for those aspects that were not covered by the survey.

3.14.1 Broader ecosystem of DAS operation

Figure 28: Heading concepts values in France



Development of social science sector

France is involved in almost all listed surveys or survey programs. On a national level, National Institute of Statistics and Economic Studies INSEE⁸⁹ collects, analyses and disseminates information on the French economy and society.

"The term "official statistics" includes all material generated by statistical surveys, as specified in the list determined every year in a ruling by the Ministry for the Economy, and the use of data collected by government administrations, public or private bodies with a public service role for purposes of general information. The design, production and dissemination of official statistics are conducted with full professional independence by the official statistical system, and by producers approved by the National Council for Statistical Information (CNIS) or the Official Statistics Authority (ASP)." [Comment to question I1.2.2, see Appendix 2]

With regards to an overarching strategy and policy to enable sustainable data access and sharing of publicly funded social science research data, there are the following measures in place. There is no handle prefix specific to PROGEDO's data but some references –Adisp's catalogue- have a persistent identifier. Indeed, some data (the Adisp's catalogue) are harvested by the search engine Isidore (<https://www.rechercheisidore.fr/>) that assigns them one identifier with its own prefix. An internal action is necessary to transfer this identifier in the DC Identifier of the OAI - PMH.

⁸⁹ See <http://www.insee.fr/en/themes/TOPICS>

In France, access is provided to bibliographic and full-text databases, to datasets, to software licenses by PROGEDO (<http://www.progedo.fr/>) for datasets and by CNRS research units' documents access, BibCnrs (<https://bib.cnrs.fr>) for publications. [Comment to question I1.1.6]

"Work is presently being conducted to reference datasets, databases & software licenses available in France but at the moment points of access of extremely numerous and other than citing two main actors, information would be incorrect. At the moment, Progedo acts as a main point of access for datasets and bib.cnrs.fr is the main point of access for databases." [Comment to question I1.1.8]

RDM Policy setting

There are two funding agencies in France. The French National Research Agency (ANR) has no information about data access or data sharing on their website.

At first glance, the main website of the National Centre for Scientific Research CNRS does not provide information about data sharing. The website from the Institute Information Sciences and Technologies lists publications that include among others strategic issues. The site itself is only available in French but some of the documents are in English as well. In the paper: "A better sharing of knowledge - An open policy for scientific and technical information of the future", published in December 2014 by CNRS conditions for supporting data sharing the following (institutional) actions are listed: "A federating text needs to be disseminated throughout the CNRS expressing unconditional support for research data" and "The organization's policy on research data needs to be defined and communicated." (p 55)

Researchers should be made aware of the benefits of data sharing and supported by training and adequate infrastructure. It is also stated within this policy that there is a "necessity for a general policy text on data which defines the conditions for uploading and sharing data and sets out the legal framework and relevant business models." (ibid)

Distributed to different institutes of CNRS there are activities towards digital management, data access and data sharing activities but an overall strategy is still lacking.

The website from the Laboratory for Analysis and Architecture of Systems (LAAS), a CNRS research unit linked with the Institute for Engineering and Systems Sciences (INSIS) and the Institute of Information Sciences and their interactions (INS2I) provides guidelines on Open Access and Data Management within the Horizon 2020 to their researchers⁹⁰.

For the Institute for Humanities and Social Sciences INSHS "sharing scientific and technological information is one of the Institute's main concerns. The INSHS supports a range of dissemination tools aimed at increasing the visibility of its researchers, such as the HAL-SHS open archives. It has also developed an open-access publishing strategy for scientific works through the portal Revues.org, the Hypotheses.org platform for academic blogs, etc. This policy is part of the European project developed through the Dariah infrastructure.

Finally, the Institute takes part in the technology transfer consortium (CVT) of the Athena Alliance in the humanities and social sciences."⁹¹

⁹⁰ See <https://www.laas.fr/public/en/open-access-and-research-data>

⁹¹ See <http://www.cnrs.fr/en/institutes/inshs-humanities-social-sciences.html>

It seems that there are some activities with regards to research data management. The requirements of Horizon 2020 proposals have led to more awareness of RDM⁹².

Data sharing culture

For this section, there are no entries in self-assessment survey.

With regards to research data some studies at single institutions exists. In their paper about a survey about data management in social sciences at the University of Lille, the authors explain: “*In France, the Conference of University Presidents put the issue of research data preservation and sharing at the top of their priorities during their annual conference in 2015, the Ministry of Higher Education and Research supports and promotes related actions, and all major public research organizations such as the CNRS (Centre national de la Recherche Scientifique) contribute to the development of data infrastructures and repositories.*”⁹³

To sum up in France there are efforts to establish a data sharing culture that could be extended to more comprehensive policies and guidelines on a national level. France is participating in cross-national studies and funding of government resource.

Enablers for data sharing

There are also some activities concerning RDM within the library community (e.g. study day on research data management in Le Havre in 2013⁹⁴, or a training session on research data at the National school for librarians in 2015⁹⁵).

3.14.2 Capability requirement areas of DAS

Organisational profile

Organisation

The acronym PRODEGO means “Production et gestion des données en sciences sociales” (Production and management of the data in social sciences). PRODEGO is located in Paris, France. [Answers from III.1.1 and information from the website (partially identical)].

PROGEDO is composed of three departments:

- CESSDA-France: to disseminate French data and give easier access to international data
- ESS-France: to coordinate and promote the use of the European Social Survey

⁹² See <http://libreas.eu/ausgabe29/09schoepfel/>

⁹³ Research data management in social sciences and humanities: A survey at the University of Lille; Joachim Schöpfel & H  l  ne Prost, available at: <http://libreas.eu/ausgabe29/09schoepfel/>

⁹⁴ See: <http://blogs.arts.gla.ac.uk/hatii/research-data-management-the-view-from-france/>

⁹⁵ See: <https://blogs.openaire.eu/?p=332>

- SHARE-France: to coordinate and promote the use of the Survey on Health Aging and Retirement in Europe.

Core services and activities

“The PROGEDO large infrastructure ensures the implementation of a public policy for social sciences and humanities. It focuses especially on research on law, economics, geography, management, history, political sciences and sociology.

Its role is based on the missions defined since 2008 in the national roadmap for large research infrastructures produced by the French Ministry for Higher Education and Research.

Reaffirmed in the national strategy for the development of infrastructures by 2020, its constitutive missions are the following:

- The development of facilities for the archive, the documentation and the availability of data and its perimeter, including a device for the secured remote access to confidential data,
- The participation to the production of large multidisciplinary studies with a national interest and selected of the European infrastructure roadmap (ESFRI),
- The animation of a national dynamics centred on the field of competence of the infrastructure. Educational & Training resources: Educational resources (curated data for educational purposes)”

The French Data Archive for Social Sciences, CESSDA France (former Réseau Quételet) in cooperation with its partners, provides researchers from France and abroad interested in data treatment with the requisite access to databases in the following domains:

- Major data, censuses and other databases of French National Statistics
- Major French research data
- Privileged access to international data

“The portal of the Réseau Quételet is presently being remodelled and information should be updated for the launch of the new website, due in the 2017 (3rd trimester).” [Comment to II.1.2.1.1]

Training sessions for programming language & software use (R/Stata);

Training sessions for users: access & handling data” [III.1.1]

Access conditions are defined and available⁹⁶.

“There has been an important turnover at Progedo in 2016 considering the main office is constituted of 4 people and two are leaving. Two new people will be hired within a month. Additional positions have been requested.” [Comment to II.1.3.4.1]

PROGEDO also offers training sessions for their users on access and handling data.

⁹⁶ See: http://www.reseau-quetelet.cnrs.fr/spip/article.php3?id_article=149&lang=en

Content current collection

More than 1500 datasets are currently available via the CNRS and its component units.

The data catalogue of the CNRS includes data from national statistics offices, from the world of academic research, as well as datasets from polling organizations. National, international and confidential data (via the Secure data access centre, CASD) is held at CNRS.

Within the frame of the Data Committee (CCDSHS) policy, CNRS coordinates the archiving, documentation and distribution of data in the humanities and social sciences from the Centre Maurice Halbwachs (ADISP), the Centre for Socio-political Data (CDSP) the INED data service and the secure data access centre (CASD).

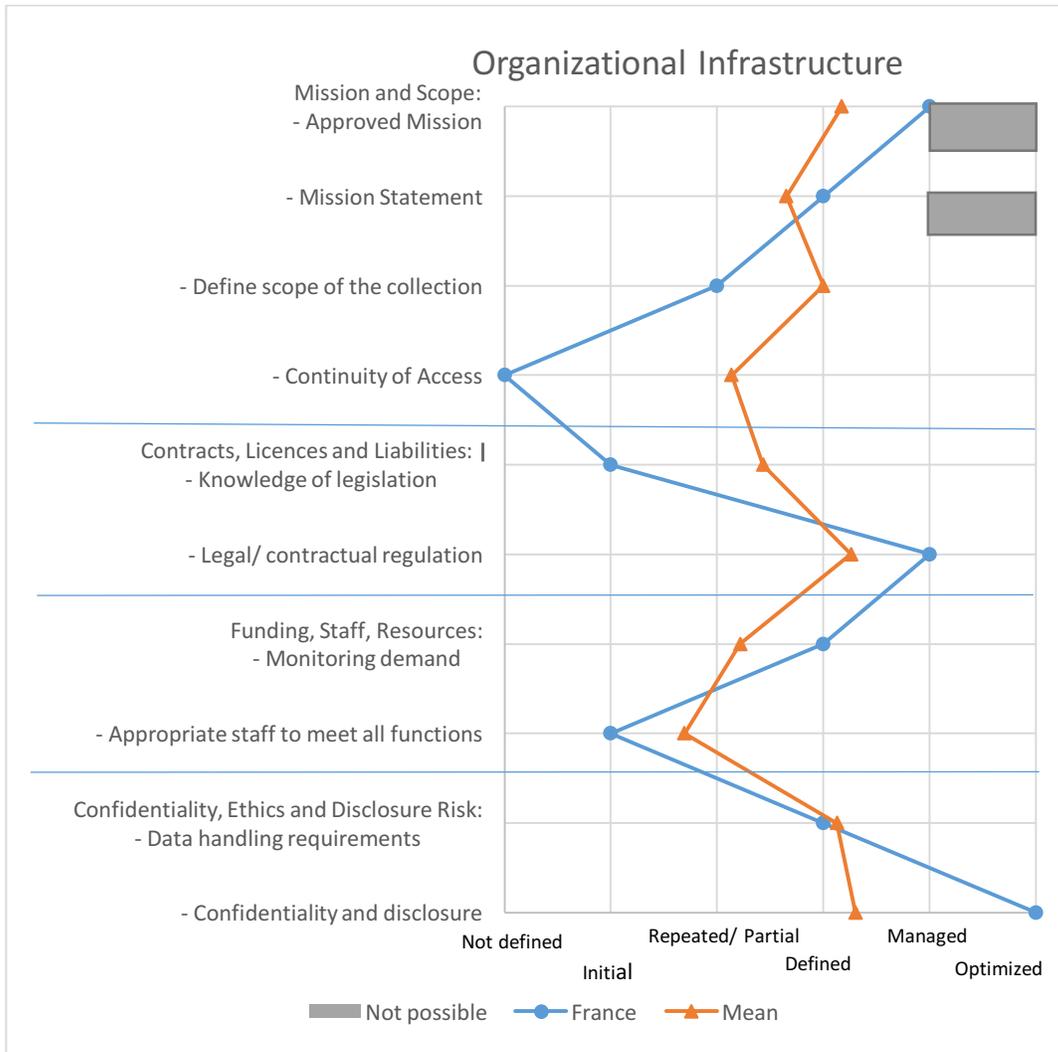
PROGEDO (with its different departments) offers various services and access to different kind of research data. Quantitative and qualitative data from the world of academic research include major socio-demographic surveys, socio-political data (e.g. Electoral results), and data on different topics such as education and training, work and employment, salaries and income, living and social conditions, health, population and demography, opinions, companies, as well as historical data. INED's surveys address a variety of topics such as the rates and evolution of fertility, familial structures, gender relations, geographical mobility, homelessness, health, mortality rates, biographical trajectories, and so on. Data from national statistics offices are also included in PROGEDO's collection. These include censuses, multiple surveys from Insee and the statistical offices of various ministries, as well as from other data producing agencies

Organisational infrastructure

Mission and scope

“PROGEDO (PROduction et GEstion des DONnées) aims at the production and management of data in social sciences and humanities. It gathers actors involved in qualitative surveys around a two-dimension common national policy: the production and the availability of SSH data. Aimed at equipping France with an infrastructure comparable with its European counterparts, PROGEDO is involved in three large consortia such as ERICs for databanks (CESSDA) and European studies (ESS and SHARE). PROGEDO is included the French large research infrastructure roadmap since 2008” [Comment to II 1.1.1.2]

Figure 29: Organizational infrastructure in France



Contracts, licences, liabilities

The CNRS (Progedo) makes agreements guaranteed by the Minister of Higher Education and Research with the holding organizations and supplies deposit licenses at a more individual level for researchers that have formed data bases. A model agreement (in French) is available at the website⁹⁷.

⁹⁷ See http://www.reseau-quetelet.cnrs.fr/spip/IMG/pdf/2013_convention_type.pdf

Funding, staff, resources

PROGEDO is funded 100% publically. [Answer III.1.2]

Confidentiality, ethics and disclosure risk

Concerning confidentiality, disclosure risk and ethics, is optimized. The Commission on Information Technology and Liberties CNIL (<https://www.cnil.fr/en/home>) provides information on data protection and offers support and advises. “On a national level, dealing with personal data is highly monitored (cf. CNIL). Within Progedo, notably CASD is specialized in this type of activity and partners benefit from their knowledge. <https://casd.eu>.” [Comment to II.1.5.1.3] CASD enables users to access confidential data in several ways.

In the field of mission, legal regulation and handling confidentiality and ethics PROGEDO is very well positioned. As many other institutions, there is no succession plan in place. The demand of resources and staff is defined even if for some functions staff is still lacking.

Digital object management (data curation) and Technical infrastructure and risk

Data object management and technical infrastructure offers a range from “not defined” to “managed” for the different aspects. Requirements and processes regarding documentation and metadata are defined and managed⁹⁸.

The CNRS makes agreements guaranteed by the Minister of Higher Education and Research with the holding organizations and supplies deposit licenses at a more individual level for researchers that have formed data bases⁹⁹.

Requirements and processes regarding documentation and metadata are defined and managed¹⁰⁰.

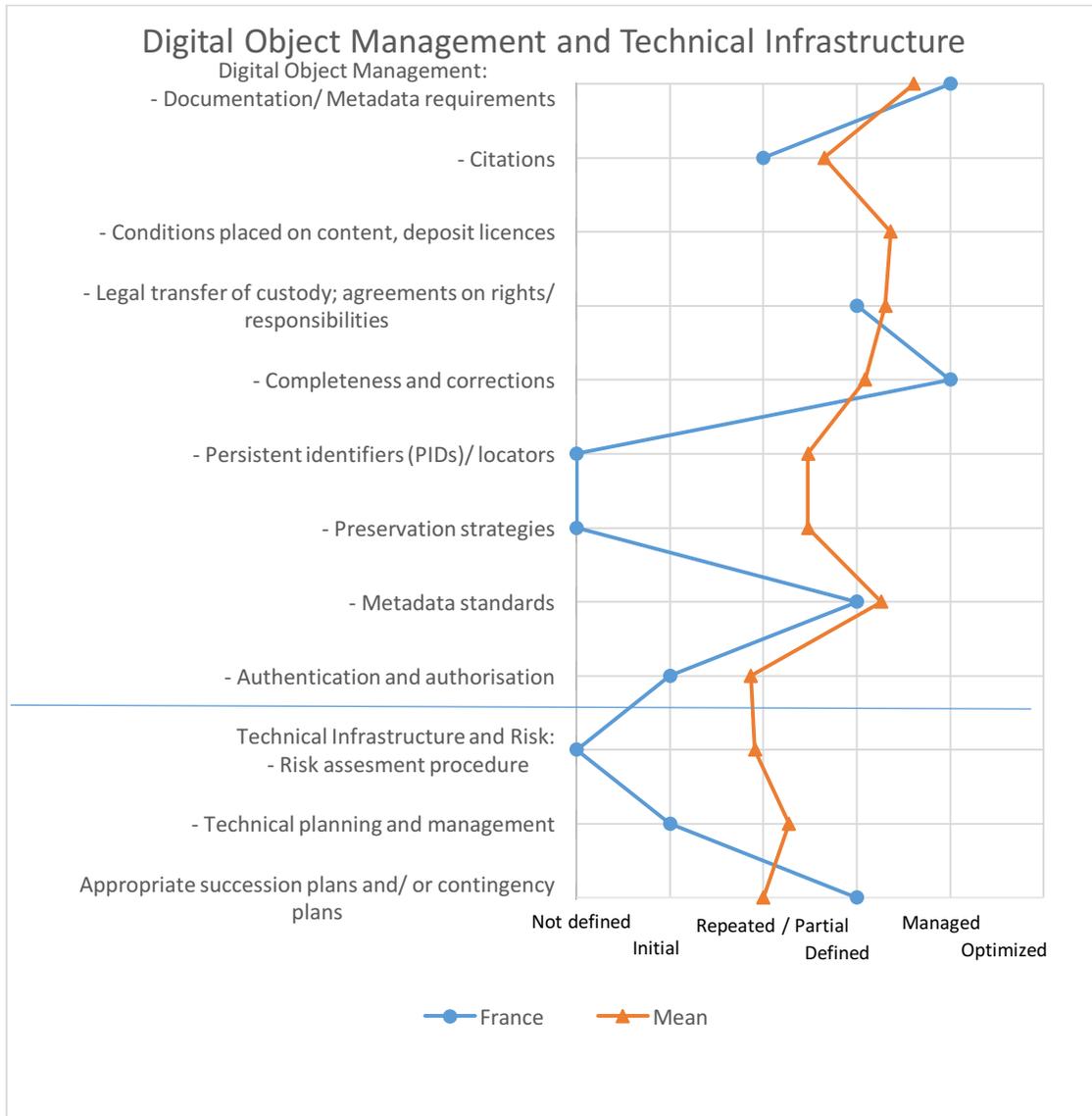
There are some persistent identifiers but not yet implemented in the CNRS units catalogue. There is lack of choice of access conditions by data depositors. Preservation strategies for the long-term and a sustainable risk management are also lacking. These issues should be addressed with regards to trust and long-term preservation.

⁹⁸ More information about it is available at: http://www.reseau-quetelet.cnrs.fr/spip/article.php3?id_article=140&lang=en.

⁹⁹ A model agreement (in French) is available at: http://www.reseau-quetelet.cnrs.fr/spip/IMG/pdf/2013_convention_type.pdf.pdf

¹⁰⁰ More information about it is available at: http://www.reseau-quetelet.cnrs.fr/spip/article.php3?id_article=140&lang=en.

Figure 30: Digital object management and Technical infrastructure in France



3.14.3 Conclusions

In general, France is on a good way. There is awareness of the necessity to optimize policies as well as infrastructure issues. Gaps are at a national level with regards to policy development and common requirements towards data management and a data sharing culture. As a good example for support researchers could be the Commission on Information Technology and Liberties CNIL that provides information and consulting on data protection.

As the social science data infrastructure in France is in a transition phase with several organizational changes these will also influence data research management and data archiving in the future. On an organizational level as well with regards to data object management, there are still some gaps that have to be optimized. It seems that there is a lack in documentation and implantation of tasks that are crucial for long-term preservation. There is not yet a mechanism for providing persistent identifiers. The institution has strength e.g. in handling confidentiality and data with disclosure risk.

3.15 Germany

Currently, in comparison with other scientific disciplines, the social sciences receive 5% of the gross domestic expenditure on research and development (GERD). German institutions and researchers have a well-established tradition in social science research data production. The production of research data by social science institutions is also frequent. The German institutions are very active in international collaborative research.

The deposit of data in an appropriate disciplinary repository is not yet required by public funders of social science research, but there is a clear awareness of the need to have discipline-specific places for deposit and support services. Significant funding for infrastructure projects enables the archiving of data and metadata.

Data sharing and reuse among social sciences researchers in Germany is very common. The proportion of researchers sharing data and the proportion of researchers able to access existing third party data they need are both estimated as high (above 30%). Further, the attitudes of researchers toward data sharing in the social science community are positive. Still, currently there are no career rewards (e.g. influence on career progression, higher success rate in obtaining research funding, better standing within the research community) related to data sharing among the academic community.

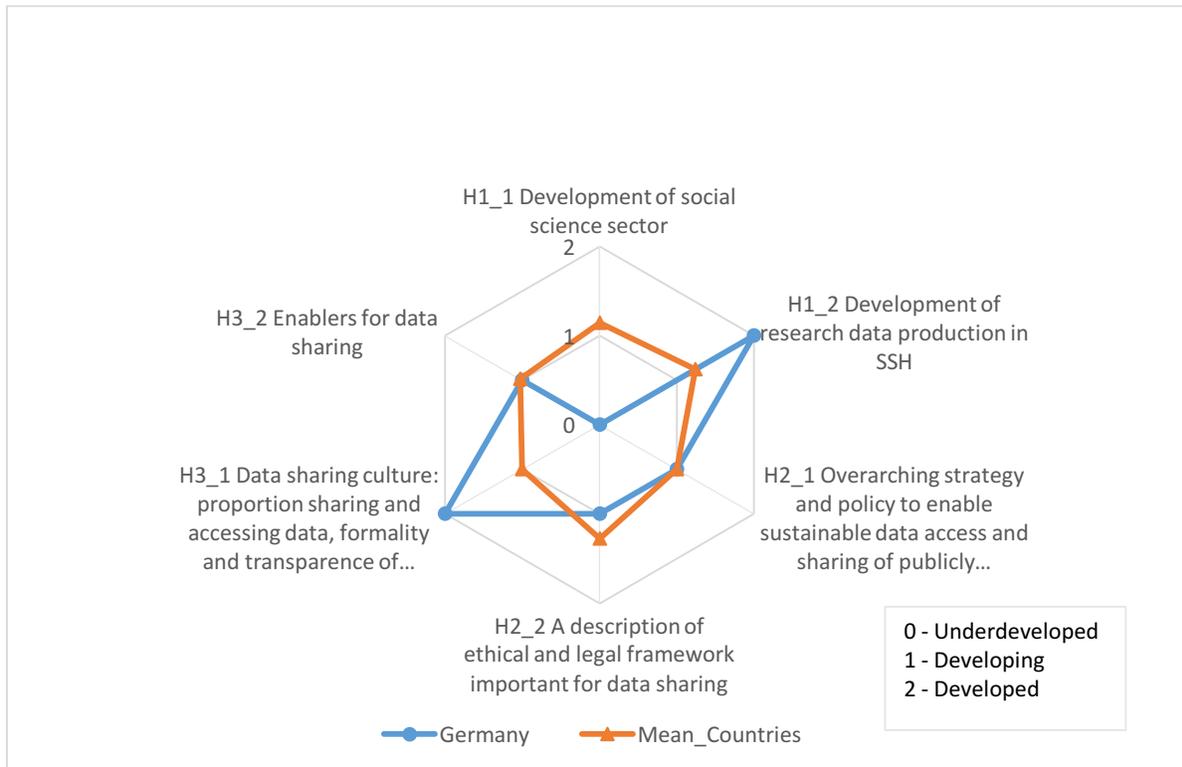
With 60 employees working at the data archive department within GESIS, staffing level is appropriate to support functions, and service needs are monitored and assessed at regular intervals. GESIS is very advanced with respect to digital object management and technical infrastructure, compared to the other European data services. GESIS has appropriate technical infrastructure resources to ensure that all functions and services of the repository are supported.

Specifics about the data collection in a country

Data were collected through the self-assessment survey. Dr. Alexia Katsanidou, head of GESIS data archive department, answered the online survey.

3.15.1 Broader ecosystem of DAS operation

Figure 31: Heading concepts values in Germany



Development of social science sector

Overall assessment of SSH development

In Germany, funding in social science research mainly comes from the government and the higher education sector sources. International and cross-border funding constitutes the second funding source. This is followed by funding from the private non-profit sector and finally from the business enterprise sector (sources: self-assessment survey).

Currently, in comparison with other scientific disciplines, the social sciences receive 5% of the gross domestic expenditure on research and development (GERD). Among the available figures, Germany is the third country that spends the least in the social sciences relative to other scientific disciplines¹⁰¹.

Access to commercial bibliographic and full-text databases (i.e. WoS, Scopus, EBSCO, JSTOR, etc.), datasets (i.e. Bankscope) and licenses for specific research software (as STATA, SPSS, etc.) is variable. Some universities provide access to these, some a combination of these, and some none. The government offers nothing directly.

¹⁰¹ Social sciences spending in GERD percentages were available for 31 countries (among the countries target within this report). The minimum is 4.1 %, the mean 13.7% and the median 12.9%.

Development of research data production in SSH

German institutions and researchers have a well-established tradition in social science research data production.

Various studies produced by social science researchers assess on a systematic basis matters of national importance (for example, German Internet Panel, SOEP, German Longitudinal Election Study, ALLBUS).

The production of research data by social science institutions is also frequent. The German institutions are also very active in international collaborative research. Indeed, Germany participates in the eight selected international surveys.

Table 13: Participation of Germany in a selection of international surveys

Comparative Candidate Survey (CCS)	Yes
Comparative Study of Electoral Systems (CSES)	Yes
European Social Survey (ESS)	Yes
European Values Study (EVS)	Yes
Generations and gender programme (GGP)	Yes
International Social Survey Program (ISSP)	Yes
Programme for International Student Assessment (PISA)	Yes
World Values Survey (WVS)	Yes

Research Data Management Policy Setting

Currently, in Germany there is growing recognition and awareness of the need to require Data Management Plans (DMPs). The German ministry, when acting as research funder, requires a DMP for some funding programs. This is not the case for the majority of other funding agencies. Most public funders do not allocate resources to cover the cost for managing data and preparing them for access during the research project lifetime.

The deposit of data in an appropriate disciplinary repository is not yet required by public funders of social science research, but there is a clear awareness of the need to have discipline-specific places for deposit and support services. Indeed, significant funding for infrastructure projects enables the archiving of data and metadata, the development of metadata standards and their use, but, according to Alexia Katsanidou, this is not transferred to the researchers.

Libraries and infrastructure organisations are offered funding to develop support services on legal and ethical aspects of data sharing. Public funders themselves do not offer any kind of support or impose requirements directly on researchers.

Data Sharing Culture

Data sharing and reuse among social sciences researchers in Germany is very common – as self-assessment results indicate. Indeed, the proportion of researchers sharing data and the proportion of researchers able to access existing third party data they need are both estimated as high (above 30%). According to the respondent of the self-assessment, in the social sciences there is more awareness on this topic than in the natural or medical sciences. This is mainly due to the fact that in the social sciences infrastructures exist that enable researchers to share and reuse data. This is not the case in other disciplines¹⁰².

There are established data sharing channels and routines. The self-assessment results indicate that the most popular data sharing channels include formal and transparent channels, with data archives or repositories ranked first and journals ranked second. Data are also shared via projects or personal websites (ranked third) and personal contacts (ranked fourth), channels that lack formality and transparency.

The attitudes of researchers toward data sharing in the social science community, as the self-assessment shows, can be characterized as positive. Indeed, according to the respondent of the self-assessment, social science researchers in Germany think in general that data sharing has benefits. They consider that reuse of existing data can answer new research questions and facilitate advancement of science, even if they think that there is a risk of data misuse and

¹⁰² A survey conducted in 2012-2013 among the researchers from all the faculties and institutes of the Humboldt University of Berlin ($n=469$) shows that 17% of the respondents have already deposited once their data in a data archive and that 14% were planning to do it, but that this possibility was not known by 34% of them. Also, the survey reveals that 27% have already downloaded data, while 10% planned to do it and 33% did not know this possibility. If the publication did not compare the social sciences researchers' answers from other researchers' answers, these figures demonstrate that in general about 30% of the researchers (all fields of study) share their data and have access to third party data in Germany.

Simukovic E., Kindling M., Schirnbacher P., 2013, *Forschungsdaten an der Humboldt-Universität zu Berlin Bericht über die Ergebnisse der Umfrage zum Umgang mit digitalen, Forschungsdaten an der Humboldt-Universität zu Berlin*, Version 1.0, available at: <http://edoc.hu-berlin.de/oa/reports/reFIYMgduNiVE/PDF/22YavRASzVauc.pdf> (accessed on 09.02.2016).

misinterpretation. However, data sharing is still perceived as a time- and resource-consuming activity.

Table 14: Attitudes toward data sharing in Germany

Data sharing has no benefits at all	Probably false
Data sharing creates healthy competition	Neither true, nor false
Data sharing creates negative competition	False
Reuse of existing data can answer new research questions and facilitate advancement of science	Probably true
Data sharing has as a risk that others may misuse and misinterpret data	True
Data sharing involves little effort and minimal costs	False

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five point scale from "5"-True to "1" False!

To summarize this aspect of the research culture in Germany, data sharing and reuse is very common, there are established (both formal and informal) channels for data sharing, and general attitudes of researchers related to sharing of data is rather positive. Thus, the data sharing culture can be characterized as developed.

Enablers for data sharing

The incentives and enablers for data sharing within the social science research community in Germany can be benchmarked as developing.

On the one hand, currently there are no career rewards (e.g. influence on career progression, higher success rate in obtaining research funding, better standing within the research community) related to data sharing among the academic community. On the other hand, data support services that facilitate data sharing and/or open access to research data are provided by different institutions like GESIS, University Library Services, and Research Data Centres. University libraries or universities often run institutional repositories, where publication and data for all disciplines can be published, and promote and support open access. Also, in Germany most data producers follow data management and data documentation standards and procedures. This facilitates data sharing and reuse.

3.15.2 Capability requirement areas of DAS

Organisational profile

Organisation

GESIS is a large organisation with five departments. The data archive is one department. It also collaborates with the following units: German Microdata Lab that is part of GESIS and provides metadata and data access to official microdata and Knowledge technologies for the social sciences that produces all technical solutions for the archive.

The archive consists of seven teams: Data registration agency da|ra, Archive operations, Archive instruments and metadata standards, International surveys, National surveys, Data linking and data security (including the Secure Data Center), and Producer relations and outreach. The archive offers also several research data centers: International Surveys, ALLBUS, PIAAC, Elections.

GESIS' primary target group includes sociologists and political scientists. The secondary user community consists of all researchers interested in social science research questions.

Funding

The data archive service is mainly funded by the Federal Ministry of Research and Education as well as the ministries of research and education of two federal states. Third party funding comes from the European Union, the German Research Council and others.

Core services and activities

Core services and activities within the GESIS archive include:

- Self-managed archiving
- Standard archiving
- In-depth added value archiving for high quality studies
- PID agency
- Secure Data Centre
- Data management training
- Anonymization and documentation checks
- Development of Metadata Standards, and
- Long term preservation

Content of the current collection

The data archive mainly offers survey data, but also aggregate data for different geographical units. The main languages of the data and metadata collection are German and English.

Organisational infrastructure

Figure 32: Organisational infrastructure in Germany

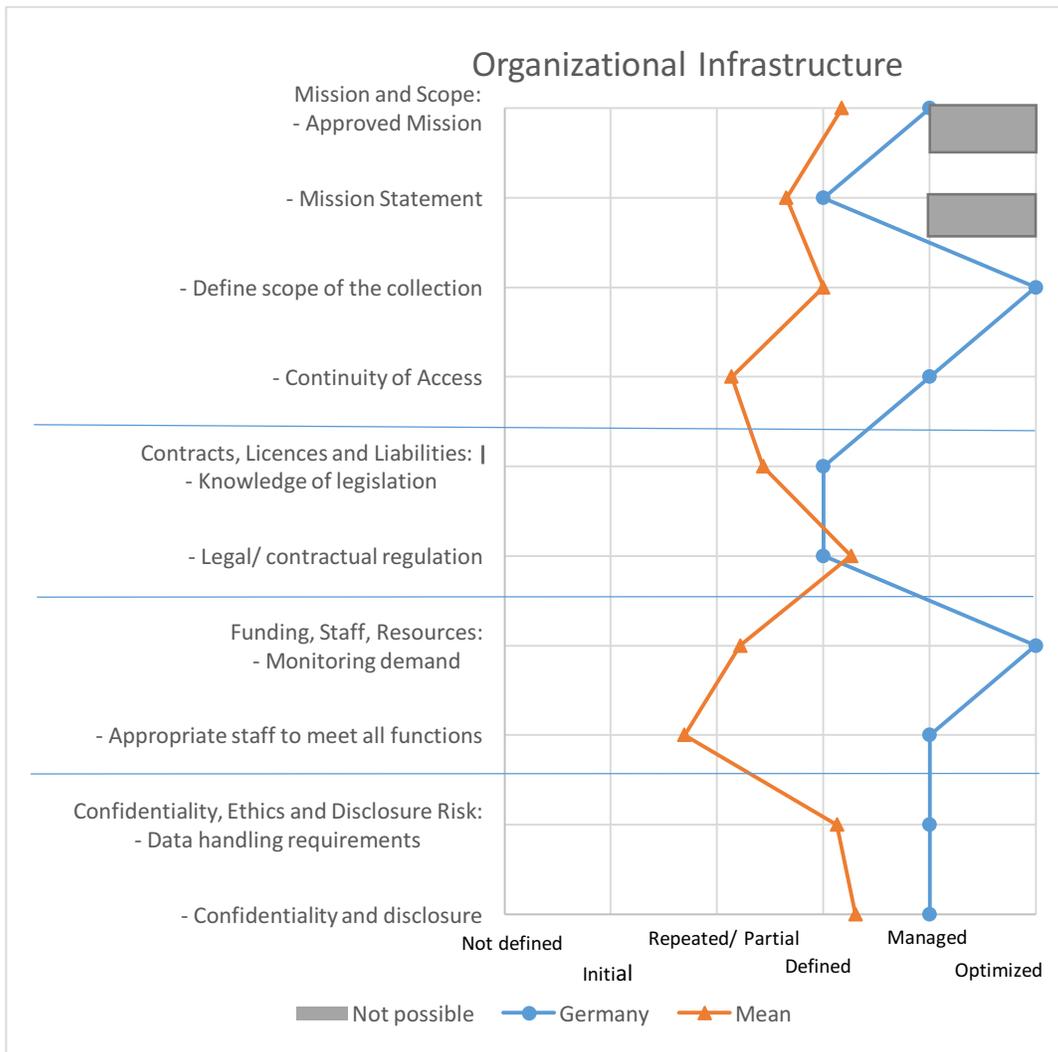


Figure 32 shows that GESIS' organisational infrastructure is mainly above the mean. The German data service has even reached the highest level in three categories. This is compliant with the aim stated in its mission statement: *“Our goal is to be the most significant and one of the leading infrastructure institutions worldwide for the Social Sciences. Thus, we focus our strengths closely in line with the thematic and structural developments in social science.”* (GESIS¹⁰³).

¹⁰³ GESIS, Mission statement, available at: <http://www.gesis.org/en/institute/the-association/mission/> (accessed 15.12.2016).

Mission and scope

GESIS has a received formal approval through clearly defined mandate. A formal and complete mission statement is available online for all relevant users, as well as the GESIS constitution document (only in German)¹⁰⁴. The GESIS constitution states that: „*The association serves the promotion of social science research. It provides basic, supra-regional and internationally significant research-based services for the social sciences. It has the task of developing and improving sociological research approaches and research tools through fundamental research*” (our translation).

The scope of the data collection is clearly defined and reviewed regularly.

Currently GESIS has no formal succession plan in place to ensure the continued availability and accessibility of data in case the organisation ceases to operate (e.g. due to funding issues or other “crisis”). According to the respondent, this issue needs to be addressed in the future even though, GESIS as the largest infrastructure for the social sciences in Germany, operates under a relatively stable financial framework. Indeed, as a member of the Leibniz Association, the umbrella organisation to currently 91 research institutions, GESIS is part of a strong network of publicly funded research institutions. In addition, GESIS has long-lasting and strong ties with German universities.

About the continuity of access, the respondent adds that: “*Due to the Archive’s systematic approach to archive and preserve research data, a transfer of data holdings back to data owners or to another institution taking over responsibility is – at least in principle – possible any time.*”

Contracts, licences, liabilities

Knowledge of legislation is widespread across the institution. Sufficient documentation on all relevant legislative aspects is available to all staff. The legal and contractual regulations are on a defined level, with publicly accessible charge and usage regulations on the GESIS website, clearly communicated to depositors¹⁰⁵ and end-users¹⁰⁶.

Funding, staff, resources

The gap between GESIS and the other data services is particularly broad among the activities monitored inside the Capability completeness area (CPA) of funding, staff, and resources according to the CESSDA-CDM. With 60 employees¹⁰⁷ working at the data archive department, staffing level is appropriate to support functions, and service needs are monitored and assessed at regular intervals. GESIS conducts every year a survey among its users to inform the future demand and adapt its services.

¹⁰⁴ GESIS constitution (only in German), available at: <http://www.gesis.org/en/institute/the-association/constitution/>, accessed 15.12.2016.

¹⁰⁵ GESIS, Data Archiving, Legal aspects, available at: <http://www.gesis.org/en/services/archiving-and-registering/data-archiving/legal-aspects/> (accessed 15.12.2016).

¹⁰⁶ GESIS, Data service for secondary analysis, Terms of use, available at: <http://www.gesis.org/en/services/data-analysis/data-archive-service/> (accessed 15.12.2016).

¹⁰⁷ GESIS, 2015 Annual Report (only in German), available at: http://www.gesis.org/fileadmin/upload/institut/Jahresbericht_2015_Web.pdf (accessed 12.01.2017).

Confidentiality, ethics and disclosure risk

Within GESIS, compliance to applicable research discipline norms and legal data handling requirements are monitored and assessed and non-compliance incidents are recorded.

GESIS requires that data depositors ensure that data meet requirements of confidentiality and non-disclosure for data collected from human subjects. For archiving purposes, the legal framework under which the data are collected must allow for data archiving and the individual's right to privacy has to be protected. Studies deposited in the archive which contain confidential data are either anonymised or access is only given through the Secure Data Center, which provides controlled and secure access to data deserving special protection. Also, there are mechanisms and procedures in place for staff training on confidentiality, disclosure risks and anonymization. For example, checks for disclosure risks take place during the ingest phase on a regular base. If problems are detected, data producers will be informed and advised, if and how data can be made accessible in a secure way.

Digital object management (data curation) and Technical infrastructure and risk

The previous Figure 32 clearly shows that GESIS is very advanced with respect to digital object management and technical infrastructure, compared to the other European data services, especially for topics like unique persistent identifiers (PID) and preservation strategies.

Digital object management

GESIS clearly specifies the information (documentation, metadata, provenance) to associate with the data for the deposit. The institution provides functions and mechanisms for proper data citations and regularly reviews and updates them. The depositor can place access conditions on the information that is being deposited.

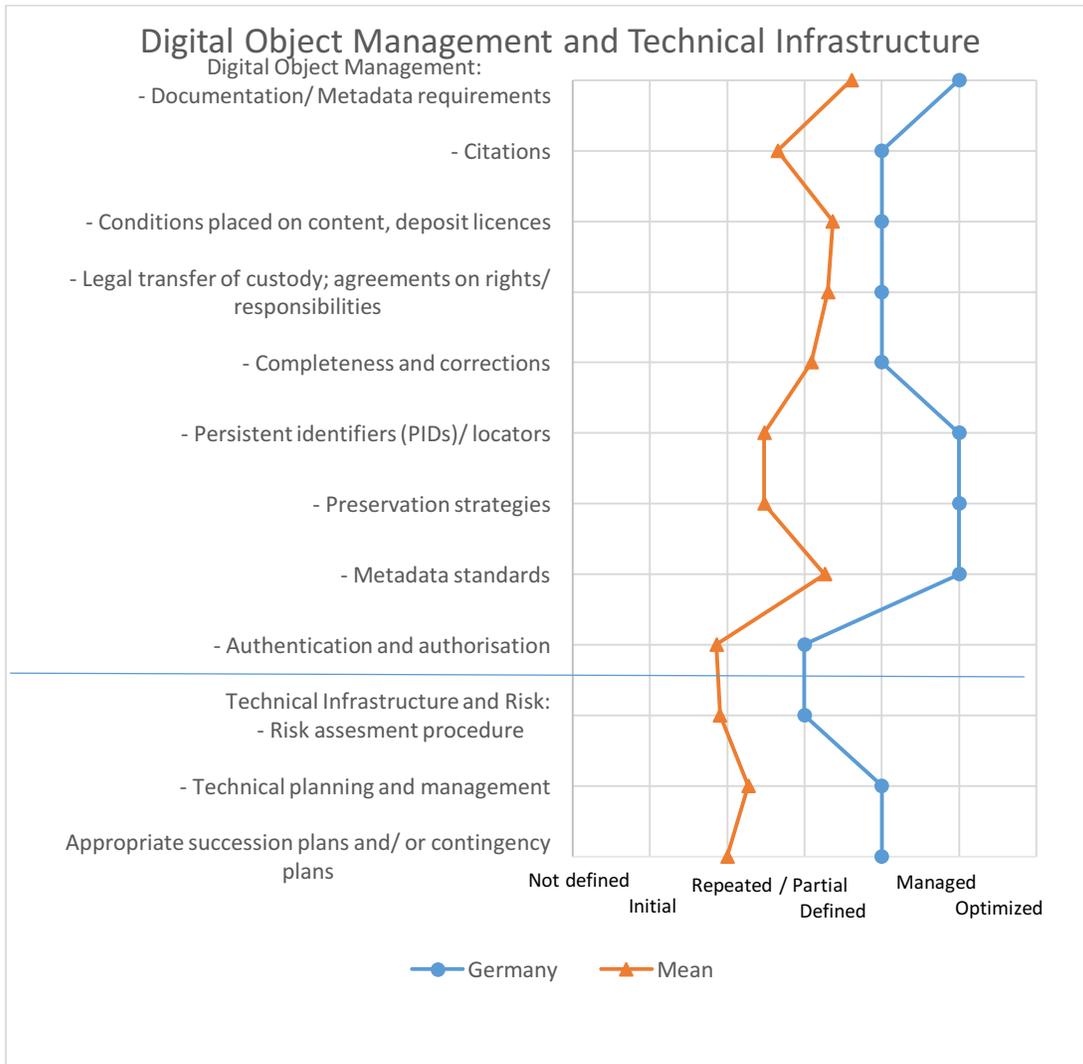
The institution has in place agreements and contracts, in accordance with higher-level policies, that confirm the legal transfer and include clear definitions of roles and responsibilities of repository and depositor. Processes and mechanisms verify the deposited material for completeness and correctness (i.e. accurate variables and values according to the documentation supplied, variables and values sufficiently labelled for reuse, similarity between the variable names in the dataset and the codebook).

da|ra (<http://www.da-ra.de/en/home>) has mechanisms in place that generate PIDs for all its data holdings. Reviews and updates of PID systems are systemised. The GESIS preservation strategy is developed in their online document "Digital Preservation Policy Principles of digital preservation at the Data Archive for the Social Sciences"¹⁰⁸. As an active member in shaping the standards through its work in the DDI Alliance, GESIS provides data and metadata that are in accordance with standards.

¹⁰⁸ DAS preservation policy. Available at:

http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/DAS_Preservation_Policy_eng.pdf

Figure 33: Digital object management and Technical infrastructure in Germany



Technical infrastructure and risk

There is a documented risk assessment methodology that is used to make a systematic analysis of security and infrastructure resilience risk factors when there are changes to the technical infrastructure, which may affect the security or resilience of any service, component or procedure.

GESIS has appropriate technical infrastructure resources to ensure that all functions and services of the repository are supported. These are monitored and analysed at regular intervals. To ensure on-going access to and preservation of its holdings, the institution has succession plans and/or contingency plans in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope/obligations. All plans are integrated into higher level policies. The relevant staff is trained in contingency issues.

3.15.3 Conclusions

German institutions and researchers have a well-established tradition in social science research data production. However, the deposit of data in an appropriate disciplinary repository is not yet

required by public funders of social science research, and there are few formal career rewards for data sharing. On the other hand, significant funding for infrastructure projects enables the archiving of data and metadata, and positive attitudes among researchers ensures that a good portion of the data produced in the country is shared and available for secondary analysis. Data support services that facilitate data sharing and/or open access to research data are provided by different institutions like GESIS, University Library Services, and Research Data Centres.

The data archive department within GESIS is well staffed and maintains a high level of competence to fulfil its mission. Its service needs are monitored and assessed at regular intervals. GESIS has appropriate technical infrastructure resources to ensure that all functions and services of the repository are supported.

3.16 Greece

The analysis shows that although the social science sector is generally well established in Greece, current level of research data production is below its potential and research data management policies are only beginning to be planned. While the ethical and legal framework important for data sharing is developed, data sharing and reuse are estimated as not very common. Incentives and enablers of data sharing are estimated more positively.

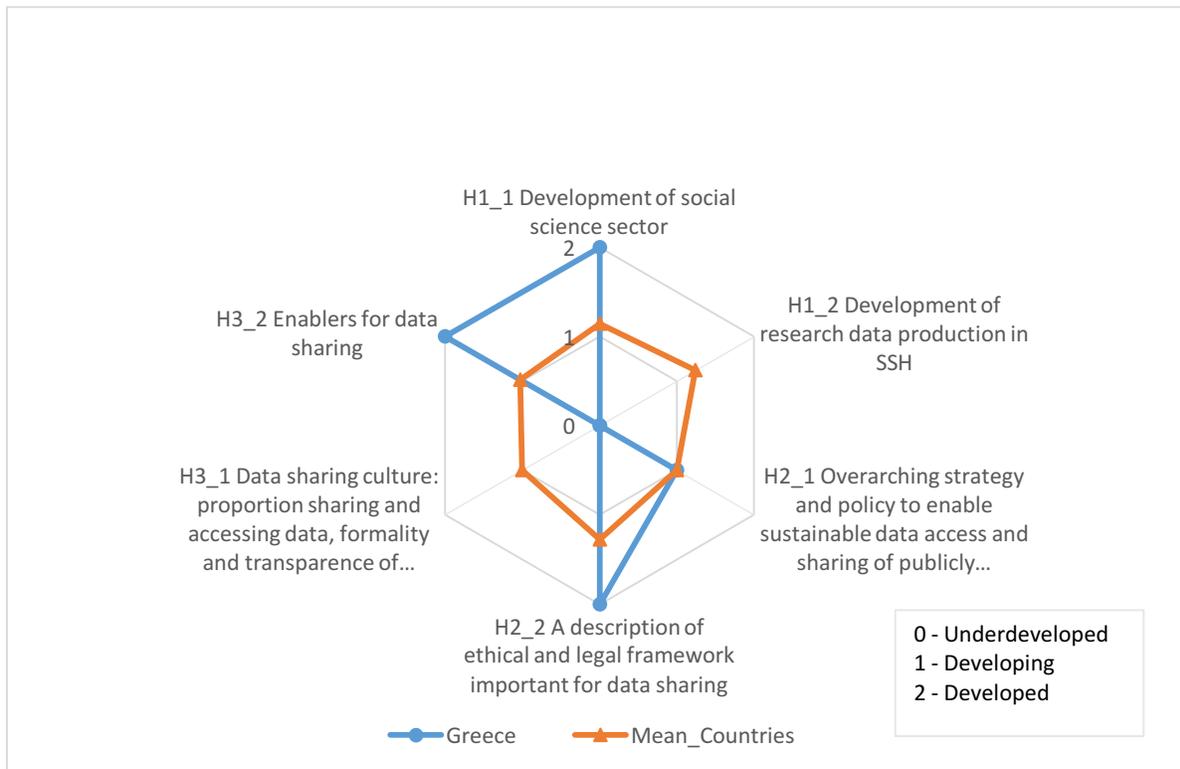
The So.Da.Net Network has a defined mission and scope, except for the continuity of access which is not defined. Contracts, licences and liabilities are also developed, both in terms of knowledge and regulations. The monitoring demand of funding, staff and resources is only partial and there is only initial developments regarding appropriate staff to meet all functions. Data handling requirements are only partial but confidentiality and disclosure risks are defined. The legal transfer of custody, agreements on rights/responsibilities and metadata standards are managed. The documentation requirements and conditions places on content are defined. Citations, completeness and corrections, and authentication and authorisation are partial. Persistent identifiers and preservation strategies are only in the initial phase. Risk assessment procedures and technical planning and management are only partial, while an appropriate succession plans are only initial.

Specifics about the data collection in a country

The information was entered beginning October 2016 by Dimitra Kondyli and Apostolos Linardis, researchers from the National Centre for Social Research (EKKE). The main method used was self-assessment. Self-assessment scores were further explained in the notes and with references to written resources, which were all extensively utilised in the preparation of current report.

3.16.1 Broader ecosystem of DAS operation

Figure 34: Heading concepts values in Greece



Development of social science sector

Overall Assessment of SSH development

The funding of the social sciences and humanities (SSH) sector are in the highest quantile among the countries. This should still be interpreted in the context of the general level of spending for research. In 2013, research and development expenditures were 0.80 of the GDP, but expenditures allocated to social sciences are even lower (0.13). However, relative to other disciplines, the intensity of investment in SSH is 19,8% and the number of researchers in SSH per capita (pop. 100,000) is 100. The investment in human resources in SSH is 21.518 (per researcher)¹⁰⁹. Moreover, as explained by informants, *»salaries and operational costs are paid by the government, so that research can be carried out without extra funding. European contribution through EU and framework programs ensures the main funding. «*

¹⁰⁹ The source of the provided information is the following report (in Greek): http://metrics.ekt.gr/sites/metrics/files/RDstatistics_2014final_Greece_el.pdf. See also the following Eurostat table: http://ec.europa.eu/eurostat/statistics-explained/images/5/53/Gross_domestic_expenditure_on_R%26D%2C_2003%E2%80%9313_%28%25_of_GDP%29_YB15.png

The situation regarding supply of science information services was described as satisfactory: »The government or university provides access to commercial bibliographic and full-text databases and to datasets but not to software licences. There is Heal-Link¹¹⁰, the Hellenic Academic Libraries portal that has full-text access to journals, e-books and to bibliographic databases.

Development of research data production in SSH

Research data production in social sciences and humanities in Greece is in the 'underdeveloped' category. This follows from the reporting as follows: Even though high quality research data are being produced on a permanent basis, data reuse has only increases slightly over the recent years. An additional constrain to the data reuse could be explain by the fact that till lately datasets produced were dispersed in several places As far as we know, there are not many studies of national importance or as many comparative surveys carried out as by the past due to the current socio-economic conjuncture. The research landscape will hopefully change because of the creation of the Greek Foundation of Research and Innovation which provides scholarships and grants to promote basic and applied research for young and senior scholars of all scientific fields¹¹¹.

Except for the PISA study, Greece is (currently) not involved in international collaborative research or cross-national studies such as CSS, ESS, ISSP, CSES, EVS, WVS, or other. The national representative reported that "*most national studies on important societal aspects are conducted ad-hoc by university department or research centres. However, there are number of studies, election surveys, that have a more permanent character and are repeated across time. There were also some studies on migration issues*" The National Centre for Social Research (EKKE) was also a partner in some international surveys:

- Until 2012, it was a member of the European Social Survey (ESS) but later abstained because of the current socio-economic conjuncture. We do hope that Greece will join ESS further to the RIS assessment in July 2016
- From 2012 to 2015, it was part of the 2nd round of PIAAC (Program for the international assessment of adult competencies).
- PISA, Eurobarometer and the World Internet Project surveys are being carried out on a permanent basis

Research institutions have tradition in producing some type of research data to a certain extent. They have experience and tradition in focusing on certain topics of surveys, rather than data. Organisations under public law like ministries etc. have specialized departments collecting data as well. However, the data collection efforts in the country are often dispersed and hard to discover by scientists.

The conclusion is that there are potentials for further development of the social science data infrastructure, both in addressing the discovery of dispersed data resources and in making available the research data for secondary use.

¹¹⁰ Available at: <http://www.heal-link.gr/?lang=en>

¹¹¹ Funding comes from national funds as well as the European Investment Bank. Available at: <https://apps.gov.gr/minedu/elidek-postdoc/>.

RDM Policy setting

Research data management policy settings are developing. Overall, incentives and high-level policy requirements are not yet on the level that would provide the highest level of sustainable conditions for data service provision. Yet there is some positive development planned.

The national representative noted that *“currently the Greek national science foundation does not require research data management and there are no requirements or recommendations for data management plans (DMP) to be submitted as a prerequisite when social science researchers are applying for public funding. Therefore, Horizon2020 and related open data etc. may have a positive impact in this perspective at national level as well”*.

The public research funding organisations operating in Greece have only partial requirements or recommendations about quality-assured social science research data with associated metadata, *“there is only a partial expectation or recommendation to offer or deposit data in an appropriate disciplinary repository or equivalent data archive service. There is no published policy recommendations on the matter”*. However, since *“a national repository of social science data has been incorporated to the National Road Map¹¹² there is an indication that the central government is seeking to promote repositories to store data that arises from social research”*. This could be a sign of an increasing awareness of importance of data infrastructure among policy makers.

Long-term curation for valuable research data assets, evaluated and selected in terms of reuse potential, is only partial: it is ‘expected or recommended’ to assess the value of research data and resources providers declared their motivations for continuing to invest in sustaining the assets. *“A number of consultations and papers have been published though there is no systematic official agenda with regard to long term curation etc. A particular concern seems gradually increasing within public organizations, given that collecting data or conducting surveys is becoming costly”*.

The research funding organizations in Greece do not offer any incentive for sharing research data with associated metadata that would cover the cost for managing the data and preparing it for access and there are no adequate resources during research project lifetime. Given that data archive services are well developed and professional, the incentives and support from national funders in a form of clear RDM policies could have great effects.

Legal and ethical aspects that facilitate social science data sharing (IPR, data protection) are clearly defined. They are embedded in ethical codes and/or legal documents that govern research and data management activities. Specifically, *“the Hellenic Data Protection Authority (HDPA)¹¹³ ensures that the Data protection law grants the data subjects (i.e. individuals) certain rights and imposes certain responsibilities on data controllers (i.e. anyone who keeps personal data in a file and processes it)”*. Also, *“national laws are being harmonised with the European legislation. However, there is little organised support is given besides guidance and recommendations. The HDPA publishes decisions regulating data protection issues, therefore no technical support or guidance is provided. Social science researchers have no guidance in dealing with related issues during the conduct or documentation of surveys. SoDaNet tries to provide technical assistance and or guidance when needed by members of the research/scientific community”*.

¹¹² See <http://www.tavernarakislab.gr/news/RoadMap.pdf>.

¹¹³ More information on the HDPA website ,available at:
http://www.dpa.gr/portal/page?_pageid=33,40911&_dad=portal&_schema=PORTAL.

Data sharing culture

The data sharing culture is underdeveloped in Greece. Estimating research data is not straightforward as there are no available official statistics so they are not able to provide an estimate for the period between 2011 and 2016. However, based on annual reports about scientific publications by the National Documentation Centre (EKT) and based on experience for the period of 2011-2012, data sharing and reuse among social sciences researchers in Greece is not so common – as self-assessment results indicate, the proportion of researchers sharing data is estimated at about 10%. There is no estimate for the proportion of researchers able to access existing third party data they need. A comment explained the situation further: *“Between 2013 and 2015 the Greek SoDaNet research infrastructure has been in construction and since January 2016 advertisements have been launched within the scientific community in order to disseminate the newly available services”*.

There are established data sharing channels and routines. Most popular data sharing channels, as reported by the country experts, include formal and transparent channels, with data archive or repository ranked first. Data is shared also via personal contacts (ranked second) and project or personal websites (third), that lack formality and transparency.

The attitudes of researchers toward data sharing, as reported through experts’ self-assessment, can be characterized as mostly negative. Even though social sciences researchers in Greece in general seem to acknowledge the benefits of data sharing, i.e. they consider false the statements that “data sharing has no benefits at all” and *“data sharing creates negative competition”*, they however, also are perceived to think it's probably false that *“data sharing creates healthy competition”*, that the *“reuse of existing data can answer new research questions and facilitate advancement of science”* and that *“data sharing involves little effort and minimal costs”*, while considering the statement that *“data sharing has a risk that others may misuse and misinterpret the data”* probably true.

As the country representative commented, based on their *“experience with important stakeholders as well as social science researchers working with data in the past, two issues prevail: a) everybody thinks that sharing data is a good practice but a very few among them have shared data.”* Furthermore, they reported that *“most of the time during the 80s and 90s, data sharing was only due to scientific and personal networks, etc. From 2000s and beyond, the data sharing culture changed in the country due to an extrovert attitude of social science researchers, European funding, development of research institutes in different science disciplines as well as lack of big surveys (mainly 2010 because of the economic crisis and public funding cuts) or to the high cost of collecting data on his/her own”*.

Attitudes towards data sharing

Table 15: Attitudes toward data sharing in Greece

Data sharing has no benefits at all	False
Data sharing creates healthy competition	Probably false
Data sharing creates negative competition	False
Reuse of existing data can answer new research questions and facilitate advancement of science	True
Data sharing has as a risk that others may misuse and misinterpret data	Probably true
Data sharing involves little effort and minimal costs	Probably false

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from "5"-True to "1" False!

To summarize data sharing culture in Greece, data sharing and reuse is not very common, there are established, both formal and informal channels for data sharing, but general attitudes of researchers related to sharing of data are neutral or negative, so in summary data sharing culture can be characterized as underdeveloped. However, the national representative mentioned that there is an "emerging data sharing culture in the latest years, particularly among younger social sciences research and academic staff".

Enablers for data sharing

Enablers for data sharing are on a higher level (developed) in Greece. To some extent there are reported to exist career rewards related to data sharing among social science researchers: "The fact is that there is an increasing need for data experts in Greece but the research community doesn't acknowledge these specific skills. They only acknowledge people who have both research and data competencies. Based on their experience, their work has been recognized because they were combining specialities, researcher competencies as well as data curator skills". As evidence based material, the national representative cites "SoDaNet members and Clarin staff since they both combine researchers' skills and they hold researchers' positions within broader research institutions". Professionals combining data management skills along with field research skills are popular in the Greek labour market. Many post graduate social science students ask for stages within SoDaNet institutions.

In Greece, there are data support services provided to social science researchers that facilitate data sharing and/or Open Access to research data. They exist mainly for socio-economic data (e.g. www.sodanet.gr) but also for textual and social media data (www.clarin.el) and for historical data (<http://www.askiweb.eu/index.php/el/>.) However, data support services are only provided by SoDaNet and Clarin.

Some data producers follow data management and data documentation standards and procedures. It depends on the legal status in teams or individuals or organizations. Independent researchers or research teams usually use their own documentation book or plan, except when conducting panel surveys or European/international surveys. Institutions as data producers use

metadata standards when linked to Hellenic Statistical Authority or when they produce data on a systematic way. There is a number of data producers of the private sector as well (public opinion research companies etc.) which also have their own data plans (e.g. Association of Greek Market & Opinion Research Companies, www.sedea.gr). The association was established in 1990 aiming at the promotion of scientific market research and social research in Greece.

In order to achieve these objectives, SEDEA adopted codes of ethics and professional practice, that members must respect and has developed surveillance mechanisms for their members. SEDEA has established communication with public and private bodies to better inform them on the usefulness of the scientifically ongoing market research, and to promote its work.

In summary, the incentives and enablers for data sharing within social science research community in the County can be benchmarked as developed.

3.16.2 Capability requirement areas of DAS

Organisational profile

Organisation

The So.Da.Net network is the Greek research infrastructure for the social sciences. So.Da.Net (<http://sodanet.gr/en>) has been incorporated along with the European Social Survey (ESS) into the 'National Road Map of Research Infrastructures 2014-2020' under the title CESSRI¹¹⁴.

The National Center of Social Research (EKKE) was established in 1959 under the auspices of UNESCO and is the only public institution in Greece dedicated to the social sciences. EKKE is a public law legal entity supervised by the General Secretariat of Research and Technology of the Ministry of Education and religious matters (Source: <http://www.ekke.gr/>).

"EKKE as the central node of So.Da.Net is the hyper administrator of services. Two researchers and one IT specialist are full time involved, while academic departments' staff also invests some time to the workload. Usually one or two external collaborators under fixed term contract are hired per year, depending on funded projects".

"The academic and research community are the main direct suppliers of data. In addition, public opinion/marketing enterprises have special agreements with a partner within So.Da.Net".

"Primary and secondary user communities are academic, research staff, students, consulting offices, and journalists".

Funding

The coordination and organization of the So.Da.Net and EKKE repository network are under public law. Public funding (Ministry of Education, State Secretariat for R&T) covers staff remuneration. Thus, any improvement design and use of the RI should be funded on a project-funding base.

¹¹⁴ For the first time the functions of Repository of EKKE will be included to its new institutional framework following the New Law 4386/2016 on R&D in Greece. Available at: <http://sodanet.gr/files/Regulation.pdf>.

Core services and activities

SoDaNet core services and activities (as well as EKKE) consist of the following:

- 1) The Data Catalogue of the Greek Research Infrastructure So.Da.Net_GR which includes quantitative and qualitative studies as well as cubes¹¹⁵.
- 2) The scientific dictionary of social terms has been designed by the Laboratory of Social Informatics, Statistics and Research Infrastructures of the Department of Sociology of the University of the Aegean, with the purpose to be a part of the national research infrastructure So.Da.Net, which supports the documentation of empirical studies. It consists an online application system in both Greek and English language, which contains definitions of scientific terms of the social sciences, which relate to their citations¹¹⁶.
- 3) System of Secondary Data Production and Indicators Management (in Greek)¹¹⁷
- 4) e-learning platform¹¹⁸

Content current collection

The current collection mainly contains quantitative data but also some qualitative data (its amount is increasing), cubes and metadata packages. Thematically the topics cover geographical data, economic data, social inequalities, social policy, immigration, cultural consumption, housing, political-electoral behaviour and politics. "All collections are in Greek with metadata about collections in English (1/3 of the total). English translation is in progress".

¹¹⁵ See <http://sodanet.gr/en/sodanet-services/katalogos-dedomenwn-en>

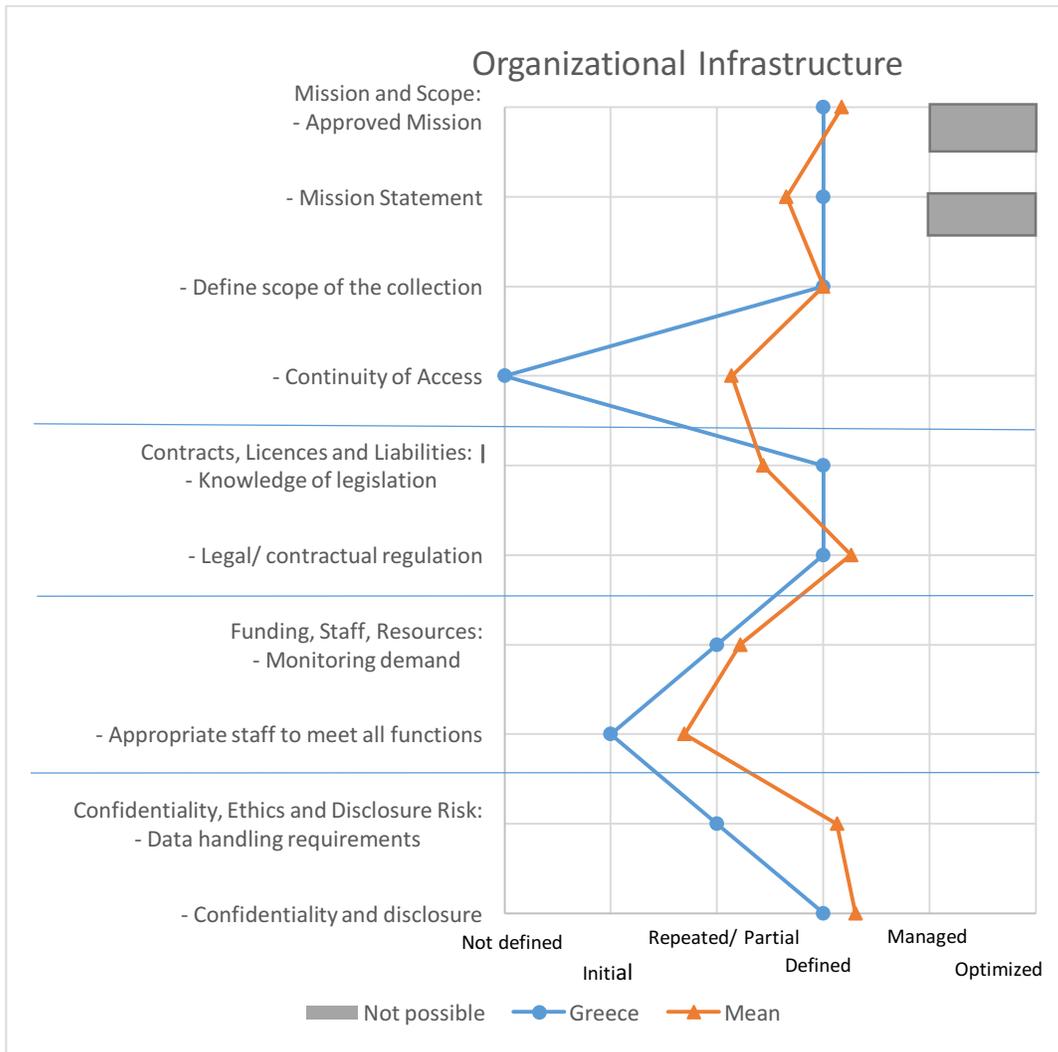
¹¹⁶ See <http://sodanet.gr/en/sodanet-services/vocabulary-terms>

¹¹⁷ See <http://sodanet.gr/en/sodanet-services/sythma-paragwghs-kai-diaxeirishs-deyterogenwn-data-en>

¹¹⁸ See <http://sodanet.gr/en/e-learning>

Organisational infrastructure

Figure 35: Organizational infrastructure in Greece



Mission and scope

The organisation has a 'repeated/partial' formal approval for its mission statement. There is some evidence to support the notion that the organisation and its activities receives approval from stakeholders: it is encouraged to continue its activities on an ad hoc basis but no formal mandate or formalised approval exists.

Explicit written statements on the role, mandate, purpose and mission of the organisation are 'defined', i.e. it provides a written, formal and complete mission statement that is available to all relevant users. The SoDaNet research infrastructure mission statement is:

"So.Da.Net supports multidisciplinary research and promotes the acquisition, exchange, processing as well as dissemination of data deriving from and related to social science research. Social science research datasets and metadata are being stored and documented based on a model compatible with DDI3 and DDI2 using Nesstar software, thus rendering them available to any interested user. Data is sought and retrieved locally, via the

repositories of its members, or through the So.Da.Net portal itself. In addition to this, the research infrastructure of So.Da.Net provides e-learning services with regard to social research methodology, development of research infrastructures, as well as, data management–use of the research infrastructure data. The services provided by So.Da.Net can be accessed freely online or under specific terms if need be depending on the request.”

The organisation has a defined scope for the repository and their collection, which guides the selection and appraisal of data. The scope of the collection/repository is explicitly defined and promulgated. Statements on scope are either integrated into or connected to policies. Policies are connected to specific processes and procedures.

However, the organisation has not defined medium or long term plans in place to ensure the continued availability and accessibility of data in case the organisation ceases to operate. There is no evidence of a continuity of access or is not applicable, which needs to be addressed within a wider composition of stakeholders in the future.

Contract, licences and liabilities

The organisation has defined sufficient knowledge and documentation on how the legislation applies to and affects the holdings and procedures of the organisation. All relevant legislative aspects are available to all staff¹¹⁹. They dispose of a researcher at EKKE dealing with data legitimate issues and institutional issues in cooperation with a legal advisor, member of the network, professor at the University of Aegean.

The services access procedures are based on legal or contractual regulations that are settled in agreements between the service users and the repository and the contractual and legal regulations are defined and ensure that the parties do not infringe any intellectual property rights. Contracts and/or agreements are standardised and implemented according to written procedures; contracts and regulations which are made publicly available. The So.Da.Net research infrastructure and its members, EKKE included, can make use of standardized agreements for institutions or individual users either for depositing data, or obtaining datasets under restricted use. These agreements are common and the same for all So.Da.Net partners’ institutions.

Funding, staff, resources

The organization has ‘repeated/partial’ monitoring of demand for its repository services, growth and funding. The organisation occasionally reviews changes in demand for the repository services and only reacts to significant changes. This strategy is changing presently since the research infrastructure is fully operating. Thus, monitoring becomes necessary mainly for funding requests for which relevant documentation is required as well as for improvement or adjustments to data services provided.

The organisation has only the ‘initial’ appropriate number of staff to support repository functions and services. There is evidence staffing is partially complete, either due to lack of resources or unable to employ qualified staff. The organisation has not defined the appropriate staffing level to support all repository functions and services.

¹¹⁹ The document is available at: <http://sodanet.gr/en/what-is-sodanet/institutional-framework> (available in Greek).

There is a need for hiring more staff, however, public sector budgetary constraints impose the rules. Funded projects allow the Repository to operate based on specialized staff employed on limited duration work contract.

Confidentiality, ethics and disclosure risk

The organisation has only 'repeated/partial' compliance with applicable research discipline norms and legal data handling requirements: It complies with norms consistently through repeated action, however there are no written procedures or procedures for exceptions and errors.

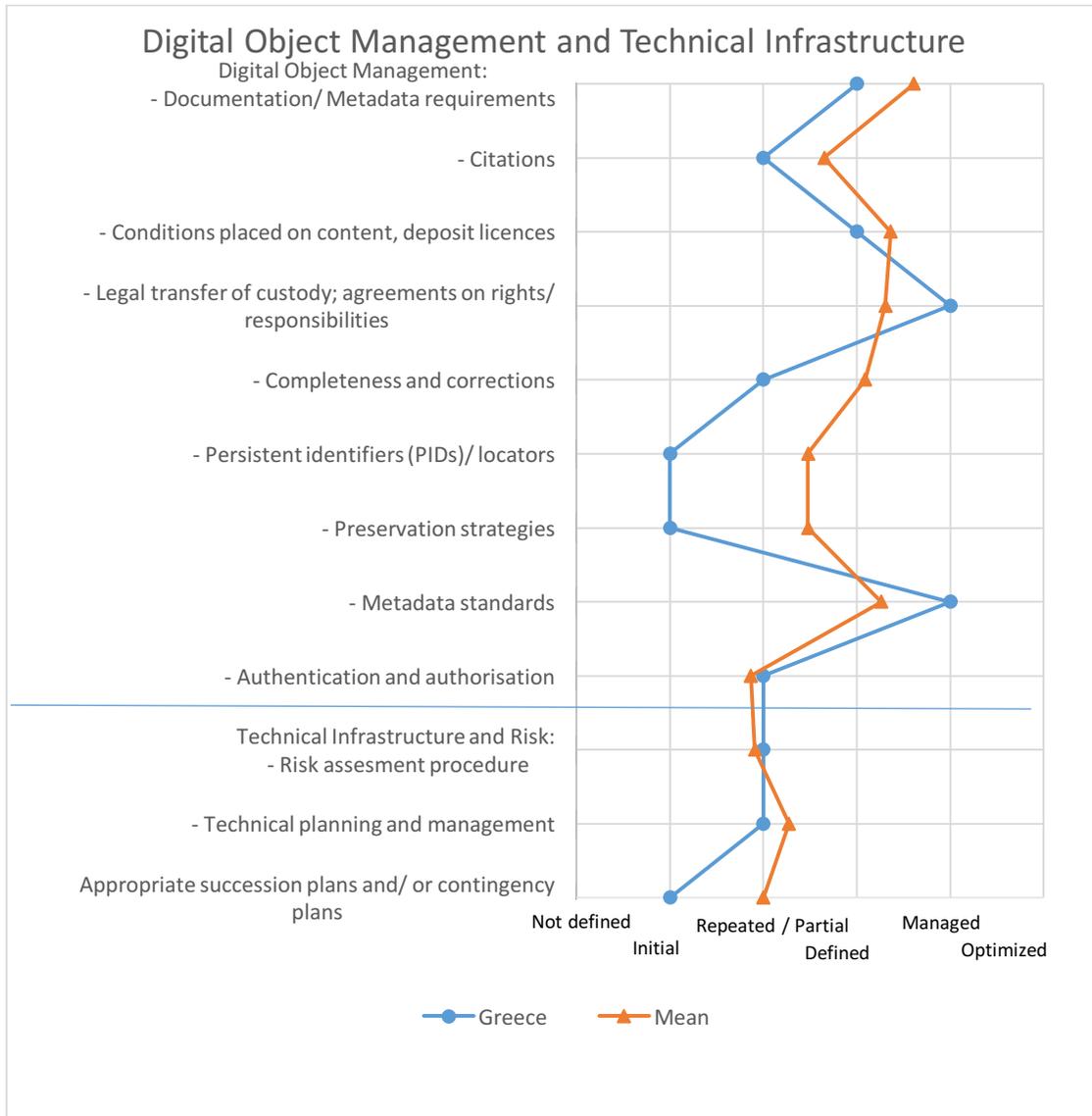
The repository defines a requirement that data depositors ensure that data meets requirements of confidentiality and non-disclosure for data collected from human subjects: Processes and procedures are in place and; standardised information is provided to the depositor prior to the deposit. Checks are performed on data after deposit; process and procedure descriptions for handling and altering sensitive data are in place.

Digital object management (data curation) and Technical infrastructure and risk

A written formal specification of required information is explicitly 'defined' and Documentation/Metadata requirements are compliant with metadata standards that are used and can be understood by the designated Community. Metadata requirements are accessible and communicated to users/depositors¹²⁰.

¹²⁰ A draft Contract between So.Da.Net_GR and Data Depositors (in Greek) is available at: <http://sodanet.gr/en/what-is-sodanet/institutional-framework>.

Figure 36: Digital object management and Technical infrastructure in Greece



The organisation offers and provides ‘repeated/partial’ functions and mechanisms for proper data citations. Citation practices are being repeated and offered regularly but lack formalisation and systemisation¹²¹. In every data set there is the way the reference citation to the dataset by the user”.

The organisation has ‘defined’ mechanisms and functions in place that allow the depositor to place access conditions on the information that is being deposited: All depositors are offered the opportunity to set access conditions on the information that is being deposited; a set of access conditions are formally defined in categories or a template. Access conditions by dataset and object are included in the Draft Contract between Service depositors and So.Da.Net¹²².

¹²¹ A draft contract between service depositors and So.Da.Net as well Data Request form where reference to citations are included are available at: <http://sodanet.gr/en/what-is-sodanet/institutional-framework>

¹²² Available at: <http://sodanet.gr/en/what-is-sodanet/institutional-framework>

The organisation has 'managed' to put in place agreements that confirm the legal transfer of the information that is being deposited: It can also make changes in order to provide appropriate digital content to its users. The agreement includes clear definitions of roles and responsibilities of repository and depositor.

"There is monitoring of the usage of agreements and contracts as actions are taken where contracts/agreements appear not to be working effectively or are not in accordance with higher level policies; reviewed and updated regularly. The contractual and legal regulations also make sure that the deposited material does not infringe on any intellectual property rights (IPR) of any other person(s) or institution(s)¹²³.

The organisation has 'repeated/partial' processes and mechanisms in place that verify the deposited material for completeness and correctness: Specifically, there are non-systematised (manual) checks of deposited material in place. Processes and procedures are repeated, but they are not formalised or documented. Rectifications are performed repeatedly, either by the repository or by returning data to depositor

The organisation has initial mechanisms in place that generates persistent, unique identifiers for all its data holdings. There is some awareness of the need for persistent identifiers and locators, but actions are sporadic and ad hoc; there are no formalised systems, processes or procedures in place.

The organisation has 'initiated' some documented preservation strategies that are relevant to its holdings: Informal, ad hoc 'contingency action points' are in place but a full comprehensive strategy is lacking. Action points are not formalised or connected to a policy.

The organisation has 'managed' to provide data and metadata that are in accordance with standards that are used and understood by the Designated Community. Format usage and enquiries are measured and assessed. The format strategy is regularly reviewed and updated.

The organisation uses 'repeated/partial' authentication and authorization infrastructure to control access to data. An authentication infrastructure emerges by repeated use of authentication approaches but it lacks standardisation and formalisation.

The organisation undertakes 'repeated/partial' risk assessments and overall risk analysis when there are changes to the technical infrastructure which may affect the security or resilience of any service, component or procedure. *"Risk assessments are undertaken when significant changes are made to the technical infrastructure but these are not analysed"*.

The organisation has 'repeated/partial' technical infrastructure resources to ensure that all functions and services of the repository are supported. There is evidence that there is likely to be an appropriate level of technical infrastructure resources to support all repository functions and services. However, the organisation has not defined what this level should be.

In particular, *"the organization has sufficient and appropriate suitable software and technical services. However, there is a weakness of appropriate management plans"*.

The organisation has only 'initial' awareness of the issue that on-going access to and preservation of its holdings should be ensured by Appropriate succession plans and/or contingency plans: There is low institutional commitment to contingency issues and no written, formal processes,

¹²³ More information is included in the "Draft Contract between Service depositors and So.Da.Net", a detailed text of 8 pages. Available at: <http://sodanet.gr/en/what-is-sodanet/institutional-framework>.

procedures, plans or other documents exist in case the repository ceases to operate or the governing or funding institution substantially changes its scope/obligations.

3.16.3 Conclusions

Regarding the broader ecosystem of data archive services operation the conclusion is, overall, that the culture and support for data sharing is expected to continue to rise also due to the institutional consolidation of national research data infrastructure So.Da.Net, the service provider of recently acquired Greek membership in CESSDA.

Regarding the capability requirement areas of data archive service of the So.Da.Net Network the conclusion is that most of the indicators of organisational, technical and digital curation management issues shows that the organisation performs all the users' services and communication activities on satisfactory level ('defined' or 'partial'). There is more room to improve regarding supply of sufficient human resources and long-term continuity and stability of digital preservation services provided, which are mainly on 'initial' stage.

3.17 Hungary

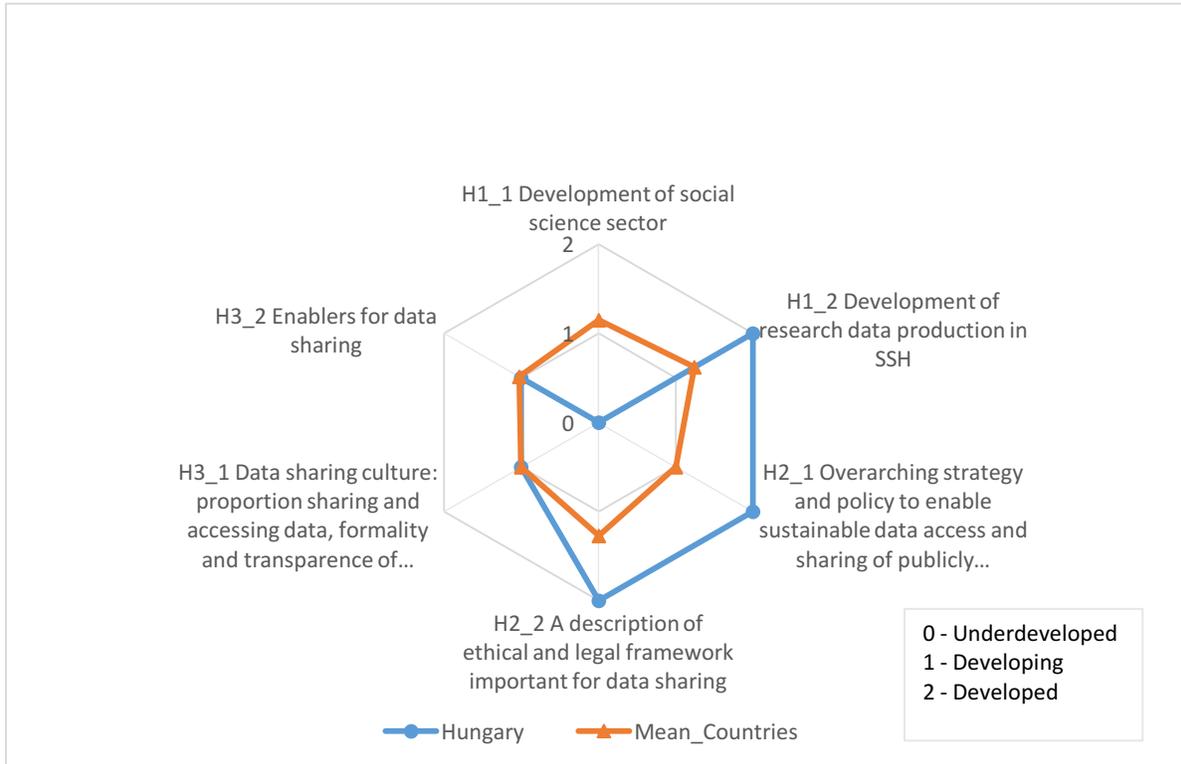
A long-term member of the formal CESSDA consortium, TÁRKI Data Archive fulfils some of the digital object management and technical infrastructure requirements. As the archive is currently underfunded and understaffed, the current decision of Hungary to pronounce TÁRKI Data Archive as its national service provider is significant for the archive.

Specifics about the data collection

Data collection in Hungary was done via Peter Hegedus from the TÁRKI Data Archive, who provided most of the information in the web survey between 28.9.2016 and 4.11.2016.

3.17.1 Broader ecosystem of DAS operation

Figure 37: Heading concepts values in Hungary



Development of social science sector

The Government and Higher education and Private non-profit sector are the main sources of funding in of social sciences research in Hungary. The Funding of social sciences and productivity of the researchers are among the mid quantiles.

RDM Policy setting

The RDM policy setting in Hungary is described as partial. There is the expectation or recommendation to have DMP in place, as well as the expectation or recommendation to offer or deposit data in an appropriate disciplinary repository or equivalent data archive service.

Long-term curation for valuable research data assets is partial. The cost for managing the data and preparing it for access are not resourced adequately during research project lifetime.

Clarification and support provided on legal and ethical aspects that facilitate social science data sharing is partial.

The publicly financed social science research should be shared between the research community, but the monitoring system is not satisfying.

Data sharing culture

Data sharing and reuse among social sciences researchers in Hungary is low – as self-assessment results indicate, the proportion of researchers sharing data is estimated as low (0-10%), and the proportion of researchers able to access existing third party data they need is low too (0-10%). The most usual way to share data is via informal and not transparent channels - through project or personal websites. Less preferred methods include data shared as supplementary data in a journal and archiving in a TÁRKI Data Archive or another repository.

Attitudes towards data sharing

Data sharing in Hungary is not common, few researchers share data although all the researchers are able to access the archived data through the TÁRKI Data Archive. This is the main reason why the number of available dataset is limited. The self-assessment survey indicates there is awareness that the reuse of existing data can answer new research questions and facilitate advancement of science, however, the benefit of data sharing is not acknowledged. The risk of misuse and misinterpretations of shared data is also perceived.

Table 16: Attitudes toward data sharing in Hungary

Data sharing has no benefits at all	Probably true
Data sharing creates healthy competition	Neither true, nor false
Data sharing creates negative competition	False
Reuse of existing data can answer new research questions and facilitate advancement of science	True
Data sharing has as a risk that others may misuse and misinterpret data	Probably true
Data sharing involves little effort and minimal costs	Neither true, nor false

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from "5"-True to "1" False!

Enablers for data sharing

There are no career rewards - related to data sharing in Hungary. Also, there are no data support services provided to social science researchers in Hungary facilitating data sharing and/or Open Access to research data. Despite that, some data producers follow data management and data documentation standards and procedures.

3.17.2 Capability requirement areas of DAS

Organisational profile

Organisation

The TÁRKI Data Archive operated by the TÁRKI Foundation what is part of the TÁRKI Group as well as TÁRKI Inc., TÁRKI Survey Unit.

Funding

The Data Archive receives its funding from grants and research contracts.

Core services and activities

The core services and activities of TÁRKI Data Archive are: 1. long-term preservation of digital research datasets from domestic and international studies; 2. keeping pace with technological change and participation in the development of data archiving standards; 3. providing access to data collections of empirical studies for users communities; 4. facilitating effective data use by providing access to our own and to our partners' collections.

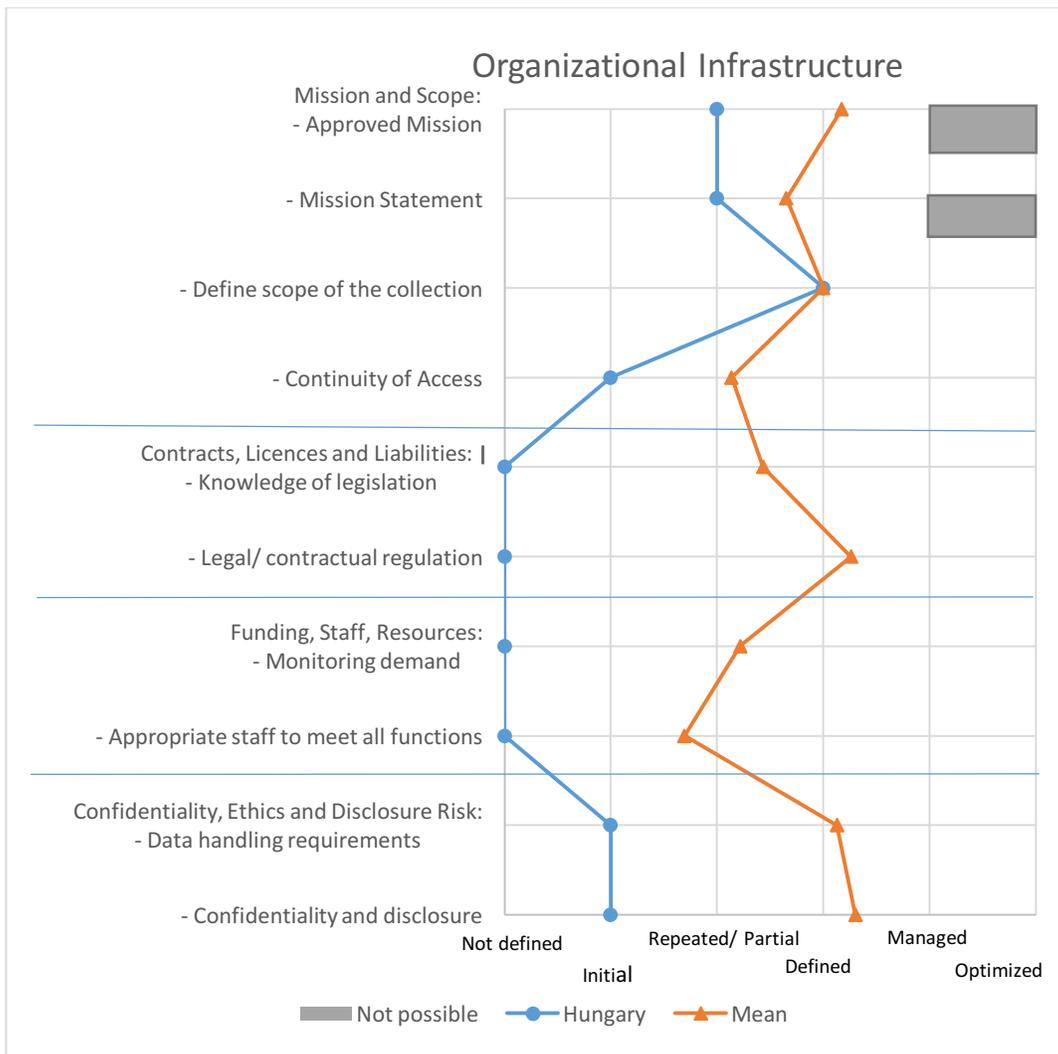
Content current collection

TÁRKI Data Archive stores data collections of empirical social research. The data collections contain factographic (i.e. numerical) data. The research data are mostly from survey research, but the archive also archives other factographic data collections (e.g. historical data, census, etc.).

Organisational infrastructure

The TÁRKI Data Archive is a smaller organization without separate departments.

Figure 38: Organizational infrastructure in Hungary



Mission and scope

The mission of the archive is not formally approved, but some informal support from external stakeholders, approval or contractual obligations exist.

Mission statement of TÁRKI Data Archive long-term is the preservation of digital research datasets from domestic and international studies; keeping pace with technological change and participation in the development of data archiving standards; providing access to data collections of empirical studies for users' communities; facilitating effective data use by providing access to our own and to our partners' collections. The mission of the archive is to provide infrastructure service, and support for all stakeholders in social research.

The scope of collection / repository is explicitly defined and promulgated. But there are no contingency plans or only informal intent/agreements.

Contracts, licences, liabilities

With regards to contracts, licences, liabilities, there is no evidence of knowledge or documentation about national legislation and how it applies to the organisation or infrastructure.

There are no service access procedures based on legal or contractual regulations that are settled in agreements between the service users and the repository.

Funding, staff, resources

There is no evidence of monitoring of demand for repository services. There is also no evidence that the organisation has defined a level for the appropriate number of staff to support all functions and services. The current function can be provided by the staff, but if the Data Archive will have more roles, the number of staff appears to be less adequate to support all functions and services. If the tasks become more complex some competencies are missing.

Confidentiality, ethics and disclosure risk

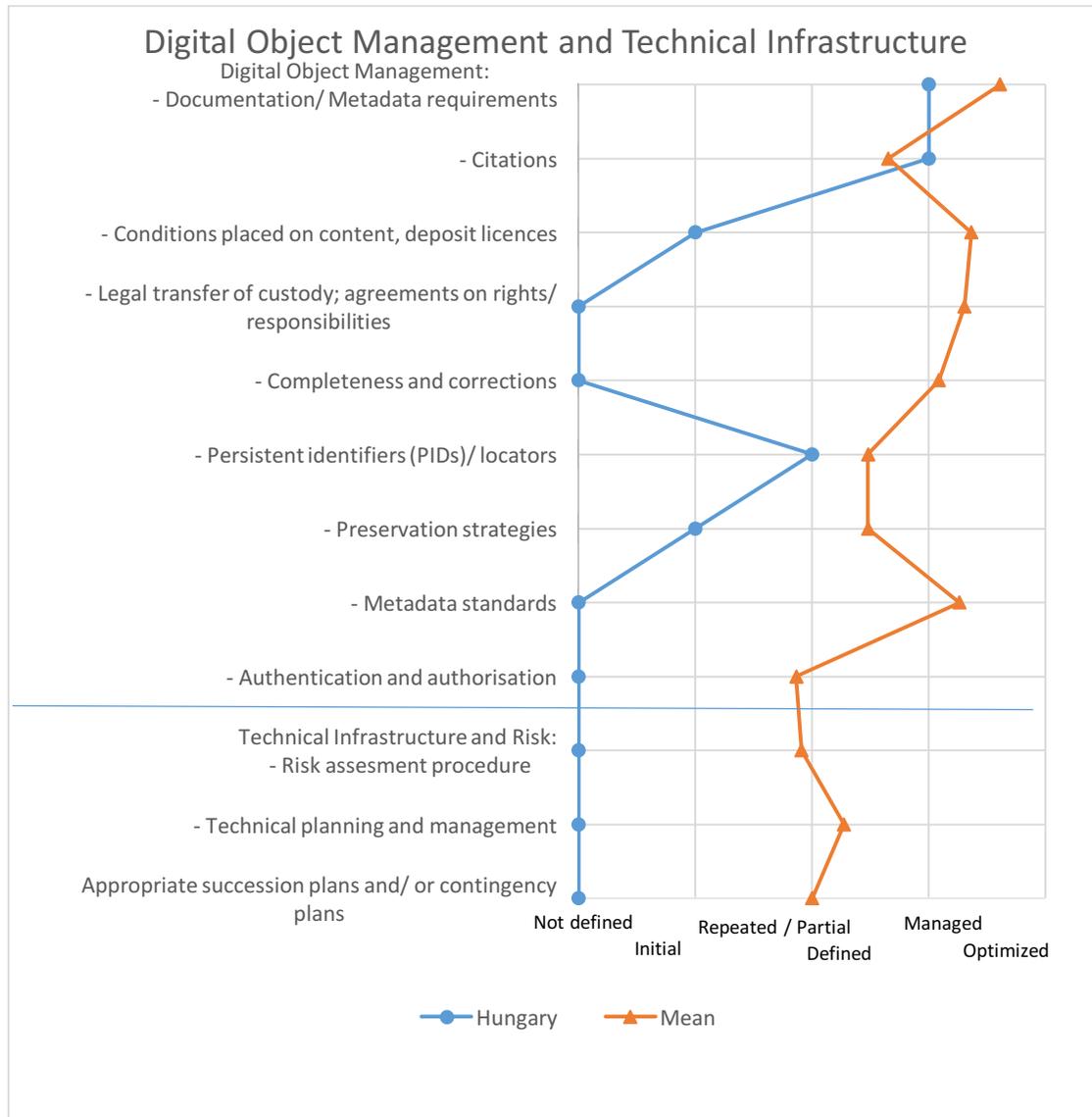
The adherence to legal and research discipline norms for handling confidential data is informal and inconsistent.

There is some awareness of the confidentiality and non-disclosure for data collected from human subjects, but decisions and procedures are ad hoc and performed on a case-by-case basis; there are no written processes and procedures in place for dealing with confidentiality, disclosure and data protection issues; no predefined criteria or non-disclosure agreements/statements are available for depositors/users. The data collections that contains personal data about natural persons – i.e. about real people - are anonymized in the archive.

Summary: Currently, the infrastructure of TÁRKI Data Archive seems to suffer from two crucial weaknesses. First is the legal dimension, where the archive scores low. No agreements on legal transfer of custody exist. Second is the dimension of funding and appropriate staff. The archive is underfunded and understaffed.

Digital object management (data curation) and Technical infrastructure and risk

Figure 39: Digital object management and Technical infrastructure in Hungary



Documentation/ Metadata requirements

A written formal specification of required information is explicitly defined (e.g. in a collection policy); requirements are compliant with metadata standards that are used and can be understood by Designated Community (e.g. DDI); metadata requirements are accessible and communicated to users/depositors¹²⁴.

¹²⁴ Available at: http://www.tarki.hu/en/services/da/docs/ddi_elements_description.pdf

Citations

Citations are required and offered to all depositors; formalised through templates or other written documents; processes and procedures are documented.

Conditions placed on content, deposit licences

The depositor of a data collection base can choose from four data access categories: A – Free access without any limitation. B – Free access for Hungarian scientific research institutions and public bodies; for others with the owner's permission only C – Free access for institutions that are members of TÁRKI; for others with the owner's permission only D – Delayed dissemination. Access with the depositor's permission only (for max. 5 years).

Besides defining the data access category, most deposits are done without any conditions on use.

Legal transfer of custody; agreements on rights/ responsibilities

The primary investigator or the owner of the dataset signs a data depositor form and decides which data access categories can be applied.

Completeness and corrections

In the case of the TÁRKI Data Archive the data depositors provide the necessary information for archiving. This information is requested on a data deposition sheet and this is the base of the investigation regarding completeness. The staff of the Data Archive verifies the archiving datasets and makes corrections for example in the case of anonymization.

Persistent identifiers (PIDs)/ locators

Mechanisms and systems for identification and location are partly in place (e.g. there may be a certain directory structure or hierarchy to make the locating of data easier), but does not comply with formalised DOI systems; mechanisms are being repeatedly used, but there is lack of formalisation and written procedures.

Preservation strategies

Only informal, ad hoc 'contingency action points' are in place, but a full comprehensive strategy is lacking; action points are not formalised or connected to a policy.

Metadata standards

During archiving TÁRKI Data Archive creates a metadata sheet based on DDI standards for each data collection. Datasheets are stored in an SQL database, which serves as the TARKI data catalogue. TÁRKI has developed the National Digital Archive metadata scheme that is based on DDI and Dublin Core and the National Digital Archive's catalogue. There is no further fixed strategy for standards.

Authentication and authorisation

If a user is eligible for a free data access, a user declaration form should be completed, signed and send to the data archive. No further authentication approach is in place.

Technical Infrastructure and Risk

Risk assessment procedure

There is no evidence that risk assessments are undertaken.

Technical planning and management

There is no evidence or awareness that the organisation has appropriate technical infrastructure resources to support all functions and services. Available skills and technology are not fully adequate to support all functions and services (for example the online dataset availability).

Appropriate succession plans and/ or contingency plans

No succession plans, contingency plans, and/or escrow arrangements are in place at the TÁRKI Data Archive.

Summary: The digital object management at TÁRKI Data Archive is not fully defined, the documentation requirements, citations and persistent identifiers are used to reach this purpose, but more definitions is required.

3.17.3 Conclusions

There are expectations of DMP in the social sciences sector in Hungary. Also, the publicly financed social science research should be shared within the research community. This request is, however, not combined with an appropriate monitoring system. There are recommendations and guidance provided on how to respect the legal requirements while sharing data as well. Both those incentives pose Hungary in the position of a relatively well developed environment with regards to data sharing in Central Europe. This is also a good starting point for formalizing the existing infrastructure of the TÁRKI Data Archive as the national service provider.

3.18 Iceland

No contact was achieved with representatives of the Icelandic research infrastructure. E-mail was sent to Rannis - the Icelandic Centre for Research - but the enquiry was never responded to.

We will, however, continue to try again to establish contact with representatives of the Icelandic research infrastructure in the next phase of the CESSDA SaW project. In task 3.3, it was agreed to try to contact Rannis, but on a lower level.

Another potential contact point is the Social Science Research Institute, which is an academic research institute under the auspices of three faculties of the University of Iceland. SSRI performs research and provides consultation in the field of social sciences, and may provide information on the general status, and broader picture of, the conditions for social science research in Iceland.

3.19 Ireland

This report shows that overall, Ireland has an average maturity level.

In recent years, we have been told, the economy has been hard on the research community. Recently the research community seems to be on the move, stimulated by open access events and international or European requirements. It is assumed that while the economic situation is improving there will be more resources available again for research.

From the conversations with the ISSDA, the Health Research Board and the Irish Research Council it seems that the key (funding) institutes and key persons are in contact with each other and striving to improve services. There are also strong links with the UK, for example with the Wellcome Trust, iDCC, and other organisations in the field. Furthermore, there is participation in a number of European projects and initiatives.

There is a strong need for a national policy and strategy with regards to RDM and RDM support. The Irish Social Science Data Archive (ISSDA) scores relatively high on the capability requirement areas. Ireland is doing well already, is on the move, and will be even doing better in the near future.

Specifics about the data collection

The data collection started September 22nd and finished December 1st 2016. We had several calls with the contact persons in Ireland; Dr John B. Howard and Julia Barrett of the ISSDA in Dublin. <https://www.ucd.ie/issda/>. Dr Howard and Julia Barrett created a short report based on the questions of the online survey. Later in the collection period they shared their views also in the online tool.

Additionally, we interviewed several research experts:

- Dr Patricia Clarke, Health Research Board (HRB), Programme Manager–Policy and EU Funding, National Delegate for H2020 Health, Dublin. (one-hour interview held at 20 October)
- Dr Eucharia Meehan, director, and Peter Brown, deputy director Irish Research Council (IRC), Dublin. (one-hour interview held at December 1st)

It was difficult to plan the interviews at short notice.

The experts agreed to the interviews on the basis to share their views with us and the contact persons, but not to share the descriptions of the interviews to a broader audience.

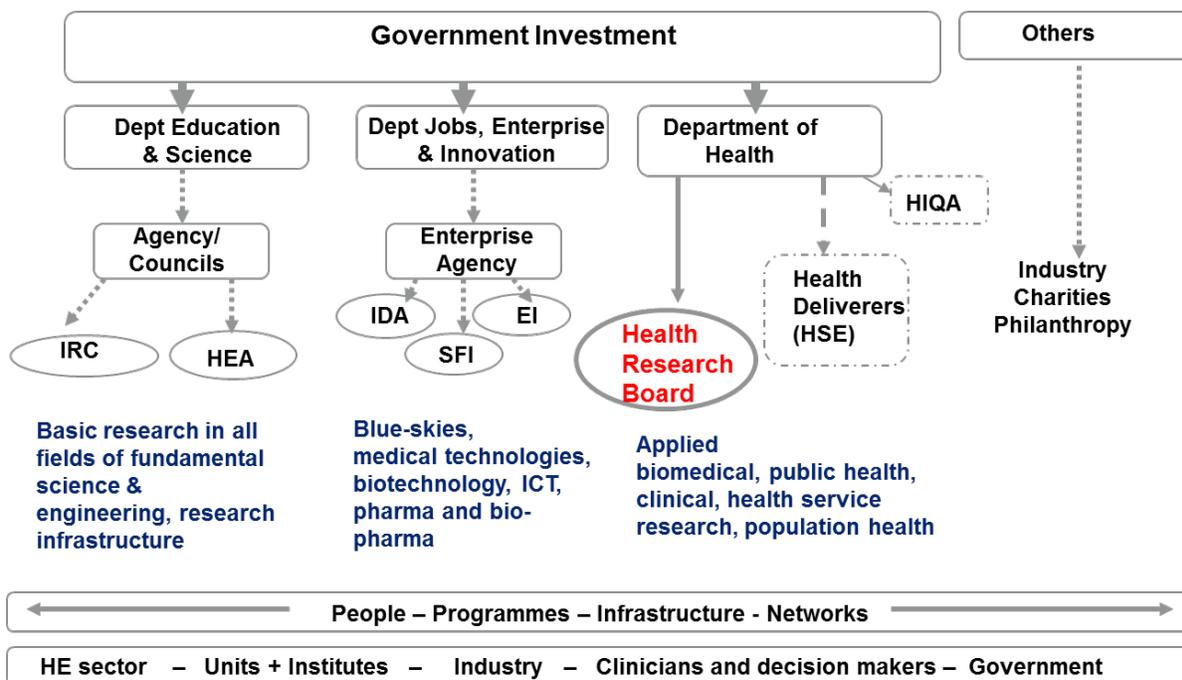
Dr Patricia Clarke (HRB) shared several resources:

- Slides on Irish Research Landscape, budget and health research areas
- HRB Data Project report ‘Proposals for an Enabling Data Environment for Health and Related Research in Ireland’ (2016). This gives a good overview of the health research environment for research data and Appendices 4 and 5 list the key people who contributed and the key Health and Social Care Data Collections in Ireland.
- Call guidelines for new Emerging Investigators Award scheme with explicit inclusion of Section 7. Data Management and Sharing Plan.

Health Research Board (HRB) (<http://www.hrb.ie>) funds health research in applied biomedical and clinical research, population health and health services research areas. In building a strong enabling environment for this research the HRB has a strong interest in research data (development and Open Access).

The Irish Research Council (IRC) (<http://www.research.ie/>) established in mid-2012 under the Government’s Public Sector Reform Plan, the Irish Research Council (‘the Council’), a merger of two former councils (the Irish Research Council for Humanities and Social Sciences, IRCHSS, and the Irish Research Council for Science, Engineering and Technology, IRCSET), is an associated agency of the Department of Education and Skills (DES) and operates under the aegis of the Higher Education Authority (HEA).

Figure 40: Irish research landscape, picture provided by HRB



In this country report the responses by the ISSDA are complemented by the valuable information provided by the HRB and IRC.

3.19.1 Broader ecosystem of DAS operation

Figure 41: Heading concepts values in Ireland

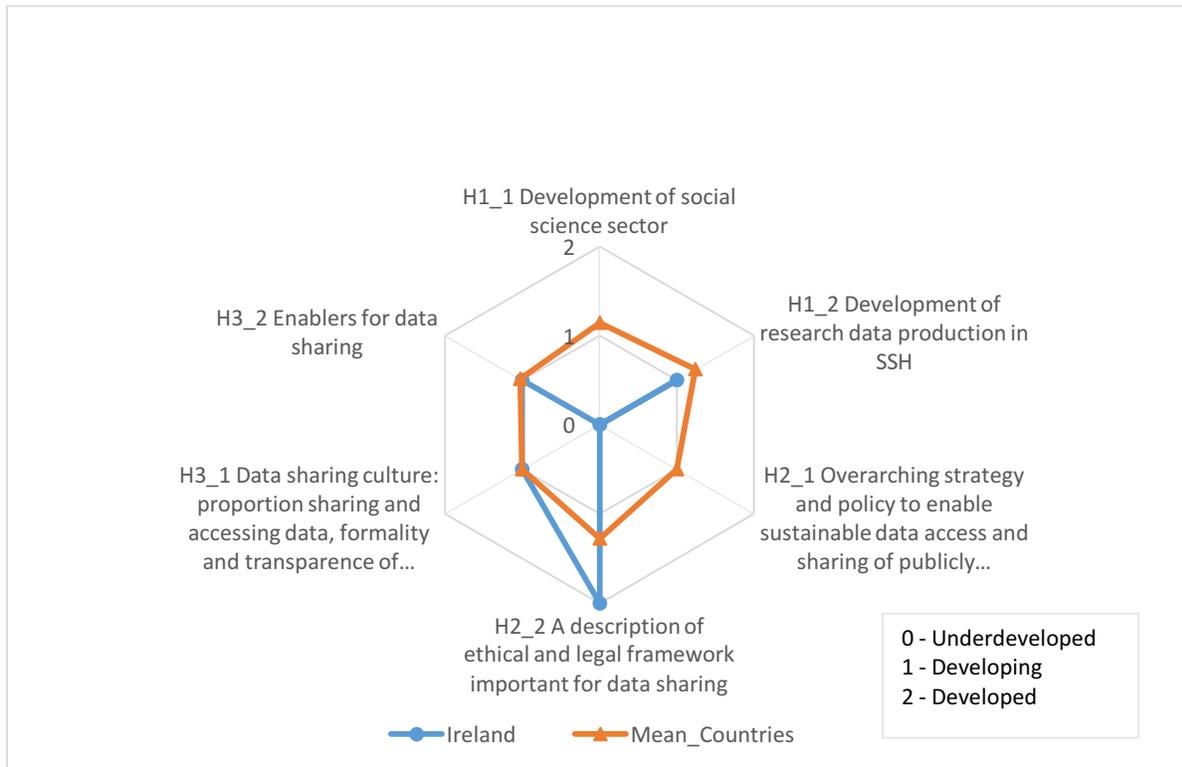


Figure 41 shows that Ireland has a mixture of development values and scores average or lower than average, but higher than average regarding the ethical and legal framework important for data sharing. Ireland is on a developing level in sharing culture, enablers for data sharing and development of research data production in SSH. There is an 'underdeveloped' level for the social science sector and overarching strategy and policy to enable sustainable data access and sharing of publicly financed social science research data.

Development of social science sector

Overall assessment of SSH development

Funding of SSH and productivity of the researchers are in the lowest quantile; impact on designated community is not clear. ISSDA does not have data on overall research productivity; research award data from all state funders is not publicly available to assess levels of funding awarded.

Due to the absence of data it is difficult to assign a value to the level of overall assessment of the SSH development.

Sources of research funding stem primarily from research grants awarded by public-sector agencies or organizations. The Irish Research Council has the primary funding remit, with a non-exclusive focus on social sciences and humanities research. The Irish Research Council has also recently undertaken a commitment to assessing need of and vehicles for participation in European

research infrastructures. Other funding for social sciences stems from the European Commission and private non-profit grant makers.

Access to commercial bibliographic and full-text databases is provided through the higher-education and health sectors through the agency of Higher Education Institutions and the Health Services Executive. The seven Irish universities coordinate access to major resources through a programme called the Irish Research e-Library (IReL), with limited participation by other Irish Higher Education Institutions. Funding is provided for HSS resources via the agency of the Higher Education Authority, while funding for STM resources is provided by the Department of Jobs, Enterprise and Innovation. The university libraries also coordinate licensing of other resources, such as Science Direct, while Scopus and other resources commonly used in assessments are licensed by individual Higher Education Institutions.

Development of research data production in SSH

The development of research data production in SSH in Ireland is slightly below average.

There is a frequent production of research data, part of these are deposited in the Irish Social Science Data Archive¹²⁵ and see also the HRB report. There are some examples of research excellence:

The majority of such surveys are undertaken by the Central Statistics Office and include: Household Budget Survey & Quarterly National Household Survey (QNHS); National Travel Survey (NTS); EU-Survey on Income and Living Conditions (EU SILC). Other significant ongoing surveys include Growing up in Ireland (GUI); National Longitudinal Study of Children and The Irish Longitudinal study on Ageing (TILDA), both undertaken by state-funded research groups. TILDA is also connected at a European level through 'SHARE'. Ireland also participates in the European Social Survey. Diverse other studies have been undertaken but fewer are longitudinal. Electoral studies: trying to establish a more longitudinal type of research instead of project based, environmental studies and health studies monitoring alcohol abuse and intake, drugs misuse (HRB). The HRB has funded a number of key projects with SSH researchers, typically under the Population and Health Services Research schemes. Key examples include the HRB Centre for Health, Diet and Obesity, the HRB Centre for Primary Care.

Ireland is a successful partner in a number of Horizon 2020 European projects that involve a focus on research data. For example, under the Big Data and Public Health Topic Horizon 2020 Health Work Programme 2016 Ireland is involved in the top two ranked projects.

Additional source: HRB Data Project report 'Proposals for an Enabling Data Environment for Health and Related Research in Ireland' (2016). Appendix 5 list the key Health and Social Care Data Collections in Ireland.

RDM policy and support setting

The overarching strategy and policy to enable sustainable data access and sharing of publicly financed social science research data is at an underdeveloped level and below the average.

Policy for RDM has not been established by Irish funders or generally by the Higher Education Institutions sector, although discussions of the basis for potential policy development have

¹²⁵ An overview is available at: <https://www.ucd.ie/issda/data/>

become active during 2016. A national Open Access policy statement has been articulated by a high-level stakeholder group (2012), but has not been updated to address principles relating to research data. Subscription to the policies articulated by this stakeholder group is not binding on state funding agencies.

An Open Government Data Portal has been established with a focus on administrative information and promotion of a culture of data sharing amongst public- sector bodies.

Promotion of Data Management Plans as part of the research lifecycle is limited to individual Higher Education Institutions and research organizations, where the state of development is uneven, but awareness is increasing. The funding agencies, such as the IRC and HRB, have loose recommendations because there is no national position. It mainly represents to good practice and international guidance.

The HRB has started to ask for DMPs under two new schemes under its Strategic Plans 2016-2020 – the HRB Investigator Awards and the HRB Emerging Investigator Awards (EIA).

The EIA is the first scheme to include specific questions on a Data Management and Sharing Plan (DMP). This is in line with the growing international consensus on the need to comply with good practice in the use and reuse of research data, with the goal to preserve and share research data and datasets in a manner that maximises their long-term value and enhances new opportunities for future research. Additionally, it is advised, insofar as possible, to make research data compliant with 'FAIR' principles by following the guiding principles for data sharing and stewardship, which are now internationally recognised (e.g. EC, NIH, UK Concordat on Open Research Data). Although it will not be a mandatory requirement at this stage and it will not be part of assessment criteria, applicants are encouraged to complete the DMP in the application process. Applicants can opt out from the sharing of their research finding due to legal, ethical, Intellectual Properties or other contractual reasons.

The ISSDA is unaware of public-sector organizations that require deposit of social sciences data with ISSDA or other repositories. There is growing awareness of the desirability of making research data available for secondary use, promoted particularly by the emerging Open Science Policy Framework, EC policies, and international funders such as the Wellcome Trust. The ISSDA is unaware of specific policies regarding grant supports amongst Irish funders for data archiving activities.

Hence, there is growing recognition and awareness about the value of research data produced and about the need for long-term preservation; scarce or no investment and support for long-term curation provided.

Ethical and legal framework

The ethical and legal framework, important for data sharing is developed and above average. There are recommendations and guidance provided on how to respect the legal requirements while sharing data.

The legal and ethical framework for data sharing is that which governs other research publications, but of increased importance are regulations regarding data protection at both national and EU level. Some special provisions also apply to official data from the Central Statistics Office. Training support in data protection policy and practice is offered by the Office of the Data Commissioner. Within the Higher Education Institutions sector, organizational ethics review boards have a responsibility for vetting research proposals for compliance to the regulatory framework.

There have been various initiatives and discussions on legislation in Ireland. Knowledge Transfer Ireland is responsible for the National IP Protocol 2016 that provides a framework for best practice, guiding on the expected norms for research-related engagements between industry and State research performing organisations (RPOs) (see <http://www.knowledgetransferireland.com/>).

During the preparation of the new EU General Data Protection Regulation (GDPR) the HRB was a member of the EU Lobby group led by Wellcome Trust UK working to ensure that the voice of research was heard. The EU Lobby Group is now looking at how to coordinate a common implementation that will best support research across the EU. We welcome the continued involvement and leadership of the Wellcome Trust following the UK Brexit decision.

Within Ireland the Department of Justice is responsible for preparing national legislation that will see the DPR implemented in Ireland. There is a national coordinating group that works across the departments here. There have been detailed discussions in Ireland related to key new legislation from the Department of Health-the Health Information Bill and Patient Safety Bill.

Data sharing culture

The data sharing culture is developed and on average level. But as is also explained below, this is more of an assumption and not based on data.

The ISSDA has informed us that there are no available data to provide an evidence base for assertions about data sharing culture and practice. Based on local experience, sharing of research data is not widespread in the social sciences. Also, based on local experience, Open Access principles are often met with concerns over constraints on intellectual property and translation of research findings to innovation and entrepreneurship.

The IRC mentions that researchers often use project or personal websites to store their data, and with no national policy nor national repository they are hardly to blame for choosing informal and less transparent data sharing routines. If there would be a national service and a policy shared between institutions researchers could plan and store through proper protocols.

Data support services are available to social scientists in various contexts. The ISSDA engages with researchers primarily around surveys of national significance. Individual Higher Education Institutions often provide document-oriented institutional repositories, some of which have begun to accept datasets. Other major heterogeneous repositories also accept deposits of data from non-affiliated researchers: the UCD Digital Library (on behalf of ISSDA) and the Digital Repository of Ireland (on behalf of the Irish Qualitative Data Archive).

The Data Producers with which ISSDA commonly interacts are largely professional organizations well versed in best practices of sharing social sciences data, but a growing number of new contributors to ISSDA's data archive require special guidance with regards to a range of best practices.

Major data producers with which ISSDA interacts make use of 'standard' statistical data management software and provide data in SPSS portable format or similarly ubiquitous data formats. There is some unevenness in quality of survey datasets, often relating to the practices of the survey unit that deployed the survey and assembled results.

HRB is involved in data training events. It held the Open Science conference in June 2016 and will hold FAIR data training in January 2017. HRB also participates in numerous Open Access events, including acting as a National Node under the EU Pasteur4OA project.

Irish librarians are involved in the National Open Access Committee.

The ISSDA attends and presents at various events e.g. the Winter School that is run in the University of Limerick every January. It is also starting to organise data training events e.g. it co-hosted (with The Children's Research Network for Ireland and Northern Ireland) a workshop on Anonymising Research Data, last June.

Attitudes towards data sharing

The self-assessment for Ireland done by the ISSDA the scores on statements regarding attitudes of social science researchers on data sharing are not predominantly negative or positive. Data sharing is seen as somewhat beneficial, can create a negative competition and it is probably not a risk that others may misuse or misinterpret data. Data can be reused and facilitate advancement of science.

Enablers for data sharing

Enablers for data sharing is on an average developed level.

There is not much information available to substantiate the average developed level that is shown in the diagram.

In the survey the enablers for data sharing are: career rewards relating to data sharing and publishing in journals that expect data used in the publication to be available for reuse from a trusted digital repository. So far, we have not heard from career rewards relating to data sharing. With regards to publishing the researchers publish through the big publishing houses.

3.19.2 Capability requirement areas of DAS

Organisational profile

The ISSDA (Irish Social Science Data Archive) is administratively and physically at University College Dublin. ISSDA is a single unit residing administratively within UCD Library; ca six regularly contributing staff. The primary and secondary user communities are the academic and other professional research organisations; graduate-level students at Irish universities; a broad range of other national and international researchers.

Funding

The major funding for the ISSDA comes from the University College Dublin and partial funding from the Irish Research Council.

Core services and activities

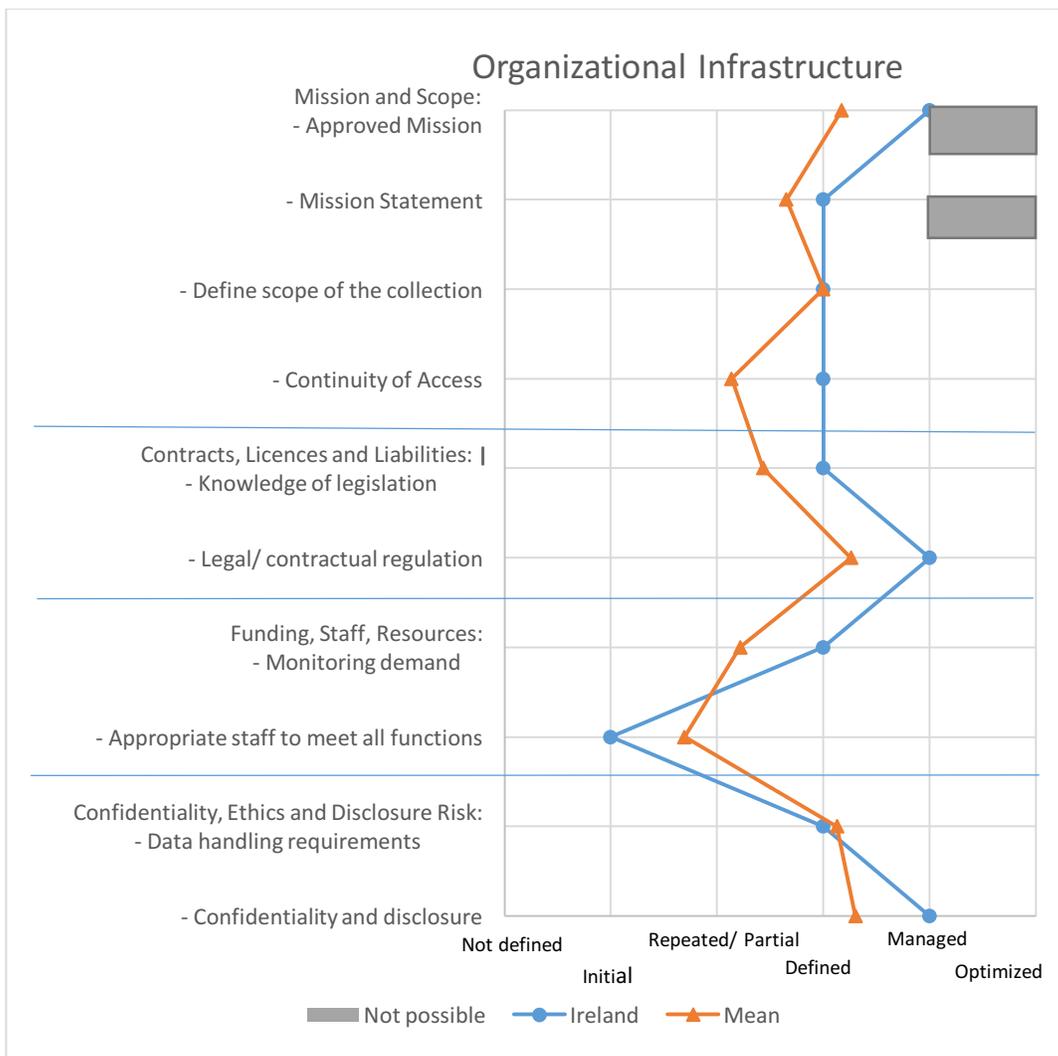
Intake of quantitative datasets from Data Producers; review and normalisation of dataset; cataloguing of datasets; publicity and support of dissemination of datasets; management of agreements with Data Providers and other services provider; community outreach and training programmes.

Content current collection

The overall content of the current collection are longitudinal studies from the national statistical bureau, major Irish longitudinal studies of ageing and children; surveys on a broad range of additional social topics. The studies are primarily in the English Language; accommodation is made in some surveys for Irish language.

Organisational infrastructure

Figure 42: Organizational infrastructure in Ireland



The ISSDA has a long-standing mission statement that has not been subject to approval by other organizations. Mission and related policy documents, including collections policy, license templates, sustainability arrangements, etc., are available on the ISSDA web site. Agreements with Data Producers have been reviewed and agreed by the host institution's legal office as well as counsel of major Data Producers. These agreements are detailed and cover the range of issues identified in the survey.

ISSDA tracks use of studies in its archive and regularly reports on deposits and usage of data.

Staffing of ISSDA is funded primarily by the host organization, University College Dublin. A tranche of external funding from the Department of Education, via the Irish Research Council, is expected from 2017. Staffing is limited, but as the scope of services and demand is also proportionately small, it will be adequate with the additional 1 FTE staff member in 2017.

Technical functions of the data repository are handled by professional staff of the Library of University College Dublin.

Strenuous efforts are made to assure that ISSDA is following all aspects of the regulatory environment and international best practices with regards to handling of confidential data. Processes are defined and documented internally and are subject to continuous review and improvement.

Digital object management (data curation) and Technical infrastructure and risk

The procedures and policies identified in this area are adhered to by ISSDA in general.

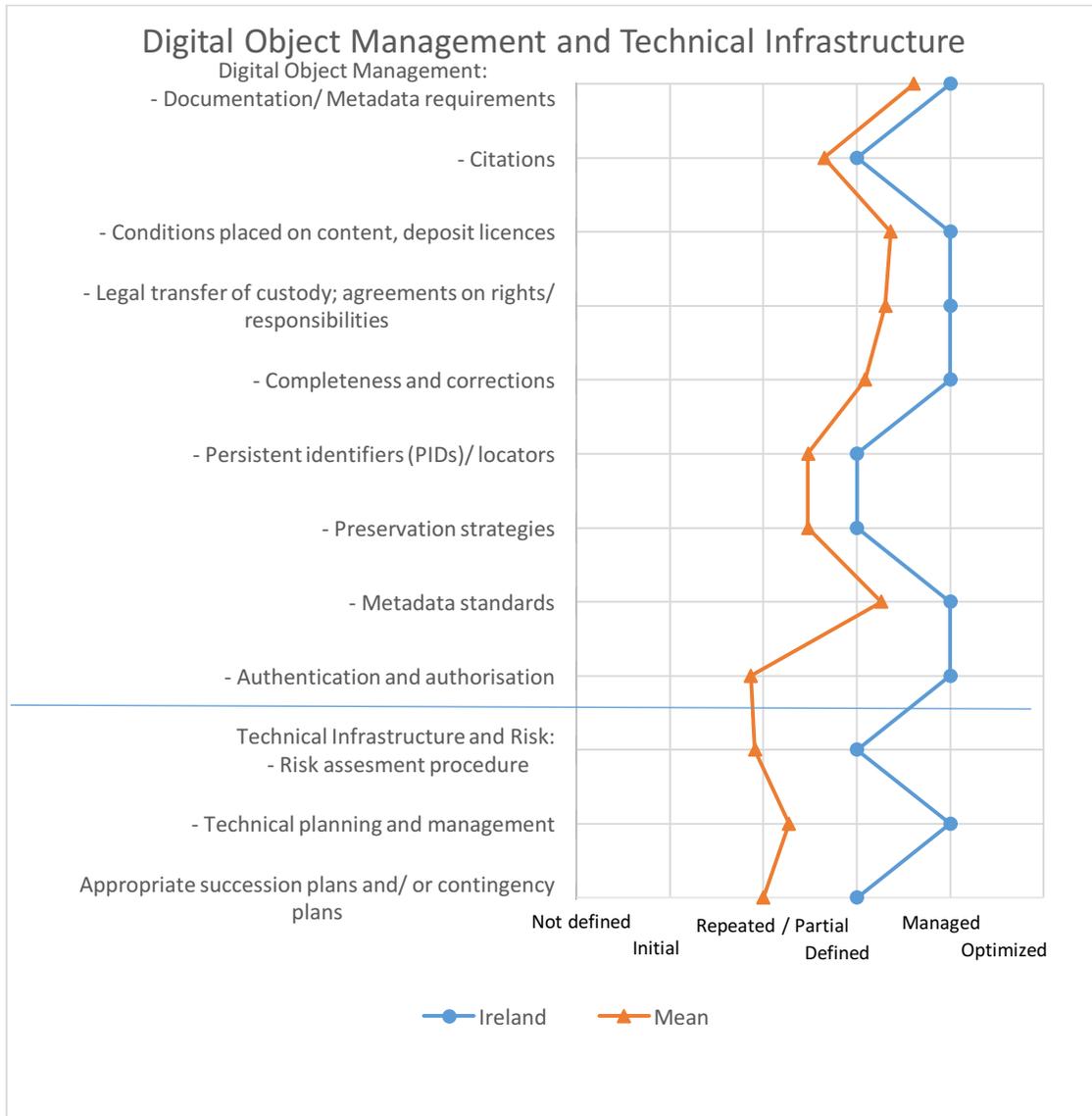
ISSDA has capacity for assignment of DataCite DOIs to datasets and associated documentation, but the practice will be implemented in 2017. It is hoped that guidance from CESSDA will become available, for example, around naming procedures for versions of statistical studies.

ISSDA follows the data preservation practices in place for the UCD Digital Library, which has been certified by the Data Seal of Approval process.

Descriptive metadata is produced in DDI version 2.1 format; additional guidance on best practices amongst CESSDA Service Providers is awaited.

The area covered by this portion of the survey are covered by the policies and practices of UCD Digital Library.

Figure 43: Digital object management and Technical infrastructure in Ireland



3.19.3 Conclusions

This report shows that overall, Ireland has an average maturity level.

The maturity levels that appear in the spider diagrams, seem to be different than the level of maturity that is shown in the textual responses in the survey.

From the conversations with the ISSDA, the Health Research Board and the Irish Research Council it appears that this maturity could be improved by a national policy and strategy with regards to RDM and RDM support. All initiatives currently undertaken would be greatly supported by such a national and not fractional support.

The research community has not been able to finance the nice-to-haves, such as a good RDM support due to a lack of resources, due to the overall economy.

So far Ireland has profited from strong links with the UK for example in the legislative field. With Brexit, it is unsure how these relationships will be impacted.

There are other strong links with the UK, for example with the Wellcome Trust, iDCC, and other organisations in the field. Furthermore, there is participation in a number of European projects and initiatives.

The Irish Social Science Data Archive (ISSDA) scores relatively high on the capability requirement areas.

Ireland is doing well already, is on the move, and will be even doing better in the near future.

3.20 Israel

In Israel, the social science sector is well developed. The main source of research funding comes from national and international institutions. There is a well-established production of research data both at international level, participating in the main international surveys, and at national level (Israeli Central Bureau of Statistics and other topic related repositories).

Israel has not a clear strategy regarding to requirements or recommendations about the management of the research data as a part of on-going project activity. So far there are no national public research funding organizations that provide incentives for sharing data. Consistently, the data sharing and reuse among social sciences researchers in Israel is rather low (0-10%), especially for data collected in small projects, and it is usually shared via informal channels (own network of peers and colleagues).

The attitudes of social science researchers toward data sharing are positive. Nevertheless, the risk of data misuse and misinterpretation is taken in high regard.

There are no career rewards related to data sharing in the academic community in Israel. Some support services that facilitate data sharing are available, provided by the Israel Social Sciences Data Center (ISDC). So far it is not recognised as a national DAS, but a national data archive in the social sciences was included in the recommendations of the Council for Higher Education in order to give financial support to new or existing research infrastructures.

According to the Global Open Access Portal (GOAP) Open Access is a new initiative in Israel, with few activities, publications, and policies that promotes Open Access in the country.

Specifics about the data collection in a country

The data collection started in September 9th and concluded in October 26th. Two persons were contacted to provide information on the country: prof. Micha Mandel, from the Department of Statistics, Faculty of Social Sciences at the Hebrew University of Jerusalem and chair of the Israel Social Sciences Data Center (ISDC), a DAS established by the Faculty of Social Sciences at the Hebrew University of Jerusalem; prof. Noah Lewin-Epstein, professor of Sociology and Dean of the Faculty of Social Science at the Tel Aviv university.

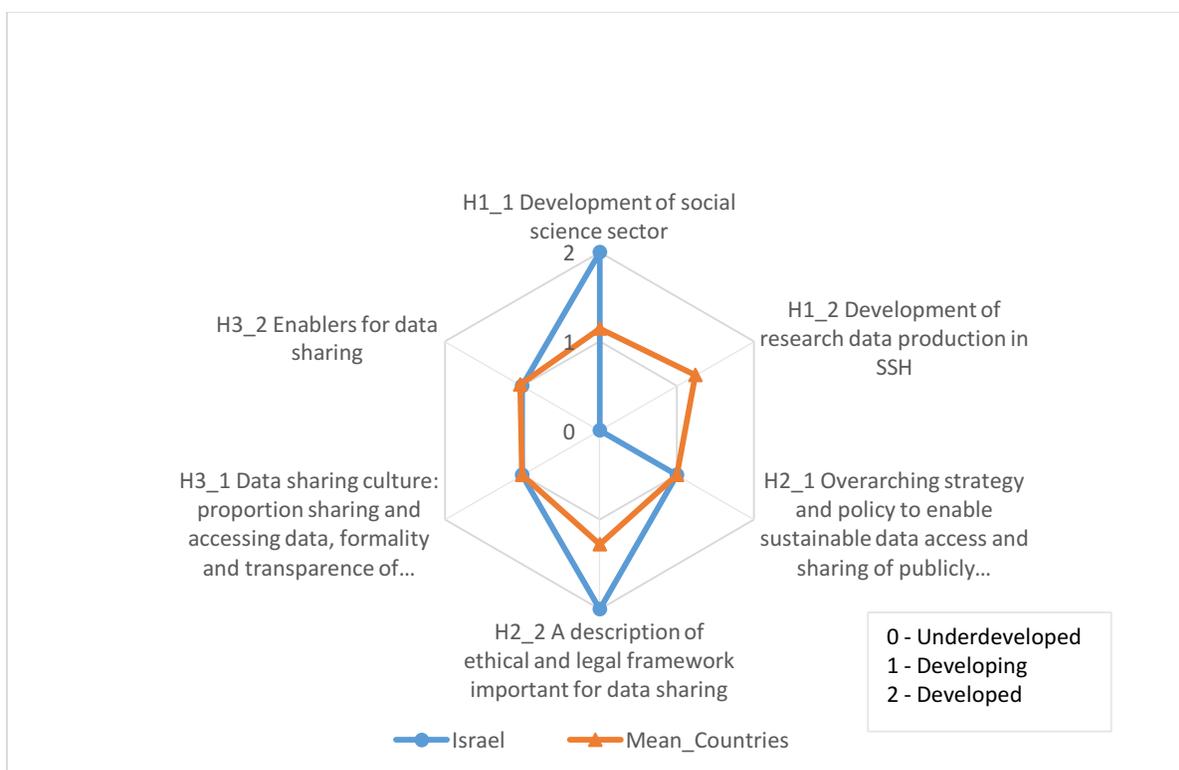
The data was collected using the self-assessment survey provided by the CESSDA SaW project. Some information was provided by e-mail, on request of the partner responsible for the data collection in the country; other questions were directly answered by the contact persons using the web form. The first part of the survey (sections I.1., I.2., I.3., see Appendix 2) was filled with the help

of prof. Lewin-Epstein; the second one (section 4.1., “DAS proto-activities”) was filled with the support of prof. Micha Mandel.

During the data collection, some problem arose about finding the right e-mail address to contact the head of the DAS. With the assistance of the ISDC support staff, however, a contact with prof. Micha Mandel was established after few days. Since he is a statistician and not a social science researcher, he was only able to answer the questions related to the DAS activities. So it was agreed to contact an social science expert to have information about the SSH context in Israel. He suggested prof. Noah Lewin-Epstein, the Israeli contact person for the European Social Survey. Prof. Lewin-Epstein provided details on the broader ecosystem of DAS operation as required in the first part of the self-assessment survey.

3.20.1 Broader ecosystem of DAS operation

Figure 44: Heading concepts values in Israel



Development of social science sector

In Israel, the social science sector is well developed, with a score above the average of the all countries, in a context where in 2014 GERD reached 4.11% of GDP compared to the EU average of 2.03% in the same year. The main source of research funding comes from the government and the higher education sector, followed by international institutions, private non-profit sector and business enterprise sector. The number of social science publications is high, as Israel occupies the eighteenth position upon the World ranking for the 2008-2013 period, with the 0.98% and 1.1%

in Scopus and Web of Science respectively. Universities have access to tools for research (i.e. commercial bibliographic databases, datasets, software).

The high level of development in social science is coherent to a well-established production of research data (even if below the average of the all surveyed countries), both at international and national levels. In fact, Israel is involved in the main international surveys, like:

- European Social Survey (ESS)
- International Social Survey Programme (ISSP)
- Comparative Study of Electoral Systems (CSES)
- Programme for International Student Assessment (PISA)
- Survey of Health, Ageing and Retirement in Europe (SHARE)

In addition, at national level the Israeli Central Bureau of Statistics conducts several annual surveys in topics like education, health, immigration, labour, living conditions and welfare.

RDM Policy setting

There's no a clear strategy regarding to requirements or recommendations about the management of the research data as a part of on-going project activity. As the expert stated, there is a growing recognition and awareness of need to require a Data Management Plan for public funded research, as well as there's the need to have a recognised disciplinary place where research data can be deposited and curated for a long-term preservation. However, in his experience, there are not national public research funding organizations that provide incentives for sharing data, i.e. considering part of the budget for managing the data and preparing it for access.

Both contact persons confirm that so far there is not a recognised national repository for the social science research data. They also confirm that the topic is on the agenda of committees dealing with research infrastructure but so far, no concrete steps have been taken yet.

Regarding the legal and ethical framework, researchers should follow recommendations and guidance while they share their own data, but firstly at an institutional level. In fact, university ethical committees require information on how research data will be handled and how the anonymity of the respondents will be preserved.

Data sharing culture

Based on the experience of the expert, the data sharing and reuse among social sciences researchers in Israel is rather low (0-10%). However, a distinction should be made between projects that collect research data. Large scale projects (like international collaborations such as ESS, SHARE, etc.) are intended from the start for a wide use, so the data sharing and reuse is very common. Instead, data sharing in smaller projects is very low. Consistently, the informal contacts (peers and colleagues) are the most used channel to share data.

The attitudes of social science researchers toward data sharing, in the expert's opinion, can be described in general as positive, but with some remarks. There is a positive evaluation of the benefits of sharing data, especially in reusing existing data for answering new research questions and facilitating advancement of knowledge, and it's believed that this practice involves little effort and minimal costs. Nevertheless, the risk of data misuse and misinterpretation is also underlined.

Probably that is a consequence of a lack of formal and transparent places of deposit that can guarantee an appropriate context and documentation to the shared data. Lastly, the position on the relationship between data sharing and competition in research is neutral, so no negative or positive opinions are expressed on this topic.

Attitudes towards data sharing

Table 17: Attitudes towards data sharing in Israel

Data sharing has no benefits at all	False
Data sharing creates healthy competition	Neither true, nor false
Data sharing creates negative competition	Neither true, nor false
Reuse of existing data can answer new research questions and facilitate advancement of science	True
Data sharing has as a risk that others may misuse and misinterpret data	Probably true
Data sharing involves little effort and minimal costs	Probably true

Source: expert interview. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from "5"-True to "1" False!

To summarize data sharing culture in Israel, data sharing and reuse is not common, especially for data collected in small projects, and it is usually shared via informal and not transparent channels, inside of the own network of peers and colleagues. The general attitudes of researchers related to sharing of data are positive, but the concern about misuse of research data is high and should be taken in consideration in order to increase the awareness on this topic.

The lack of a clear RDM policy setting and no presence of a developed DAS are probably responsible for the gaps in data sharing culture, like low sharing of research data and absence of formal channels to disseminate data.

However, the attitudes in the research community about the benefits of sharing data are positive, but the fear that data can be misinterpreted is still strong, so relevant actions should be promoted in that direction.

Enablers for data sharing

There are no career rewards related to data sharing in the academic community in Israel. There are some support services that facilitate data sharing and/or Open Access to research data, provided by the Israel Social Sciences Data Center (ISDC). The services include data management, data preservation, and data access. Its data are described and published using the DDI standard and NESSTAR software. The main sources of data include the Central Bureau of Statistics (CBS), the National Institute of Insurance (NIOI), central and local government agencies, research institutes as well as independent researchers from affiliated institutes. In that case, data management and documentation standards and procedures are followed.

The presence of the ISDC guarantees a set of data support services to social science research community. Data management and data documentation standards and procedures that facilitate data reuse are mainly followed in handling institutional data (i.e. from the Israel Central Bureau of Statistics). The lack of career-related rewards, no incentives for sharing data from national public research funding organizations and the concern about possible data misuse probably slow down the development of strong data sharing culture in the social science research community, even in a context where some data support services are available.

3.20.2 DAS proto-activities

DAS activities

Currently in Israel there is not a recognised Data Archive Service (DAS). The Israel Social Sciences Data Centre (ISDC) is not a national centre but a unit established by the Faculty of Social Sciences at the Hebrew University of Jerusalem. Its user community includes Israeli universities, colleges and research institutes. The ISDC now houses approximately 1000 datasets including national sample surveys, local studies, census micro-data and government records in selected fields as well as macro-economic and regional series. Sources of data include the Central Bureau of Statistics (CBS), the National Institute of Insurance (NIOI), central and local government agencies, research institutes as well as independent researchers from affiliated institutes. Since 1995, the ISDC serves also as an authorized distributor of Public Use Files from Israel Central Bureau of Statistics. Major sources of foreign data are ISSP and ICPSR Archive in Michigan, where ISDC serves as the Official Representative of Israel.

There are other topic related repositories in the country. Among these:

- the Israel National Election Studies (INES): a repository storing data from the pre- and post-election surveys in order to investigate voting patterns, public opinion, and political participation in Israel;
- the Israel Gerontological Data Centre (IGDC): a comprehensive infrastructure for conducting research on various aspects of population aging, including its social, health-related and economic implications.

In the 2010, the Ministry of Science and Technology started to work on the Research Infrastructure for Israel. The purpose was to compile a database, which would constitute the basis for a road map setting out the establishment of National Research Infrastructure in the country. Accordingly, in 2012 the Planning and Budgeting Committee (PBC) of the Council for Higher Education decided to initiate a long-term planning process for its financial support of major research infrastructure in the public-budgeted academic institutions. An advisory committee was appointed to advise the PBC on the subject by mapping the existing research infrastructure, examining and prioritizing needs for new infrastructure, and forming a roadmap. A national data archive in the social sciences was included in their recommendations, but so far no decision has been taken yet. Given its activities and goals, the ISDC is surely the potential centre that could act as a national DAS for the social sciences. Its staff can guarantee the appropriate skills in long-term preservation and data access, according to the international standards of social sciences data archiving.

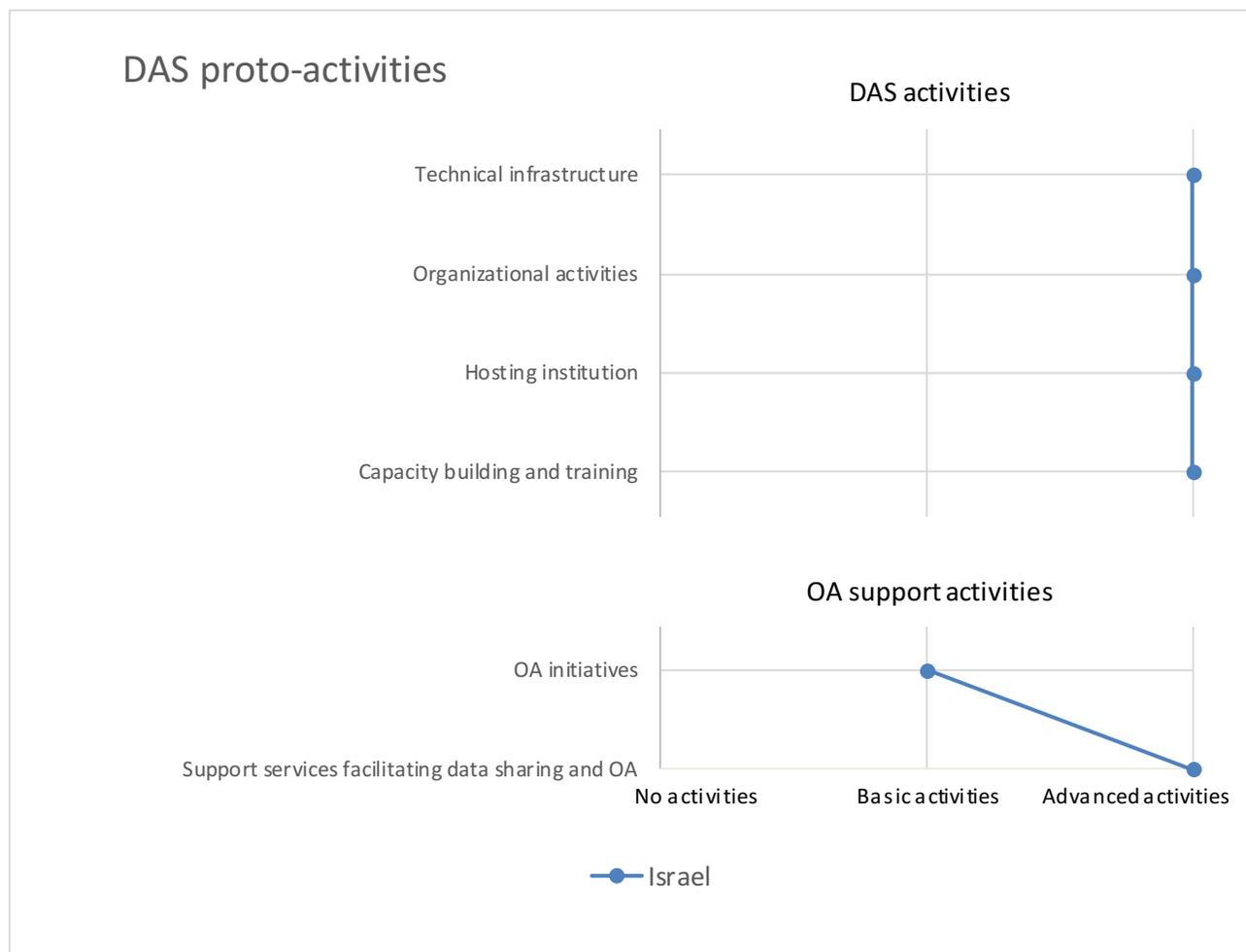
Open access (OA) support activities

According to the information available in the Global Open Access Portal (GOAP) Open Access is a new initiative in Israel. There are two Open Access repositories in the country. The first is Israeli Scholar Works: it was established in 2006 with the aim of collecting the scholarly output of Israel Academic Institutions and Jewish scholars all around the world in several topics (mainly in medical field). The second one is the Weizmann Institute of Science, an institutional repository with reports on Computer Science and Applied Mathematics. However, as of December 2016 they are no more listed in the OpenDOAR. Currently there are five Open Access journals published in Israel, as reported in the Directory of Open Access Journals (DOAJ). So far, no Open Access policies are registered in the Registry of Open Access Repository Mandates and Policies (ROARMAP).

According to the GOAP, a key organization in dealing with Open Access publications is MALMAD, the Inter-University Centre for Digital Information Services, established in 1998 by the Committee of University Heads ("VERA") as a support unit of IUCC (Inter University Computation Centre). Its role is to serve as a consortium for the acquisition, licensing and operation of information services to universities and colleges in Israel, included subscriptions for online publications and bibliographical databases. Currently, Israel is not involved in OpenAIRE project.

As already mentioned, ISDC provide data support services that facilitate data sharing but there's no a clear indication of promoting Open Access. According to the ISDC website, data can be deposited for free usage. Data are available for free only for the users from the Faculty of Social Sciences at the Hebrew University and from ISDC affiliated institutions. For other users, charges apply.

Figure 45: DAS and OA activities implementation type in Israel



3.20.3 Conclusions

In a well-established social sciences context, the main gap that should be underlined is the lack of a solid data sharing culture inside the social sciences research community. At the moment, that culture has no support at the national institutional level: in fact, so far there are no national public research funding organizations that provide incentives for sharing data, as well as there's not a recognised national Data Archive Service that provides services that facilitate data sharing. At the same time, a developed data sharing culture is not settled in the social sciences community: the data sharing and reuse among researchers is rather low, and initiatives that promotes Open Access are not common inside the country.

In the last years, the situation is evolving, as there is a growing recognition and awareness of need to require a Data Management Plan for public funded research as well as a national data archive in the social sciences. These actions should be supported in order to improve the data sharing culture and increase the awareness of the importance of sharing data inside the social sciences research community.

Currently in Israel there is not a recognised Data Archive Service (DAS). The ISDC, however, is surely the potential centre that could act as a national DAS for the social sciences. Its staff, in fact, can guarantee the appropriate skills in long-term preservation and data access, according to the international standards of social sciences data archiving.

3.21 Italy

The analysis shows that although the high level of development of the social sciences sector is visible from country macro indicators in Italy, research data production is characterised as developing and research data management is characterised as underdeveloped. Ethical and legal frameworks are developing and data sharing and reuse, although not very common, can be characterised as developed, while incentives and enablers for data sharing are still developing.

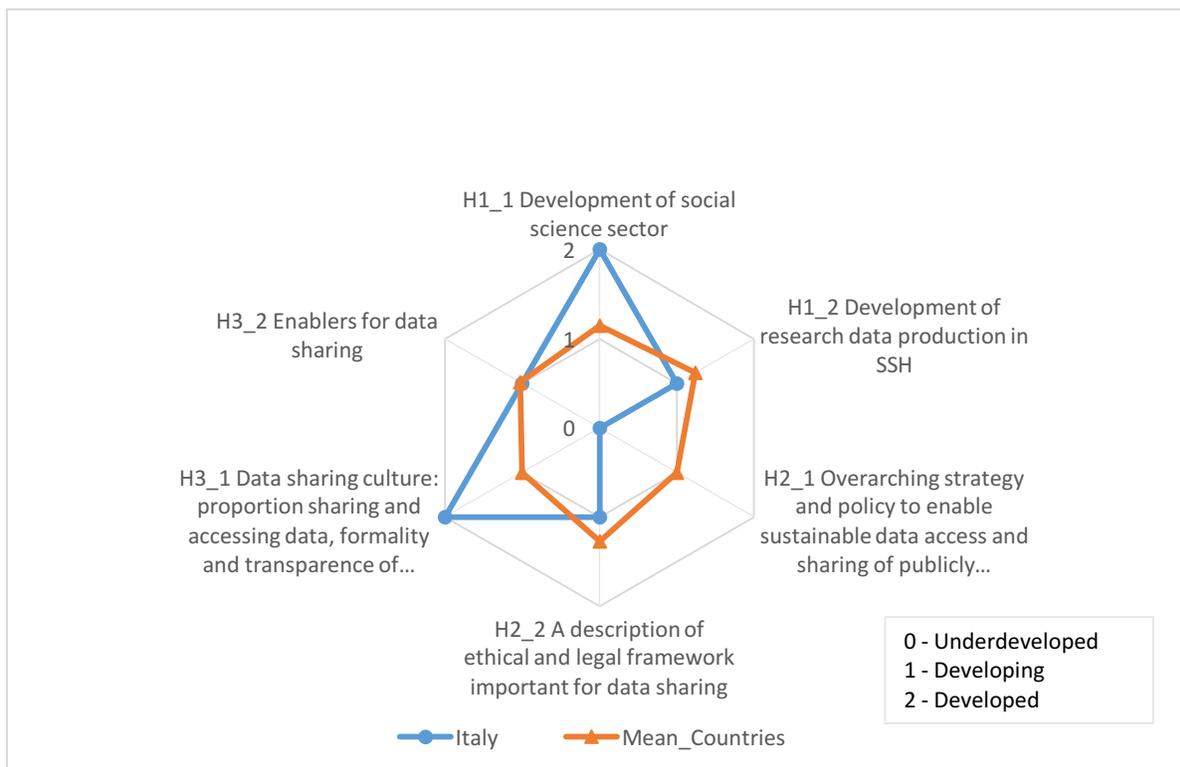
The data archive service of the UniData - Bicocca Data Archive has a defined mission and scope and legal regulation but only initial development in the area of knowledge of legislation and funding, staff and resources. Data handling requirements are only partial, but confidentiality and disclosure risks are defined. The legal transfer of custody, agreements on rights/responsibilities are managed. The documentation requirements, conditions places on content, deposit licences, persistent identifiers, metadata standards and risk assessment procedures are defined. Citations, completeness and corrections and technical planning and management are only partial, while authentication and authorisation and appropriate succession plans are not defined.

Specifics about the data collection in a country

Sonia Stefanizzi, head of UNIDATA was first contacted on 26/09/2016. The information was finally entered by Domingo Scisci and Carlo Pisano, researchers from the UniData – Bicocca Data Archive beginning of October 2016. The method used was thus mainly self-assessment.

3.21.1 Broader ecosystem of DAS operation

Figure 46: Heading concepts values in Italy



Development of social science sector

Overall Assessment of SSH development: High level

The funding of the social sciences sector and humanities (SSH) sector and the productivity of the researchers in Italy are in the highest quantile. The impact on the designated community is strong. Overall, the level of development is high.

In Italy, there were 1,616 researchers per million inhabitants in 2011¹²⁶, while there are no data on the general and relative intensity of investment. Thus, it is not possible to estimate the financial stability, research capacities and results achieved in the field of social sciences in Italy. However, there is at least information on the ranking of the sources of research funding. Considering all disciplines (not only social science), the most important source is the business enterprise sector (45.2%), followed by the government and higher education sector (42.4%), foreign funding (international or cross-border) (9.7%) and last, the private non-profit sector (2.7%). Limiting to only social science, however, the government and higher education amount to 79.85%, while non-profit organisations fund 2.3%, based on the report of the Italian National Open Access Desk (2012)¹²⁷.

The Milano-Bicocca provides access to commercial bibliographic and full-text databases, datasets and software licences¹²⁸.

Development of research data production in SSH

In Italy, research data production is developing; however, high quality research data with high potential for reuse is rare. There are only a few research projects that produce data. The existing data is low quality and dispersed as data is held at many different institutions. There are not many studies of national importance and no current examples of international collaborative research.

Italy was involved in some international collaborative research or cross-national studies such as EVS (all waves), PISA (all waves), ESS (only some waves), ISSP (only some Waves) and CSES (only one wave) but not in the CSS and WVS. In addition, many surveys are conducted by private research institutes and non-profit organizations. In political science, the most remarkable survey is conducted by ITANES (Italian National Election Studies), which promotes a research programme on voting behaviour in Italy¹²⁹.

RDM Policy setting

Research data management in Italy is developing. Currently Italy's National Science Foundation does not require RDM and "there are no requirements or recommendations about long-term curation". Moreover, the public research funding organizations do not offer any incentives for

¹²⁶ Source available at: http://www.uis.unesco.org/ev.php?URL_ID=3755&URL_DO=DO_TOPIC&URL_SECTION=201

¹²⁷ Italian National Open Access Desk report. Available at: http://eprints.rclis.org/29671/1/What%20do%20Italian%20Researchers%20think%20about%20Open%20Research%20Data_rev-charts.pdf

¹²⁸ For more information visit the website of the University of Milano-Bicocca Library, available at:

<https://www.biblio.unimib.it/go/Home/Home-English>.

¹²⁹ Information on how ITANES data is disseminated can be found on its website, available at:

<http://www.itanes.org/en/data/>

sharing research data with associated metadata, such as providing adequate resources for managing the data and preparing it for access.

Data sharing culture

The data sharing culture in Italy is not developed. The country representative for was not able to provide estimates for the proportion of researchers sharing the data and the proportion of researchers able to access existing third party data. However, according to the Italian National Open Access Desk (2012) report there is initial growing awareness about the importance of providing support on legal and ethical aspects of data protection but no organised support is given.

However, there are established data sharing channels and routines. The most popular data sharing channels in Italy are supplementary data in journals (alongside papers) which are ranked first. Data is shared also via informal personal contacts (ranked second), data archives or repositories (third) and via project or personal websites.

Most attitudes of Italian researchers toward data sharing that were self-assessed by the country representatives could be characterized as neutral. They consider false the statement that “data sharing has no benefits at all”, are neutral regarding the statements that “data sharing creates healthy competition”, that the “reuse of existing data can answer new research questions and facilitate advancement in science” and that “data sharing involves little effort and minimal costs”, while they consider the statements that “data sharing creates negative competition for the researcher” true and “data sharing has a risk that others may misuse and misinterpret data” probably true.

Attitudes towards data sharing

Table 18: Attitudes towards data sharing in Italy

Data sharing has no benefits at all	False
Data sharing creates healthy competition	Neither true, nor false
Data sharing creates negative competition	True
Reuse of existing data can answer new research questions and facilitate advancement of science	Neither true, nor false
Data sharing has as a risk that others may misuse and misinterpret data	Probably true
Data sharing involves little effort and minimal costs	Neither true, nor false

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from “5”-True to “1” False!

To summarize data sharing culture in Italy, data sharing and reuse are probably not very common but formal channels are more important than informal ones. Although they recognize the potential

of data for advancement in science and other benefits, Italian researchers are a bit sceptic on data sharing because of the risk of misuse and interpretation and because of negative competition.

Enablers for data sharing

Enablers for data sharing are developing in Italy; however, here are no career rewards related to data sharing in the academic community. In fact, “there are no clear rules in national law and national funding programmes on data sharing, so that’s an indicator that the data sharing is not relevant in career progression.”

On the other hand, there are data support services available to social science researchers that facilitate data sharing and/or Open Access to research data. For instance, “UniData – Bicocca Data Archive (<http://www.unidata.unimib.it/>) offers services on data management planning, long-term preservation, and access to data. In this period, they are working on archiving research data produced by the researchers and research groups of the University of Milano-Bicocca in several disciplines (sociology, economy, psychology).”

3.21.2 Capability requirement areas of DAS

Organisational profile

Organisation

Unidata – Bicocca Data Archive (<http://www.unidata.unimib.it/>) is an interdepartmental centre of the University of Milan-Bicocca, established in 2015. It is housed in the Department of Sociology and Social Research. The centre is the Italian point of reference for the research data archiving and dissemination, based on the example of the National Archives located in major European countries and beyond.

UniData is structured in five Units: Archive, National and International Funding, Public Relations, Training, and Legal Issues. The Archive isn't structured in units or departments. The director of UniData is responsible for the Archive and the staff consists of two full time persons and other temporary employees (students, interns, etc.).

The Designated community of UniData – Bicocca Data Archive is the Italian academic community. Some (open) data gathered are accessible to all users who make a request.

Funding

UniData is 100% publicly funded, through membership of seven Departments of the University of Milano-Bicocca.

Core services and activities

The core services and activities of UniData are data curation, data preservation, review of data sources, support in the Data Management Plan, data dissemination, and management of data requests.

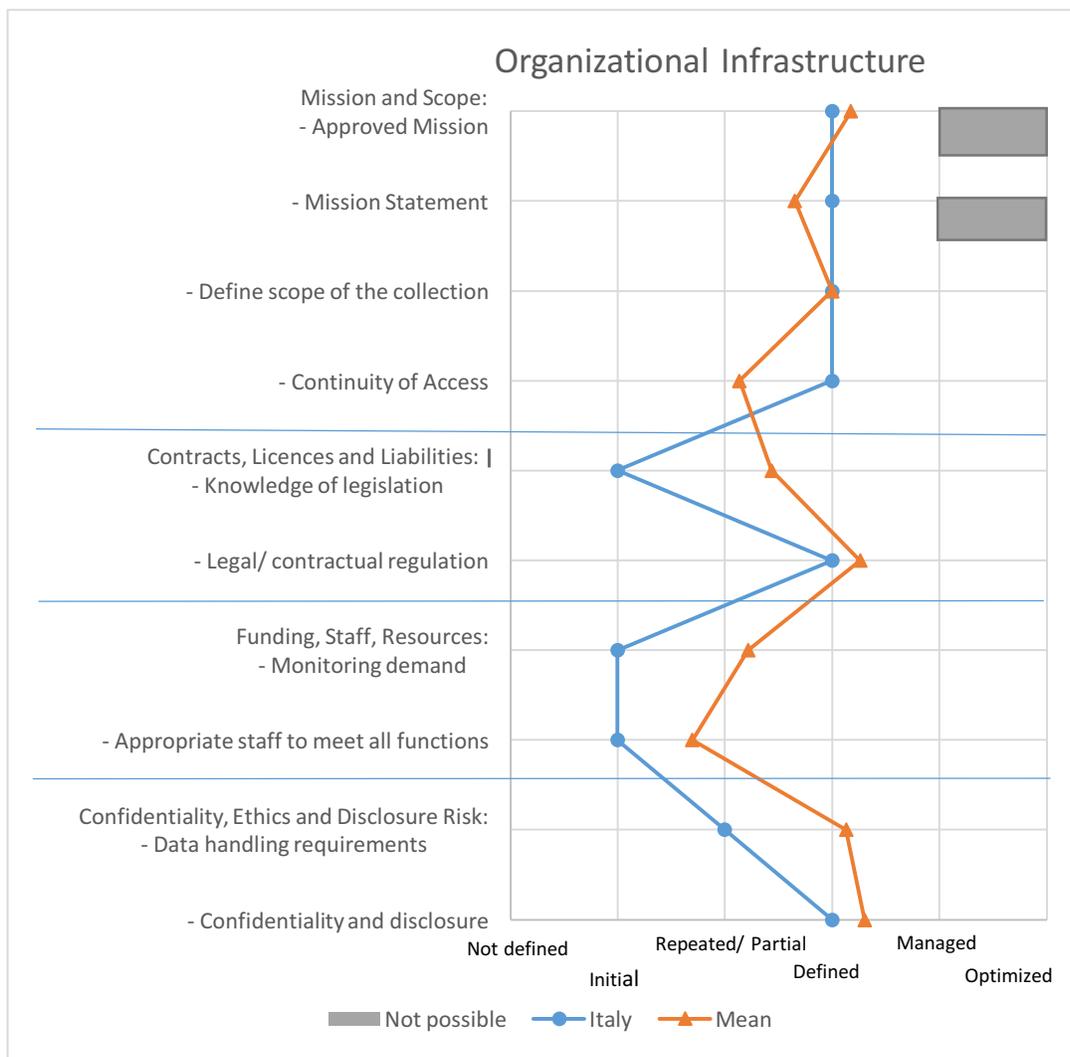
Content current collection

UniData archives about 1,000 datasets (both micro and macro data). However, most of those datasets are from the National Institute of Statistics (ISTAT) or from Eurostat, while data from academic research projects are rare. The collections deal with labour force survey, household budget, household income and wealth, health conditions, family and social subjects, security, school to work transition, time use, trips and holidays. The “Aspects of Daily Life” is the most important survey conducted by the National Institute of Statistics, and it collects information concerning to the social, domestic and urban environment; health condition; cultural habits; satisfaction ratings on some public services; ICT use. Moreover, UniData disseminates the Eurobarometer in Italy, which is already distributed by GESIS.

All metadata produced by UniData are both in Italian and English. The most of data collected are in Italian language (variables name, labels, etc.), whereas the international data (Eurobarometer, ESS, etc.) are in English language.

Organisational infrastructure

Figure 47: Organisational infrastructure in Italy



Mission and scope

The organisation has repeated/partial approval for its mission statement, i.e. there is some evidence to support the notion that the organisation and its activities receives approval from stakeholders; the organisation is on an ad hoc basis “encouraged” to continue its activities, but no formal mandate or formalised approval exist. An explicit written, formal and complete mission statement on the role, mandate, purpose and mission of the organisation is defined and is included in the Statute of UniData¹³⁰ – Bicocca Data Archive where it is available to read for all relevant users but only in Italian¹³¹.

The organisation has defined and promulgated an explicit scope for the repository and their collection and formal, written succession/contingency plans and agreements with external organisational frameworks. Specifically, “UniData has a written succession plan that ensures the data accessibility involving other internal organizations (other Departments of the University of which UniData is part).” In the event of cease, the Statute of UniData “establishes a plan to ensure the continued availability and accessibility of the data, according to the cause of cessation. If only one Department remains, the archive will cease its activities and the University of Milano-Bicocca will redistribute the holdings among the Departments that have been involved in the center. If the Steering Committee and the Participants Assembly will detect the impossibility to carry on with the activities, the archive will cease their activities and the Departments involved in the center will decide how to allocate the holdings. Lastly, if UniData will change their structure (shifting in a new institution or changing its mission), the holdings will be made available to the new structure. «

Contract, licences and liabilities

The organisation has very limited knowledge and/or documentation on how the national legislation applies and affects the holdings and procedures of the organisation. Specifically, “the staff of UniData has a sufficient knowledge on all relevant legislative aspects but is not documented.”

Contracts and/or agreements are defined, i.e. they are standardised and implemented according to written procedures and made publicly available. For instance, the agreement for data access is available on UniData website for each dataset¹³².

¹³⁰ Available at: <http://www.unidata.unimib.it/wp-content/uploads/Statuto-UniData.pdf>

¹³¹ The English translation of the mission statement is the following: “The main purpose of UniData – Bicocca Data Archive is to provide a solid and multidisciplinary infrastructure to the scientific community in compliance with the directives of the “European Strategy Forum on Research Infrastructures” (ESFRI) in the field of secondary analysis for the social sciences. UniData promotes the development of three areas. The first is to strengthen the sharing and dissemination of research data, widening the role of ADPSS-Sociodata that has worked to collect, archive and distribute a large amount of data in the last decades. The objective is to establish a new Italian Data Archive for the social sciences. The second area concerns enhancement of the secondary analysis, with attention to the promotion of a responsible use of the empirical data, to the sharing and dissemination of such resources, and the development of methodological tools that facilitate their analysis. The last area concerns the promotion of a framework in order to develop interdisciplinary empirical researches. UniData will be a place of convergence, integration and collaboration among different academic institutions.”

¹³² Rules for access to the data distributed by UniData and specific data user restrictions are available at: <http://www.unidata.unimib.it/wp-content/uploads/Regolamento-Accesso-Dati.pdf>

Funding, staff, resources

The organisation only reacts to significant changes in demand for the repository services when the changes occur. It does not attempt to adjust its funding, i.e. “UniData adapts their services in compliance with significant changes of demand or with specific requests from designated community.”

The staffing is only partially complete as the organisation has not defined the appropriate staffing level to support all repository functions and services, i.e. “The staff of UniData – Bicocca Data Archive has the core competencies but the number of staff is inadequate to support all functions and services due to lack of resources.”

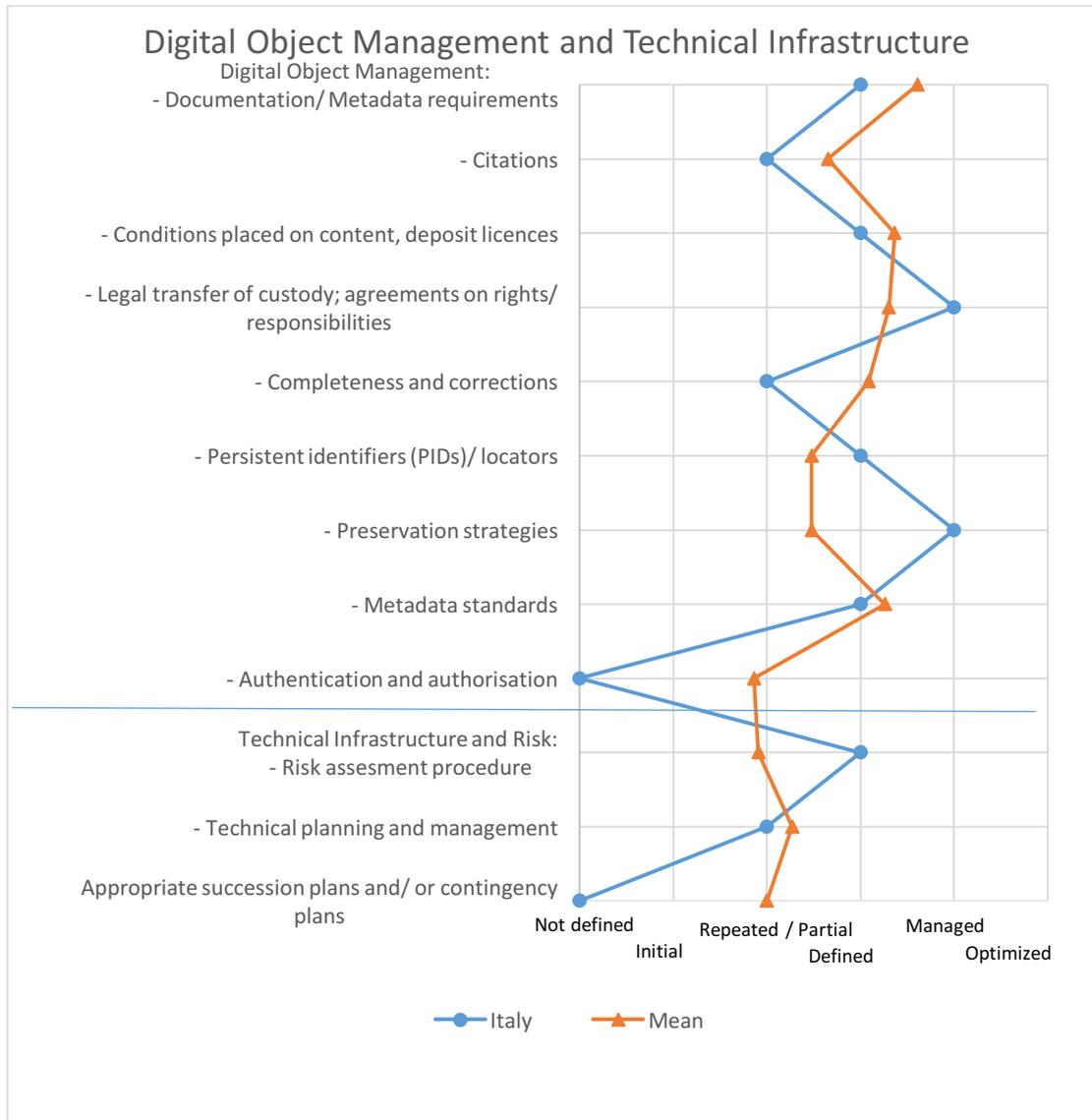
Confidentiality, ethics and disclosure risk

The organisation partially complies with norms through repeated action. However, there are no written procedures or procedures for exceptions and errors, i.e. “UniData complies with social science norms, e.g. respecting the confidentiality and ethics requirements, but there are not formal written procedures.”

Processes and procedures are defined in place and standardised information is provided to the depositor prior to the deposit. Checks are performed on data after deposit. Process and procedure descriptions for handling and altering sensitive data are in place. Specifically, “UniData doesn't accept confidential data for the deposit. Therefore, the depositor should create anonymised data to deposit own data in UniData archive. UniData informs the depositor on data confidentiality, and carries out checks in order to assess the disclosure risks.”

Digital object management (data curation) and Technical infrastructure and risk

Figure 48: Digital object management and Technical infrastructure in Italy



Documentation/Metadata requirements

A written formal specification of required information is explicitly defined and requirements are compliant with metadata standards that are used and can be understood by Designated Community (e.g. DDI); metadata requirements are accessible and communicated to users/depositors. "At present, UniData is producing a formal specification of required information (that we have already defined) in compliance with DDI standard and ingest procedures."

Citations

Citation practices are being repeated and offered regularly but there is no culture of citation or the procedures lack formalisation and systemisation, i.e. “UniData requires and offers a model of data citation for all data collected. The indication of data citation is available on website for each dataset, but the procedures are not formalised and documented.”

In addition to the data citation, they also have a “**deposit requirement**” statement: the users are required to quote all data and documents disseminated by UniData and used in their own publications. In addition, they also should send bibliographic citation related to the publications where the requested data and documents are used.

Conditions placed on content, deposit licences

All depositors are offered the opportunity to set access conditions on the information that is being deposited. A set of access conditions are formally defined in categories or a template. The Data Deposit Agreement includes the possibility to set the access conditions for data deposited¹³³. UniData encourages the Creative Commons 4.0 license, but the depositor can choose, if needed, some limitations in the data access according to purpose and type of users.

Legal transfer of custody, agreements on rights/responsibilities

There is monitoring of the usage of agreements and contracts that are reviewed and updated regularly. Actions are taken where contracts/agreements appear not to be working effectively or are not in accordance with a higher level of policies. With this statement the archive accepts the rights of data editing and archiving.

The Data Deposit Agreement¹³⁴ is available on the UniData website but only in Italian. The document includes the purpose of contract, the conditions of data use and of data access, the definition of roles (depositor, archive and users), the actions that can be performed on data, etc.

Completeness and corrections

There are non-systematised (manual) checks of deposited material in place; processes and procedures are repeated but they are not formalised or documented. Rectifications are performed repeatedly, either by the repository or by returning data to depositor. “At present, UniData is producing a formal definition of procedures and processes formalized in written documents. The documents are not yet available, but are in implementation.”

¹³³ The document is available at: <http://www.unidata.unimib.it/wp-content/uploads/Accordo-Deposito-Dati.pdf> (only in Italian)

¹³⁴ Data Deposit Agreement, available at: <http://www.unidata.unimib.it/wp-content/uploads/Accordo-Deposito-Dati.pdf>

Persistent identifiers (PIDs)/locators

There are mechanisms and systems in place to persistently identify and locate data and metadata (either by following external systems like DOI, or by internal PID systems). All processes and procedures are documented and formalised, i.e. “Each data archived has an internal ID number (unique identifier) and a DOI in order to generate a persistent identifier of data collected.” As the Italian representatives further explained, “UniData has the ability to assign a DOI to all data. It is assigned automatically to all new studies for which UniData is the primary distributor: Thus, the National Institute of Statistics’ survey don’t have DOI. In the future, once it is available, the new system of archival management, DOI will also be associated with all the studies already available in the archive. «

Preservation strategies

The strategy is periodically reviewed and updated, i.e. “UniData is supported by the IT Department of the University of Milano-Bicocca in order to guarantee an effective preservation strategy. “

Metadata standards

A data/metadata format strategy is explicitly defined, formalised, and communicated to users. Data and metadata are provided in formats that are commonly in use and understood by users, i.e. “In accordance with the license, UniData provides data collected in open format (CSV) or in the formats of the most used statistical software by the Italian academic community (Stata, SPSS, SAS, R). As far as we know, metadata standard used by the Designated Community doesn't exist in Italy. UniData provides metadata comply with the DDI-C 2.5 standard for each data collected. All information related to metadata are clearly communicated to the users through the website.”

Authentication and authorisation

There is no authentication approach in place, i.e. “At present, UniData uses an internal authentication system that allows to control the access to data. In the next months, the CAS (Central Authentication System) will be implemented in order to obtain the same system of the University of Milano-Bicocca, and to ensure a more effective control of the data access managed by IT Department of University. “

Risk assessment procedure

There is a documented risk assessment methodology (and tool) that is used to make a systematic analysis of security and infrastructure resilience risk factors when there are changes to the technical infrastructure, i.e. “UniData is supported by IT Department of the University of Milano-Bicocca in order to guarantee a systematic risk analysis of the infrastructure.” The documentation is not publicly available.

Technical planning and management

There is evidence that there is likely to be an appropriate level of technical infrastructure resources to support all repository functions and services. However, the organisation has not defined what this level should be, i.e. "UniData is supported by IT Department of the University of Milano-Bicocca in order to guarantee an appropriate technological infrastructure, but there isn't a defined and documented technological level to ensure the supply of the services."

Appropriate succession plans and/or contingency plans

There are no succession plans, contingency plans, and/or escrow arrangements, i.e. "UniData is supported by IT Department of the University of Milano-Bicocca in order to guarantee the data preservation, but there isn't a plan in place in order to ensure the data access in case the repository ceases to operate." In the event of cease, there is not a "technical" plan in order to ensure the data access. As reported in the Statute of UniData, »according to the causes of the cessation, the holdings will be redistributed among the Departments that have been involved in UniData or it will be decided how to allocate the holdings. For example, if the holding will be redistributed among different Departments there is not a clear instruction about how the data accessibility is technically ensured.

3.21.3 Conclusions

Regarding the broader ecosystem of data archive services operation, we can conclude that:

- Funding of SSH and productivity of the researchers in Italy are at the high level of development and that impact on designated community is strong.
- Research data production in SSH is developing. High quality research data with high potential for reuse is rare and data quality is low. Italy is currently only involved in some cross-national studies.
- Research data management policy is not required by the national institutions and data management plans are underdeveloped.
- Ethical and legal framework important for data sharing is developing. There is a little or no awareness regarding the support needed.
- Data sharing and reuse is not very common in the social sciences research community; there is no information on its channels, while attitudes can be characterized as neutral and therefore developed.
- Incentives and enablers for data sharing within the social science research community in Italy are currently not existent but there are data support services available for researchers that would facilitate data sharing. Thus, it can be benchmarked as developing.

Regarding the capability requirement areas of data archive service of the UniData - Bicocca Data Archive we can conclude that:

- Organisational infrastructure: It has a defined mission and scope and legal regulation but only initial development in the area of knowledge of legislation and funding, staff and resources. Data handling requirements are only partial but confidentiality and disclosure risks are defined.

- Digital object management and technical infrastructure: The legal transfer of custody, agreements on rights/responsibilities are managed. The documentation requirements, conditions placed on content, deposit licences, persistent identifiers, metadata standards and risk assessment procedures are defined. Citations, completeness and corrections and technical planning and management are only partial, while authentication and authorisation, and appropriate succession plans are not defined.

3.22 Kosovo

Broader ecosystem of data service operation in Kosovo can be considered generally underdeveloped, although some positive developments were identified in the country.

The majority of research in social science in Kosovo is currently financed by international funds and some important studies are carried out by the civil society sector (NGOs, Think Tanks). Although the government of Kosovo is willing to increase investments in science, the amount projected by the lows was not yet fully realised. The lack of communication and cooperation between different research units was recognised as one of the problems and the lack of knowledge about think tanks and their activities.

A notable effort towards establishment of data services for social sciences in Kosovo has been made during the SEEDS project¹³⁵. Issues related to RDM policy setting were discussed during the interviews with policy makers and other relevant stakeholders conducted during the SEEDS project, so we can say that overarching strategy and policy to enable sustainable data access and sharing of research data is developing.

Data sharing and reuse among social sciences researchers in Kosovo is very rare, but the core group of researchers could be identified who are willing to share their data and who recognise the benefits of such practices.

DAS proto-activities in Kosovo started with the SEEDS project. Most of the issues related to the process of establishment of data services for social sciences were tackled through the project, such as policy development, governance, human resources, and technical issues. The skills and knowledge were developed through the project which can be used to further influence the development of data services in the country.

Specifics about the data collection

The data was collected during the first week of November 2016. The online self-assessment instrument was used to collect information from the country representative, and this was supplemented by information gathered via desk research using mainly the reports from the SEEDS project, and the RRPP Western Balkans programme¹³⁶. Self-assessment questionnaire was filled up by Nitë Bylykbashi-Deliu, a project manager at the Centre for Political Courage (CPC) in Prishtina, Kosovo. Because of her involvement in the SEEDS project, in which the CPC is one of the partners, and her experience in coordinating several social science research projects, she can be considered as a good informant for providing data to this survey. Most of the information, no

¹³⁵ SEEDS - South-Eastern European Data Services (2015-2017). <http://seedsproject.ch>

¹³⁶ Regional Research Promotion Programme - Western Balkans. <http://www.rrpp-westernbalkans.net>

matter the source, was recorded into the online survey tool, and additional information was pulled out directly from desk research sources.

3.22.1 Broader ecosystem of DAS operation

Figure 49: Heading concepts values in Kosovo



Development of social science sector

Overall assessment of SSH development

There is a lack of readily available comparative statistics for Kosovo on gross domestic expenditure for research and development. According to the RRPP Policy Dialogue - Briefing Paper 2/2014¹³⁷, the estimate of GDP allocated to the research and development is around 0.1%, which is the smallest level in the region of South Eastern Europe. This is also far less than the amount of 0.7% of GDP defined in the Kosovo's Law on Scientific Research Activities¹³⁸. On the other hand, the growth of GDP by 3.8% was recorded in the third quarter of 2016 compared with the third quarter of 2015¹³⁹.

From the results of the self-assessment survey, ranking of research funding sources in the social sciences by the amount they provide, research funding is provided mostly by international programs. Other funding - government and higher education, business enterprise sector and private non-profit sector are reported to be extremely low. This is confirmed by the results of the project SEEDS *Survey on production, preservation and use of research data*¹⁴⁰ conducted in 2015 among researchers in Kosovo, where it was found that more than 50% of surveyed researchers used international funds for their research. The second and third common source was the institution where the respondent is working (27%) and personal funds (21%). Public funding from national science funding bodies was reported by 6% of the respondents, other public sources by 11%, and private by 6%.

Research capacities were assessed through existence of basic e-infrastructures providing access to scientific information (full-text and bibliographic databases, datasets) and software licences for the software used in research. According to the National Library of Kosovo website¹⁴¹, a very basic set of bibliographic and full-text databases and datasets are available for researchers in Kosovo. Software licences for specific research software (e.g. STATA, SPSS) are not provided on a national or university level.

Development of research data production

Kosovo was not involved in any of the international collaborative research or cross-national studies, which are high quality research with a high potential for reuse.

According to the self-assessment survey results, there are some studies that assess matters of national importance (public opinion survey, election survey, etc.) produced by the social science

¹³⁷ RRPP (2014). *The State of Implementation of the National Research Programme of the Republic of Kosovo*. Available at: http://www.cpc-ks.org/repository/docs/Briefing_Paper_National_Science_Program_356224.pdf

¹³⁸ Available at:

<http://www.kuvendikosoves.org/common/docs/ligjet/Law%20on%20Scientific%20Research%20Activities.pdf>

¹³⁹ Kosovo Agency of Statistics (2016). *Gross Domestic Product (GDP), Q3 2016*. <http://ask.rks-gov.net/en/kosovo-agency-of-statistics/add-news/gross-domestic-product-gdp-q3-2016>

¹⁴⁰ SEEDS (2016). *D3 – Report on evaluation of research and legal conditions: Kosovo*. http://seedsproject.ch/wp-content/uploads/2015/06/Report_Kosovo_SEEDS_final_rk_FINAL.pdf

¹⁴¹ National Library of Kosovo Website available at <http://www.biblioteka-ks.org/> (accessed 23.1.2017.) lists the following databases: EBSCO research databases, OECD Library - books, journals, statistics, Edward Elgar e-books and e-journals, Cabridge journals, IMF elibrary - publications and statistics, Duke University Press books and journals.

researchers, but they are not systematic¹⁴². The studies are usually carried out by the civil society sector (NGOs, Think Tanks). Our respondent estimated that the average production of research data by the social science institutions in Kosovo is rare, and the data are produced ad hoc rather than systematically. SEEDS Survey showed that 88% of the researchers did produce data in the last 5 years, although this is for sure an overestimate since the survey sample was self-selected with those who engage in data gathering being much more interested in participating.

As stated in the RRPP Policy Brief report¹⁴³, it seems that civil society sector produces research of higher quality than in public and private sector, although the general level of analysis is weak, due to low quality of higher education development in general. In the same report, main problems were identified in the current state of research in different institutions. Among other, the lack of communication and cooperation between different research units was recognised and the lack of knowledge about think tanks and their activities.

RDM Policy setting

Funders' data management and sharing strategy and/or policy

Public funders in Kosovo currently do not have any policies related to research data management and preservation and no incentives for sharing research data with associated metadata. It was estimated in self-assessment survey that there is a growing recognition and awareness among public funders about the need to require Data Management Plans and Open data, and about the need to have a disciplinary specific repository for research data together with support services. There is also a growing awareness about the value of research data produced and the need for its long-term preservation, but no investment and support is provided for this.

It is important to note that these estimates were based on the fact that Kosovo was involved in the SEEDS project, which was the first and only project in Kosovo that started to raise awareness about all of the specific issues related to RDM policy identified in our survey. These issues were discussed during the interviews with representatives of the Ministry of Education, Science and Technology and other relevant stakeholders. The full report of these activities can be found in SEEDS project deliverables¹⁴⁴.

Another initiative worth mentioning is Open Data Kosovo that aims to promote open access and data sharing in general. They are developing innovative technological applications and

¹⁴² Here are some examples of studies pointed out in the survey:

"Public opinion surveys such as:

- <http://levizajafol.org/folnew/wp-content/uploads/2015/09/CorruptionSCAN-Public-Opinion-Survey.pdf>
- https://www.ndi.org/files/Kosovo_BaselinePresentation_121010.pdf
- http://www.uboconsulting.com/opinion_polling.html

UNDP also carries many public opinion surveys

- <http://www.ks.undp.org/content/kosovo/en/home/publications.html>

Election surveys are usually carried by NGO "Gani Bobi", <https://web.facebook.com/Gani-Bobi-Center-141175572576013/info/?rdr> "

¹⁴³ RRPP (2013). Policy brief: Social Science Research in Kosovo. http://www.rrpp-westernbalkans.net/en/policy-dialogue/Kosovo/mainColumnParagraphs/05/text_files/file/Policy_Brief_English.pdf

¹⁴⁴ SEEDS (2016). D3 – Report on evaluation of research and legal conditions: Kosovo.

http://seedsproject.ch/wp-content/uploads/2015/06/Report_Kosovo_SEEDS_final_rk_FINAL.pdf

programmes to enhance good governance, government transparency and accountability¹⁴⁵. The Portal of Open Data Kosovo¹⁴⁶ was created by the government of Kosovo with professional help from Open Data Kosovo. Through the Portal, administrative data from municipalities and ministries are made available for the general public.

Legal and ethical framework

Since there are no policies related to data management and sharing, it is understandable that there is also no awareness of existing clarification and support provided on legal and ethical aspects that facilitate social science data sharing. Anyway, all of the relevant laws are existing in Kosovo, and some issues related to data sharing and preservation are identified in these laws¹⁴⁷.

Data sharing culture

Prevalence of data sharing and reuse

Data sharing and reuse among social science researchers in Kosovo is very rare. In the self-assessment survey, the estimate of proportion of social science researchers who have shared their research data was on the lowest level (0-10%), the same as the estimate of proportion of researchers who have been able to access existing third party data.

Existing routines of data sharing

Current practices are that only 15% of researchers, as shown in SEEDS Survey, stored their data in a data archive/repository, while the vast majority stored the data on their own or their colleague's personal computers, and fewer than 30% stored the data on a server at their local institution/university. According to the same survey results, the majority of research data is not shared further than the research team members, and the small percentage of respondents do share their data among members of their institution or publicly. Most used sources where researchers obtain the data for secondary analysis were the national and international projects websites, National Statistical Office and data archives in other countries. In our self-assessment survey, it was reported that the researchers are using the following routines for data sharing: project or personal websites and informal contacts (peers and colleagues). No other offered choices (data archives, supplementary data in journals, and other sources) were used in ranking of sharing routines.

The SEEDS survey also identified the barriers to conducting secondary analysis in Kosovo. The main barrier is perceived to be the lack and inaccessibility of relevant data. Also, the same survey revealed that poor documentation of data and lack of training in secondary analysis contributed to

¹⁴⁵ Some of the applications that have been developed by Open Data Kosovo include Kosovo Municipality Procurement Visualizer, Kosovo Election Monitoring Visualizer, Gender Corruption Survey Visualizer.

¹⁴⁶ Portal of Open Data Kosovo. <http://data.rks-gov.net/en/>

¹⁴⁷ SEEDS (2016), p. 10

the obstacles of conducting secondary analysis and that secondary analysis was not an important part of research culture in Kosovo.

Attitudes towards data sharing

From the results of our self-assessment survey we can say that data sharing is perceived as having some benefits for social scientists and that they might think that data sharing involves little effort and minimal costs. In the SEEDS survey, it was shown that most of the researchers perceived considerable benefits to their research and teaching from better access to data produced in Kosovo and elsewhere, and more that 80% of researchers think that sharing data is very important within their discipline.

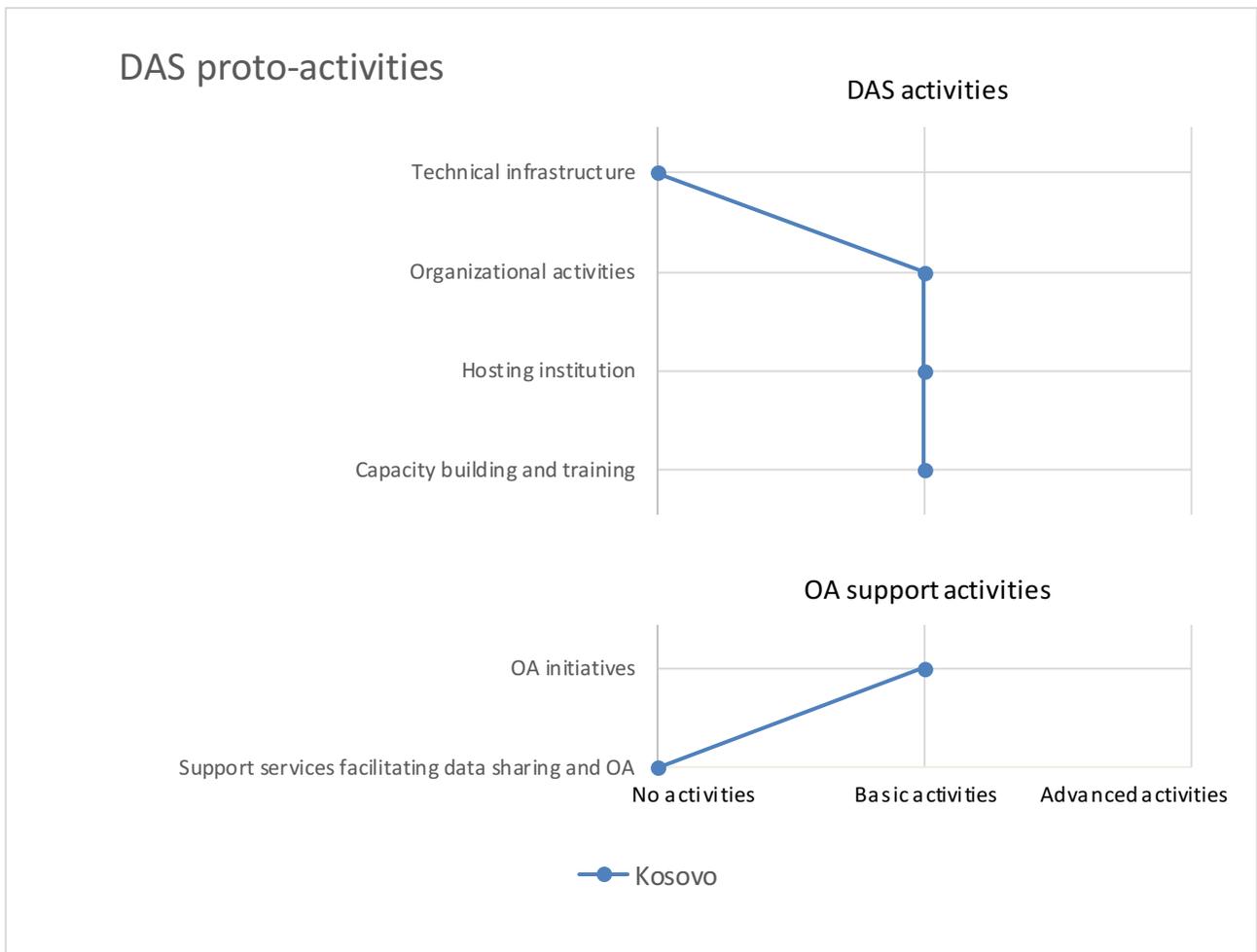
One of the reason for the low proportion of researchers sharing their data could be the lack of appropriate infrastructure for data archiving in the country. The results of the SEEDS Survey showed that the majority of respondents (70%) had a positive attitude about sharing their data in a social science data archive, provided that their data are preserved for the long term in a secured environment, and 58% of respondents answered that the ideal access to their data should be public or available to the broader scientific community.

Enablers for data sharing

According to the self-assessment survey results, there are no career rewards related to data sharing in the academic community of Kosovo. Also, there are no support services provided to social science researchers that facilitate data sharing and/or Open Access to research data, and most of the researchers are not following data management and international data documentation standards and procedures that facilitate data reuse, due to lack of archiving and data sharing culture. The SEEDS Survey showed that some researchers did retain well-documented data with metadata, and that some of them did apply some internal/institutional standard for describing the data.

3.22.2 Data archive service (DAS) proto-activities

Figure 50: DAS and OA activities implementation type in Kosovo



DAS activities

There is currently no established data service for the social sciences or for other disciplines in Kosovo. No specific technical infrastructure exists that could be used for or applied to a new data services in social sciences or in other scientific disciplines.

Some notable efforts towards establishment of data services for social sciences have been made during the SEEDS project¹⁴⁸ coordinated by FORS - Swiss Centre of Expertise in the Social Sciences. Partner from Kosovo in this project is the Centre for Political Courage in Prishtina. The SEEDS project tackled most of the issues related to the process of establishment of data services for social sciences such as policy and governance, human resources and technical issues. It also helped to promote the importance and benefits of data sharing among relevant stakeholders, researchers and policy makers.

¹⁴⁸ SEEDS - South-Eastern European Data Services. <http://www.seedsproject.ch>

A suitable institution that could host the future data service could be the University of Prishtina, but capacities of human resources in hosting data services would have to be built from the beginning.

Through the SEEDS project, CPC staff did gain knowledge, skills and experience, which should be used to influence future developments in the area of social science data management and preservation. CPC was also involved in the RRPP Data Rescue project¹⁴⁹ during which the descriptions and the data from all Regional Research Promotion Programme (RRPP) projects were combined on one single platform called SEEDSbase¹⁵⁰ hosted at FORS - Swiss Centre of Expertise in the Social Sciences.

In the area of technical infrastructures, which are currently lacking in Kosovo, cooperation could be established with relevant local institutions such as the Agency of Information Society (AIS) responsible for developing, maintaining and preserving the IT infrastructure and administrative data of public institutions¹⁵¹. Also, as it was demonstrated with the RRPP Data Rescue project, the technical infrastructures and expertise could be available under some conditions through international cooperation with CESSDA members' service providers.

Open Access (OA) support activities

Initiatives and projects promoting open access to scientific information are rare in Kosovo. Open access to research data was promoted during the SEEDS project. There is also an NGO "Open Data Kosovo", which was already mentioned in the chapter about RDM policy setting, which promotes Open Access principles with primary focus on open government data.

As mentioned above, there are no support services provided to social science researchers that facilitate data sharing and/or Open Access to research data.

3.22.3 Conclusions

- Social sciences sector in Kosovo is still in development. No comparative indicators are available for the development of SSH in Kosovo but the willingness of the government of Kosovo to increase investment in science in general sounds promising. Also promising is the fact that in the year 2016 Kosovo recorded the growth of GDP. That might have an influence also on development of research data production in Kosovo that is currently estimated as underdeveloped.
- Funders' data management and sharing strategy and policies are currently non-existent in Kosovo. Through the SEEDS project awareness have been raised about specific issues related to RDM policies, so we can say that the strategy and policy to enable data sharing has started to develop. Ethical and legal framework important for data sharing is still underdeveloped.
- Data sharing culture is developing in Kosovo, but enablers of data sharing are completely underdeveloped.

¹⁴⁹ RRPP Data Rescue project. http://seedsproject.ch/?page_id=618

¹⁵⁰ SEEDSbase. <https://seedsdata.unil.ch/>

¹⁵¹ SEEDS (2016). *D3 - Report on evaluation of research and legal conditions*.

http://seedsproject.ch/wp-content/uploads/2015/06/Report_Kosovo_SEEDS_final_rk_FINAL.pdf, page 37.

- DAS proto-activities exist in Kosovo through the involvement in the SEEDS project. Open access initiatives are rare and not coming from the government.

3.23 Latvia

Development level of SSH in Latvia can be relatively seen as average among all ERA countries audited (except for research data production in social sciences that is evaluated as more developed than on average). RDM policy setting in Latvia is slowly emerging, and data sharing culture and enablers for data sharing within social science research community in Latvia are slowly developing as well. Implementation of Data Archive Service (DAS) and Open access (OA) support activities, however, is still basic.

There is currently no social sciences data archive (DAS) in Latvia. The strengths and potentials for DAS development in Latvia lie in positive attitudes of researchers towards data sharing; previous experience of a data archive; positive plans regarding policy setting for RDM. Finally, there is a variety of initiatives facilitating and/or providing access to social sciences research data for reuse: there are university repositories; OA repository Academia in National Library of Latvia; a database aiming at creating overview of research projects commissioned by state and municipal institutions maintained by Cross Sectoral Coordination Centre; plans to develop another more comprehensive scientific information system by Ministry of Education and Science; as well as Open data initiative supported by the Ministry of Environmental Protection and Regional Development in accordance with e-government policy guidelines. However, these top-down attempts do not always consider the specifics of discovery and reuse of, and access to social sciences research data. Given the technical infrastructure and involvement with social sciences research data, currently Cross-Sectoral Coordination Centre and OA repository Academia in National Library of Latvia could be considered as possible partners for developing DAS.

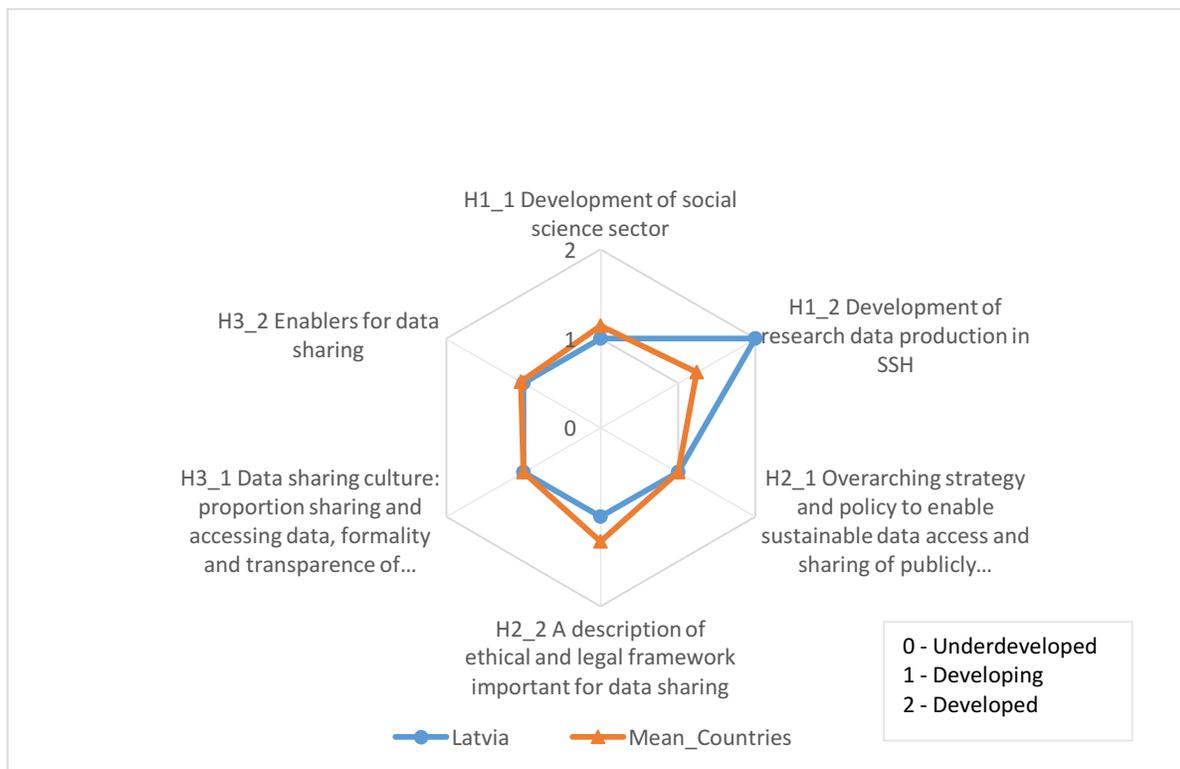
Lack of clear and explicit declaration of RDM policy principles and funding that comes with it, as well as lack of support of different aspects of RDM are among barriers and conditions limiting DAS development in Latvia, together with the fact that CESSDA is not identified among European RI's in the Ministry of Education and Science shortlist in Guidelines for Science, Technology Development and Innovation for 2014-2020.

Specifics about the data collection in a country

Data was collected mainly by contact person filling in the self-assessment survey, filled in by Baiba Bela, Associated professor of University of Latvia, Chief of Board of Latvian Sociological Association. Additional information on Latvian Social Sciences Data Archive (LSZDA) was provided by Ausma Tabuna and Professor Aivars Tabuns in e-mail correspondence. Desk research, where possible and appropriate, was used as well. Report was complemented by information acquired as response to follow-up questions (via e-mail correspondence) to Ministry of Education and Science, and OA repository Academia in National Library of Latvia.

3.23.1 Broader ecosystem of DAS operation

Figure 51: Heading concepts values in Latvia



Social science sector and ethical and legal framework for data sharing in Latvia is, according to the results of audit, slightly less developed than average, while development of research data production in SSHG seems to be above the average, the rest of dimensions – strategies and policies to enable sustainable data access and sharing, data sharing culture and enablers for data sharing – being on the same level.

Development of social science sector

As part of the introduction to results section, basic profile of the country regarding scientific development is described. Data for this section is collected centrally.

Overall assessment of SSH development

General intensity of investment in SSH-sector in Latvia is slightly below medium, if compared to other countries included in the audit – 0.1% (GERD in SSH as percentage of GDP). Intensity of investment in SSH (relative to other scientific disciplines) is rather low - 14%, but slightly higher than average. Human resource potential in SSH can be estimated as comparatively high – 103 researchers per 100.000 inhabitants (Number of researchers (head count) in SSH per 100.000 inhabitants), while investment in human resource in SSH is below the average – 11092.

Most important sources of funding in social sciences, based on self-assessment, are ranked as follows: (1) international and cross-border funding; (2) government and higher education sector; (3) private non-profit sector, and, finally, (4) business enterprise sector. Experts consider the gap

between national and international funding being considerable, and the existing funding model unsustainable, as public national level funding for SSH is scarce (SSH are not prioritized receiver of public research funding), but international - unpredictable and with growing competition rates. It should also be noted that Latvia ranked 27 in 2012 with regards to public funding for research among EU member states, and in 2012-2013 only 10% of national funding was awarded via institutions, while 90% was awarded on competitive basis¹⁵², that might create barriers for continuity of institutional affiliation.

Researchers in academic sector have publicly funded, free of charge access to commercial databases with scientific papers (i.e. EBSCO, JSTORE) and software, purchased via universities.

Overall assessment of development of social science sector is medium (funding of SSH and productivity of the researchers are in the mid quantile; impact on designated community is limited).

Development of research data production in SSH

There are well established traditions of international collaborative research or cross-national studies - Latvia has participated in European Social Survey (ESS) – 2 rounds, 2006&8; International Social Survey Program (ISSP); Comparative Study of Electoral Systems (CSES); European Values Study (EVS) – 3 waves, World Values Survey (WVS) – one wave 1996; Programme for International Student Assessment (PISA) – 5 waves, starting from 2000, – Latvia has at some point participated or is still participating in six out of eight international studies enlisted.

According to self-assessment, the average production of research data in the SSH institutions in Latvia can be characterized as frequent - institutions have well established tradition in data production (at least, one large scale nationally representative survey is usually conducted by university institutes, while market research and public opinion research companies produce data frequently). There are also several studies of national importance, for example, biannual surveys for Human Development Report, produced by University of Latvia¹⁵³.

Overall, research data production in Latvia can be regarded as developed, as there are well-established traditions of producing both national and international research data.

RDM Policy setting

Overall incentives and high-level policy requirement are not yet on the level that would provide acceptable level of sustainable conditions for data service provision. However, there are some positive developments to be expected.

Currently Ministry of Education and Science in national funding calls does not require a DMP. There is, however, growing recognition and awareness of need to require DMP, and it is expected

¹⁵² European Commission 2015. European Research Area. Facts and Figures 2014. LATVIA, pp 396; 399. Available at: https://ec.europa.eu/research/era/pdf/era_progress_report2014/country_fiches/era-lv.pdf (last accessed 2017-02-02)

¹⁵³ Data is not available at the moment, but for reports see homepage of University of Latvia, Faculty of Social sciences, available at: <http://www.szf.lu.lv/eng/research/publications-and-current-research-projects/regional-and-human-development-in-latvia/>.

that DMPs and Open Access to research data will be included as request in next period of national research project calls in Latvia. However, funding authority's understanding of the meaning of Open Access to data and DMP could be an issue.

One of the central laws regarding research in Latvia, the Law on Scientific Activity, specifies that information on scientific research financed by the state or municipality institutions should be made publicly available, and access to the research outcomes (if access to this information is not limited by the law) should be provided by the institution that commissioned the research¹⁵⁴.

Right now, the public research funding organizations have no recommendations for publishing research data with associated metadata. There is, however, growing recognition and awareness of the need to store and provide access to data. Thus, according to the previously mentioned Law on Scientific Activity, Ministry of Education and Science, considers beneficiaries of public scientific funding as responsible to provide access to research results. That in practice means university library repositories are expected to store and provide access also to SSH data, so no support to SSH-repository or discipline- specific support services can be expected.

Awareness of need for sustainable long-term curation of research data within public research funding organizations operating in social sciences is estimated as very initial.

Cost for managing the data and preparing it for access are resourced adequately during research project lifetime are not financed by public research funding organizations in Latvia. There are plans to do so in the next funding periods, but it is not clear yet even for research community, how these costs should be estimated.

Thus, overall RDM policy setting in Latvia can be defined as emerging, as there is growing awareness as well as intentions to formulate policy principles supporting data sharing.

Ethical and legal framework important for data sharing in SSH in Latvia can be characterized as developing. The awareness of these issues among professionals working in the field is growing. In case of questions, State Data Protection Inspection can be consulted. Support regarding legal aspects in research should be provided in relation to next period national research project calls.

Data sharing culture

Data sharing and reuse among social sciences researchers in Latvia is estimated as rare – based on self-assessment survey, the proportion of researchers sharing data is estimated as low (0-10%). The proportion of researchers able to access existing third party data they need is estimated in self-assessment survey as low (0-10%) as well.

As self-assessment shows, data in Latvia is mostly shared via channels that lack formality and transparency: via informal contacts (peers and colleagues) - ranked 1st, and via project or personal websites - ranked 2nd. The formal and transparent channel, data archive or repository coming 3rd. As expert comments, personal contacts dominate, and, even though some data is shared via university repositories, there is the issue of discoverability as data in a concrete example mentioned by expert could only be found, if searched within the repository (but not in any other way, even though it should be indexed in Google).

¹⁵⁴ Gundars Kulikovskis, Diana Petraityte, Blagoy Stamenov; RIO Country Report 2015: Latvia; EUR 27867 EN; doi:10.2791/917332.

The attitudes of researchers towards data sharing, as self-assessment shows, can be characterized as rather positive, though two statements related to competition could not be assessed due to lack of experience and data. There is a positive attitude toward benefits of data sharing and advancement of science, and perceived risk of misuse and misinterpretation of data is low. According to the expert, there is a common understanding both by Ministry of Education and Science officials and within social sciences research community that minimal level of data sharing ('just drop data at repository') means little effort and minimal costs. This might indicate lack of commitment to, and experience or knowledge of the resources needed for preparing data for sharing, that might in long-term turn to be a challenge to discoverability and reusability of social science research data.

Attitudes towards data sharing

Table 19: Attitudes towards data sharing in Latvia

Data sharing has no benefits at all	False
Reuse of existing data can answer new research questions and facilitate advancement of science.	True
Data sharing has as a risk that others may misuse and misinterpret data	False
Data sharing involves little effort and minimal costs	Probably false
Data sharing creates healthy competition in research	Estimate not available for 2011-2016
Data sharing creates negative competition (for example, being scooped and therefore reduced publication opportunities) for the researcher	Estimate not available for 2011-2016

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from "5"-True to "1" False!

To summarize, data sharing and reuse is not common at all, there are mainly informal channels for data sharing, but general attitudes of researchers related to sharing of data are rather positive, so in summary data sharing culture can be characterized as developing, same as average level in audited ERA countries.

Enablers for data sharing

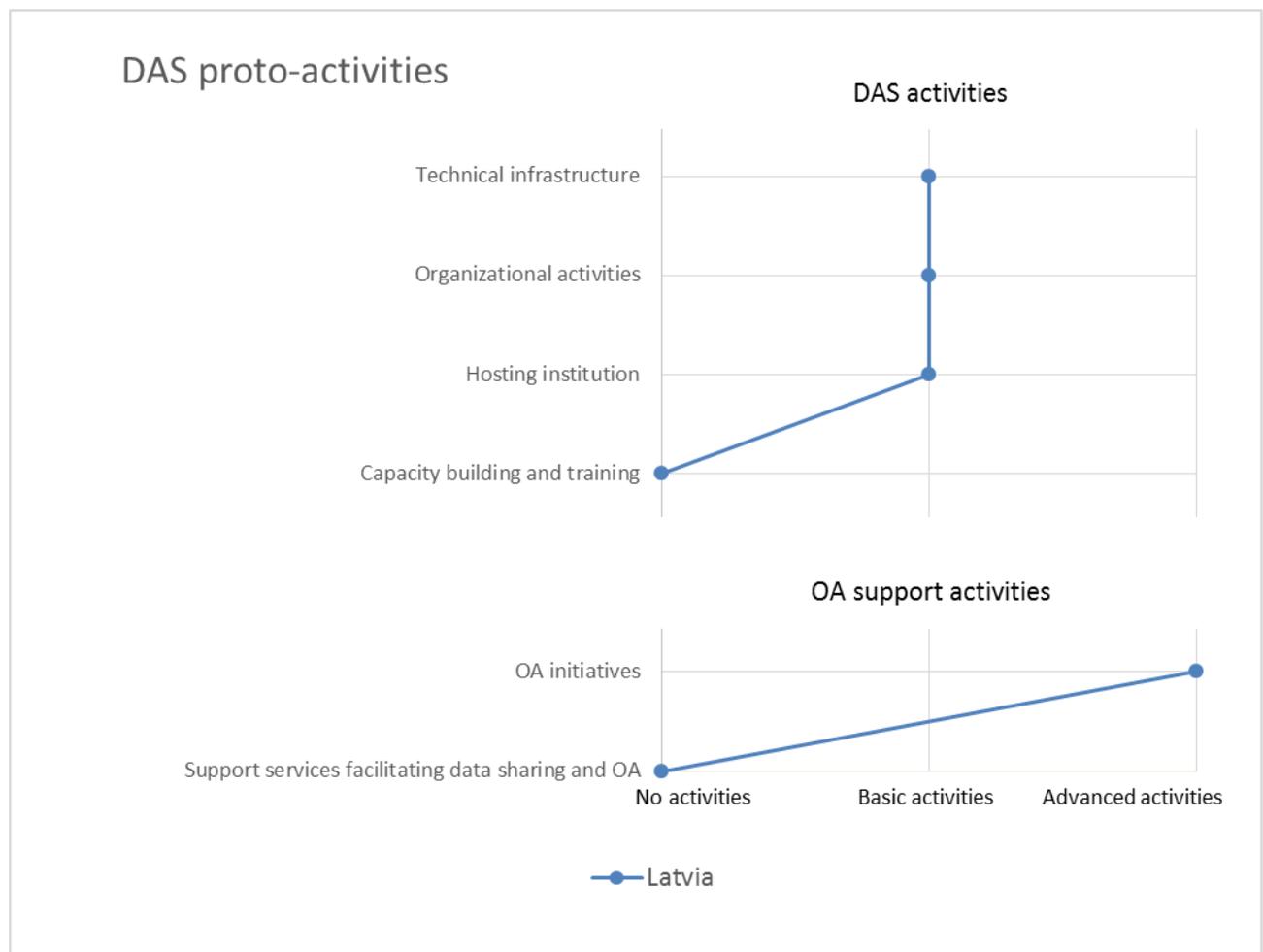
There are, as self-assessment survey indicates, no career rewards related to data sharing the social sciences academic community in Latvia, and neither data support services available to social science researchers that facilitate data sharing and/or Open Access to research data.

There is, however, an OpenAire Node¹⁵⁵ hosted by University of Latvia Library, and some basic information is available on the webpages. It has been organizing events related to Open Access, Open Data and Open Science, although not with social science research data focus. In cooperation with University of Latvia, an e-course "Open Science" based on courses available as part of FOSTER¹⁵⁶ is developed to be used in the Latvian context¹⁵⁷.

According to self-assessment survey, some data producers, however, follow data management and data documentation standards and procedures. In summary, the incentives and enablers for data sharing within social science research community in Latvia are developing.

3.23.2 Data archive service (DAS) proto-activities

Figure 52: DAS and OA activities implementation type in Latvia



¹⁵⁵ More information available (Latvian only) at: <http://www.napd.lu.lv/> (last accessed 2017-02-22).

¹⁵⁶ See: <https://www.fosteropenscience.eu/courses> (last accessed 2017-02-22)

¹⁵⁷ It can be, however, observed that main focus in OA related documents and guidelines, and support activities focus more on article publications rather than research data, and research data is more seen as belonging to Open Data together with big data and publicly available municipal and governmental information systems.

DAS activities

The overall assessment of Data Archive Service (DAS) activities implementation in Latvia is basic.

Latvia is one of the few members of the European Strategy Forum on Research Infrastructures (ESFRI) that has approved national research infrastructure (RI) roadmap rather late: an informative report accepted in spring of 2016, indicates that during 2016, Latvia was participating in following research infrastructures: BBMRI ERIC, CLARIN, ESS ERIC, JIV ERIC, EATRIS ERIC¹⁵⁸ (ESSneutron and EU-OPENSREEN, included in previous plans¹⁵⁹, were not in the list). Guidelines for Science, Technology Development and Innovation for 2014-2020 in Latvia, accepted in 2013, refer also to involvement or other cooperation with ELIXIR, INSTRUCT and PRACE. As indicated in the Guidelines, CESSDA is not in the list of strategically important ERICs identified by the researchers' communities in Latvia¹⁶⁰. The shortlist was prepared based on consultations with representatives of scientific disciplines coordinated by Latvian Academy of Sciences¹⁶¹. Next wave of follow-up and discussions regarding results and needs of scientific community of Latvia to participate in research infrastructures is expected in 2018.

Latvia is participating in European Cloud Initiative¹⁶², the leading institution being Ministry of Environmental Protection and Regional Development, while Ministry of Education and Science is involved in activities related to European Open Science Cloud and research infrastructures.

At the moment of the audit, there are no existing national and/or institutional technical infrastructures in social sciences in Latvia¹⁶³, and no active data archive service (DAS). There have, however, been similar initiatives before. The Latvian Social Sciences Data Archive (LSSDA), founded in 1996 and hosted by the Institute of Philosophy and Sociology of Latvian Academy of Sciences, was initially started to provide access to national ISSP data to a wider scientific community and even general public, with a strong focus on online analysis¹⁶⁴.

¹⁵⁸ Documents (in Latvian only) available at: <http://tap.mk.gov.lv/lv/mk/tap/?pid=40385056&mode=mk&date=2016-04-05> and <http://tap.mk.gov.lv/mk/mksedes/saraksts/protokols/?protokols=2016-04-05> (last accessed 2017-02-22)

¹⁵⁹ European Commission 2015. European Research Area. Facts and Figures 2014. LATVIA, pp 396; 399. Available at: https://ec.europa.eu/research/era/pdf/era_progress_report2014/country_fiches/era-lv.pdf (last accessed 2017-02-02) Gundars Kulikovskis, Diana Petraityte, Blagoy Stamenov; RIO Country Report 2015: Latvia; EUR 27867 EN; doi:10.2791/917332

¹⁶⁰ Guidelines for Science, Technology Development and Innovation for 2014-2020, pp 85-88. Full text in Latvian, available at: <http://polsis.mk.gov.lv/documents/4608> (last accessed 2017-02-09)

¹⁶¹ Report on Capacities and requirements of scientific institutions in Latvia for participation in ESFRI list of research infrastructures [Latvijas zinātnisko institūciju kapacitāte un nepieciešamība dalībai ESFRI Eiropas Pētniecības infrastruktūrās. Situācijas novērtējums un rekomendācijas] is available (in Latvian only) at: http://www.izm.gov.lv/images/statistika/zinatne/ESFRI_petijums_LZA.pdf (last accessed 2017-02-22)

¹⁶² The European Cloud Initiative, available at: <https://ec.europa.eu/digital-single-market/en/%20european-cloud-initiative> (last accessed 2017-02-15)

¹⁶³ List of Research Infrastructures in the homepage of Latvian Council of Science (institution responsible for developing national RI roadmap) show that social sciences and earth & environmental sciences in Latvia do not have any RIs, available at: <http://www.lzp.gov.lv/ri/> (last accessed 2017-02-09).

¹⁶⁴ It was meant as an alternative to statistical data analysis packages requiring considerable skill and licenses to use them (e-mail communication with former LSSDA representatives, Professor Aivars Tabuns and Ausma Tabuna, 2017-02-05).

Through participation in the East European Data Archive Network and via cooperation with the Swedish National Data service, LSSDA managed to acquire necessary skills and support to make accessible data from several national and international survey data series in Latvian with funding from several national applied social sciences and international research projects¹⁶⁵. However, due to unsustainable funding, human resources policy and changes of institutional affiliation as a result of major transformations of social sciences landscape in Latvia around 2006-2009, after several unsuccessful attempts to include LSSDA activities in national and international research and infrastructure project proposals, the archive is not active anymore after 2009¹⁶⁶. In the final years, LSSDA web catalogue included national ISSP data; data from several other national surveys; information on data collected in social science projects funded by Latvian Council of Science, and a collection of not yet published social sciences data accumulated during years that might be lost now¹⁶⁷.

As a spin-off to LSSDA, a database of social sciences research data and metadata with potential for reuse and relevant to regional development from different national and regional universities in Latvia was developed¹⁶⁸ as a part of project "Promoting the utilization of the research potential of local universities for regional development in Latvia" active in 2009-2011, funded by EEA Norway grants.

There are also more recent initiatives, most of which are related to developments of Open Science, Open Access and Open Data, aimed at better administration of research processes and outputs, including research data by different stakeholders - Ministry of Education and Science, Ministry of Culture and National Library of Latvia, Cross Sectoral Coordination Centre, Ministry of Environmental Protection and Regional Development, as well as single university repositories.

The Law on Scientific Activity specifies that access to outcomes of publicly funded scientific research should be, as long as allowed by the law, provided by the institution having commissioned the research. Following this Law, the Cross-Sectoral Coordination Centre (National development planning institution associated with the Cabinet of Ministers) maintains a database of results of research projects commissioned by governmental and municipal institutions¹⁶⁹. The database currently includes available descriptions of finished and planned research projects as well as reports and publications, but no data and/or metadata¹⁷⁰.

According to recent changes in Law on Scientific Activity, the Ministry of Education and Science is responsible for developing a National Scientific Information System that would integrate existing registers of scientific information, and have the aim of providing information needed for planning of scientific activities; reducing administrative resources needed for administration of research as well as contributing to transparency on public spending on research¹⁷¹. This System is expected to

¹⁶⁵ Tabuna, Ausma. (2002). Latvian social science Data archive. In Hausstein Birgitte and Paul de Guchteneire (ed.) Social Science Data archives in Eastern Europe, Bergisch Gladbach: E.Ferger Verlag, pp. 181-184.

¹⁶⁶ E-mail communication with LSSDA representatives, Professor Aivars Tabuns and Ausma Tabuna, 2017-02-04, 2017-02-05 and 2017-02-06.

¹⁶⁷ See above.

¹⁶⁸ <http://www.petnieciba.lv/> was available 2010-2013, hosted by University of Latvia, Faculty of Social Sciences.

¹⁶⁹ Gundars Kulikovskis, Diana Petraityte, Blagoy Stamenov; RIO Country Report 2015: Latvia; EUR 27867 EN; doi:10.2791/917332, pp 66.

¹⁷⁰ Database of research projects and publications [Pētījumu un publikāciju datu bāze], available in Latvian here: <http://petijumi.mk.gov.lv/> (last accessed 2017-02-15).

¹⁷¹ Changes in Law on Scientific Activity, December 8, 2016. Available online in Latvian, <https://www.vestnesis.lv/op/2016/249.6> (last accessed 2017-02-09).

cover information on outputs of scientific activity – publications, patents and other IP units, conferences and workshops (research data are not included), and it is not expected that the coverage of types of scientific outputs might expand. The Ministry of Education and Science is also developing Academic network (ERDF funded project “Establishment of unified national importance academic network in Latvia for provision of research activities”) that, among other things, aims to provide or facilitate the development of “virtual laboratories, the availability of digital libraries, online discussions and conferences, research integration in the higher education system, integrated learning solutions, interoperable and integrated administrative systems of research, academic institutions, and universities”¹⁷².

The Ministry of Culture of the Republic of Latvia has in 2015 approved the Strategy for Digital Cultural Heritage as part of Guidelines for Cultural Policy 2014-2020¹⁷³. One of the main activities described in these guidelines is the development of infrastructure for any kind of digital content, so focus lies on libraries, museums and other cultural heritage institutions. Scientific publications and research data, their long-term preservation and online availability are, however, also included in the strategy under Activity 4.2. Facilitation of use of Digital cultural heritage. National Library of Latvia has been assigned the task and resources to develop and maintain Open Data repository for scientific information¹⁷⁴. It is expected that implementation of these guidelines will facilitate the preservation of scientific information, including research data¹⁷⁵.

National Library of Latvia has developed an Open Access academic repository, ACADEMIA (<https://academia.lndb.lv/>), opened in December 2016. ACADEMIA aims at publishing scientific publications, data and other material produced by researchers in higher education and research institutions in Latvia, including research data. Right now, however, repository does not include any research data (only reports from research projects). The contents of the repository at the time of the audit was made up mainly from PhD dissertations, other student papers and publications from two higher education and research institutions, University of Latvia and Latvia Academy of Culture. There are, however, different new developments to expect including the research data, depending on the discussions with universities and research communities. It is planned that ACADEMIA will assign DOI to all digital objects from 2018. Considering the prospective developments and expressed interest, the repository should be considered as an optimal solution for prospective partner for establishing DAS.

University repositories are starting to develop, and researchers are starting to use them for storing and sharing data. Even though these repositories theoretically accept different types of materials, including research data, and materials are indexed, expert mentions that for researchers it can, however, be difficult to find and access data in university repositories, as they are now¹⁷⁶.

¹⁷² Gundars Kulikovskis, Diana Petraityte, Blagoy Stamenov; RIO Country Report 2015: Latvia; EUR 27867 EN; doi:10.2791/917332, pp 65.

¹⁷³ Kultūrpolitikas pamatnostādnes 2014.–2020. gadam „Radošā Latvija”. Digitālā kultūras mantojuma attīstības stratēģija. Available in Latvian via homepage of Ministry of Culture, http://www.km.gov.lv/lv/dokumenti/planosanas_doc.html or directly available a https://webcache.googleusercontent.com/search?q=cache:A90SnibDiNcJ:www.km.gov.lv/lv/doc/dokumenti/nozaru_s_trategijas/Strategija_digitalais_mantojums_final.pdf+&cd=1&hl=en&ct=clnk&gl=se (last accessed 2017-02-15)

¹⁷⁴ Kultūrpolitikas pamatnostādnes 2014.–2020. gadam „Radošā Latvija”. Digitālā kultūras mantojuma attīstības stratēģija, p. 23

¹⁷⁵ Gundars Kulikovskis, Diana Petraityte, Blagoy Stamenov; RIO Country Report 2015: Latvia; EUR 27867 EN; doi:10.2791/917332, pp 67.

¹⁷⁶ Information about E-resource repository of the University of Latvia, can be found here, available at: <http://www.biblioteka.lu.lv/home/e-resources/ulrepository/> (last accessed 2017-02-17). It is built in

University repositories have also been among the most active in disseminating information Open Access principles and creating and shares knowledge about Open Access in the academic and scientific environment¹⁷⁷. However, the universities in Latvia are facing decline in funding because of consequences of declining student population related to demographics.

Open Data initiative, coordinated and offered by Ministry of Environmental Protection and Regional Development, as part of e-government policy guidelines, has focus on data from public registers and publicly available parts of governmental and municipal information systems, aggregated statistics and also to some extent, research data (mainly focusing on big data, not social sciences)¹⁷⁸.

There have also been some attempts in social sciences research community to facilitate user-friendly solutions for access to research data for reuse. Association of Latvian sociologists (LSA) started under 2015-2016 discussion on accessibility of publicly funded social science research data for reuse with several stakeholders, including of Ministry of Education and Science, but, as the expert indicate, there seems to be, by Ministry of Education and Science as well as other institutions, support to generalist solutions described above instead of discipline specific, that would be more suitable to the needs social science research community has.

Currently Cross-Sectoral Coordination Centre has expressed some preliminary interest in adapting the database for making datasets as available together with scientific reports, so it should be considered as a possible partner for DAS development. Another possible partner could be the National Library of Latvia Open Access academic repository ACADEMIA established in 2016, that has databases among types of possibly publishable scientific outputs (even though majority of material published until now consists of reports, thesis and articles).

Open access (OA) support activities

Currently, there are many OA activities in Latvia as we have seen in previous sections. However, the developments in OA to research data has until now been rather general and not taking into specific needs of social sciences and researchers' perspective. OA activities in general, as can be judged from activities and support available in OpenAire NOAD in University of Latvia, policy documents, university repositories - seem to have focus more on scientific articles and working papers than data.

As mentioned above, there is no data support services available to social science researchers to facilitate data sharing and/or Open Access to research data.

DSpace and allows to deposit, store and access also research data. Discoverability of the resources, however, might be an issue, as it is difficult to find research data, if precise title of the project or resource is not known.

¹⁷⁷ Gundars Kulikovskis, Diana Petraityte, Blagoy Stamenov; RIO Country Report 2015: Latvia; EUR 27867 EN; doi:10.2791/917332.

¹⁷⁸ More information on Open data in Latvia and examples of data included in the Open Data initiative, (in Latvian) can be found here, http://www.varam.gov.lv/lat/darbibas_veidi/e_parv/atvertie_dati/?doc=20449 (last accessed 2017-02-15)

3.23.3 Conclusions

Maturity of dimensions of broader ecosystem of DAS operations in Latvia can be characterized by the same level or slightly below than average (except for development of data production in SSH). Policy setting for RDM is developing slowly – no clear requests, obligations or procedures related to sharing research data are formulated in policy documents yet (except for statement that results of publicly funded research have to be made available for general public). Clearer requests related to data sharing (DMP, obligation to share data after the project is finished etc.) can be expected in near future. General attitudes towards data sharing in research community seem to be positive, even though there are no enablers in terms of career incentives, and not much support available regarding legal, technical and other issues related to sharing data. The general direction towards data sharing requirements and DMP requirements in national public funding calls could also be considered a positive development. Lack of understanding of work involved in preparing data for sharing can lead to underestimation of resources and frustration. Support with regards to legal, ethical as well as other issues should be provided.

Thus, overall assessment of Data Archive Service (DAS) activities implementation is basic, and so is implementation of Open access (OA) support activities. Positive attitude of research community is manifested as research community-driven discussions related to access to publicly funded research data for reuse, initiated by Latvian Sociological association (LSA). Together with the top-down activities within the framework of science and innovation policy, aimed at creating an effective system of administering research processes (Guidelines for Science, Technology Development and Innovation for 2014-2020) and culture policy, aimed at ensuring preservation and dissemination of digital cultural heritage, including research data (Strategy for Digital Cultural Heritage, part of Guidelines for Cultural Policy 2014-2020), these indicate growing awareness of importance of data sharing in different stakeholder groups in Latvia.

It seems, however, reasonably promising to assume that with some integration of discipline-specific approach in documentation and dissemination of SSH data, the existing initiatives and systems mentioned in the report could be partially adapted for the purpose and meet the needs of the research community for access to high quality social sciences research data with high potential for reuse. Integration of social sciences research community perspective – the need to share and be able to find and access for reuse high quality data - with already ongoing initiatives in Cross-Sectoral Coordination Centre, National Library of Latvia OA repository ACADEMIA, National Scientific Information System as well as university repositories should be explored while identifying a possible partner for DAS development.

3.24 Lithuania

Broader ecosystem of DAS operations in Lithuania is on an average maturity level with developed social science sector and self-reported data sharing culture. The rest of dimensions are on “developing” level. In 2016, Research Council of Lithuania adopted Guidelines for Open Access to Research Results.

Lithuanian Data archive for Humanities and Social Sciences (LiDA) was founded in 2006 and has been member of CESSDA since 2012. Development of the organisational infrastructure of LiDA shows a mature DAS with insufficient funding and hence not fully operational. Self-assessed digital object management and technical infrastructure in general are above the average. At the moment of audit, LiDA is performing only basic functions to provide access to existing data. There are no new decisions about future funding, but it is expected that it might be allocated from the European Structural funds allocated to Lithuania. Still amount of this funding might be very limited.

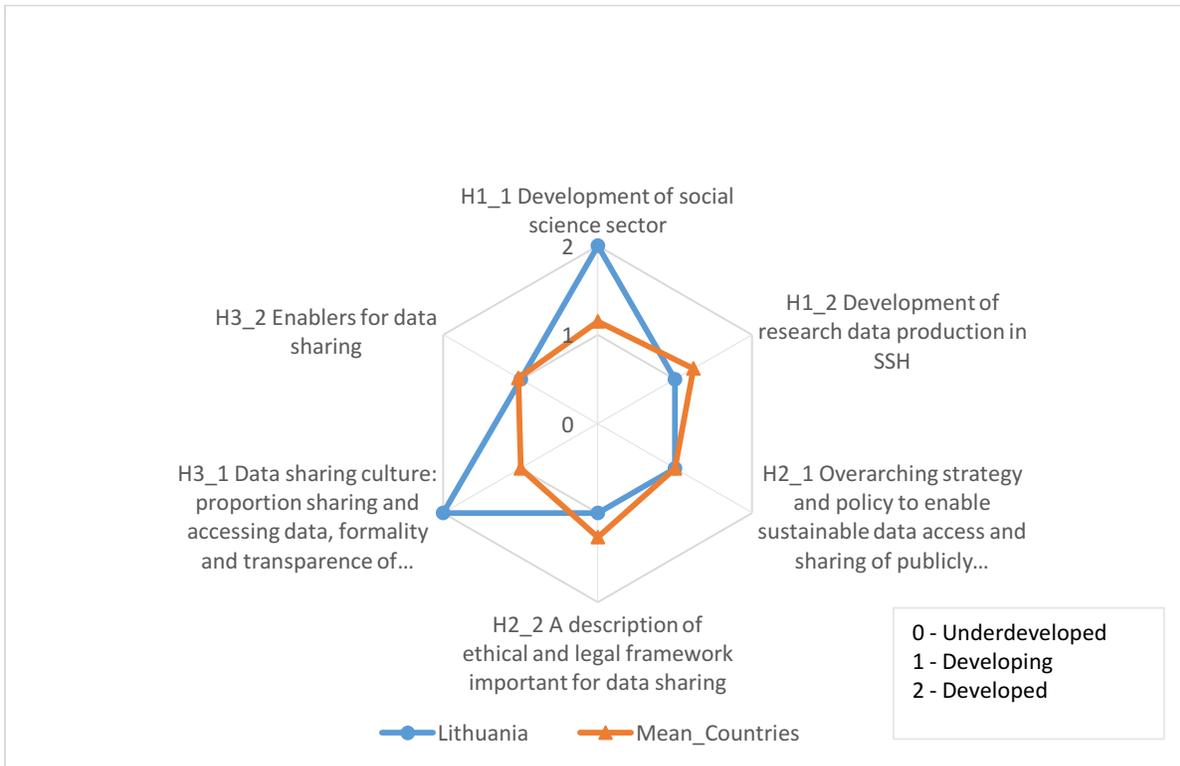
Specifics about the data collection in a country

Data was collected mainly via self-assessment survey, information was provided by Algis Krupavičius from LiDA, and supported by documents and LiDA’s webpage, where indicated. Some indicators (see introduction) were collected centrally.

3.24.1 Broader ecosystem of DAS operation

While two of broader DAS ecosystem dimensions – development of social science sector and data sharing culture - are developed on the highest level (and above average in the audited ERA countries), two other –development of research data production and ethical and legal framework important for data sharing are developing (and below the average in the audited ERA countries). RDM policy settings dimension in Lithuania is developing and on the same level as average in the audited countries.

Figure 53: Heading concepts values in Lithuania



Development of social science sector

In this section, basic profile regarding scientific development in Lithuania is described thus giving insight on general conditions influencing the DAS operation.

Overall assessment of SSH development

General intensity of investment in SSH-sector in Lithuania is above medial value, if compared to other ERA countries audited – 0.21% (GERD in SSH as percentage of GDP) – only five other countries - UK, Norway, Netherlands, Portugal and Iceland has a higher proportion. Intensity of investment in SSH (relative to other scientific disciplines) is also above average – 20.5% -only five other countries invest proportionally more in SSH.

Human resource potential in SSH in Lithuania can be estimated as high – with 247 researchers per 100.000 inhabitants it comes 1st among the audited countries that have data on this indicator (I1.1.3 Number of researchers (head count) in SSH per 100.000 inhabitants). Investment in human resources in SSH is, however, rather low – 9966, below median value among audited ERA countries with data available.

Most important sources of funding in social sciences and humanities, based on self-assessment by a DAS expert, are ranked as follows: (1) government and higher education sector; (2) international and cross-border funding; (3) business enterprise sector, and, finally, (4) private non-profit sector.

Researchers have publicly financed access to bibliographic and full-text databases, datasets and software licences.

Altogether overall development of social science sector in Lithuania is on a high level.

Development of research data production in SSH

There are well-established traditions of international collaborative research or cross-national studies in Lithuania. The country has participated in European Social Survey (ESS) rounds 4 to 7 (2008-2014); International Social Survey Program (ISSP); Comparative Study of Electoral Systems (CSES) Module 1 (1996-2001) and Module 4 (2011-2016); European Values Study (EVS) waves 1990, 1999 and 2008; Programme for International Student Assessment (PISA) –2015, 2012, 2009, 2006: and Generations and gender programme (GGP) Waves 1 and 2 – Lithuania has at some point participated or is still participating in six out of eight international studies enlisted.

In addition to participation in above-mentioned international studies, social scientists in Lithuania produce a range of systematic studies of national importance. An example included the web catalogue of Lithuanian Data Archive (LiDA)¹⁷⁹ is a study Social Policy, rounds I to VI (2010-2015, aiming to examine the attitudes of Lithuanian population towards social policy.); another example is a collection of thematic public opinion studies Social Economic Barometer (2004-2010, several thematic surveys during a year), aimed at investigating the opinions of Lithuanian population regarding the most pressing socio-economic issues.

According to self-assessment, the average production of research data in the SSH institutions in Lithuania can be characterized as periodical - institutions have tradition in producing some type of research data to a certain extent, and in recent years they have been building up a tradition of data production.

So overall, research data production in Lithuania is developing and as such this dimension lies slightly below the average in all audited countries.

RDM Policy setting

Overall incentives and high-level policy requirements in Lithuania have not reached a level that provides appropriate level of sustainable conditions for data service provision. Of the four policy setting requirements audited, all are on “Initial” level.

There is no explicit and clear request for a DMP, but there is a growing recognition that such would be needed.

Recommendations for depositing data in an appropriate disciplinary repository are on “Initial” level as well. According to self-assessment data, a few attempts to recommend a repository were initiated 3-5 years ago by Research Council of Lithuania (RCL), but were discontinued later.

OpenAire 2016 description of OA and the National Research Environment in Lithuania provides a more detailed description of contract agreements with the Research Council Lithuania (RCL), for example, in 2012 the article 82 of contract agreement said: The principal investigator and the implementing institution shall undertake, in the manner and under the terms set forth by the institution, to ensure a proper storage and the use of the data received in the course of the project implementation, and following three years from the end of the project implementation shall ensure an access to such data¹⁸⁰.

¹⁷⁹ LiDAs Nesstar server, available at: <http://www.lidata.eu/webview/index.jsp> (last accessed 2017-02-09).

¹⁸⁰ OpenAire 2016. OA in Lithuania. The National Research Environment, <https://www.openaire.eu/en/support/guides/88-rep-man-toolkit-usage-stats> (last accessed 2017-02-09)

There is growing awareness about the need for research data for long-term curation to be carefully evaluated and selected, in terms of reuse potential. Managing of the data and preparing it for access can be implicitly covered in the overall research project budget. There is a discussion to have some additional budgeting for data management, but no decisions are taken yet.

The Law on Higher Education and Research in Lithuania does imply that result of all publicly funded research must be publicly available (unless limited by intellectual property or other requirements). The implementation is evaluated as problematic, though, because of lack of support on institutional level as well as fragmented technical infrastructure, lack “critical mass” as they include just a fraction of research outputs (publications and data), and even an initiative of 2011 towards more general project ‘National open access archive of research information (MIDAS)’ with the aim to provide infrastructure for preservation and open access to research data had not succeeded by 2015. The target of the Programme for Development of Lithuanian Research and Studies Informational Infrastructure for 2013-2016 was to have at least 10% of collected data should be publicly available free of charge by 2016¹⁸¹.

In the beginning of 2016, Research Council of Lithuania adopted Guidelines for Open Access to Research Results¹⁸², including open research data policy that are fully aligned with the Horizon 2020 OA mandate and Open Research Data Pilot and European Commission recommendations on access to and preservation of scientific information¹⁸³.

Therefore, overall RDM policy setting in Lithuania can be characterized as developing, as there can be seen growing awareness and initial developments towards establishing the key aspects enabling data sharing.

Ethical and legal framework important for data sharing in SSH in Lithuania can be characterized as developing. According to self-assessment, there is growing awareness of importance of clarification and support on legal and ethical aspects of data sharing (IPR, data protection), but no organized support is available, and it is not regulated by any legal framework and the discussions are fragmented.

Data sharing culture

Data sharing among social sciences researchers in Lithuania is estimated in self-assessment as being on medium level (10-30%). The proportion of researchers able to access existing third party data they need – is also estimated medium (10-30%). This could be explained by an initiative of Research Council required to deposit the research data at the Lithuanian Social Science Data Archive for few years. Temporarily it increased data sharing, but this practice was changed to a new rule that each institution individually deciding about access to research data.

There are, as self-assessment shows, established channels and routines for sharing data in Lithuania: most often used data sharing channel is formal and transparent - data archive or

¹⁸¹ Agnė Paliokaitė, Pijus Krūminas, Blagoy Stamenov; RIO Country Report 2015: Lithuania; EUR 27882 EN; doi:10.2791/049440

<https://rio.jrc.ec.europa.eu/en/country-analysis/Lithuania/country-report> (last accessed 2017-02-09)

¹⁸² Guidelines for Open Access to Research Results. In Lithuanian: Atvirojos prieigos prie mokslo publikacijų ir duomenų gairės,

<https://www.e-tar.lt/portal/lt/legalAct/dceeeb10e05711e59cc8b27b54efaf6e> (last accessed 2017-02-09)

¹⁸³ OpenAire 2016. OA in Lithuania. The National Research Environment,

<https://www.openaire.eu/en/support/guides/88-rep-man-toolkit-usage-stats> (last accessed 2017-02-09)

repository is ranked first. Data is shared also via personal contacts (ranked second), that lack formality and transparency.

The attitudes of researchers towards data sharing, based self-assessment, can be characterized as mainly neutral or positive. The social science research community in Lithuania is perceived in self-assessment as acknowledging benefits of data sharing to some extent, for example, advancement of science, and not being worried about effort and costs related to data sharing, but left undecided regarding concerns about competition in science and publication opportunities as well as misinterpretation or misuse of data.

Attitudes towards data sharing

Table 20: Attitudes towards data sharing in Lithuania

Reuse of existing data can answer new research questions and facilitate advancement of science.	True
Data sharing has no benefits at all	False
Data sharing involves little effort and minimal costs	Probably false
Data sharing creates healthy competition in research	Nor true, nor false
Data sharing has as a risk that others may misuse and misinterpret data	Nor true, nor false
Data sharing creates negative competition (for example, being scooped and therefore reduced publication opportunities) for the researcher	Nor true, nor false

Source: self-assessment survey. Question asked: *Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from "5"-True to "1" False!*

To summarize data sharing culture in Lithuania, data sharing and reuse is rather common, there are established, both formal and informal channels for data sharing, and general attitudes of researchers related to sharing of data are neutral or positive, so in summary data sharing culture can be characterized as developed.

Enablers for data sharing

There are no career rewards related to data sharing the social sciences academic community in Lithuania, as self-assessment indicates.

There are data support services available to social science researchers that facilitate data sharing and/or Open Access to research data, provided by LiDA. LiDA was organizing workshops, trainings, using Nesstar for data publishing etc. Due a lack of funding these services are not available at the moment. Self-assessment indicates that there are also data producers that follow data management and data documentation standards and procedures that facilitate data reuse, for example, data sets available via LiDA are documented according to DDI (1.2) metadata standard and study description metadata is documented bilingually, in English and Lithuanian -basically, all

the CESSDA recommended fields of DDI are filled out. Each data set in LiDA has its unique PID which is constructed to reflect the main attributes of the data set¹⁸⁴.

On institutional level, however, it is estimated that less than 20% institutions have internal procedures relating to open access and preservation of scientific information/data¹⁸⁵.

In summary, the incentives and enablers for data sharing within social science research community in the Lithuania can be benchmarked as developing.

3.24.2 Capability requirement areas of DAS

Organisational profile

Organisation

LiDA was established in June 2006 by the Policy and Public Administration Institute at Kaunas University of Technology in partnership with Vilnius University, Institute for Social Research, the Republic of Lithuania Ministry of Education and Science.

In 2008 LiDA joined the European Social Survey (ESS) and in 2010 the International Social Survey Programme (ISSP). In 2011 LiDA was recognized as a national research infrastructure and has since then been included into the Lithuanian Roadmap for Research Infrastructures. In 2012 LiDA became a member of the Council of European Social Science Data Archives (CESSDA). Sociologists and political scientists and students are the most active users.

Funding

During the first phase of development, 2006-2008, LiDA was funded from EU Structural funds. The second phase of the LiDA development, 2009-2012, was funded by the Ministry of Education and Science and by EU Structural funds. Since 2012, LiDA has no permanent funding and it is surviving on extremely small university funding. International experts proposed in June 2016 to transfer LiDA from the previous host university, Kaunas University of Technology, to a new host institution, Vytautas Magnus University. Because of lack of permanent funding LiDA lost most staff after 2012/2013, and at the moment LiDA has only part-time staff. At the moment of the audit, there were no new decisions about future funding, but it is expected that it might be allocated from the European Structural funds allocated to Lithuania. Still amount of this funding might be very limited.

¹⁸⁴ See for example, available at: www.lidata.eu/data/quant/LIDA_ESS_0240

¹⁸⁵ Tautkevičienė, 2011 in Agnė Paliokaitė, Pijus Krūminas, Blagoy Stamenov; RIO Country Report 2015: Lithuania; EUR 27882 EN; doi:10.2791/049440, pp 57. Available at: <https://rio.jrc.ec.europa.eu/en/country-analysis/Lithuania/country-report> (last accessed 2017-02-09)

Core services and activities

LiDA is responsible for:

- acquisition and dissemination of national and international survey data;
- access to international data archives;
- data analysis training;
- publication of data analysis teaching materials.

Content current collection

LiDA provides access to survey data from the humanities and social sciences collected at Lithuanian higher education and research institutions. The key providers of the data are researchers from Kaunas University of Technology, Vilnius University, Vytautas Magnus University, the Lithuanian Social Research centre, the Institute of Lithuanian History, as well as public opinion survey organisations (RAIT, Vilmorus).

Currently, LiDA catalogues contain more than 500 data sets on qualitative social surveys, historical statistics and the political system. The archive stores the data from the European Social Survey, the European Values Studies, the International Social Survey Programme, the European Election Studies, Lithuanian national election studies and other social studies of national importance.

Most collections from social surveys are on: Politics (100), Economics (59), Regular opinion polls (36), Media, communication and language (29), Social stratification and groupings (25), Social welfare policy and systems (16), Health (21).

Furthermore, there have been attempts to provide access to the official statistics and the survey data stored at the Statistics Lithuania through the LiDA catalogue. All the data in the system are linked to search engines, such as the Lithuanian virtual library.

Study descriptions are documented bilingually, i.e. in English and Lithuanian. One part of collections is in English (international studies), but the majority of studies are national and in Lithuanian.

Organisational infrastructure

Figure 54: Organizational infrastructure in Lithuania

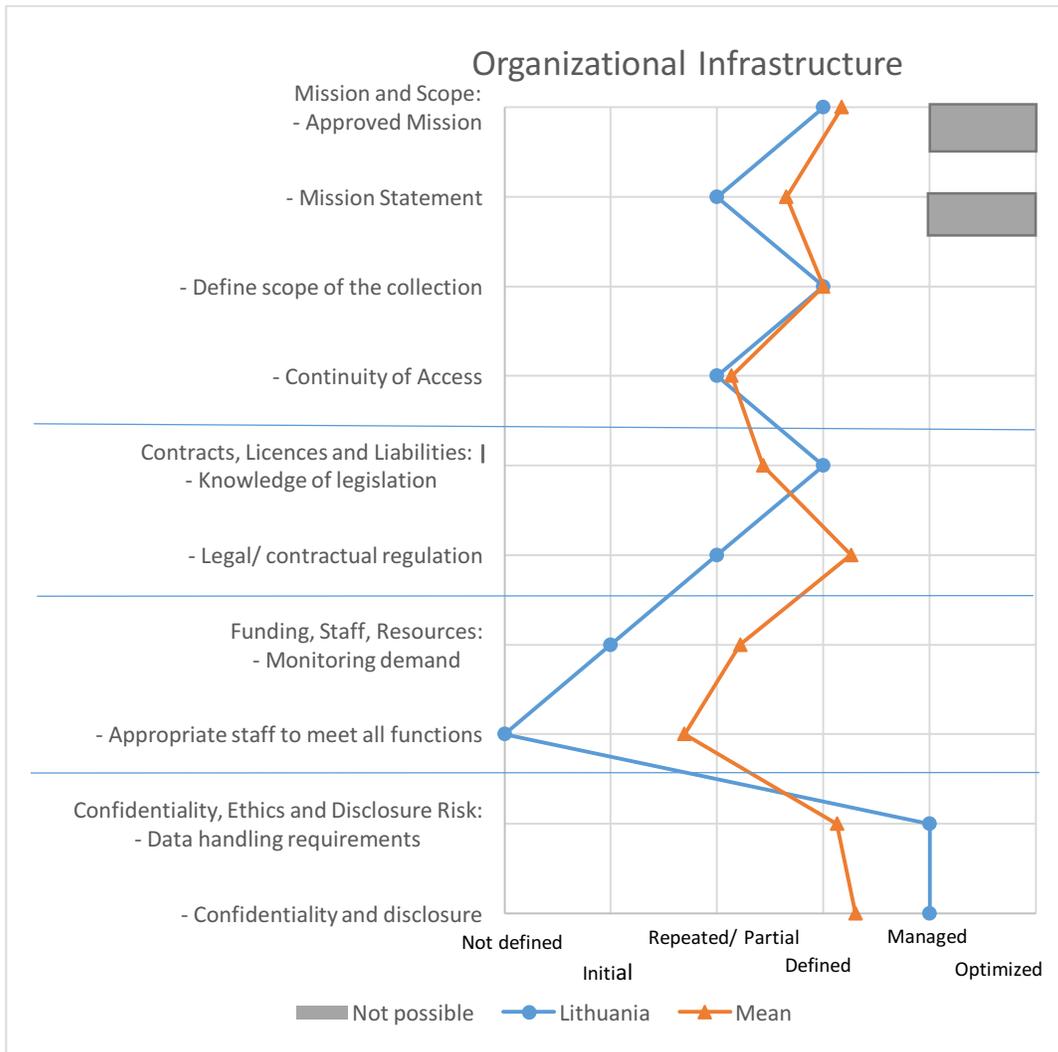


Figure 54 shows that the level of the organizational infrastructure varies. In some aspects LiDA is placed above the average, but in most aspects LiDA is placed below, or even much below, the average. A lot of problematic issues can be attributed to lack of funding that limits both daily operation and development.

Mission and scope

LiDA's Mission Statement states that the organization is a digital social science and humanities data service allowing to collect, search, browse, analyse and download empirical data.

Contract, licenses and liabilities

On knowledge of legislation (developed) the self-evaluation shows a level above the average. The staff has sufficient knowledge of the legislation, and LiDA has set clear rules of data acquisition and management according to the CESSDA requirements. The LiDA has clear terms and instructions of data use. They are available both in English and in Lithuanian. There are also rules for a data deposit in English and Lithuanian¹⁸⁶.

Funding, staff, resources

On the aspects of funding, staff and resources, LiDA is placed much below the average. The lack of permanent funding led to the loss of most of the staff after 2012/2013. There are no new decisions about future funding. It is expected that funding for operational costs will come from the European Structural funds allocated to Lithuania, but there is a long-time delay with starting this funding. The amount of this funding might also be very limited. At the moment LiDA has only part-time staff and therefore the organization does not have the appropriate number of staff to support all repository functions and service, and there is no capacity to monitor demands for its services, growth and funding. Presently the archive is only performing basic functions to provide access to existing data.

Confidentiality, ethics and disclosure risk

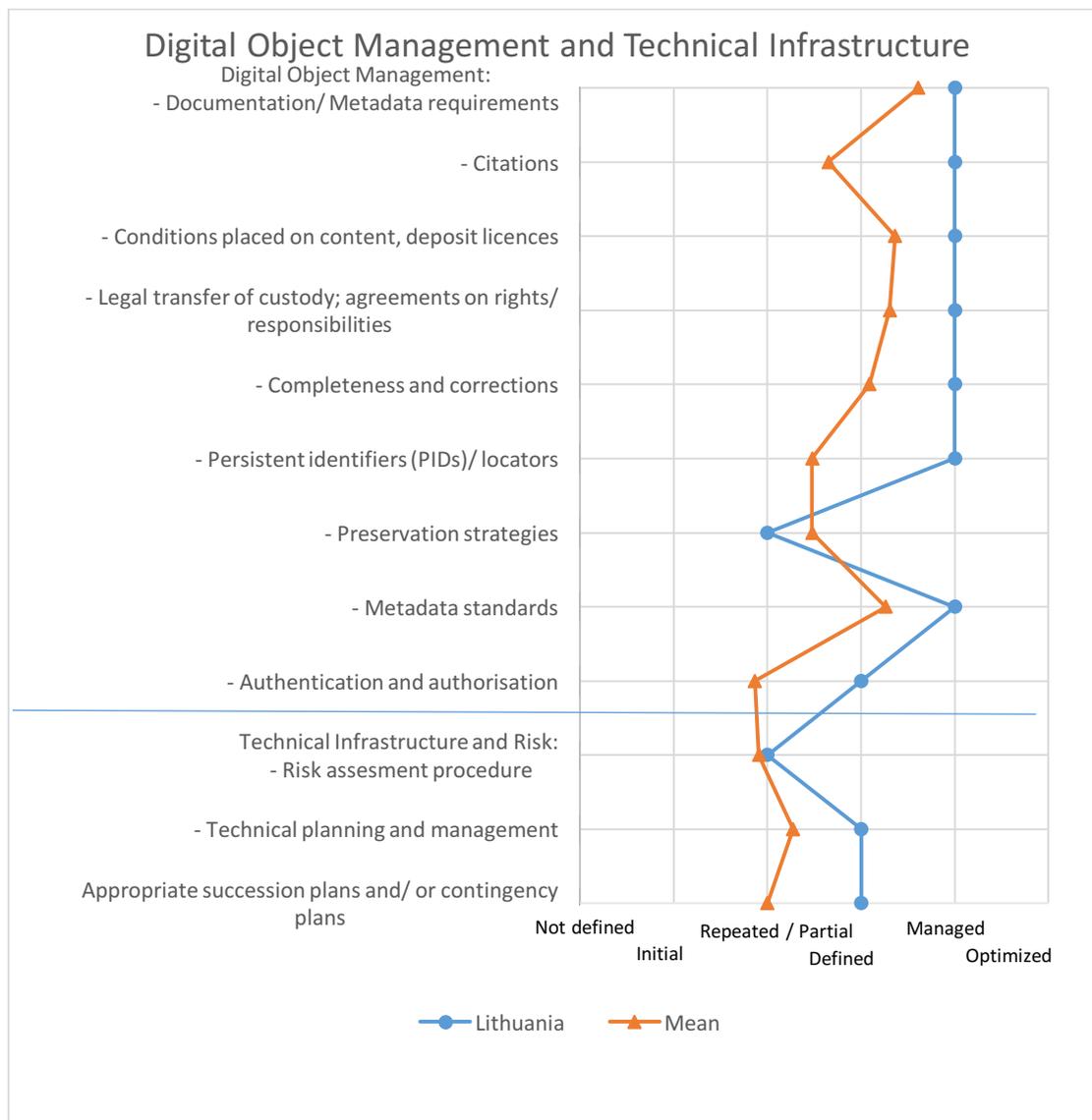
On issues like data handling requirements (managed) and confidentiality and disclosure (managed), the self-evaluation shows a level above the average.

Digital object management (data curation) and Technical infrastructure and risk

Figure 50 shows that self-assessed Digital object management and Technical infrastructure in general are above the average, and for many aspects it is on a managed level. Only one aspect - preservation strategies - have a score below the average.

¹⁸⁶ See: http://www.lidata.eu/en/index.php?file=files/eng/data/terms_for_use_data.html and <http://www.lidata.eu/en/index.php?file=files/eng/data/deposit.html>.

Figure 55: Digital object management and Technical infrastructure in Lithuania



Documentation/Metadata requirements

It is required to provide all the essential information about the study and the data being deposited according to the CESSDA recommended fields of DDI. Datasets in the catalogue of LiDA are documented according to DDI (1.2) metadata standard. Study description is documented bilingually, in English and Lithuanian. Basically, all the CESSDA recommended fields are filled out. Metadata (DDI, DC and MARC21) of all datasets are also exposed via OAI-PMH protocol.

Citations

Publications referring to the data stored in the archive should acknowledge this by a reference to the archive. To ensure that such source attributions are captured for social science bibliographic utilities, citations must appear in the footnotes or in the reference section of publications. Since 2013, due to lack of human resources, citation mechanisms are not regularly reviewed and updated.

Conditions placed on content, deposit licences

Depositing data at LiDA consists of two main stages: 1) preparing the data file so that it is correctly anonymised, clean, and edited, with appropriate labelling and descriptions; 2) submitting the data file to LiDA with all relevant supporting documentation. Once the data file is prepared with all relevant supporting documentation (questionnaires, codebooks, coding schemas, etc.) and the data deposit form is filled in, the files can be transferred via FTP server to LiDA. In some cases, depositors of data can negotiate with LiDA staff to place their data under embargo, which means that while their data is preserved at LiDA, it is not accessible to users for a limited period¹⁸⁷.

Legal transfer of custody, agreements on rights/responsibilities

Users of the data stored at LiDA must comply with the terms for use of data and with generally accepted ethical principles of research related to the correct exchange of information and knowledge¹⁸⁸.

Completeness and corrections

After being accepted for publishing, data is cleaned at the archive in cooperation with the researchers who are depositing data. The changes made to the data are clearly documented (as human readable SPSS syntax files) and deposited with the data.

Persistent identifiers (PIDs)/locators

Each data set has its unique PID which is constructed to reflect the main attributes of the data set. All the files of the data set can be accessed following the standardized rules. For example, data in SPSS format can be accessed by adding 'SPSS.SAV' to the PID and DDI file – by adding 'DDI'. So, external users or other infrastructures can easily access data and metadata stored in LiDA catalogues.

Preservation strategies

Data archiving is based on the FEDORA repository.

¹⁸⁷ See: <http://www.lidata.eu/en/index.php?file=files/eng/data/deposit.html>

¹⁸⁸ See: http://www.lidata.eu/en/index.php?file=files/eng/data/terms_for_use_data.html

Metadata standards

Data sets in the catalogues of LiDA are documented according to DDI (1.2) metadata standard. Study description is documented bilingually, in English and Lithuanian. Metadata (DDI, DC and MARC21) of all data sets are also exposed via OAI-PMH protocol.

Authentication and authorisation

LiDA has a single entry point through its main data portal www.lidata.eu, which is also available in English at www.lidata.eu/en. The LiDA portal has clear terms and instructions of data use. They are available both in English and Lithuanian (can also be accessed from the front page of the main portal).

Risk assessment procedure

Due to lack of staff this function is not performed properly.

Technical planning and management

Because of lack of funding there is no permanent upgrade of technical infrastructure.

Appropriate succession plans and/or contingency plans

There are no detailed plans and no mechanisms to implement them.

3.24.3 Conclusions

Broader ecosystem of DAS operations in Lithuania shows altogether an average maturity level (self-assessed data sharing culture and statistics-based development of social science sector are above average, while development of data production and ethical and legal framework was somewhat below the average). Overarching strategy and policy to enable sustainable data access to data is developing. In 2012, Research Council of Lithuania tested request for open access to data in contract agreements (it was discontinued later). In 2016, Research Council of Lithuania adopted Guidelines for Open Access to Research Results.

Lithuanian Data archive for Humanities and Social Sciences was founded in 2006 and has been member of CESSDA since 2012. The organisational infrastructure has in general average maturity level, items like knowledge of legislation, data handling requirements and confidentiality and disclosure being above the average. However, with regards to funding, staff and resources LiDA is placed far below the average, as a result manifests result of unsustainable funding arrangements, and this can be potentially threatening maturity level of other aspects. Self-assessed digital object management and technical infrastructure in general are above the average. The report shows that during the 2002-2012 a considerable effort and resources have been invested to build a DAS with an acceptable maturity level. However, there is risk of the investment being lost due to insufficient funding. At the moment of the audit, there were prospects of possibility of getting some funding, at the same time it might not be sufficient.

3.25 Luxembourg

As we have not gained information through the survey, we can only refer to desktop research. It seems that there is not yet established a common infrastructure for research data in Luxembourg. This does not only apply for the social sciences and the humanities but for all research disciplines. There are open access strategies implemented but they consider publications (paper, articles etc.) and do not refer to research data.

Specifics about the data collection in a country

The Luxembourg Institute of Socio-Economic Research LISER (formerly CEPS/INSTEAD) was contacted twice to support the survey or help to find persons who could do it instead. We received no answer to our mails. Thus, we do not have any information gathered by the survey.

Desktop research resulted in a few statements with regards to the data sharing and data management.

3.25.1 Broader ecosystem of DAS operation

Development of social science sector

Luxembourg has participated in cross-national studies like CCS, ESS, EVS and PISA.

RDM Policy setting

The University of Luxembourg and the University of Liège has started a common repository “Open Repository and Bibliography” ORBi in 2013. Members of the Universities should deposit all publications and bibliographic reference in the repository. Research data is not part of this policy.

In a report¹⁸⁹ the authors outline that, „*Luxembourg reported that even though it ‘is currently focusing on open access to publications’, the ORBilu open repository and bibliography system of the University of Luxembourg already links publications to research data*“ (p.21). On the ORBi website there are no explicit hints about referring to research data and it is not possible to search for publications regarding research data deposited in a repository.

The report also states that “*apart from ‘the ORBilu repository that links publications to research data, no common infrastructure is currently in place’*” (p.41)

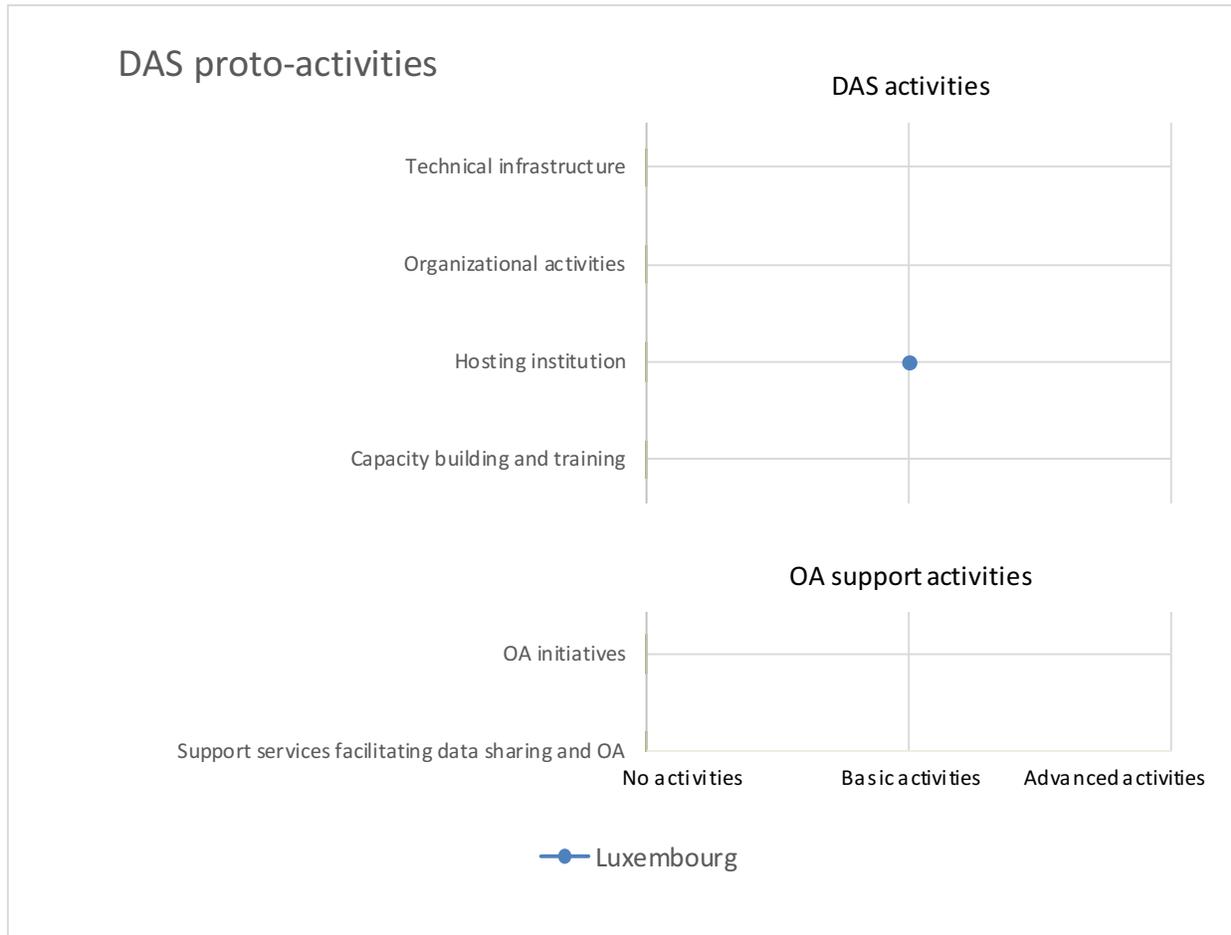
Data sharing culture

As no information is available concerning data sharing and enablers for data sharing no estimations could be made.

¹⁸⁹ From: Access to and Preservation of Scientific Information in Europe Report on the implementation of Commission Recommendation C(2012) 4890 final Edited by Maria Monica TARAZONA RUA, Daniel SPICHTINGER, Celina RAMJOUE, Jean-Francois DECHAMP.

3.25.2 Data archive service (DAS) proto-activities

Figure 56: DAS and OA activities implementation type in Luxembourg



DAS activities

Currently, there is no DAS in Luxembourg, even if in the past Luxembourg was member of CESSDA. LISER (formerly CEPS/INSTEAD), a Luxembourgish public research institute under the jurisdiction of the Ministry of Higher Education and Research, could possibly host and be used for or applied to a new DAS.

LISER is active in the social science research. Indeed, its research focus lies in the field of social and economic policy including the spatial dimension. This comprises topics like poverty, inequality, education, social inclusion, employment, unemployment, health, housing, mobility, and regional convergence. The aim is to improve the understanding of causal relationships and to provide sound evidence for the impact of institutional settings and policy options on outcomes. Based on empirical evidence, the Institute wants to provide well-grounded and clear-cut answers to policy relevant questions. The results are published in form of research reports, monographs and scientific articles.

The institute employs 122 people, of whom 65 are researchers and PhD candidates. The research team is representing a high degree of interdisciplinarity and includes demographers, economists, geographers, legal experts, political scientists, psychologists, social scientists, and statisticians.

LISER has been integrated into a unified legal framework for the Luxembourgish research institutes based on the law of 3 December 2014. It has an annual budget of approximately 18,5 million euro, of which 60% is allocated by the government, and 40% are coming from project related funding resources. Governmental funding is based on a multiannual performance contract¹⁹⁰.

3.25.3 Conclusions

See Summary above.

¹⁹⁰ See <https://www.liser.lu/?type=module&id=53>

3.26 Former Yugoslav Republic of Macedonia (FYRM)

Macedonia does not have an established data service for the social sciences or any data archive at the national level that is used by the science research community. The conducted analysis shows that main elements of broader ecosystem of DAS operations in Macedonia are underdeveloped, with indicators' values trailing the average of European countries included in the survey. However, DAS proto-activities are identified through efforts made by Institute for Sociological, Political and Juridical Research to establish DAS and develop relevant infrastructure within SEEDS project.

Specifics about the data collection in a country

The data collection started 21st September 2016 and was finished 30th September 2016. The information was provided by the Saints Cyril and Methodius University, Institute for Sociological, Political and Juridical Research (ISPJR), Skopje, potential service provider in Macedonia. The contact person who filled the survey was Aneta Cekik, a researcher from ISPJR, engaged in data archive proto activities within SEEDS project¹⁹¹. The data collection process went well without any problems.

3.26.1 Broader ecosystem of DAS operation

The Ministry of Education and Science of the Republic of Macedonia (MES) is the main national institution in the area of research and education, and the main actor responsible for designing and implementing national policies related to scientific research. In fulfilment of these tasks, MES is supported by the National Council for Higher Education, Science, Innovation and Technology (NCESIT), an advisory/expert body of the government. Within the Ministry of Education and Science, there is the Sector for Science, and Technical and Technological Development, specifically responsible for the development of policy for research activities, including research infrastructure.

Currently, there are five public universities, nine private universities, and several faculties or other higher education institutions in the register of higher education institutions of the Ministry of Education and Science. Most of the research in the social sciences in Macedonia is however performed at public research institutes, at public universities (especially their research institutes), and the Macedonian Academy of Sciences. In 2008 the public sector employed more than 90% of research staff in the social sciences and humanities (Josimovski, 2011, p. 1)¹⁹². With regards to social science research, the NGO sector and think tank organizations appear to also be large producers of research data in the last several years, especially considering the recent cuts of national funds for research for public institutions.

¹⁹¹ SEEDS (2015) Report on the Evaluation of Research and Legal Conditions for the Establishment of a Social Science Data Archive in Macedonia.

¹⁹² Josimovski, S. (2011) Social Sciences and Humanities in the FYR. Macedonia. Metrix 2011 Report, European Commission, DG Research.

Figure 57: Heading concepts values in FYRM

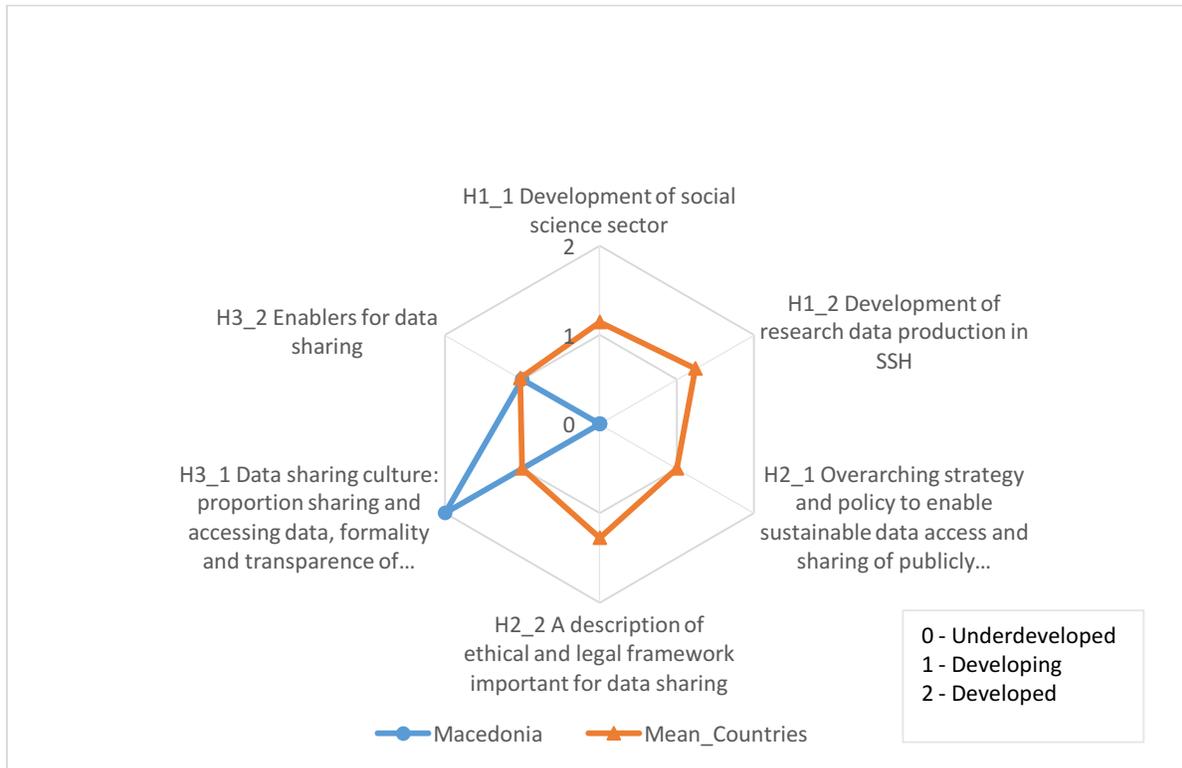


Figure 52 shows that Macedonia is a country with mostly underdeveloped elements of broader ecosystem of (potential) DAS operations, with indicators' values trailing the average of European countries included in the survey. Each of these heading concepts of broader ecosystem is further discussed in more details.

Development of social science sector

Overall assessment of SSH development

According to the self-assessment survey “academic research in Macedonia is currently financed by 0.22% of GDP, which is among the lowest levels of funding in Europe (Erawatch, 2014)¹⁹³. The main financing of research projects is done by the Ministry of Education and Science. In 2011 the share of public funding in the gross expenditure for research and development was 44.2% (which was an actual increase compared to 2000-2007), the share of international funding was 43%, while the share of private sector funding was 12.8% (Erawatch, 2014). Regarding the social sciences, in 2008, these disciplines received only 8.6% of the total gross domestic expenditure for research. 65.7% of these sources came from the public sector, 15.8% from the private sector, 18.4% from foreign sources, and only 0.1% came from the non-profit sector (Josimovski, 2011, p. 21). What is striking when discussing research funding in Macedonia is the fact that more than 90% of national expenditure for research consists of salaries for the employed researchers and current expenditures for research institutions (Josimovski, 2011; State Statistical Office of the Republic of Macedonia, 2014).”¹⁹⁴

In addition to generally unfavourable financing conditions of SSH, researchers in Macedonia also lack the other forms of the infrastructural support for the productive research. Centralized access to commercial databases or datasets, or centralized procurement of the software licenses for scientific analysis, are non-existing or only occasionally provided by the funding entities.

Development of research data production in SSH

According to the SEEDS survey from 2015, a significant number of social science staff in Macedonia is involved in research activities. In the last 5 years prior to the survey, 75% of respondents have produced or helped in producing research data. The mean number of produced datasets stands at 9.23, while the median, being the more representative indicator because of the several outliers, is 5. On average, every year, one dataset per researcher is produced. In addition, some of the SSH research institutions “have tradition in periodic production of research data to certain extent”, as self-assessment reveal. However, Macedonia is one of the countries with the lowest participation in international collaborative research or cross-national studies (exemptions are participation to EVS in 2008 and PISA in 2000 and 2015)

¹⁹³ Erawatch (2014) Erawatch Country profile Macedonia

¹⁹⁴ State Statistical Office of the Republic of Macedonia (2014) Statistical Yearbook of the R. Macedonia.

RDM Policy setting

Funders' data management and sharing strategy and/or policy

According to the results from the SEEDS survey of social science researchers' practices in data production and preservation, there is a clear need for the establishment of a social science archive in Macedonia, and researchers strongly support the idea. However, in the legal acts, as well as in some recent projects related to research infrastructures, a social science data archive is not envisioned at all. The Ministry of education and science, as the leading funder of SSH research, is "unaware of the growing importance of Data Management Plans as an integral part of research projects", and "do not provide incentives for sharing research data with associated metadata". Also, most international donors did not apply this criterion over the data set produced under their funded projects in the past, but since June 2016 ISPJR is a participant of RRPP Data Rescue project funded by Swiss National Science Foundation. This project has the aim to collect and archive all data sets generated within past RRPP projects in the period 2007-2016.

Legal and ethical framework

Legal framework regarding data protection and dissemination consists of several laws regulating issues of copyrights, archiving, safety and access to public information. Law on copyrights and related rights (2010) regulates the copyrights of authors over their work, including datasets and databases. Law on archival material (2012) regulates the protection, storage, processing and use of archival material. Law on protection of personal data (2005) regulates secrecy and protection in processing the personal data of citizens of the Republic of Macedonia. Law on free access to information of public character (2006) transparency in the work of the state and public bodies, and to provide free access to information that is produced and/or possessed by such entities.

While legal framework mostly covers particular issues relevant for data archiving, dissemination and protection in general, there is still considerable lack in development of specific recommendations and guidance aiming to clarify legal and ethical aspects that facilitate social science data sharing.

Data sharing culture

Sharing of the data between researchers in social sciences is not widespread practice in Macedonia. The SEEDS survey reveals that only 12% of the research data is preserved in data archives or repositories, while the rest is stored on local computers and servers. Accordingly, access to research data is very limited. As stated in the report, "the dominant practice in providing access is very narrowly focused". The respondents provides access to data from the last research project mostly only to "members of the research team" (44%), then "members of my institution" (19%) and "broader scientific community" (14%). This is in line with self-assessment survey, where informal contacts and project/personal websites are ranked as the most important routines of data sharing between 2011 and 2016.

Attitudes toward sharing the data examined in the SEEDS survey show some contradictions. While vast majority of researchers (98%) consider sharing the data within own discipline as important and express the willingness to deposit own data in DAS (83%), still 50% of them consider that ideal level of access to data from the last project should be limited only to project leader, project members or institution staff. The possible explanation for such contradiction is fear of the researchers about misuse of their data if shared under existing circumstances with no clear rules

on use of secondary data. This is supported by the self-assessment survey, where attitudes “Data sharing creates negative competition” and “Data sharing has as a risk that others may misuse and misinterpret data” are assumed to be probably true. The full attitudes toward data sharing in self-assessment survey is presented in Table 1. In addition, according to the SEEDS survey high percent of researchers (33%) have indicated that their data were used by other researchers for secondary analysis. However, expert in self-assessment survey expresses scepticism that such big share comes from inappropriate understanding of this question: *“based on my personal experience and from the interview data collected for the SEEDS project, I think that there might be a possibility for confusion of data sharing with simple referencing of publications”*.

Taking everything into account, it can be concluded that social sciences researchers in Macedonia have positive attitudes toward data sharing and understand the benefits of it, but practice of data sharing under current infrastructural and legal conditions are underdeveloped and fear of data misusing and negative competition stemming from data sharing is present.

Table 21: Attitudes towards data sharing in FYRM

Data sharing has no benefits at all	False
Data sharing creates healthy competition	True
Data sharing creates negative competition	Probably true
Reuse of existing data can answer new research questions and facilitate advancement of science	True
Data sharing has as a risk that others may misuse and misinterpret data	Probably true
Data sharing involves little effort and minimal costs	Probably true

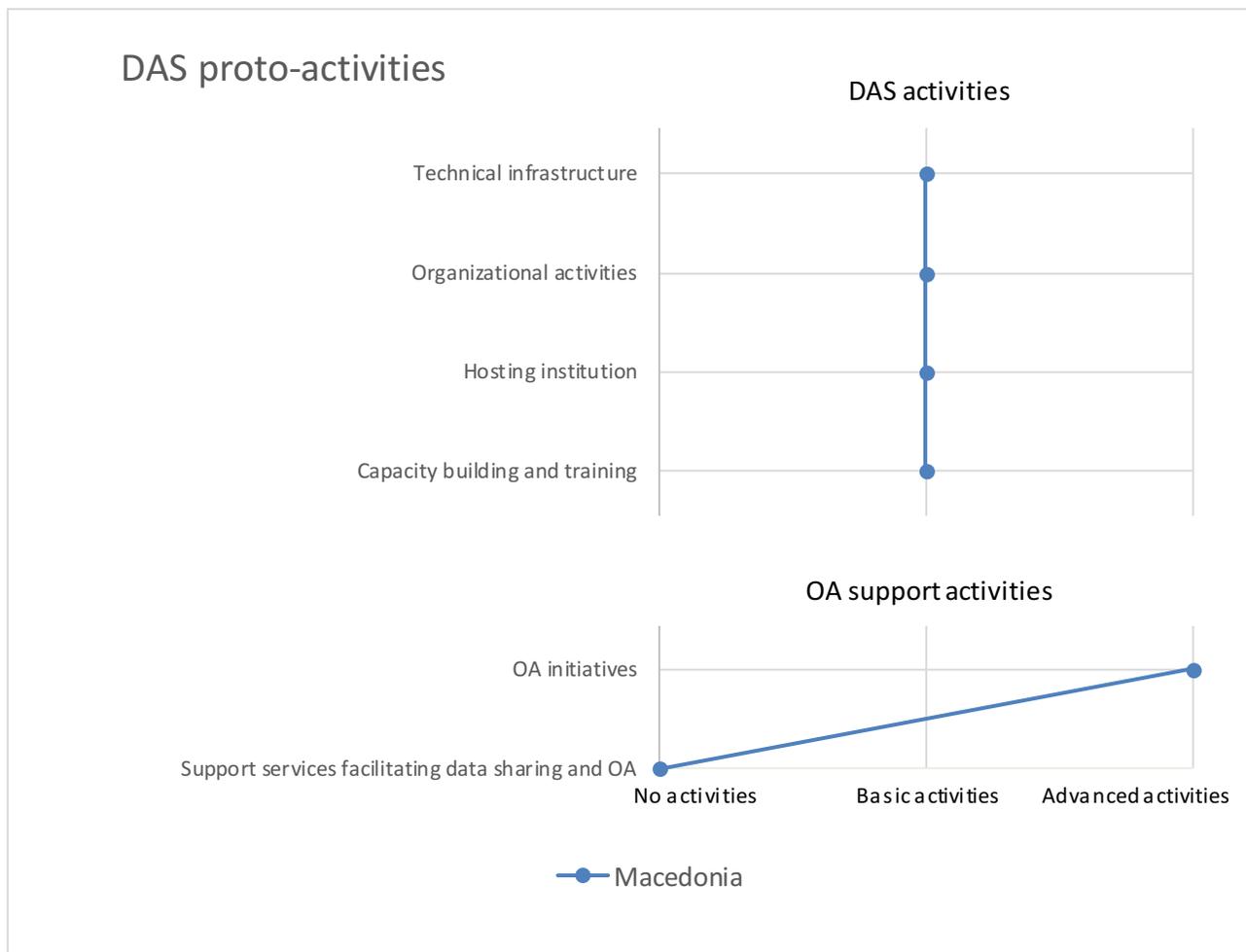
Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from “5”-True to “1” False!

Enablers for data sharing

In the self-survey assessment, no significant enablers for data sharing are identified in terms of frequent publishing in journals that require depositing data, career rewards or data support services provided to social science researchers that facilitate data sharing or open access. To certain extent, only data management and data documentation standards and procedures are identified as data reuse facilitators. As claimed in self-survey, “Some interviewees in the SEEDS project have reported use of internal standards for preservation of data, but without following the relevant international standards; one important exception is the e-repository of Goce Delchev University in Shtip used by the staff of that University”.

3.26.2 Data archive service (DAS) proto-activities

Figure 58: DAS and OA activities implementation type in FYRM



DAS activities

Macedonia currently has no established Data Archive Service (DAS), nor developed technical infrastructures that could possibly be used for or applied to a new DAS. SEEDS report refers to a couple of strategic documents related to research activities; however, issues of DAS establishment are not discussed within these documents.

Yet, the Institute for Sociological, Political and Juridical Research (ISPJR) is currently involved in the international project SEEDS, which aims to establish a DAS in Macedonia. SEEDS enables ISPJR to build the appropriate technical infrastructure (servers, software, website), and to develop the organisation (policy documents and procedures) and the specific knowledge and skills on preservation and dissemination of data, that is needed to establish a DAS. The more detailed explanations on the anticipated organizational structure, human resources and technical infrastructure of the DAS within ISPJR could be found in Establishment Plan at the SEEDS website. The Ministry of Education and Science of Macedonia supports the activities undertaken in the SEEDS project. So, the Ministry is aware and supportive of the initiative, which is important for the establishment of a DAS.

Besides ISPJR, some DAS proto-activities are also identified in the public university “Goce Delchev” in Stip, which has a repository powered by EPrints for keeping academic information and data in the social sciences and humanities. Currently 9,000 units are deposited in it, mostly research publications. It is an open access repository; data are publicly available and can be used for non-profit purposes. It is allowed for the items to be reproduced, displayed or performed, and given to third parties in any format or medium, as well as to be used for personal research or study, educational, or not-for-profit purposes without prior permission or charges. The open access is conditional upon referencing of the source, including the hyperlink and/or URL for the original metadata page. Commercial use is not allowed in any format or medium without formal permission of the copyright holders. The repository contains and preserves journal articles; conference and workshop papers; theses and dissertations; books, chapters and sections; datasets; multimedia and audio-visual materials and patents. The documentation of the material is available in Macedonian and English. Currently only staff employed at this university can deposit data and publications. Deposited data are kept on servers equipped with a data storage system and protected by a disaster recovery system. Data are kept in a standard file format (PDF) and are enriched with the appropriate metadata. The repository uses Dublin Core and Learning object metadata (LOM) and is compliant with OAIS with certain modifications in accordance with their needs. Visibility of their publication items is promoted through indexing in Google Scholar. The OAIPMH protocol is used for exchange of metadata with other repositories.

Open access (OA) support activities

Since 2012, there is a growing awareness about open access to research data thanks to the project “Initiative for Open Educational Resources” (OER), carried out by the Foundation Metamorphosis, a working group established for promoting the idea of free access to educational resources in Macedonia.

The overall project goal is to contribute to the development of critical thinking and democratization in Macedonia through the constructive use of new technologies as tools to increase the quantity and quality of educational, scientific and academic e-content. Particularly, the project has a specific aim to raise the awareness and build the capacity of the academia for creating and using open educational resources. Expected results are to:

- Raise the awareness about the importance of the concept of open educational resources and contribute towards the improvement of the legal and institutional frameworks to make them more convenient for e-content development;
- Improve the visibility of academic publications and create new resources that would be used to promote and aggregate publications under the Creative Commons licenses and with open access;
- Increase the quality of educational content relevant for the pupils and students in Macedonia;
- Increase the quality of local content freely available online by increasing the public awareness and the level of IT skills, based on the application of best practices.

The project started in 2013, as an informal network of citizens and different institutions (schools, faculties, NGOs). It publicly supports the declaration of UNESCO for Open Educational Resources. The formal web site can be found at <http://oer.mk/>, but the content is at Macedonian and Albanian language, only.

As mentioned above, no data support services facilitating data sharing or open access to social sciences researchers were identified.

3.26.3 Conclusions

According to the conducted analysis, it can be concluded:

Funding of SSH and productivity of the researchers in Macedonia are at the low level of development and that impact on designated community is small or non-existing. Development of research data production is underdeveloped. While researchers and research institution in SSH produce some data, they do not participate up to a large extent in international and cross-country collaborative research.

RDM policy setting is underdeveloped. SSH funders do not provide incentives nor impose requirements for DMP. Specific recommendations and guidance aiming to clarify legal and ethical aspects that facilitate social science data sharing do not exist.

Social sciences researchers in Macedonia have positive attitudes toward data sharing and understand the benefits of it, but practice of data sharing under current infrastructural and legal conditions is underdeveloped.

Basic DAS activities in Macedonia exist and they are mostly initiated through SEEDS and RRPP Data Rescue project.

3.27 Malta

No report. No productive contact was established during the survey.

3.28 Moldova

The analysis shows that Moldova does not have an established data service for the social sciences. No clear policy on Research Data Management exists in the country. In general, the social science sector of Moldova is underdeveloped. The main source of social science research funding is the public sector and the main funding institution is the Supreme Council for Science and Technological Development (SCSTD).

Specifics about the data collection in a country

Information was collected mainly via webpage of Academy of Sciences of Moldova. Another major source of information was *Peer Review of the Moldovan Research and Innovation System*¹⁹⁵, written by Toivo Räm, Brigitte Weiss, Krzysztof Gulda, George Bonas, Manfred Spiesberger, Daniel Funeriu, and Francien Heijs in 2016. Unfortunately, there is no data available from the self-assessment survey provided by the CESSDA SaW project.

3.28.1 Broader ecosystem of DAS operation

The science system in Moldova consists of higher education, research and innovation development institutions. It includes not only the public and non-public institutions of higher education and basic research, but also private enterprises that operate in the fields of research, development and innovation. Research is also conducted by a number of non-governmental organizations (NGOs).

The Academy of Sciences of Moldova (ASM) is the main coordinator of scientific and innovation activities in the country. Most of the research capacity in Moldova is also under the ASM or institutions closely related to it. Still some research institutes are directly administratively subordinated to different Ministries as the former so-called “*branch research institutes*”. As it was noted in the Peer Review of the Moldovan Research and Innovation System, the current governance of the Moldovan research and innovation (R&I) system is based on the ASM, which de facto fulfils the role of a Ministry of Science. However, the ASM is at the same time is the policy developer, the manager of major share of public R&I funds through the Center for Fundamental and Applied Research Funding (CFARF) and the main research performer. Eventually all this results into a clear institutional conflict of interests for the ASM. The largest share of the state R&D budget is allocated to the operating costs of ASM’s research institutes.

Another important feature of Moldovan science system is a strong separation between teaching and research activities. Such a system is considered outdated and removing those barriers in Moldova would contribute to a better use of its human resources and to the optimal use of

¹⁹⁵ Toivo Räm, Brigitte Weiss, Krzysztof Gulda, George Bonas, Manfred Spiesberger, Daniel Funeriu, Francien Heijs (2016) *Peer Review of the Moldovan Research and Innovation System*. Luxembourg: Publications Office of the European Union.

research funding and research infrastructure. However, negative trends within Moldovan science system are reinforced by shrinking research community and intense brain-drain, which occurred internally and externally as research careers are not attractive for young researchers, the research community is rapidly ageing and the overall numbers for R&D personnel have decreased five-fold since the country's independence in 1991. The number of researchers per 1 million people is 4.5 times lower than in EU¹⁹⁶.

There are different problems related to scientific infrastructures, but the major one that research infrastructures are outdated and funding for its renewal is not sufficient.

Development of social science sector

Overall assessment of SSH development

The Moldovan Academy of Sciences holds most of the country's research capacity and research infrastructure. It is a key public research organization and the focal point of research activities, and a hub of collaboration for higher education institutions and research institutes. However, social science sector is underdeveloped within the ASM system. There only few institutions with social science drive as National Institute of Economic Research, Institute of Ecology and Geography or University of European Political and Economic Studies” Constantin Stere”. As already was mentioned about research, including social science, has low profile in higher education institutions as these organizations are almost exceptionally oriented to teaching. Moldova as a country has underdeveloped social science sector. Still researchers in Moldova have access to some commercial bibliographic and full-text databases, i.e. EBSCO etc.

Development of research data production in SSH

The National Bureau of Statistics of Moldova is major producer of empirical data. This organization is collecting not only national statistical data, but conducting major national surveys as the Labour Force Survey, Household Survey and Household Budget Survey.

Among private companies in public opinion and marketing research are Independent Sociological and Informational Service (OPINIA), CIVIS Centre for sociological, psychological and political analysis and investigations, which conducted surveys on migration, life styles, labour market, social exclusion and so on. However, in the field of studies that systematically assess matters of national importance there are sporadic initiatives, but not systematic or harmonized.

About academically generated research data there is no reliable information. But Moldovan participation in main international collaborative research and cross-national studies is very low. Moldova was a participant only in few such studies, including few waves of European Values Study (EVS) and World Values Survey (WVS), Fertility and Family Surveys (FFS).

¹⁹⁶ Report available at: https://rio.jrc.ec.europa.eu/sites/default/files/report/Moldova-PSF_PR-KIAX16004ENNOP.pdf

RDM Policy setting

Funders' data management and sharing strategy and/or policy

Since sharing of research data is not common in Moldova, requirements about preparing Data Management Plans are non-existent for social science research data. However, there is growing recognition and awareness about the value of research data produced and about the need for long-term preservation. The funding agencies do not provide incentives for sharing research data with associated metadata.

Legal and ethical framework

In 2013 Moldovan government passed new National Strategy for the development of information society „Digital Moldova 2020”. Among three pillars of this strategy at least two are talking about infrastructures and data, i.e. first pillar is focused in access and infrastructure – improvement of connectivity and network access, and, second pillar is on digital content and electronic services-promoting digital content and generating services.

Still the legal and regulatory framework is not yet completely adjusted to the realities of the digital environment, and development of legal and ethical framework important for data sharing is in the initial phase.

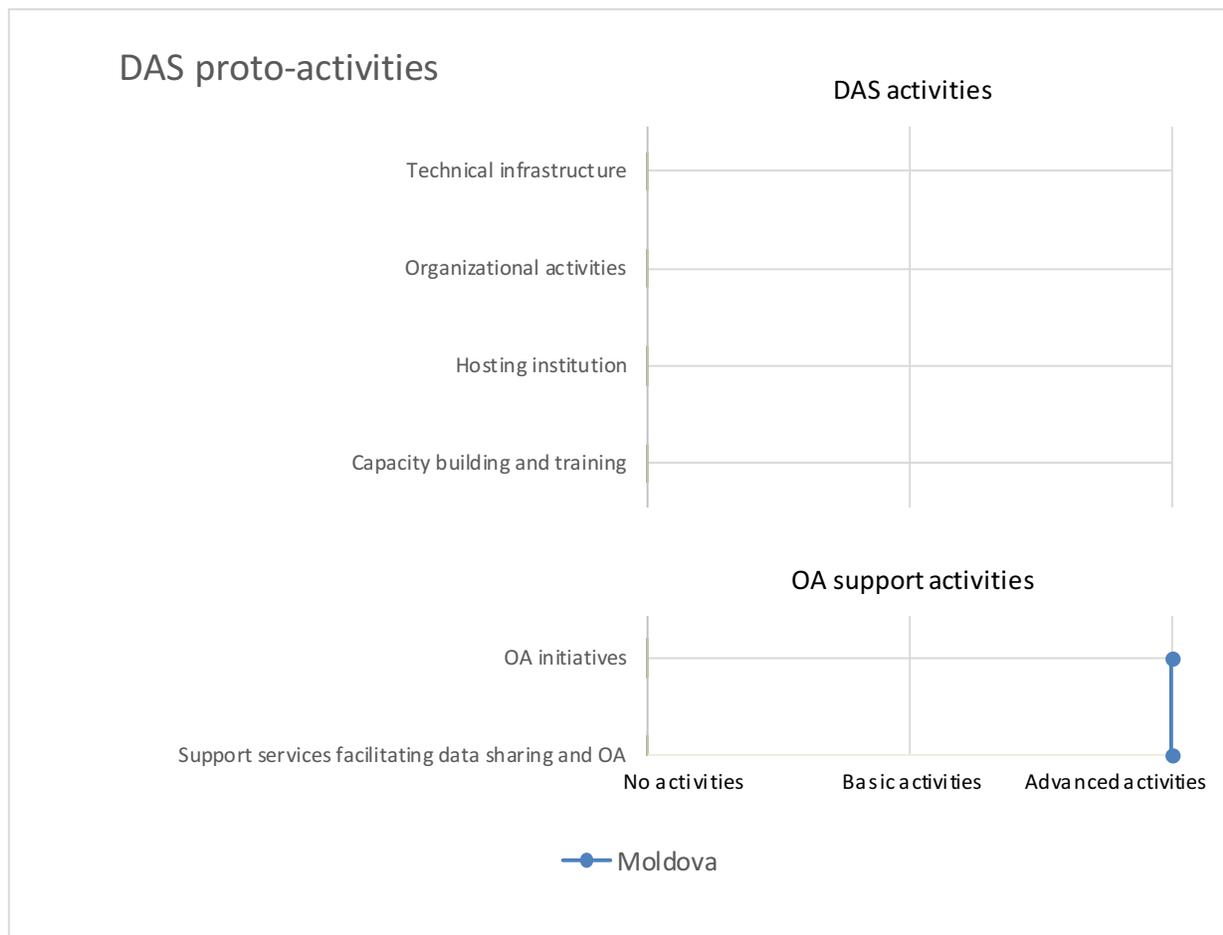
Data sharing culture

There is no detailed data available on data sharing practices in Moldova. Indirectly it is possible to point out that a demand for data sharing is increasing. The open data portal- date.gov.md- launched within a national open data initiative between the launch of the portal in April 2011 and 31 October 2014 the portal registered as many as 141,982 visits with 104,949 unique visitors. This means that the data is used more often for secondary analysis.

As usual in cases there are no institutionalized data service providers the necessary data is acquired directly from project and personal websites with individual permissions, i.e. via informal contacts to use the available data. In Moldova, researchers do not have career rewards related to data sharing. On the other hand, relatively strong promotion of the institutional and national Open Access policies helps to build up data sharing culture as well.

3.28.2 Data archive service (DAS) proto-activities

Figure 59: DAS and OA activities implementation type in Moldova



DAS activities

There are no existing national data archiving infrastructures in Moldova. However, it might be considered to use the joint EU- Moldovan programme „Connection of the Moldovan Centres of Excellence to the EU Research Infrastructure” to encourage DAS activities.

Open access (OA) support activities

Another even more important channel to promote DAS activities is an active Open Access (OA) movement in Moldova through Electronic Resources for Moldova - Resurse Electronice pentru Moldova (REM) in partnership with European Federation for Intercultural Learning (EFIL) and International Network for the Availability of Scientific Publications (INASP). The REM consortium brings together 19 organizations and aims to provide the access to e-resources, subscription of databases, training and development institutional repositories.

In 2013-2016, nine institutions in Moldova were a part of Tempus funded project Modern Information Services for Improvement Study Quality. One of the main project objectives was the creation of institutional OA repositories.

Since 2011, Moldova also was implementing a national open data initiative as part of its Governance e-Transformation Agenda. This initiative as well is encouraging the data sharing as government agencies and ministries are responsible for publication of government data, and for releasing at least 3 datasets on a monthly basis. In early 2017, as many as 944 datasets were published on a webpage of this initiative. The most active publishers of open data were the Ministries of Health, Interior, Economy and Education, and the National Bureau of Statistics.

3.28.3 Conclusions

Development of research data production in SSH is at the initial phase. There is a huge lack of international collaborative research and low participation in international research programs.

Ethical and legal framework for data sharing is underdeveloped. Data sharing and reuse is new and still rare in the social sciences research community.

One of the major problems that can persist, and even worsen is the shortage of funds for social science research and research infrastructures. Quite intensive advocating for institutional and national Open Access policies might be helpful instrument to increase awareness about data management and consolidation of the data infrastructure of research organizations and major universities making it available for collaborative research.

3.29 Montenegro

Broader ecosystem of data service operation in Montenegro can be considered generally underdeveloped, although some positive developments were identified in the country.

Most research in social science in Montenegro is currently financed by international funds. Research data production has a potential for development, and research capacities could be improved.

A notable effort towards the establishment of data services for social sciences in Montenegro has been made within the SEEDS project¹⁹⁷. Issues related to RDM policy setting were discussed during the interviews with policy makers and other relevant stakeholders conducted as an activity the SEEDS project, so we can say that overarching strategy and policy to enable sustainable data access and sharing of research data is developing.

Data sharing and reuse among social science researchers in Montenegro exist as a practice among researchers, and they mostly share their data inside their networks of colleagues. Using the data in secondary analysis is hard because the data is purely documented so some improvements can be achieved by educating researchers in data management and secondary analysis.

DAS proto-activities in Montenegro started with the SEEDS project. Most of the issues related to the process of establishment of data services for social sciences were tackled through the project, such as policy development, governance, human resources, and technical issues. The skills and knowledge were developed through the project which can be used to further influence the development of data services in the country.

¹⁹⁷ SEEDS - South-Eastern European Data Services (2015-2017). <http://seedsproject.ch>

Specifics about the data collection

The data was collected during the first week of November 2016. The online self-assessment instrument was used to collect information from the country representative, and this was supplemented by information gathered via desk research using mainly the reports from the SEEDS project. Self-assessment questionnaire was filled up by Nikoleta Tomović, a teaching assistant at the Humanistic studies of University of Donja Gorica, Department of International relations and Diplomacy and the executive director of one of the most influential think tanks in Montenegro, Centre for Monitoring and Research (CeMI), whose main goal is to provide expert support for continuous monitoring of the overall process of transition in Montenegro. She was also involved in the SEEDS project, in which the CeMI is one of the partners, so she can be considered as a good informant for providing data to this survey. Most of the information, no matter the source, was recorded into the online survey tool, and additional information was pulled out directly from desk research sources.

3.29.1 Broader ecosystem of DAS operation

Figure 60: Heading concepts values in Montenegro

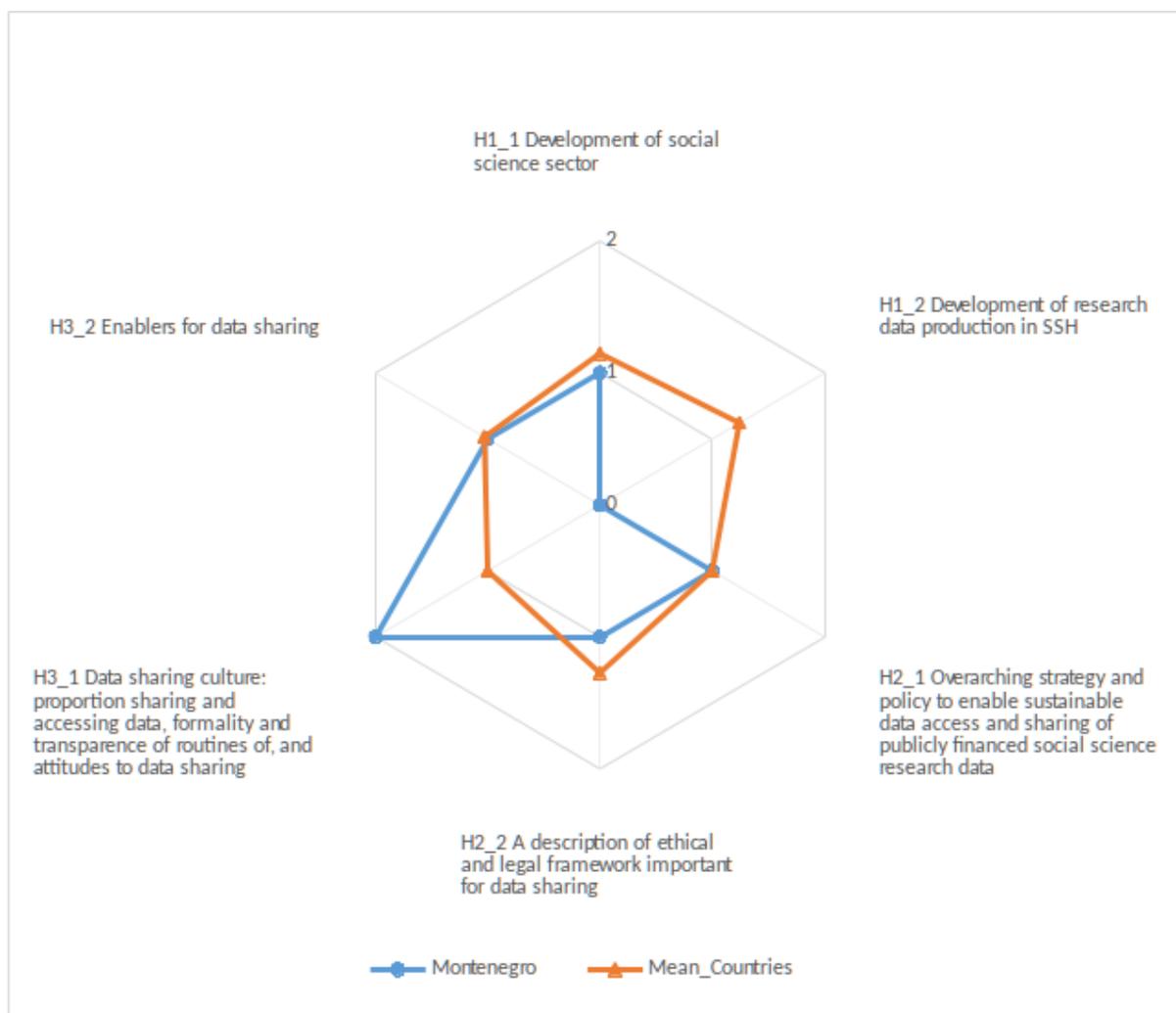


Figure 55 shows that the broader ecosystem of data archive service operation in Montenegro has been categorized as developing for most of the indicators, except for development of research data production in SSH which was estimated to be underdeveloped and data sharing culture, which might be considered developed. What is shown in this picture are the estimates concluded from generic methodology based on data from the self-assessment survey, but by adding the information from other sources into the picture, we see that this was an overestimate in most cases.

Development of social science sector

Overall assessment of SSH development

The general intensity of investment in SSH is 0.1 of the GDP. Relative to other disciplines, the intensity of investment in SSH is 25.2% and the number of researchers in SSH per 100.000 capita is 97. The investment in human resources in SSH is 5604 (GERD in SSH per researcher in SSH).

The results of the project SEEDS Survey on production, preservation and use of research data¹⁹⁸ conducted in 2015 among researchers in Montenegro, found that 60% of surveyed researchers used international funds for their research. The second common source was the institution where the respondent is working (21%), public funding from national science funding bodies was reported by 16% of the respondents, and in 9% of cases the research was financed by researcher's own funds.

Research capacities were assessed through existence of basic e-infrastructures providing access to scientific information (full-text and bibliographic databases, datasets) and software licences for the software used in research. According to the University of Montenegro Central University Library¹⁹⁹ list of databases there are no commercial full-text and bibliographic databases available for researchers in Montenegro. In the self-assessment survey, it was indicated that the access to databases is provided, but seems to be limited to just two EBSCO full-text databases. No commercial datasets are available for researchers. Software licenses for SPSS are provided on a national level.

Development of research data production

Montenegro was involved in some international collaborative or cross-national research studies, which are high quality research with a high potential for reuse. These were the following: Comparative Study of Electoral Systems (CSES), Module 4 [2011-2016], European Values Study (EVS) [2008], and Programme for International Student Assessment (PISA) [2015, 2012, 2009, 2006, 2003].

According to the self-assessment survey results, there are many studies that assess matters of national importance (public opinion survey, election survey, etc.) produced by the social science researchers. Our respondent stated that "these studies are published by different faculties, research centers, think tanks, institutes, Montenegrin Academy of Science and Art, etc. For

¹⁹⁸ SEEDS (2016). D3 – Analysis of existing potentials for the establishment of a social science digital data archive in Montenegro. http://seedsproject.ch/wp-content/uploads/2015/06/Report-for-Montenegro-SEEDS_bk-CEMI-Final_rk1-August-2016_FINAL.pdf

¹⁹⁹ University of Montenegro Central University Library. <http://www.ucg.ac.me/me/o-univerzitetu/centralna-univerzitetska-biblioteka>

example, the issue of corruption is very often treated from think tanks and research centers through publishing policy studies, research analysis and public opinion surveys".

It was estimated that the average production of research data by the social science institutions in Montenegro is periodical, e.g. that institutions have tradition in producing some type of research data to a certain extent. As described by our respondent, "depending on available funds, research centers and faculties have published studies and analysis through implementation of policy and research projects. The funds for research activities are the most evident problem, regarding the fact that there are no sustainable channels for project funding from state institutions, while donations from foreign donors are mostly oriented on policy activities, more than research activities". SEEDS Survey showed that 71% of the researchers did produce data in the last 5 years, although this is for sure an overestimate since the survey sample was self-selected with those who engage in data gathering being much more interested in participating.

RDM Policy setting

Funders' data management and sharing strategy and/or policy

Public funders in Montenegro currently do not have any policies related to research data management and preservation and no incentives for sharing research data with associated metadata. It was estimated in self-assessment survey that there is a growing recognition and awareness among public funders about the need to require Data Management Plans and Open data, and about the need to have a disciplinary specific repository for research data together with support services. There is also a growing awareness about the value of research data produced and the need for its long-term preservation, but no investment and support is provided for this.

It is important to note that these estimates were based on the fact that Montenegro was involved in the SEEDS project, which was the first and only project in Montenegro which started to raise awareness about all the specific issues related to RDM policy identified in our self-assessment survey²⁰⁰. These issues were discussed during the interviews with representatives of the Ministry of Science, Ministry for Information Society and Telecommunications, Ministry of Education, Montenegrin Academy of Sciences and Arts and many other relevant stakeholders²⁰¹.

Legal and ethical framework

There is growing awareness about the problem and about the need to provide clarification on legal aspects, but no organised support is given to research community to facilitate data sharing. This is understandable since there are no policies related to data management and sharing. Anyway, all the relevant laws are existing in Montenegro, and some issues related to data sharing and preservation are identified in these laws²⁰².

²⁰⁰ Comment from the respondent: "I really do not know for some other similar activities regarding growing awareness DMP. In the other hand, through interviews held with key stakeholders through SEEDS project, I concluded that the idea of creation of DMP was, at this moment, unknown in Montenegro. Through SEEDS project, CEMI has promoted these goals."

²⁰¹ SEEDS (2016). D3 – Analysis of existing potentials for the establishment of a social science digital data archive in Montenegro. http://seedsproject.ch/wp-content/uploads/2015/06/Report-for-Montenegro-SEEDS_bk-CEMI-Final_rk1-August-2016_FINAL.pdf

²⁰² Ibid., p. 7

Data sharing culture

Prevalence of data sharing and reuse

The estimate of proportion of social science researchers who have shared their research was on the medium level (10-30%) in our self-assessment survey, the same as the estimate of proportion of researchers who have been able to access existing third party data.

Existing routines of data sharing

Current practices are that 19% of researchers, as shown in SEEDS Survey, stored their data in a data archive/repository, while the vast majority of them stored the data on their own or their colleague's personal computers. According to the same survey results, the majority of research data is not shared further than the research team members (51%), but the data are also shared between members of the same institution (21%), and with broader scientific community (23%). When asked about sources where they obtain data for secondary analysis, researchers reported network of colleagues inside and outside their research unit as the most common source, followed by websites of projects, data archives from other countries, and National Statistical Office. In our self-assessment survey, it was reported that the researchers are using the following routines for data sharing: project or personal websites and informal contacts (peers and colleagues), and supplementary data in journals.

The SEEDS survey also identified the barriers to conducting secondary analysis in Montenegro. "The data exist but are poorly documented and unusable", was the most common answer (43%) given in the survey, followed by "researchers are not trained well enough in secondary analysis" (35%) and "data exist but are not accessible" (32%). However, some researchers think that "not enough data exists" (27%) and that secondary analysis "is not part of the research culture" (17%).

Attitudes towards data sharing

From the results of our self-assessment survey we can say that data sharing is perceived as having benefits for social scientists, that data sharing creates healthy, not negative competition in research, and that reuse of research data can answer new research questions and facilitate advancement of science. In the SEEDS Survey, it was shown that most of the researchers perceived considerable benefits to their research and teaching from better access to data produced in Montenegro and elsewhere, and that majority of them consider sharing data important within their discipline.

One of the reason for the low proportion of researchers sharing their data publicly could be the lack of appropriate infrastructure for data archiving in the country. The results of the SEEDS Survey showed that the majority of respondents (84%) had a positive attitude about sharing their data in a social science data archive, provided that their data are preserved for the long term in a secured environment, and 60% of respondents answered that the ideal access to their data should be public or available to the broader scientific community.

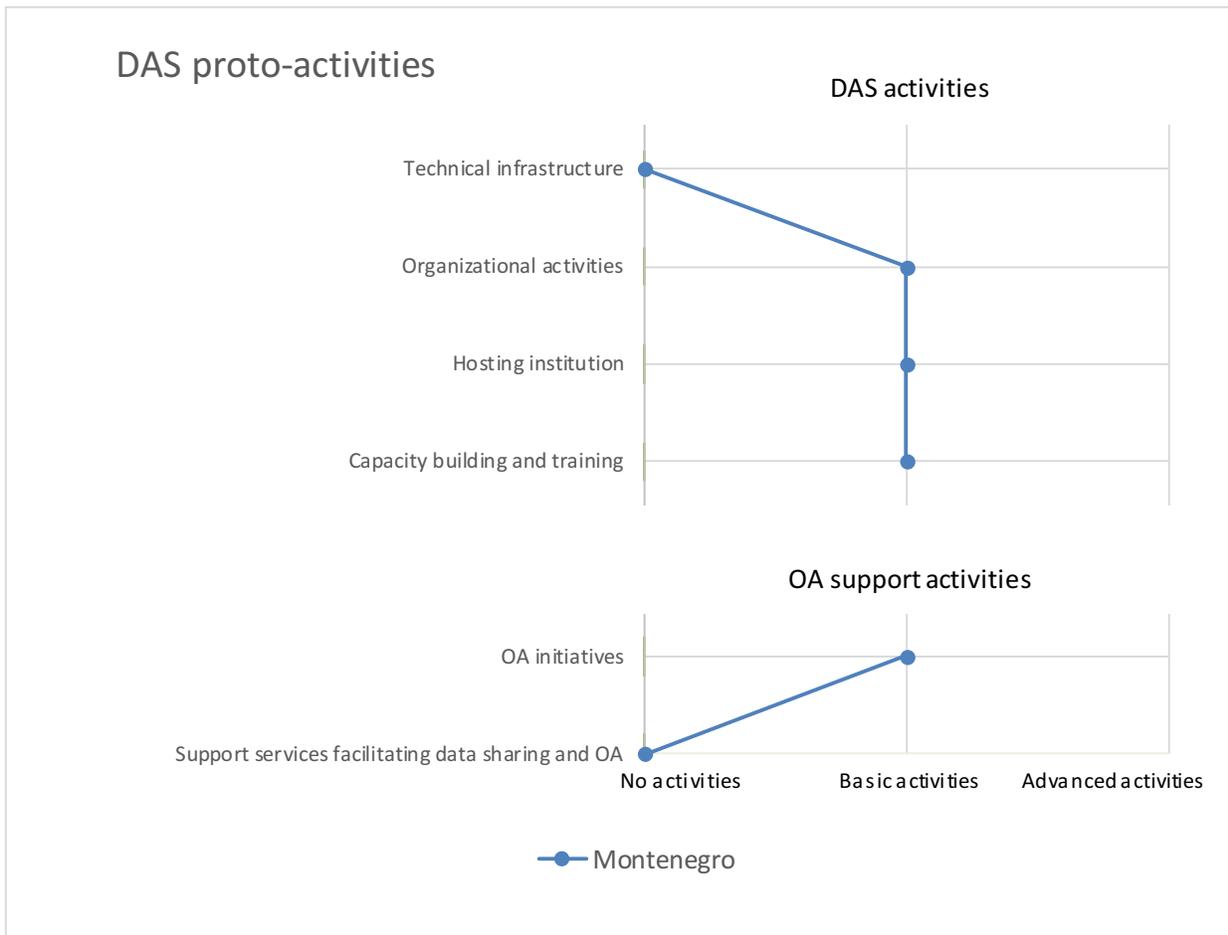
Enablers for data sharing

Based on the self-assessment survey results, there are no career rewards related to data sharing in the academic community of Montenegro, and no support services provided to social science

researchers that facilitate data sharing and/or Open Access to research data. According to the self-assessment survey, some data producers do follow management and data documentation standards and procedure, but the SEEDS Survey showed that this practice very rare.

3.29.2 Data archive service (DAS) proto-activities

Figure 61: DAS and OA activities implementation type in Montenegro



DAS activities

Some notable efforts towards establishment of data services for social sciences in Montenegro has been made during the SEEDS project²⁰³ coordinated by FORS - Swiss Centre of Expertise in the Social Sciences. Partner from Montenegro in this project was the Centre for Monitoring and Research (CeMI). The SEEDS project tackled most of the issues related to the process of establishment of data services for social sciences such as policy and governance, human resources and technical issues. It also helped to promote the importance and benefits of data sharing among relevant stakeholders - researchers and policy makers.

²⁰³ SEEDS - South-Eastern European Data Services. <http://www.seedsproject.ch>

There is currently no established data service for the social sciences or for other disciplines in Montenegro. No specific technical infrastructure exists that could be used for or applied to a new data services in social sciences or in other scientific disciplines.

A suitable institution that could host the future data service could be Ministry of Science and/or Montenegrin Academy of Science and Art, as recognised through the SEEDS project.

Through the SEEDS project, CeMI staff did gain knowledge, skills and experience, which should be used to influence future developments in the area of social science data management and preservation.

Open Access support activities

Open Access has not been widely promoted in Montenegro. Only two projects that promoted open access and open data were identified. One of them is the SEEDS project that has been dealing with open access to research data. Another one is the project in the area of open data HOMER - Harmonising Open Data in the Mediterranean through Better Access and Reuse of Public Sector information²⁰⁴.

3.29.3 Conclusions

- Social sciences sector in Montenegro is still in development.
- Funders' data management and sharing strategy and policies are currently non-existent in Montenegro. Through the SEEDS project awareness have been raised about specific issues related to RDM policies, so we can say that the strategy and policy to enable data sharing has started to develop together with the ethical and legal framework important for data sharing.
- Data sharing culture is developed in Montenegro, but enablers of data sharing are still underdeveloped.
- DAS proto-activities exist in Montenegro through the involvement in the SEEDS project. Open access and open data initiatives are rare.

²⁰⁴ HOMER project. <http://www.homerproject.eu/>

3.30 The Netherlands

All in all, the Netherlands has a very high maturity level. The country is in general above the average. The Netherlands scores high in relation to both the broader eco-system of DAS-operation and the capability requirement areas of DAS.

This high maturity level reveals a lot of strengths in relation to the broader eco-system, such as participation in all major (inter)national projects and having requirements/recommendations for DMP's. There is a potential for further recommendations/guidance for ethical aspects in relation to legal framework for data sharing and for developing the data sharing culture. A barrier for developing the last can be the attitude of researchers; it might be difficult to move the researchers toward data sharing.

The high maturity level also reveals a lot of strengths in relation to the capability requirement areas. There is for example an internal policy document about identity and access to data, while internal software development and maintenance is monitored by management plans and quarterly reports. DAS has received certificates of DSA, WDS as well as the NESTOR seal (DIN 31644).

The conclusion is that the Netherlands is doing well and will continue to develop in the future.

Specifics about the data collection in a country

The data collection started 21st September 2016 and was finished 31st October 2016. The information was provided by Data Archiving and Networked Services (DANS). DANS is an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO)²⁰⁵.

The data collection process went well without any problems.

²⁰⁵ See: <https://dans.knaw.nl/en/about/organisation-and-policy/organisation-and-policy>

3.30.1 Broader ecosystem of DAS operation

Figure 62: Heading concepts values in The Netherlands

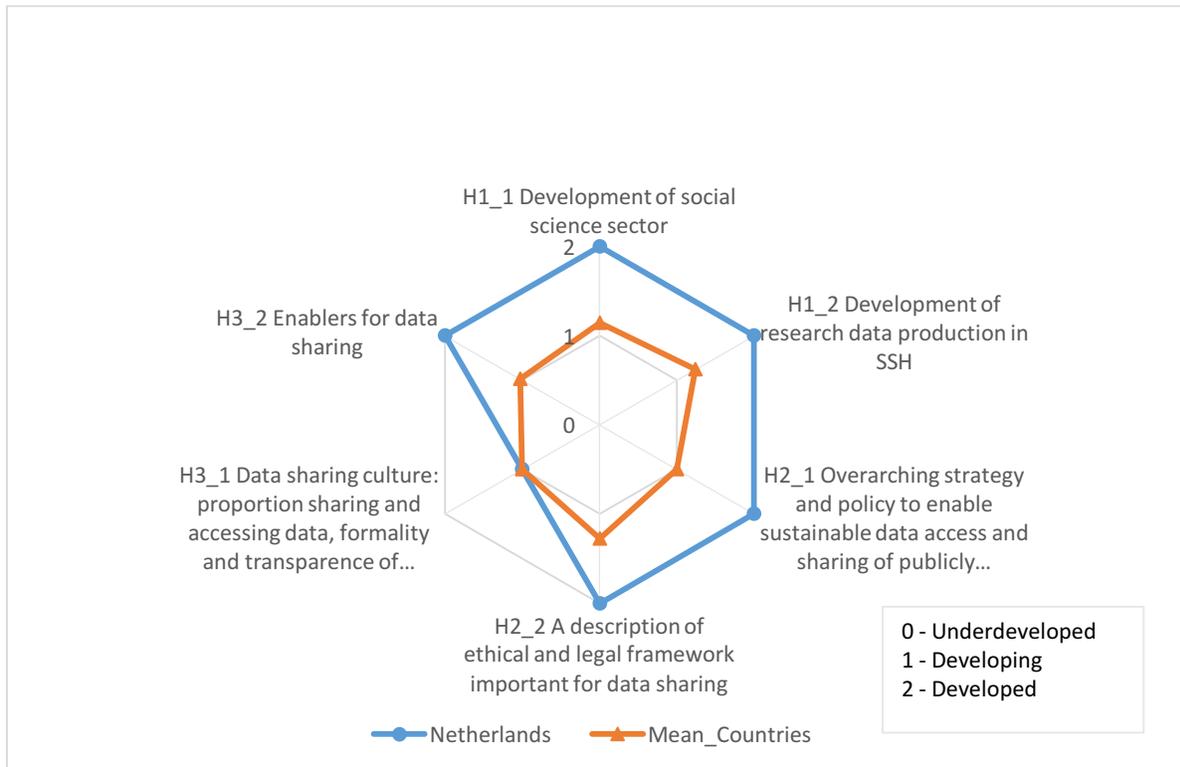


Figure 62 shows that the Netherlands is on the developed level at all areas with exception of 'Data sharing culture'. Here is the country on the developing level.

Development of social science sector

Overall assessment of SSH development

The financial stability, research capacities and results achieved in the field of social sciences are in the highest quantile (on the developed level) and above the average.

Every University Library in the Netherlands has contracts with commercial bibliographic and full-text databases and some of them provide access to commercial data sets. All the universities and other higher education institutes are affiliated with SURF, the collaborative ICT organisation for Dutch education and research. SURF provides national software licenses²⁰⁶. DANS is the national representative of ICPSR and provides by this membership free access to ICPSR data for all universities. Every university can obtain remote access to data of Statistics Netherlands (CBS); costs are involved for the connection to CBS, not for the data itself. DANS gives access to scientific use files of Statistics Netherlands.

²⁰⁶ See: <https://www.surf.nl/en/services-and-products>

Development of research data production in SSH

The research data production in SSH is on the developed level and above the average.

The Social Science scholars in the Netherlands have a longstanding tradition in survey research. They are initiator or take part in all major international and national survey projects, like:

- ESS <http://www.europeansocialsurvey.org/>
- ISSP <http://www.issp.org/>
- EVS <http://www.europeanvaluesstudy.eu/>
- SHARE <http://www.share-project.org/>
- PISA <https://www.oecd.org/pisa/>
- GGP <http://www.ggp-i.org/>
- HBSC <http://www.hbsc.org/>

Examples of other important survey projects are:

- LISS <http://www.lisspanel.nl/website/>
- DPES <http://www.dpes.nl/en/>
- TRAILS <https://www.trails.nl/en/home>
- NTR <http://www.tweelingenregister.org/en/>
- Dynamics of youth <http://www.uu.nl/en/research/dynamics-of-youth/organisation>

There is a frequent production of research data by the social science institutions. The institutions have well established tradition in data production. In the online archiving system, EASY contains for the various sub disciplines in the Social Sciences the following amount of data sets:

- Behavioural and educational sciences: 1237 datasets
- Economics and Business Administration: 221 datasets
- Law and public administration: 785 datasets
- Social sciences: 4565 datasets

One should also consider the fact that data from international projects are often archived outside the Netherlands, for example ISSP, ESS and EVS.

RDM Policy setting

Funders' data management and sharing strategy and/or policy

The overarching strategy and policy to enable sustainable data access and sharing of publicly financed social science research data is on the developed level and above the average.

The requirements or recommendations for DMPs in most cases are defined:

- Requirements of the Netherlands Organisation for Scientific Research (NWO): <http://www.nwo.nl/en/policies/open+science/data+management>
- Guidance: <http://www.nwo.nl/en/policies/open+science/data+management+chapter>

- NWO recommends: checklist by RDNL: <http://www.researchdata.nl/en/services/data-management/>.
- In the Netherlands there is also the National Coordination Platform LCRDM: <https://www.surf.nl/en/lcrdm>.
- DANS is involved in the OpenAire project, <https://www.openaire.eu/89-elly-dijk> with a focus on Open data and Data Management Planning: <https://www.openaire.eu/opendatapilot>.

The way of depositing data in an appropriate disciplinary repository is defined (formal requirement). NWO states that *“the data should preferably be archived at a national or international data repository. If that is not possible, the data should be archived by the institutional repository.”* So there is no requirement or recommendation about discipline-specificity²⁰⁷.

In relation to data appropriate for data curation, there is a requirement to assess those. It is understood that the best use of resources involves making choices based on value judgements and selecting material for curation. Investment and support for long-term curation is in place, based on contractual arrangements.

- Information on data contracts between DANS and depositors²⁰⁸
- *“The importance and the value of reuse on the one hand and the costs and feasibility of data storage on the other should be in reasonable balance with each other and have a bearing on the volume of the data to be stored.”*

The public research funding organizations in the Netherlands provide, on the partial level, incentives for sharing research data with associated metadata. There is explicit recognition that additional cost for preparing the data for access are legitimate project cost that can partially cover the RDM cost up to a certain limit. On NWO site²⁰⁹ it is stated that costs are covered for the data that are suitable for reuse (costs are not fully covered).

Legal and ethical framework

The ethical and legal framework important for data sharing is placed at a defined level and above the average in Figure 57. However, the correct ranking should be between ‘Partial’ and ‘Defined’.

There are recommendations and guidance provided on how to respect the legal requirements while sharing data but this does not exist for ethical aspects (mentioned at the ‘Defined’ state).

- There is a code of conduct on the use of personal data in science that is based on the law from the Association of universities in the Netherlands (VSNU)²¹⁰
- We are not aware of a specific code of conduct on ethical aspects in the social sciences
- There is a more general code of conduct for all scientific disciplines (VSNU)²¹¹
- Other documents, published by the Royal Academy of Sciences²¹²

²⁰⁷ See <http://www.nwo.nl/en/policies/open+science/data+management+chapter>

²⁰⁸ See: https://dans.knaw.nl/en/deposit/information-about-depositing-data/data-contract/data-contract?set_language=en and <http://www.nwo.nl/en/policies/open+science/data+management+chapter>

²⁰⁹ See: <http://www.nwo.nl/en/policies/open+science/q+and+a#kosten>

²¹⁰ See: <http://www.vsnu.nl/code-pers-gegevens.html>

²¹¹ See:

http://www.vsnu.nl/files/documenten/Domeinen/Onderzoek/Code_wetenschapsbeoefening_2004_%282014%29.pdf

²¹² See https://www.knaw.nl/en/news/publications/zorgvuldig-en-integer-omgaan-met-wetenschappelijke-onderzoeksgegevens?set_language=en

Data sharing culture

The data sharing culture in the Netherlands is – like the average – on the developing level.

There are no estimates for the number of scholars in the Netherlands that share their data. However, there is a difference in sharing research data among disciplines. According to self-assessment, scholars in quantitative Political Sciences and Sociology are more accustomed to sharing their research data and to use research data of others than in other disciplines.

The proportion of researchers able to access existing third party data they need is estimated in self-assessment as high (>30%). It can be explained by the fact that DANS is well known among the scholars in the Social Sciences and the Humanities (DANS and its predecessors have existed more than 50 years). In addition, the use of data from the DANS services is free of charge (<https://easy.dans.knaw.nl/ui/home>). Via DANS researchers can access the data of ICPSR (<https://www.icpsr.umich.edu/icpsrweb/>). Researchers from universities can also make use of the data of Statistics Netherlands (CBS) via remote access²¹³.

There are established data sharing channels and routines. As self-assessment indicates, most popular data sharing channels include formal and transparent channels, with data archive ranked first. Data is also shared via informal contacts (ranked second), project websites (ranked third) and supplementary data in a journal (ranked fourth). There are not any exact figures for evidence and the ranking is therefore based on DANS experience.

The attitudes of researchers toward data sharing are depending on the research background and discipline. There are researchers who only make use of survey data collected by international survey programmes like ESS, SHARE, EVS and ISSP. These researchers advocate the advantages of data sharing. Other researchers, often more qualitative oriented, only collect their own data.

Table 22: Attitudes towards data sharing in The Netherlands

Data sharing has no benefits at all	False
Data sharing creates healthy competition	Neither true, nor false
Data sharing creates negative competition	Neither true, nor false
Reuse of existing data can answer new research questions and facilitate advancement of science	Neither true, nor false
Data sharing has as a risk that others may misuse and misinterpret data	Neither true, nor false
Data sharing involves little effort and minimal costs	False

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from “5”-True to “1” False!

The scores indicate that data sharing in the Netherlands is beneficial but resource demanding. It is difficult to assess the consequences – good as well as bad – in relation to data sharing.

²¹³ See <https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research>

Enablers for data sharing

Enablers for data sharing are on the developed level and above the average.

There are to some extent career rewards related to data sharing in the academic community in the Netherlands. There are researchers whose articles are more cited when the associated data are publicly available, according to the self-assessment. Because citations are important for career opportunities of researchers it can be stated that data sharing is career rewarding. Also, data sharing is an obligation for the high profile longitudinal surveys to get continued funding. Implicit this influence the career opportunities of researchers.

DANS provides guidance and archiving services based on the front office/back office model. Research institutes are the first point of contact for the researchers, but DANS provides information if needed as well. EASY is the online self-archiving system by DANS for researchers upon completing their research²¹⁴. DANS provides deposit instructions for various research disciplines. It is not designed solely for SSH data but also for other disciplines²¹⁵.

- DataverseNL is the Dataverse network in the Netherlands. With DataverseNL, researchers and lecturers can store, share and register research data online, both during research and for up to ten years afterwards. DataverseNL is a shared service jointly offered by over 10 participating institutions and managed by DANS. DataverseNL is built upon the software application Dataverse developed by Harvard University. As of January 2016, DataverseNL offers access to more than 300 published studies (<https://dataverse.nl/dvn/>).
- SurveydataNL: The goal of the Survey Data Netherlands website is to bring survey data and metadata from all different sources into one website. In addition to the website, Survey Data Netherlands provides a one-stop service for research projects for the dissemination and long-term preservation of survey data (<http://www.surveydata.nl/>).
- The Essentials 4 Data Support course is offered by DANS, 4TU and SURFsara, as part of the Research Data Netherlands alliance. The introductory course 'Essentials 4 Data Support' offers data supporters the know-how in supporting researchers in storing, managing, archiving and sharing research data, in order to foster professionalization and cooperation among data supporters²¹⁶.
- The Research Data Journal for the Humanities and Social Sciences (RDJ) is a peer-reviewed journal, which is designed to comprehensively document and publish deposited datasets and to facilitate their online exploration. RDJ is e-only and open access, and focuses on research across the Social Sciences and the Humanities²¹⁷.
- DANS provides background information on RDM and DMP²¹⁸: a NARCIS - National Academic Research and Collaborations Information System (www.narcis.nl). NARCIS is the main national portal for those looking for information about researchers and their work. Besides researchers, NARCIS is also used by students, journalists and people working in educational and government institutions as well as the business sector.

²¹⁴ See: <https://easy.dans.knaw.nl/ui/deposit>

²¹⁵ See: <https://dans.knaw.nl/en/deposit/information-about-depositing-data>

²¹⁶ See: <http://datasupport.researchdata.nl/en/about-the-course/>

²¹⁷ See <http://www.brill.com/products/online-resources/research-data-journal-humanities-and-social-sciences>

²¹⁸ See: <https://dans.knaw.nl/en/deposit/information-about-depositing-dat>

Finally, it should be mentioned, that some data producers follow data management and data documentation standards and procedures²¹⁹.

3.30.2 Capability requirement areas of DAS

Organisational profile

Organisation

Data Archiving and Networked Services (DANS) is located in The Hague, The Netherlands.

DANS is part of the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO).

KNAW: As the forum, conscience, and voice of the arts and sciences in the Netherlands, the Academy promotes quality in science and scholarship and strives to ensure that Dutch scholars and scientists contribute to cultural, social and economic progress. As a research organisation, the Academy is responsible for a group of outstanding national research institutes. It promotes innovation and knowledge valorisation within these institutes and encourages them to cooperate with one another and with university research groups.

NWO: The Netherlands Organisation for Scientific Research ensures quality and innovation in science and facilitates its impact on society. Its main task is to fund scientific research at public research institutions in the Netherlands, especially universities. NWO focuses on all scientific disciplines and fields of research. The funds are allocated by means of a national competition on the basis of quality and independent assessment and selection procedures. NWO plays several roles as a broad, national research organisation that actively contributes to various elements of national science and innovation policy.

In 2016 DANS the staff consists of 42 full-time equivalent in total and 8,4 full-time equivalent for the Archive department. DANS has six departments: Archive, Software development, Policy & Communication, Research and Innovation, Project Support, and Office Support²²⁰.

The designated communities of DANS consist of scholars in the humanities and the social sciences. Initiatives in the life sciences and medicine are being explored.

DANS's primary target group includes scientific researchers and public and private organisations that carry out research commissioned by the government. They are both data producers and data users. Within the federative infrastructure (Front Office - Back Office Model), DANS also focuses on research organisations for acquiring data. Data with a scientific significance collected by companies are also welcome. Researchers in training (PhD students) belong to this primary target group as well. DANS has, in collaboration with the Cultural Heritage Agency (Rijksdienst voor het Cultureel Erfgoed), a special task of managing archaeological research results. Besides scientific researchers and research organisations, companies and the interested public are welcome to use DANS services.

²¹⁹ DDI is a metadata standard that facilitate data reuse. There is one organization in the Netherlands that makes use of DDI; CentERdata (<http://www.lissdata.nl/dataarchive/>). For the majority of researchers DDI is too complicated, most people document their data very elaborate in PDF codebooks.

²²⁰ Information about the organizational chart is available at: <https://dans.knaw.nl/en/about/organisation-and-policy/organigram>.

Funding

DANS is an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO). In order to fulfil her task to provide sustained access to digital research data in the DANS Archive, DANS receives structural lump sum financing from KNAW and NWO (Public funding). This covers 2/3 of the annual budget. The other 1/3 comes from international and national projects²²¹.

Core services and activities

DANS promotes sustained access to digital research data. For this, DANS encourages scientific researchers to archive and reuse data in a sustained form, for instance via the online archiving system EASY and DataverseNL. With NARCIS, DANS also provides access to thousands of scientific datasets, publications and other research information in the Netherlands. The institute furthermore provides training and consultancy and carries out research on sustained access to digital information. Driven by data, DANS ensures the further improvement of access to digital research data with its services and participation in (inter)national projects and networks (<https://dans.knaw.nl/en/about>).

Content current collection

Content of collection per discipline at 31st October 2016:

- Behavioural and educational sciences - 1237 datasets
- Economics and Business Administration - 221 datasets
- Humanities - 32051 datasets
- Interdisciplinary sciences- 149 datasets
- Law and public administration - 785 datasets
- Life sciences, medicine and health care - 402 datasets
- Science and technology - 83 datasets
- Social sciences - 4565 datasets

Daily updates are available at <https://easy.dans.knaw.nl/ui/browse>.

The archive's holdings consist of data from various disciplines:

- Data from the humanities domain: texts, spreadsheets, databases, images, transcripts (including time based transcripts), audio and video.
- Data from the social sciences: mainly quantitative (statistical) data, questionnaires, codebooks, test responses and some qualitative datasets which include field notebooks, interviews and interview transcripts.

²²¹ Strategy Policy 2015-2020 (in Dutch), available at <https://dans.knaw.nl/nl/over/organisatie-beleid/informatiemateriaal/DANSstrategienotaNL.pdf>.

English summary (including a DANS Budget Plan for the multi-annual budget 2014-2020), available at <https://dans.knaw.nl/en/about/organisation-and-policy/information-material/DANSstrategienota20152020UK.pdf>

- Archaeological data: reports, data from excavations and trial trenching and post excavation analysis: texts, databases, spreadsheets, GIS files and images (vector and raster images).
- Data from geospatial sciences: vector graphics, CAD drawings, GIS files.
- Data from the life sciences and health research: statistical data, spreadsheets, databases, laboratory notebooks and texts.
- Other data: Linked Open Data, models, algorithms, scripts, and executables.

The main language of the data and metadata collection is Dutch and English.

Organisational infrastructure

Figure 63: Organizational infrastructure in The Netherlands

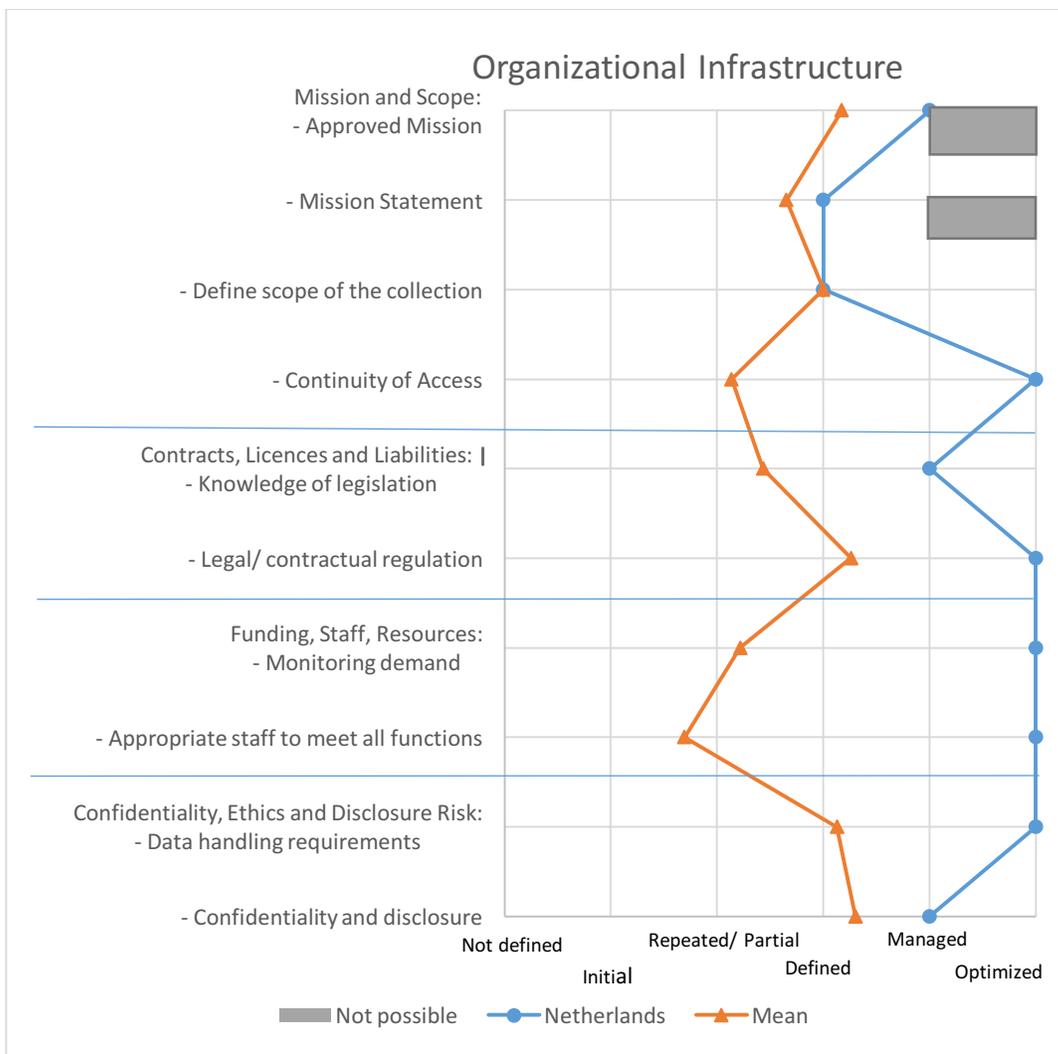


Figure 63 shows that the organizational infrastructure in general is on the managed or optimized level and above the average.

Mission statement

The mission of DANS is to promote sustainable access to digital research data. For this, DANS encourages scientific researchers to archive and reuse data in a sustained form, for instance via the online archiving system EASY and DataverseNL. With NARCIS, DANS also provides access to thousands of scientific datasets, publications and other research information in the Netherlands. The institute furthermore provides training and consultancy and carries out research on sustained access to digital information. Driven by data, DANS ensures the further improvement of access to digital research data with its services and participation in (inter)national projects and networks.

DANS is an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO) and is governed by the "Cooperation Agreement (Samenwerkingsovereenkomst) DANS" between NWO and KNAW from 2005, which was updated in January 2015. These organizations commit themselves to the following (unauthorised translation): *"In the event of termination of the agreement NWO and KNAW will make arrangements regarding the discontinuity or modified continuity of DANS. In that case KNAW and NWO will accept in particular the responsibility to store the data files archived at DANS elsewhere in a most responsible manner and under equivalent technical conditions."* The Collaboration Agreement is available on request. In DANS strategy policy 2015-2020, the mission, target groups and strategy of DANS are described²²².

Knowledge of legislation is on the managed level. One of the policy advisors of DANS is also the legal officer. His explicit task is to monitor all the development in (inter)national legislation concerning all aspects of data archiving and disseminating. He reports new developments to the management of the data archive. This information is documented on the internal DANS site, externally on the page on legal information of the DANS website²²³.

Legal/contractual regulation is on the optimized level. Both licenses and conditions of use are not infringing any intellectual property rights. These rights are not transferred to DANS except for datasets deposited under the CCO Waiver license where all rights are waived. This is explained on the page on legal information of the DANS website²²⁴.

In regards to monitoring demand, it is also on the optimized level.

The same is the case concerning appropriate staff to meet all functions. There is enough appropriate staff available. Updates job descriptions exist which set out the required qualifications of the digital repository personnel and contain an organizational chart and/or staff development plan based on the tasks and objectives of the digital repository. Recently, DANS has revised its system of job descriptions. The following job categories pertain to the DANS Archive in a strict sense:

- Manager (Dutch: coördinator) Archive
- Data manager: junior, mediro, senior
- Preservation officer
- Technical archivist

²²² English summary available at: <https://dans.knaw.nl/en/about/organisation-and-policy/information-material/information-material>

Full text (in Dutch) available at: <https://dans.knaw.nl/nl/over/organisatie-beleid/informatiemateriaal/DANSstrategienotaNL.pdf>

²²³ See: https://dans.knaw.nl/en/about/organisation-and-policy/legal-information?set_language=en

²²⁴ See above

- Manager (Dutch: coördinator) Software development
- Software developer: junior, medior, senior
- Systems information management (Dutch: functioneel beheerder): regular and senior
- IT management (Dutch: technisch beheerder): junior, medior, senior
- Application management (Dutch: applicatiebeheerder)
- Security officer
- Legal advisor

Most of the people fulfilling these jobs belong to the sub departments Archive and Software development. Other sub departments are Policy & Communication, Research and Innovation, Project Support, and Office Support²²⁵.

DANS uses Personal Development Plans (POP: persoonlijk ontwikkelingsplan) for the development of the staff.

Data handling requirements are on the optimized level. The licenses as well as the conditions of use are based upon the relevant legislation: relevant laws and codes of conduct are referred to in these two contracts²²⁶.

Finally, the confidentiality and disclosure are on the managed level. The handling of personal data, data depositors want to deposit within the data archive is “work in progress”. Internally, a data risk classification has been developed, aimed mainly at the security of personal data. All staff members are obliged to sign a “declaration of confidentiality”²²⁷.

Digital object management (data curation) and Technical infrastructure and risk

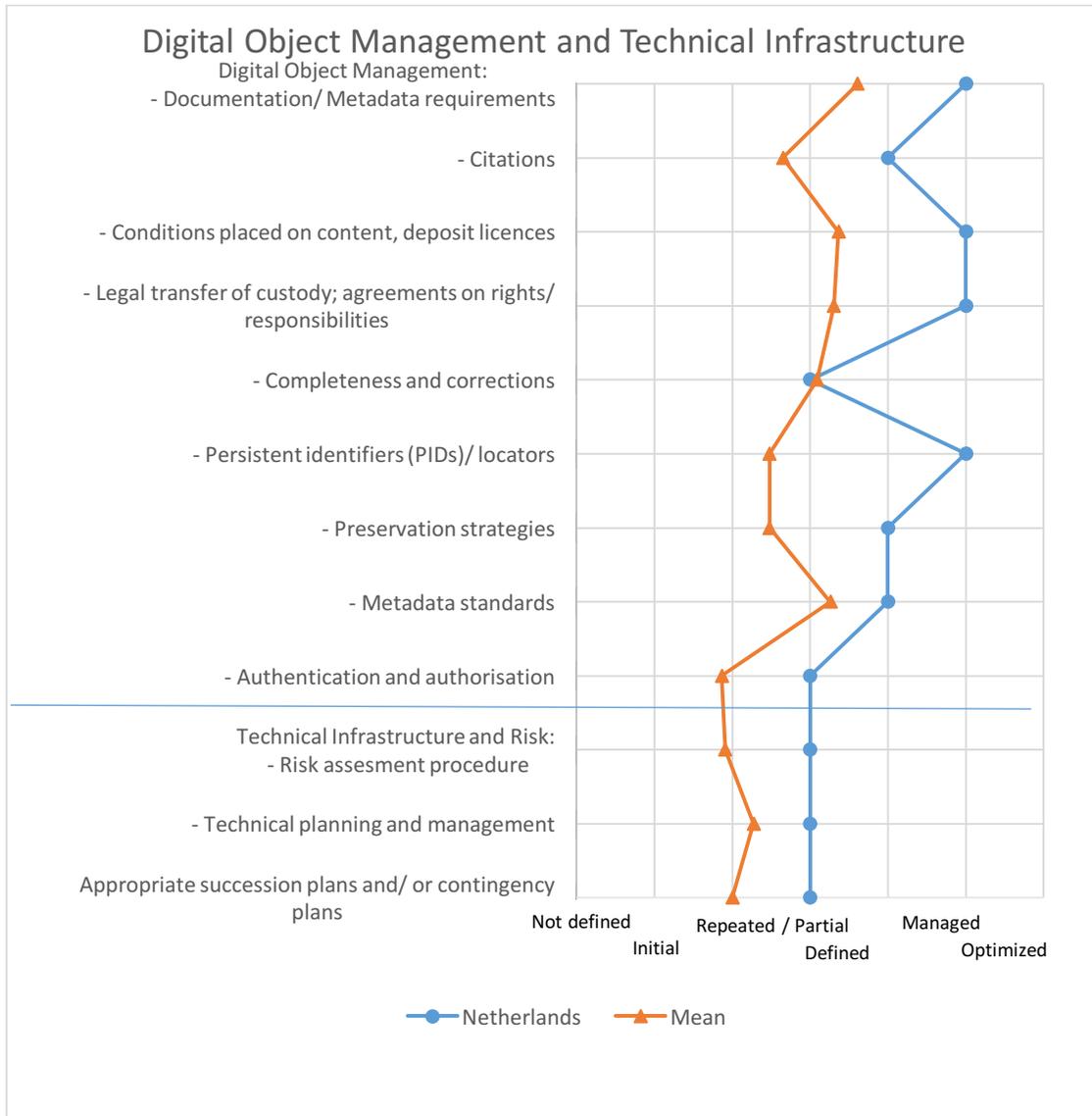
Figure 59 shows that digital object management (data curation) in general is on the managed or optimized level and above the average. The technical infrastructure and risk is defined and above the average.

²²⁵ The organizational chart of DANS, available at: <https://dans.knaw.nl/en/about/organisation-and-policy/organigram>.

²²⁶ See: <https://dans.knaw.nl/en/about/organisation-and-policy/legal-information>

²²⁷ See: https://dans.knaw.nl/en/about/organisation-and-policy/legal-information/PrivacyreglementDANSUK_april09versie2.82.pdf

Figure 64: Digital object management and Technical infrastructure in The Netherlands



Digital object management (data curation)

Completeness and corrections is on the defined level.

- Information about data management procedures²²⁸.
- Information on preferred formats²²⁹
- DANS is DSA, WDS as well as DIN 31644 (nestorSeal) certified²³⁰

²²⁸ See: <https://dans.knaw.nl/en/deposit/information-about-depositing-data/dans-provenance-document-uk.pdf>

²²⁹ See: <https://dans.knaw.nl/en/deposit/information-about-depositing-data/DANSpreferredformatsUK.pdf>

²³⁰ See <https://dans.knaw.nl/en/deposit/information-about-depositing-data/about/organisation-and-policy/certification>

Authentication and authorization is also on defined level. There is an internal policy document about identity and access to data (part of the wider policy of the Royal Netherlands Academy of Sciences and Arts).

Technical infrastructure and risk

Technical planning and management is on defined level. Data storage management has been outsourced. The Archive has a Service Level Agreement (SLA) with its data storage management provider, which includes a confidentiality statement. Storage usage and costs are monitored on a quarterly basis. Software is partial developed by DANS itself and partial by third parties (Open Source software). Internal software development and maintenance is monitored by management plans and quarterly reports. Only internal documentation is available.

In relation to appropriate succession plans and/or contingency plans, it is on the defined level. DANS is an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organization for Scientific Research (NWO). In the event of termination of the agreement, NWO and KNAW will decide regarding the discontinuity or modified continuity of DANS. In that case, KNAW and NWO will accept the responsibility to store the data files archived at DANS elsewhere in a most responsible manner and under equivalent technical conditions. DANS is DIN 31644 certified (<https://dans.knaw.nl/en/deposit/information-about-depositing-data/about/organisation-and-policy/certification>, see C12).

Finally, the risk assessment procedure is on the defined level. There is biannual risk assessment in collaboration with SURF²³¹.

3.30.3 Conclusions

The self-assessment was conducted by Data Archiving and Networked Services (DANS) and the data collection went very well. The results show that the Netherlands in general has a very high maturity level.

In relation to the broader eco-system of DAS operation, the Netherlands is on the developed level in the development of social science sector. The scholars take part in all major (inter)national projects. The same is the case in relation to the RDM Policy setting, which is assessed on the developed level. It is argued that there in most cases are requirements or recommendations for DMP's and that DANS is involved in the OpenAire project with a focus on open data and DMP. It is stressed that there is missing recommendations/guidance for ethical aspects (DANS has recommendations/guidance in relation to legal framework for data sharing). The data sharing culture is on a developing level. Here the researchers can access existing third party data and use data from the DANS service free of charge. Furthermore, there are established data sharing channels and routines. The attitude toward data sharing is however depending on research background and discipline. In other words, there is room for development but maybe it can be difficult to move the attitudes of researchers toward data sharing. Regarding enablers for data sharing there are to some extent career rewards related to data sharing because it is important for career opportunities of researchers – and because data sharing is needed to get continued funding in relation to high profile longitudinal surveys.

²³¹ See: <https://www.surf.nl/en/services-and-products/surfaudit/index.html>

In relation to the capability requirement areas of DAS, the organizational profile shows that DANS in 2016 had 42 full-time equivalent in total and 8,4 full-time equivalent for the archive department. DANS' primary target group are scientific researchers and public and private organisations that carry out research commissioned by the government. Overall, DANS promote access to digital research data, datasets, publications and other research information, provides training and consultancy, carries out research, and finally ensures further improvement of access to data. Both the organizational infrastructure and digital object management (data curation) is in general on the managed or optimized level and above the average. Finally, the technical infrastructure and risk is on a defined level and above the average.

3.31 Norway

In Norway, the broader ecosystem of DAS operations is on a high maturity level. Social sciences sector, data production in social sciences, overarching strategy and policy framework as well as ethical and legal for data sharing are developed above average. The Research Council of Norway requests all recipients of social science research funding to deposit, store and share data through Norwegian Centre for Research Data. NSD is Data protection official playing an important role in dissemination of knowledge on ethical and legal guidelines regulating research.

Data sharing culture and enablers for data sharing are developing. NSD provides services that facilitate data sharing and OA to research data. A comprehensive 5-year project NORDi started in 2016 with the aim to strengthen cooperation with research community regarding different aspects of data sharing.

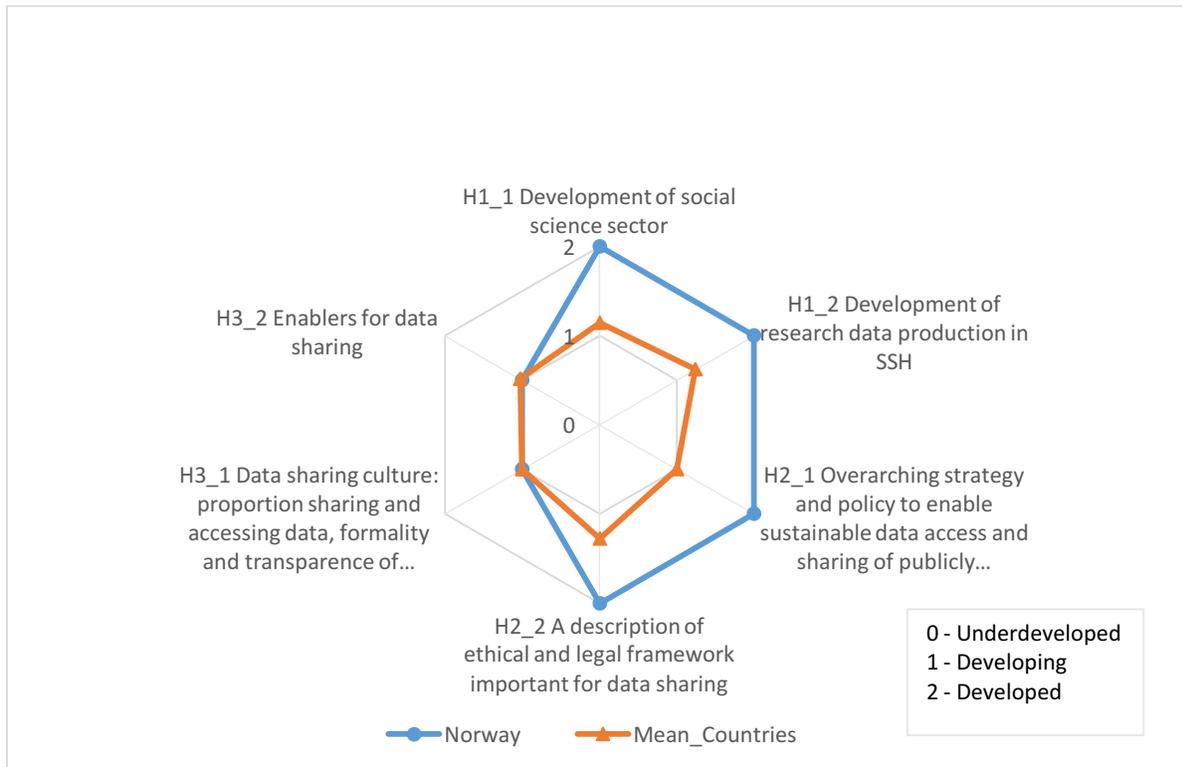
NSD is on a very high maturity level with regards to organizational infrastructure, being a result of 45 years of sustainable growth. Digital object management and technical infrastructure are also on a very high maturity level except for a few areas. However, these are under development. Further development of digital object management and technical infrastructure areas are also included in NORDi.

Specifics about the data collection in a country

Data was collected mainly via self-assessment survey completed by two experts from NSD - Dag Kiberg and Vigdis Kvalheim. Where possible, it was supported by various documents and NSD's webpage.

3.31.1 Broader ecosystem of DAS operation

Figure 65: Heading concepts values in Norway



All dimensions of broader ecosystem of DAS operation in Norway are developed on the highest level (and above average in the audited ERA countries), except for dimensions of data sharing culture and enablers for data sharing (on the same level as average in the audited ERA countries).

Development of social science sector

In this section, basic profile regarding scientific development in Norway is described thus giving insight on general conditions influencing the DAS operation.

Overall assessment of SSH development

General intensity of investment in SSH-sector in Norway is above medium, if compared to other ERA countries audited – 0.24% (GERD in SSH as percentage of GDP) – only the Netherlands, Portugal and Iceland has a higher value. Intensity of investment in SSH (relative to other scientific disciplines) is rather low – 14.6% - above average as well.

Human resource potential in SSH in Norway can be estimated as high – with 224 researchers per 100.000 inhabitants it was 3rd highest in among the audited countries (Number of researchers (head count) in SSH per 100.000 inhabitants), and so was the investment in human resources in SSH – 82632, 2nd highest among audited ERA countries with data available .

Most important sources of funding in social sciences and humanities, based on self-assessment by a DAS expert, are ranked as follows: (1) government and higher education sector; (2) international and cross-border funding; (3) private non-profit sector, and, finally, (4) business enterprise sector.

The overall development of social science sector is on a high level (funding of SSH and productivity of the researchers are in the highest quantile and thus the expected impact on designated community is expected to be rather strong).

Development of research data production in SSH

There are well established traditions of international collaborative research or cross-national studies in Norway. The country has participated in Comparative Candidate Survey (CCS), rounds 2009 and 2013; European Social Survey (ESS); International Social Survey Program (ISSP); Comparative Study of Electoral Systems (CSES); European Values Study (EVS), World Values Survey (WVS); Programme for International Student Assessment (PISA), and Generations and gender programme (GGP) – Norway has at some point participated or is still participating in all eight international studies enlisted.

In addition to participation in above mentioned international studies, social scientists in Norway produce a range of systematic studies of national importance. Some examples from NSD's web catalogue include biannual public opinion survey Norwegian Citizen Panel (2013-2016)²³²; an annual Media Survey (conducted since 1999) studying development of media through the perspective of both the journalists and the population²³³; National Election Surveys²³⁴ as a part of Norwegian Election Research Programme covering elections and political processes in Norway (since 1957); Culture and Mass Media Survey aiming at mapping the use of different types of mass media, cultural facilities and the supply of entertainment (since 1991)²³⁵.

According to self-assessment, the average production of research data in the SSH institutions in Norway can be characterized as frequent - institutions have well established tradition in data production. Survey results of researchers across different disciplines in the end of 2013 in Norway confirm the self-assessment – on average, only 4% of researchers have indicated that they do not generate research data (ranging from 1% to 7% in different disciplines)²³⁶.

²³² NSD - Norwegian Centre for Research Data web catalogue, available at: http://www.nsd.uib.no/nsddata/serier/norsk_medborgerpanel_eng.html (last accessed 2017-02-09)

²³³ NSD - Norwegian Centre for Research Data web catalogue, available at: http://www.nsd.uib.no/nsddata/serier/medieundersokelsene_eng.html (last accessed 2017-02-09)

²³⁴ NSD - Norwegian Centre for Research Data web catalogue, available at: http://www.nsd.uib.no/nsddata/serier/norske_valgundersokelser_eng.html (last accessed 2017-02-09)

²³⁵ NSD - Norwegian Centre for Research Data web catalogue, available at: http://www.nsd.uib.no/nsddata/serier/mediebruksundersokelsene_eng.html (last accessed 2017-02-09)

²³⁶ DAMVAD 2014. Sharing and archiving of publicly funded research data. Report to the Research Council of Norway, pp 31. Available at: <http://www.damvad.com/wp-content/uploads/2016/01/satellite.pdf> (last accessed 2017-02-02).

Report analyses results of survey of researchers conducted in Norway, December 2013. The Research Information System CRISIn was used as a sampling frame for randomly selecting respondents aiming at representativeness within research disciplines in research institutes, universities and university colleges. In general, representativeness was achieved, with general response rate 31%. Response rate for social sciences researchers was between 28.6% in universities and university colleges to 41.2% in research institutes. As the survey questions had focus on aspects related to OA, sharing and reusing of research data, it might be that respondents are more positively biased towards the OA issues than general population, so the results should be interpreted with precaution.

So overall, research data production in Norway is developed, as there are well established traditions of producing both national and international research data.

RDM Policy setting

Overall incentives and high level policy requirements in Norway have reached a level that provides appropriate level of sustainable conditions for data service provision.

The Research Council of Norway (RCN) implemented its "Policy for Open Access to Data" in 2014²³⁷. The policy, formulated in the form of recommendations, sets out a requirement for DMP for projects funded by the RCN, and require digital data to be stored in a research data deposit service. For the social sciences, humanities and, some of the medical and health sciences and some of the natural sciences research is explicitly required that data is deposited, stored and shared through Norwegian Centre for Research Data (NSD). This is explicitly written in paragraph 8.2 in the General Terms and Conditions of R&D project contracts between the Research Council of Norway and recipient of research funding, and applies to projects with closing date for applications after 1. April 2015²³⁸. This contractual requirement, according self-assessment, is the strongest implementation mechanism of the Norwegian policy of Open Access.

Thus, two of the RDM policy setting requirements are on "Partial" level – there is an expectation to have a DMP in place, and regarding the costs for preparing the data for access, there is an explicit recognition that these are legitimate project costs. The two other requirements are on "Managed" level: a place of data deposit is formally defined in the requirements, and sanctions for not complying with regulation are in place, as well as requirements to assess, evaluate and select data for long term curation, regarding the reuse potential. Therefore, overall RDM policy setting in Norway can be defined as developed, as there is partially operationalized strategy and policy in place regarding the key aspects enabling data sharing.

Ethical and legal framework important for data sharing in SSH in Norway can be characterized as developed, as clarification and support on legal and ethical aspects of data sharing (IPR, data protection) is provided on "Managed" level, meaning that there is an organized service provided by Norwegian Centre for Research Data (NSD) to support and encourage legally and ethically sound data sharing practice. NSD is the Data Protection Official for Research for approximately 140 research and educational institutions, including all the Norwegian universities, university colleges, several hospitals, and a number of independent research institutions. The main task of Data Protection Official is to assist the institutions in fulfilling their statutory duties relating to internal control and assurance of their own research. An important part of this work is a preliminary assessment of research projects in accordance with the Personal Data Act and the Personal Health Data Filing System Act, and following up on project changes, extensions and project completion. Furthermore, NSD provides guidance and advice to institutions, researchers and students on issues regarding privacy protection, with a focus on sustaining needs of research within the regulatory framework.

²³⁷ The Research Council of Norway 2014. Policy for Open Access to Data, available at: http://www.forskningradet.no/en/Article/Open_access_to_research_data/1240958527698 (last accessed 2017-02-09)

²³⁸ The Research Council of Norway: General Terms and Conditions for R&D Projects, available at: http://www.forskningradet.no/en/Contract_and_reporting/1138882213515 (last accessed 2017-02-09)

Data sharing culture

Data sharing and reuse among social sciences researchers in Norway is estimated in self-assessment as very common –based on the number of projects funded by the Research Council of Norway (and thus bound to share data) the proportion of researchers sharing data is estimated as medium (10-30%). The proportion of researchers able to access existing third party data they need – is also estimated in self-assessment as high (above 30%).

Results of the survey of researchers across different disciplines conducted in Norway in the end of 2013 that was already mentioned before, confirm the estimates mentioned above. When asked about sharing research data they produce in 2013, 12% of social scientists indicated that data is available for all, 7% - that data is available for other researchers, 52% share their data upon demand, but 25% have indicated that their data is not available for reuse by others. As for reuse, 63% of surveyed social sciences researchers had reuse of other researchers' data within the last three years, though the general trend across disciplines had been that most reused data produced by other researchers in their institution; majority of those not having reused data produced by other, would like to do so in future²³⁹.

There are, as self-assessment shows, established channels and routines for sharing data in Norway. Not surprisingly, as there are RDM policy requirements, most often used data sharing channel is formal and transparent - data archive or repository is ranked first. Data is shared also via personal contacts (ranked second) and project or personal websites (third), that lack formality and transparency.

The attitudes of researchers towards data sharing, based on report on survey of researchers²⁴⁰ and self-assessment, can be characterized as mainly neutral or negative. Even though social sciences researchers in Norway are perceived in self-assessment to acknowledge general benefits of data sharing and to be neutral regarding benefits of data sharing for advancement of science, they are also perceived to fear misuse of their data and negative competition as a result of data sharing in research.

Attitudes towards data sharing

Table 23: Attitudes towards data sharing in Norway

Data sharing has no benefits at all	Probably false
Reuse of existing data can answer new research questions and facilitate advancement of science.	Neither true, nor false
Data sharing has as a risk that others may misuse and misinterpret data	Probably true
Data sharing creates negative competition (for example, being scooped and therefore reduced publication opportunities) for the researcher	Probably true
Data sharing creates healthy competition in research	False
Data sharing involves little effort and minimal costs	Unable to provide estimate for 2011-2016

²³⁹ DAMVAD 2014. Sharing and archiving of publicly funded research data. Report to the Research Council of Norway, pp 34-35; 46. Available at: <http://www.damvad.com/wp-content/uploads/2016/01/satellite.pdf> (last accessed 2017-02-02).

However, as the survey questions had focus on aspects related to OA, sharing and reusing of research data, it might be that respondents are more positively biased towards the OA issues than general population, so the results should be interpreted with precaution.

²⁴⁰ DAMVAD 2014:74.

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from "5"-True to "1" False!

The survey conducted among researchers in Norway at the end of 2013, supports the self-assessment results to some extent. Asked to name three most important barriers to sharing data, 38% named data sensitivity issues; almost a third – 29% - of surveyed social sciences researchers mentioned that preparing data for sharing is time consuming ("Making data available takes away valuable time for research"); 25% of surveyed social sciences researchers named concerns related to misinterpretation of their data ("I am afraid other researchers will not understand my data"), and 18% feared reduced possibilities of future scientific publications, and, finally, lack of technical infrastructure was mentioned as one of three most important reasons by 16%²⁴¹.

To summarize data sharing culture in Norway, data sharing and reuse is very common, there are established, both formal and informal channels for data sharing, but general attitudes of researchers related to sharing of data are neutral or negative, so in summary data sharing culture can be characterized as developing.

Enablers for data sharing

There are no career rewards related to data sharing the social sciences academic community in Norway, as self-assessment indicates. In the survey conducted among researchers in Norway at the end of 2013, social science researchers indicate potential enablers for data sharing (choosing three most important from a list): better technical infrastructure is named among three most important by 36%; implementation of system of citation – by 33% more resources to OA activities – 25%; more training -24%; implementation of guidelines for OA to research data – 20%²⁴².

There are data support services available to social science researchers that facilitate data sharing and/or Open Access to research data, provided by NSD. The survey conducted among researchers across all disciplines in Norway at the end of 2013, however, indicate that support available on institutional level regarding different aspects of OA to research data is perceived as insufficient²⁴³.

The services provided by NSD include support and training in data management, support to data preservation, and data access. Self-assessment indicates that there are also data producers that follow data management and data documentation standards and procedures that facilitate data reuse.

In summary, the incentives and enablers for data sharing within social science research community in the Norway can be benchmarked as developing.

²⁴¹ DAMVAD 2014: 51.

Difficult to align the statements directly with the items used in self-assessment, but in principle three out of six statements used in self-assessment are named as barriers to data sharing by social sciences researchers, so the general negative-neutral attitudes are confirmed, even though the %-values cannot be used directly.

²⁴² DAMVAD 2014: 58.

²⁴³ DAMVAD 2014: 53-54.

3.31.2 Capability requirement areas of DAS

Organisational profile

Organisation

NSD was established in 1971 as a national research infrastructure with the name Norwegian Social Science Data Service. Until 2003 NSD was affiliated with the Research Council of Norway, and has since been a limited company wholly owned by the Ministry of Education and Research, and headquartered in Bergen. During the years NSD covered a wider range of disciplines than just the social sciences and in 2015 the name was changed to Norwegian Center for Research Data, keeping the abbreviation NSD.

The designated communities of NSD consist of scholars in the humanities and the social sciences. Initiatives in the life sciences and medicine are being explored. NSD's primary user community is researchers and students. The secondary communities consist of the universities and research institutions, as well as the governmental and public sector.

Funding

In 2015, income by source was 29.8% from the Research Council of Norway (17 percent via its basic grant), 24.6% from the government ministries, 29.3 from other public and private sector project sponsors, and 15.7% from European Union projects and other sales and contract research income.

Core services and activities

NSD is a tool for research. NSD offers a range of services that frees resources and capacity for research within its sector. NSD's strategies and instruments for fulfilling its social mission and performing its tasks on behalf of Norwegian research are continuously changing and evolving in step with the needs and frame conditions of research. Core services: Collect, curate and provide access to a wide variety of data types, provide data protection services, and provide access to relevant tools and training services.

NSD works to ensure that researchers and students enjoy access to data by acquiring, processing, archiving, maintaining and disseminating data to the research community. Research data are currently disseminated to researchers and students both in Norway and abroad. NSD is also a national centre of expertise in the protection of personal data used in research. The organisation's top-level competence in this field is unique at international level. NSD evaluates national and European regulations for the use of personal data on an ongoing basis. The aim is to contribute to the valuable and legitimate use of personal data in research to the benefit of the society. In collaboration with other national and international centres, NSD develops tools and services within its field and participates in the international supply of services to research.

Content current collection

NSD's data holdings provide information about the human society at different levels. The data are organised in four main categories: Individual Level Data, Regional Data, Data about Institutions and Data about the Political System.

NSD gives access to a vast amount of individual level data from surveys within a wide spectre of topics ranging from living conditions, health, labour, and use of media to elections, democracy, and social values. NSD's tasks as a national archive of research data financed by the Research Council of Norway has resulted in research projects in the humanities, social sciences, health and medicine, and environment and development being contractually required to deposit their data with NSD. This can be everything from text, sound, video data to the more traditional survey data.

NSD archives a number of other types of micro data generated by Statistics Norway and other producers of data, such as ESS, ISSP, universities and colleges, the institute sector and opinion polling institutes.

The Database for Statistics on Higher Education (DBH) is a central source of information on Norway's tertiary education sector, covering universities, colleges and vocational colleges, libraries and student welfare organisations.

The Municipality Database is regularly updated with new data, and at the end of 2015 it contained information about every single Norwegian municipality over a period of 250 years. The database comprises a large collection of data, including comprehensive documentation, the possibility of putting together time series with standardised units, and presenting data in thematic maps.

The Ecclesiastical Database comprises official statistics of baptisms, confirmations, marriages, funerals, etc. at individual congregation level.

NSD is responsible for the archiving, processing and distribution of data and documentation, and for running the ESS' official website.

NSD operates and administers two registers related to academic publishing; the Norwegian Register for Scientific Journals, Series and Publishers and ERIH PLUS, both of which list several thousand academic journals, book series and publishers from all over the world. NSD has operated ERIH PLUS, the European Reference Index for the Humanities (ERIH) and Social Sciences (PLUS), since July 2014, and in December 2015 it was moved to NSD. The Norwegian Register for Scientific Journals, Series and Publishers is administered by NSD, is an overview of publications and publishers that are regarded as being of academic quality

The linguistic composition of the archive's collections are Norwegian and English.

Organisational infrastructure

Figure 66: Organizational infrastructure in Norway

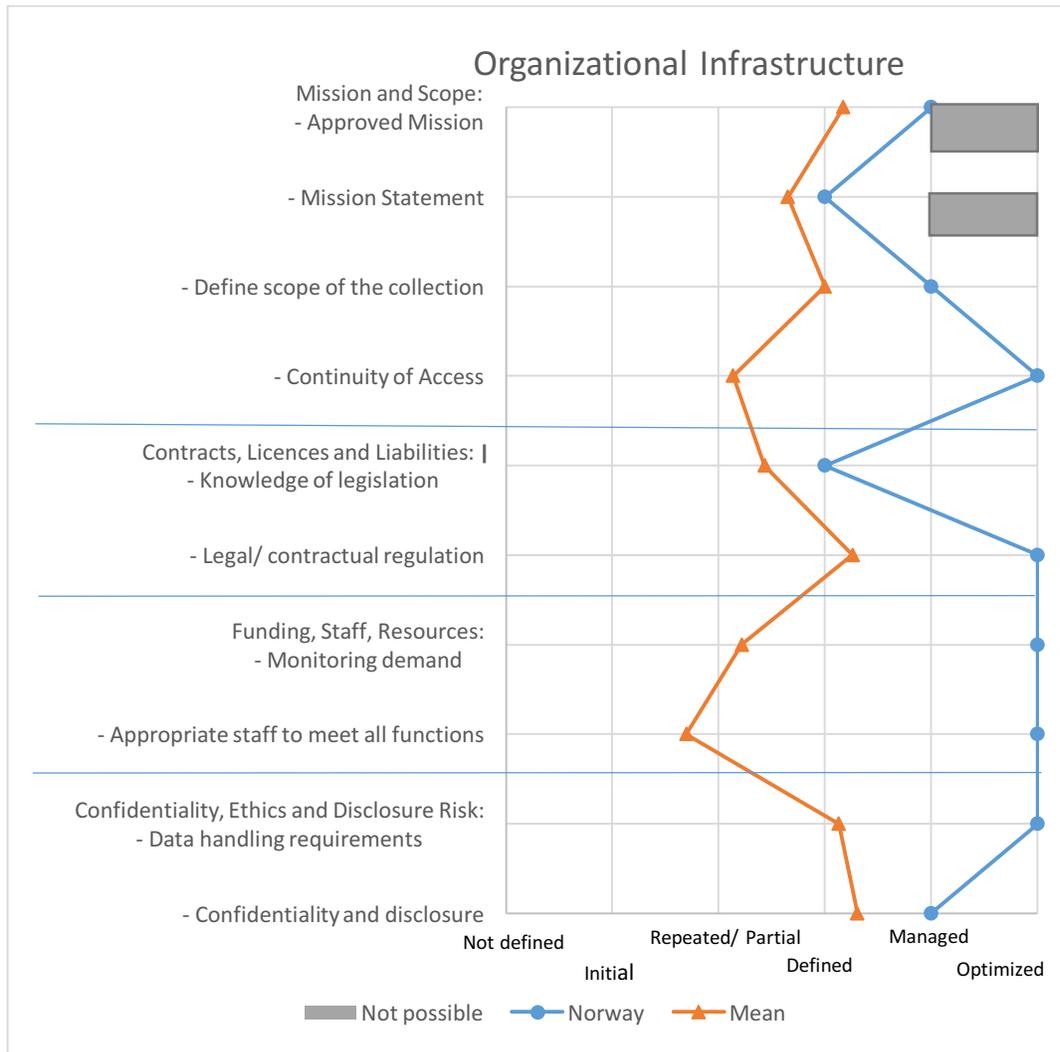


Figure 66 shows that the organizational infrastructure in general is on an optimized or a managed level. In all aspects NSD is placed above the average.

Mission and scope

The organisation has a received formal approval through a clearly defined mandate and provides a written, formal and complete mission statement that is available for all relevant users. NSD's Mission Statement is released under the Norwegian State Ownership Report. It states (translated from Norwegian): "NSD works in a long-term perspective to improve opportunities and working conditions for empirical research that is primarily dependent on access to data. This is done by collecting, processing, adapting, archiving and maintaining data and disseminating it to research communities while ensuring that the data is available in a form that enables it to be used in research without major legal, financial or practical costs being incurred by the users." The organisation has

defined an explicit scope for the repository and their collection. The definition of the scope is on a managed level²⁴⁴.

Contract, licences and liabilities

Sufficient and documented knowledge on all relevant legislative aspects is available to all staff.

The usage and success of access licenses and the access conditions framework are continuously assessed; monitoring of wider legal framework (e.g. national and EU regulations); regular and formalised contact with relevant stakeholders. All data consumers are bound by a Terms of Use with NSD. Co-workers that handle data are also committed to sign a Terms of Use. The user commits to delete the data when the analyses are finished or at the latest two years after the access. A copy of resulting reports or publications that are based on the data should be sent to NSD. The data consumer also commits to refer to the producer and distributor in a foreword or footnote in the publication. If the user is a student, the supervisor also should sign a Declaration²⁴⁵. If the user does not comply with the conditions, he or she will not get access to data. In most cases the users have to be affiliated to an approved institution in order to access the data.

Funding, staff, resources

NSD has a number of different sources of income. Its main grants are provided by the Research Council of Norway, government ministries, the university and college sector and the European Union.

In 2016 NSD had close to 100 persons employed. According to the NSD Annual Report 2015, 86 man-years of work (including that done by part-time students) were performed in 2015. NSD is organised in four major parts: Administration, Individual Level Data and Data Protection, ICT and Development and Survey and Data Services.

Confidentiality, ethics and disclosure risk

Most data that NSD manages is anonymous data. Non-anonymous data will only be stored at NSD after fulfilling relevant formalities. NSD's archival staff will check this before receiving data. The archival staff at NSD has relevant competence within privacy issues and other ethical questions. In addition, NSD's in-house Data Protection Official for Research unit will be consulted when necessary.

Datasets with disclosure risk are bound by a data processor agreement that describes the character of the data and limitations of use. Before person identifiable data can be archived at NSD data provider must submit the filing permissions so that NSD can ensure that the archiving of data is according to the law of privacy.

Data with a disclosure risk is processed and stored according to specific procedures described in "Archiving procedures" on NSD's websites. Such data will be stored in a safe-deposit box, preferably the same day as they are received or stored temporarily in a locked filing cabinet. In some cases, NSD conducts the integration between identification register and dataset on behalf of

²⁴⁴ Strategic plan, available at: http://www.nsd.uib.no/nsd/english/strategic_plan.pdf

²⁴⁵ Declaration (in Norwegian), available at: http://www.nsd.uib.no/nsddata/utlaansrutiner_en.html and in English, available at: <http://www.nsd.uib.no/data/individ/publikasjoner/tilgangsbrev-engelsk.pdf>

the researcher. In these cases, the following rules apply: 1) The identifiers are kept isolated from the data. 2) Computers used for this purpose are off-line with strong virus protection. Data with disclosure risk may be distributed if the data consumer meets all conditions for access.

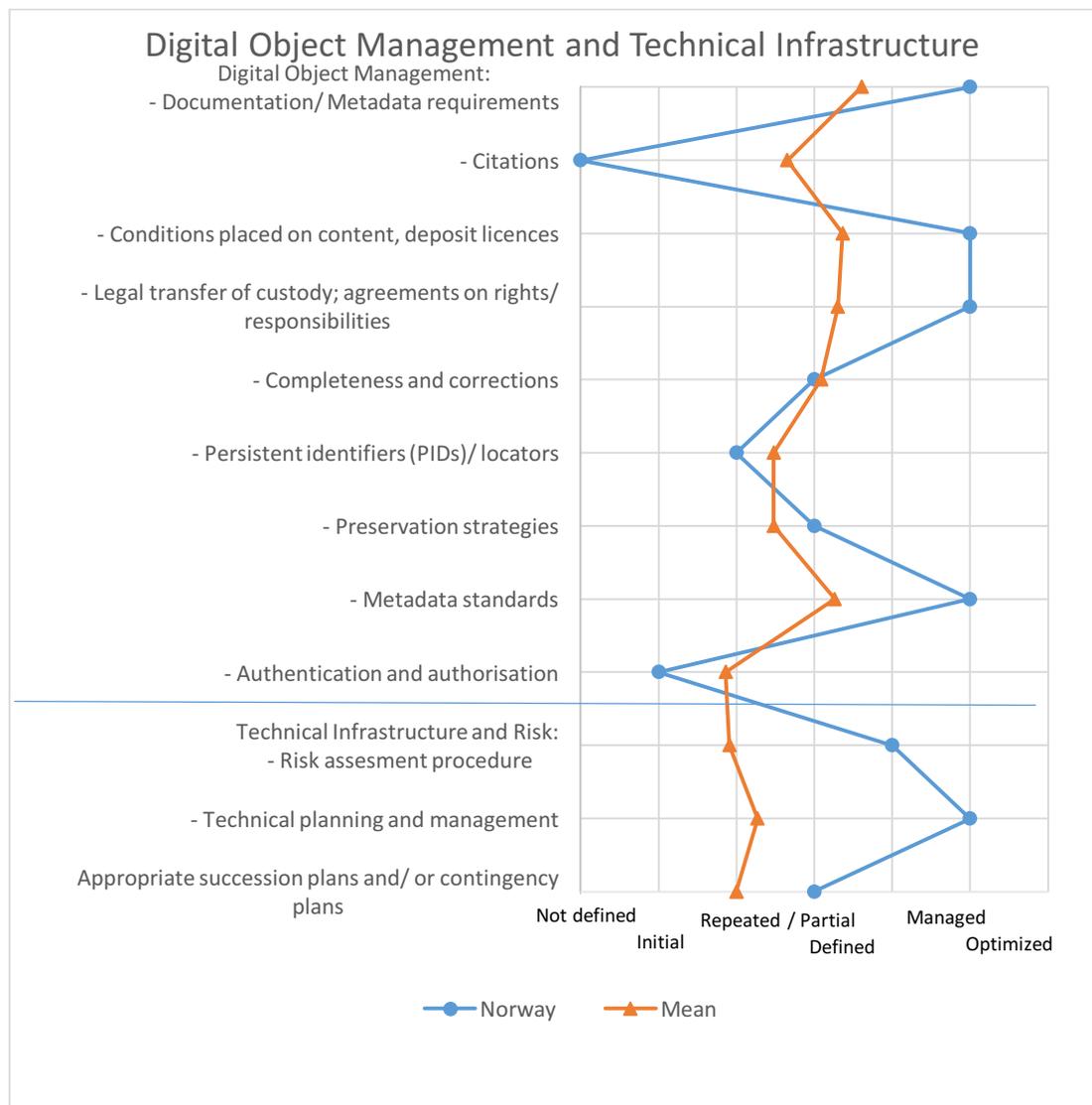
NSD always checks data to make sure that it is not indirectly identifiable through a combination of background variables. If so, an anonymous version is created, either by aggregating or deleting variables²⁴⁶.

Digital object management (data curation) and Technical infrastructure and risk

Figure 67 shows that Digital object management and Technical infrastructure in general are above the average, and for many areas on an optimized level. Three areas - citations, persistent identifiers and authentication and authorisation - have a score below the average, but these are areas where work is in progress and where it is expected that NSD will reach a higher level during 2017.

²⁴⁶ See http://www.nsd.uib.no/nsddata/utlaansrutiner_en.html and <http://www.nsd.uib.no/data/individ/publikasjoner/tilgangsbrev-engelsk.pdf>

Figure 67: Digital object management and Technical infrastructure in Norway



Documentation/Metadata requirements

Advice and guidance for data producers/depositors on metadata and documentation are provided through a web form. There is also a guide on NSD's website to help with the correct documentation²⁴⁷.

²⁴⁷ See http://www.nsd.uib.no/arkivering/014_dokumentasjonsguide.pdf

Citations

Work in progress. Agreement in place with DOI provider and procedures and practices in place by November 7, 2016.

Conditions placed on content, deposit licences

Depositors set access conditions in the “Agreement for the deposition of research data for archiving with NSD”²⁴⁸.

Legal transfer of custody, agreements on rights/responsibilities

An agreement is entered into between NSD and the data supplier with the following conditions for archiving and making research data accessible: NSD shall store and secure original data and documentation as they were transferred; NSD shall keep the data readable and accessible at all times; NSD shall ensure that data are not unintentionally changed or lost; NSD shall secure the integrity and confidentiality of the data; NSD is entitled to publish descriptions of the data on NSD's web pages and share them with international data catalogues; The right of ownership of data is retained by the data producer / data supplier at all times during the deposit period; NSD is entitled to process data in accordance with established procedures for data protection and long-term storage; If necessary, ND and the data supplier/producer can enter into an agreement on the anonymization of data; NSD can make data available to others in accordance with the conditions set out in the agreement²⁴⁹.

Completeness and corrections

Staff who manages data has a written set of instructions and routines that they adhere to and all data transformations are documented²⁵⁰.

Persistent identifiers (PIDs)/locators: repeated/partial

There is work in progress to persistently identify and locate data deposited at NSD by DOIs. An agreement is signed with a DOI provider. Procedures in practice from November 2016.

Preservation strategies

NSD has a plan for long-term preservation. This considers obsolete file formats in the future. The routines are documented in “NSD's archiving procedures”²⁵¹. A data Preservation Platform, to strengthen the current preservation practices, policies and technologies in the infrastructures, is part of the NORDi project²⁵².

²⁴⁸ Available at: http://www.nsd.uib.no/nsddata/arkivering/en/006_archiving_agreement.pdf

²⁴⁹ See http://www.nsd.uib.no/nsddata/arkivering/en/006_archiving_agreement.pdf

²⁵⁰ See http://www.nsd.uib.no/nsddata/arkivering/en/009_archiving_procedures.html

²⁵¹ See above.

²⁵² See <http://www.nsd.uib.no/nordi/english.html>

Metadata standards

NSD create metadata in Nesstar Publisher. Nesstar Publisher is DDI-C compliant.

Authentication and authorisation

There is work in progress on authentication and authorization. Part of the NORDi project.

Risk assessment procedure

NSD aims to store data and metadata in a way that protects the data from unauthorized access, violation of the data's integrity and in other ways loss of data and documentation. The procedures are set up based on a risk and vulnerability analysis. Data are always kept in different storages/servers. Backups and non-anonymous data are stored outside NSD's premises. Non-anonymous data is always prepared on offline computers. In case non-anonymous data should be transferred online or on portable devices, the files are encrypted.

Technical planning and management

Technical planning and management are optimized. The technical infrastructure resources are reviewed at regular intervals. The project Norwegian Open Research Data Infrastructure (NORDi)²⁵³ is a result of such a review. During a five-year period (2016 – 2021) NSD will develop a new infrastructure that will make it easier to locate, use and share research data. The project includes: development of a cost effective and user friendly data management and archival tool and providing support for research data management plan; increasing efficiency and quality of the infrastructure to enable it to collect, document, enhance and share more data; strengthening of the current preservation practices, policies and technologies in the infrastructure; development of a common discovery portal for all NSD-held data collections and beyond; provision of efficient and flexible access to all of NSD data holdings; provision of the ability to track and document its research activity with respect to data archiving, (re)use and citation to research institutions; strengthening and expanding guidance- and training services at NSD.

Appropriate succession plans and/or contingency plans

NSD has implemented functions and mechanisms to be adequately prepared for major institutional changes.

3.31.3 Conclusions

Broader ecosystem of DAS operations as well as capability requirement areas of DAS in Norway show a very high level of maturity.

²⁵³ See above.

Where potential gaps can be identified - citations, PIDs and authentication and authorization – work is already in progress. Potential developments with regards to gaps in broader ecosystem of DAS operations, with regard to more positive attitudes and more support to researchers with issues related to OA to data are tackled on different levels in the project NORDi.

NSD has a long standing strong position in the Norwegian social science landscape and during time in a sustainable way built-up capacities that can be used for further development.

3.32 Poland

Poland has developed social sciences and humanities sector producing high quality data. Despite the low level of funding (measured as GERD in SSH as % GDP level) the country participates in a number of international surveys. The culture of data sharing among researchers and research institution is on medium level. It was estimated that currently up to one third of the Polish researchers in SSH are sharing produced data and more than one third of them are able to access existing third party data. On the policy level the situation seems to be worse. There are nearly no open access policies, institutional strategies or subject-based initiatives for research data already in place and no steps towards more developed policy can be expected in the near future. Most important institution in area of social data preservation and open access to the research data is Polish Social Data Archive (ADS). This institution is well developed regarding internal standards of data acquisition, archiving and publishing. Problematic moment is the low level of institutionalization of this archive. It is neither independent organization nor department in organization but „only” a research program based on the current agreement between two top Polish academic institutions. Low personal resources represent another weak point for the operation of the archive.

Specifics about the data collection in a country

This report for Poland is based mainly on the data from self-assessment template filled in by Marcin W. Zieliński, head of Polish Social Data Archive - if not stated otherwise.

3.32.1 Broader ecosystem of DAS operation

Development of social science sector

Overall assessment of SSH development

Table 24: Assessment of SSH development in Poland

Indicator	Values for Poland	Rating of Poland (0 low level, 1 medium level, 2 high level)
GERD in SSH as % GDP	0.09	0
GERD in SSH as % of GERD	10.2	1
Number of researchers in SSH per capita	83	1
GERD in SSH per researcher in SSH	11284	1
Access to databases	²⁵⁴	1
WoS publications	missing	-
Average	-	0.8
Overall	-	0

According to ADS staff, the rank of the sources (sectors) of research funding in social science in Poland by the amount they provide is following: 1) government and Higher education sector, 2) abroad (international or cross-border), 3) business enterprise sector, and 4) private non-profit sector.

Bibliographic and full-text databases (mainly through Virtual Library of Science/ Wirtualna Biblioteka Nauki) and national software licenses are accessible in Poland,

The overall assessment of SSH development in Poland was evaluated as one on the low level – i. e. „funding in SSH and productivity of the researchers are in the lowest quantile; impact on designated community is small or non-existing “. Main reason for this is probably the low level of financial support from government (GDP in Poland is quite low compared to western countries and GERD in SSH as % of GDP is low even in the limited framework of Polish GDP).

²⁵⁴ Government or universities in Poland provide access to commercial bibliographic and full-text databases (as are WoS, Scopus, EBSCO or JSTOR) and national software licences.

Table 25: Development of research data production in SSH in Poland

Indicator	Situation in Poland	Rating of Poland (0 low level, 1 medium level, 2 high level)
International collaborative research	ISSP, EVS, WVS, CSES, ESS, CSS, PISA, GGP	2
Studies of national importance	Polish National Election Study (PGSW), Polish General Social Survey (PGSS), Social Diagnosis, Polish Panel Survey (POLPAN), ESS, ISSP	2
Average production of data	Periodical	1
Sum	-	5
Overall	-	5 (developing)

Poland participates in Comparative Candidate Survey (in 2013 wave), European Social Survey (all rounds), ISSP (regularly), Comparative Study of Electoral Systems (all waves), European Values Study (wave 2, wave 3, wave 4), Programme for International Student Assessment (participated in 2000, 2003, 2006, 2009, 2013), World value survey (second wave) and Generations and gender programme.

Among existing studies that systematically assess matters of national importance for Poland are the European Social Survey (ESS), Polish National Election Study (PGSW), International Social Survey Programme (ISSP), Polish General Social Survey (PGSS), Social Diagnosis, and Polish Panel Survey (POLPAN). In general, the production of SSH data in Poland can be characterized as „periodical“. Apart from many ad-hoc surveys, there are several survey programs, which are conducted periodically.

The overall result of the assessment suggests that the research data production in SSH in Poland is on the „developing stage“. There are some examples of research excellence and streams of research stand out (both qualitative and quantitative); there are also many of examples of good international collaborative research.

RDM Policy setting

There is a growing recognition and awareness of need to require DMP in Poland, but defined enforceable rules on DMP in SSH field are still missing. The same can be said about the situation regarding depositing data in an appropriate disciplinary repository. There is the recognition of the problem; but neither rules, nor formalized way for researchers exist at the moment.

The Background note on open access to scientific publications and open research data²⁵⁵ in April 2016 admits that there are nearly no open access policies/institutional strategies or subject-based initiatives for research data already in place in Poland; and no steps towards more developed policy are expected in the near future.

The research funding organizations operating in Poland explicitly recognize that additional costs are needed to prepare the data for the access and that for the project budget that these costs are legitimate, since they can partially cover the RDM cost up to a certain limit. Despite this recognition, there is no explicit formalization of named costs and no other incentives in this direction are observed at the moment.

Regarding general situation with support provided on the legal and ethical aspects of data sharing in Poland, it seems that no organized support is provided. However, there is an awareness of the problem of legal matters, especially in the field of the issue of re-use and/or re-processing of existing data and its publication.

Data sharing culture

The proportion of social science researchers in Poland that have shared the research data they produced in the period between 2011 and 2016 was estimated between 10% and 30%, which is the medium level. There is no problem with data sharing in a case of big, periodically repeated studies (like ISSP). Nevertheless, gathered data suggests that most of researchers conducting „ad-hoc“ projects do not provide open access to their data. Some social science data producers in Poland do follow data management and data documentation standards and procedures that facilitate data reuse.

According to Zieliński's estimate rather high proportion (more than 30%) of social science researchers in Poland were able to access existing third party data they need in the period between 2011 and 2016.

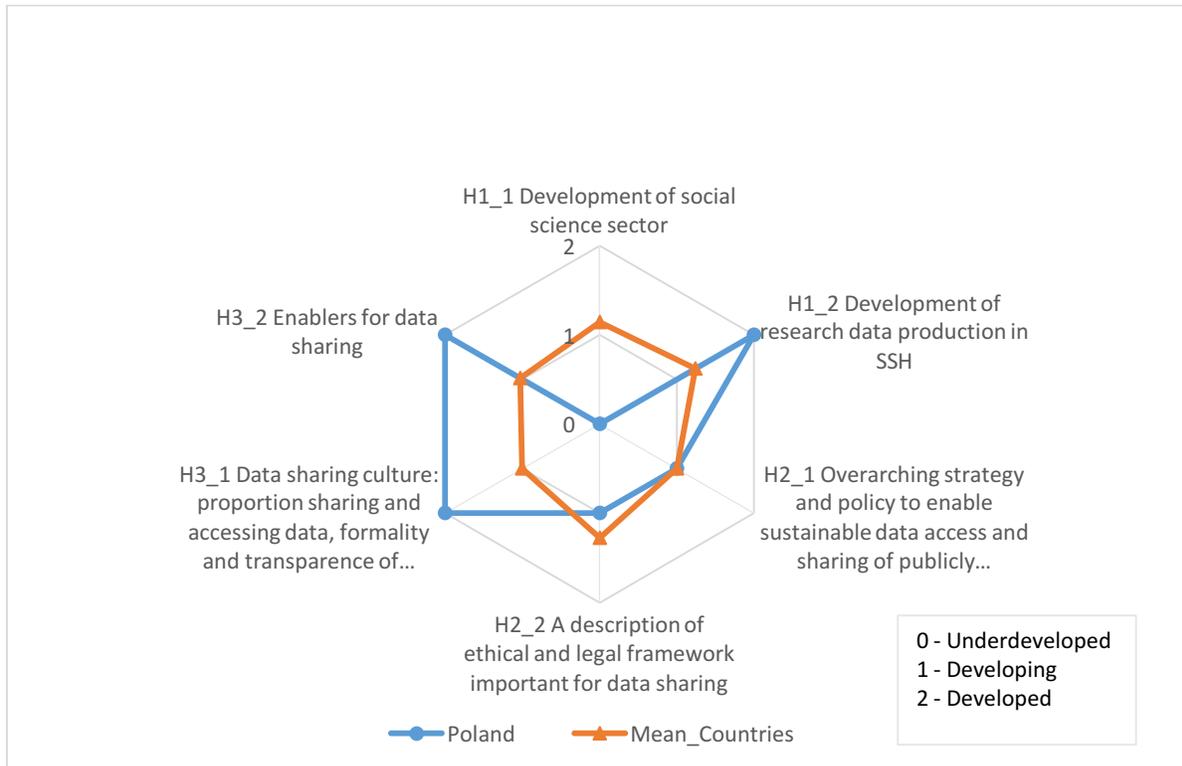
It was acknowledged that the most frequently used way to research data for secondary analysis in Poland is via informal contacts (peers and colleagues). Data archive or repository and project or personal websites are in general less frequently used channels for the data access. The exceptions are periodically conducted studies, which are accessed mostly via the data archive.

To some extent there are career rewards related to data sharing. In general, these career-rewards are related to the potential publications that use other's data in order to increase citations (except data citation itself which are not considered as citations that are counted for the career in Poland).

Polish Social Data Archive is to some extent a provider of data support services (e.g., it provides "good practices" handbook and individual consultations for potential depositors). Interdisciplinary Centre for Mathematical and Computational Modelling of the University of Warsaw provides trainings in legal matters of data sharing and data sharing culture, and it also maintains some very general repositories, e.g. for supplementary data for journals.

²⁵⁵ Published by the directorate-general for research and innovation (RTD) in April 2016, available at: https://ec.europa.eu/research/openscience/pdf/openaccess/background_note_open_access.pdf

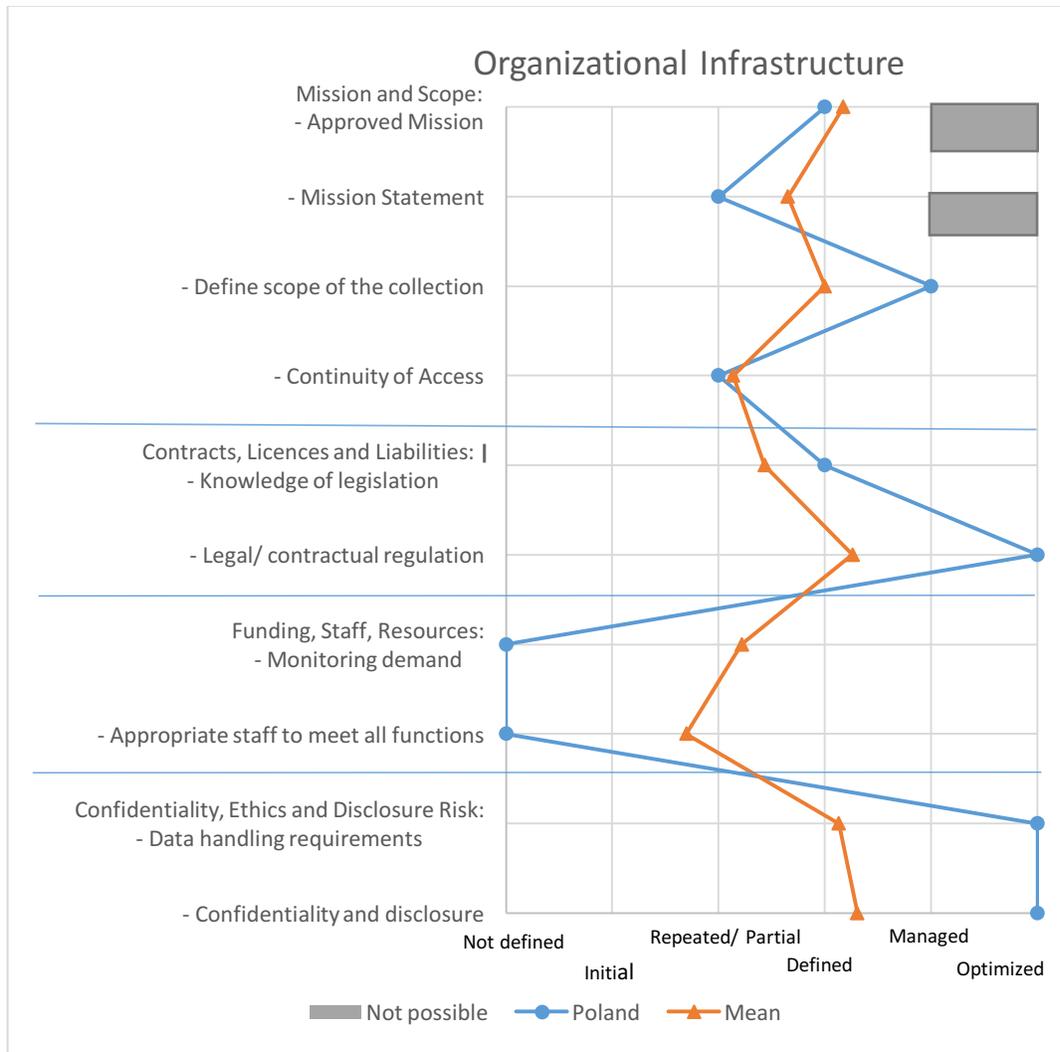
Figure 68: Heading concepts values in Poland



3.32.2 Capability requirement areas of DAS

Organisational profile and infrastructure

Figure 69: Organizational infrastructure in Poland



ADS is not fully institutionalized but rather operates as a „research program” within the University of Warsaw and Institute of Philosophy and Sociology of the Polish Academy of Science. At the University of Warsaw ADS is operated by the Robert B. Zajonc Institute for Social Studies. Its administrative and physical location is at the same Institute. The answers provided by the ADS staff are related to the archive (not the institution where it is located and operated). However, one should keep in mind that ADS is not an institution and does not have institutionalized form. Thus, for example, it could be stated that the data archive itself is not funded at all. It is not an institution and people providing archiving services are in fact the employees of the Institute for Social Studies at the University of Warsaw; while the other costs are covered by the University of Warsaw and the Institute of Philosophy and Sociology.

Polish Social Data Archive operates under the agreement between the University of Warsaw and the Institute of Philosophy and Sociology of the Polish Academy of Science. This agreement is a written statement on the role, mandate, purpose and mission of the organization. Main points in

mentioned agreement are: creating strategy of social data archiving; acquiring, developing, archiving and data sharing in the field of social sciences; implementation of training programs in the field of social data archiving; cooperation in order to raise funds for these purposes. ADS and its activities are on an ad-hoc basis “encouraged” by stakeholders to continue its activities; however, no formal mandate or formalized approval exists. Regular and formalized contact with relevant stakeholders is not yet implemented.

The definition of scope for the repository is monitored for compliance with general policy. Though, there is no direct monitoring of demand for repository services. The archive's collections are available in Polish and English (sometimes combined, some collections are available only in Polish, other – only in English). At present ADS provide access to almost 100 surveys including all general, repeated studies from Poland.

Students of various social science departments are the primary user community for the Archive. The secondary user community includes university level staff and employees of tertiary public and private institutions.

ADS has partial plans to ensure the continued availability and accessibility of data in case the organization ceases to operate, but there is no general plan about what will be with data after potential end of ADS.

ADS has sufficient and documented knowledge of all relevant legislative aspects. Monitoring of wider legal framework (both national and EU regulations) is exercised. The access licenses and the access conditions framework are continuously assessed.

The reviews of the procedures and policies for compliance to research discipline norms for confidential data handling are performed at regular intervals. However, the informant found it hard to assess the state of development in this direction. Zieliński stressed that the archive is a research program within Institute for Social Studies at the University of Warsaw and as such it does not have the direct legal support. Therefore, he stated that this issue is optimized at the level of the University but not at the level of the Institute of Philosophy and Sociology and not at the Archive itself.

The usage and success of confidentiality and disclosure mechanisms are continuously assessed, reviewed and updated; monitoring of wider legal framework (e.g. national and EU regulations); automated checks and anonymization mechanisms may be in place. Each data set is at least triple checked in order to provide anonymization of sensitive data and the access is provided only to anonymized data.

The number of staff at ADS appears to be inadequate to support all the functions and services (but the archive has not defined a level for the appropriate number of staff to support the necessary functions and services). Many core competencies were identified as missing. The archive has no departments and all the services are provided by two people working at the Institute for Social Studies at the University of Warsaw.

Core services and activities that are offered by the ADS:

ADS clearly specifies all the information (documentation, metadata, and provenance) that needs to be associated with the data that is to be deposited. These information are available at “DATA DEPOSIT IN ADS” section of ADS website (<http://www.ads.org.pl>).

This section includes specific general requirements the depositors should meet are available in English at the archive website²⁵⁶. There are also step by step instructions for depositors (forms to fill in, required media for data deposition etc.)²⁵⁷. In these documents, all required information (metadata scope, data formats) is formally specified; data and metadata requirements are accessible and are communicated clearly to users and depositors.

Providing functions and mechanisms for proper data citations is however in the initial phase. Citations are offered if requested by the depositor; however, there are no practices or general strategies written down (the approach is ad-hoc and case-by-case).

Formal, written agreements and contracts are in place. Responsibilities and legal transfer of custody are clearly defined. Contractual templates are used consistent²⁵⁸. Legal and contractual framework is regularly reviewed and updated; and all legal and contractual regulations are aligned to higher level policies. In this respect, all roles and responsibilities are identified and maintained.

All depositors are offered the opportunity to set the access conditions on the data and information that is being deposited. However, there is no set of access conditions formally pre-defined in categories or as a template.

ADS staff controls data files to ensure that all the variables and values are accurate according to the documentation supplied and sufficiently labelled for reuse. The checks are made to make sure that the variable names in a dataset match the variable names in a codebook (completeness and correctness checks are run).

Polish archive has no system for persistent identification yet. Implementation of documented preservation (for cases of degradation of storage media, the obsolescence of media drives, and safeguards against accidental or intentional digital corruption of preserved material) strategies is still in the initial phase. Only informal, ad hoc 'contingency action points' are in place, but a full comprehensive strategy is still lacking.

A data/metadata format strategy is communicated to users in the "DATA DEPOSIT IN ADS" section of ADS website (<http://www.ads.org.pl>). Data and metadata are provided in formats that can be understood by users. ADS do not explicitly claim using DDI or any other widespread standard for metadata. ADS uses Authentication and Authorization Infrastructure for access to deposited materials.

²⁵⁶ Summary listing of the requirements depositors have to meet, available at:

http://www.ads.org.pl/pdf/Summary_listing_of_the_requirements_depositors_have_to_meet.pdf

²⁵⁷ See <http://www.ads.org.pl/index.php?tresc=jakdeponowacE.html&v=E>

Data deposit form is also available in English at <http://www.ads.org.pl/pdf/FORM04E.pdf>

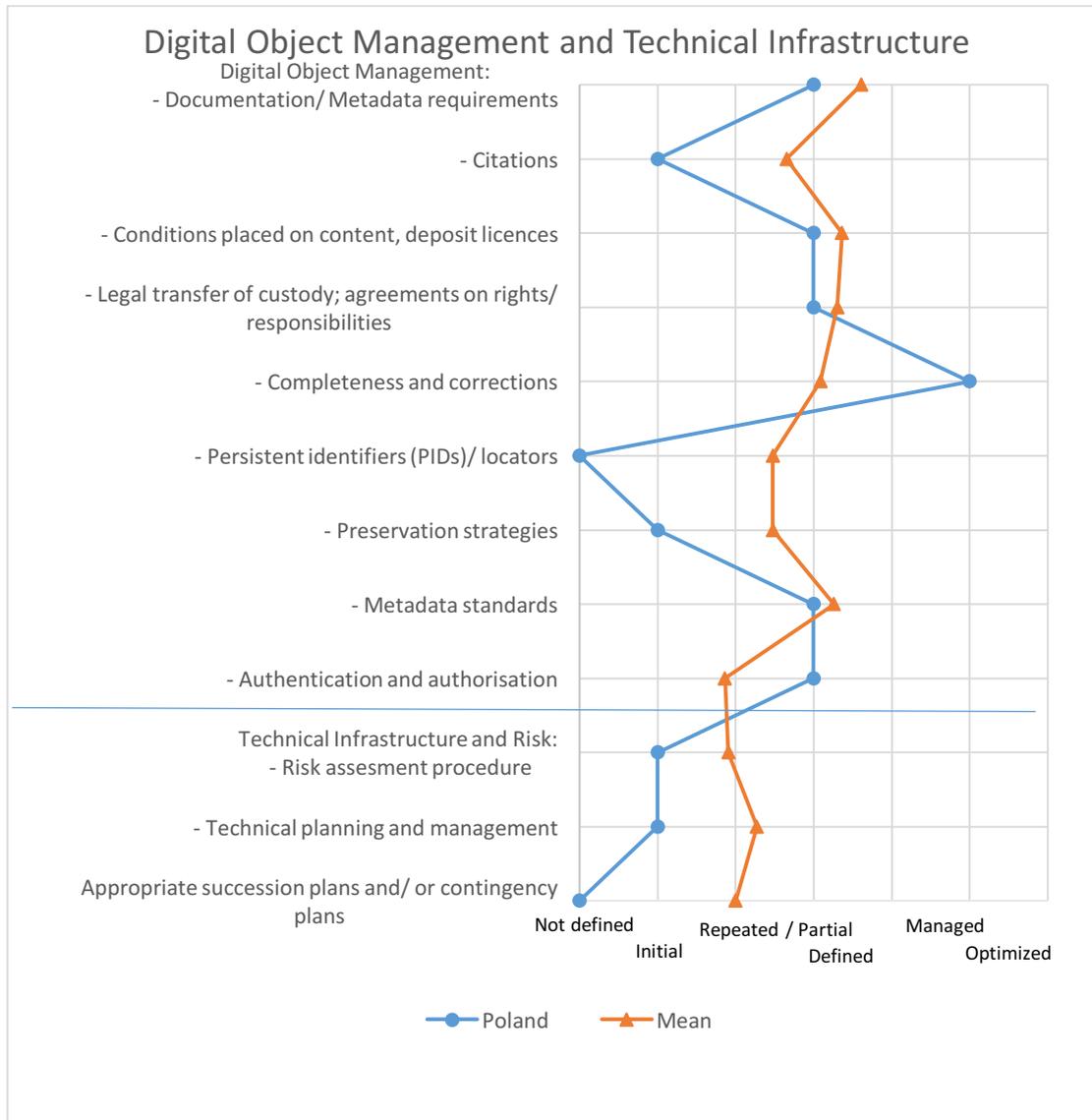
Data set registration form (which is in fact a license agreement form) is available at

<http://www.ads.org.pl/pdf/FORM02E.pdf>

²⁵⁸ ADS' data set registration form is available at the web site of the archive: <http://www.ads.org.pl/pdf/FORM02E.pdf>

Digital object management (data curation) and Technical infrastructure and risk

Figure 70: Digital object management and Technical infrastructure in Poland



There is a certain awareness at ADS of the need for the risk assessment. Some risk assessments are already undertaken, but generally, this is still in the initial phase. Some ad-hoc steps are taken upon the request of the funding organizations and they serve a specific purpose.

Technological infrastructure is only partially complete due to the lack of personal and financial resources. ADS hasn't defined the appropriate technical resources or level needed to support all the repository functions and services.

Core services and activities that are offered by ADS are: help in preparing data according to ADS standards (metadata, appropriate data and documentation preparation), checking the data and documentation prepared by data depositors (if they are in accordance to the ADS requirements), archiving data and metadata and providing access to the data for authenticated users. Other important services are aimed on providing reports on data usage and users and trainings in data preparation and use of the collected data.

3.32.3 Conclusions

Despite limited sources (both personal and financial), production of social science data in Poland is quite intensive. Poland participates in most of the important international research series fielded in Europe. The data sharing culture among Polish researchers is growing and a considerable part of them is sharing their data. Most important channels for sharing are informal; however, sharing through ADS archive become more important today. Institutional support for development of sharing culture is currently low; policies supporting open access to research data are neither under the preparation nor expected to be prepared in the nearest future. Most important institution in the field of social science data archiving and sharing is Polish Social Data Archive (ADS). This archive is not a CESSDA member but it has quite long history and it developed certain standards for its actions and provides all important services in area of data acquisition, archiving and access to archived data. The low level of institutionalization and low personal capacities were identified as the most serious obstacles for the development of the Polish archive services.

3.33 Portugal

All in all, Portugal has a low-medium maturity level. The country is in some areas above the average and in other areas below the average. Portugal scores high in relation to the broader eco-system of DAS-operation and scores low-medium in relation to the capability requirement areas of DAS.

There are a lot of strengths in relation to the broader eco-system, such as international surveys carried out in Portugal and having a growing recognition on the importance of Open Access principle and recommendations in this domain. There is potential for changing the data sharing culture and move the researchers toward data sharing. However, it can be a barrier to change the attitude of researchers.

In relation to the capability requirement areas of DAS, there is planned development in regards to implementing DOI and preservation policy. A security plan is also being discussed and there has just been written a succession plan.

The conclusion is that Portugal is doing fine and will continue to develop in the future.

Specifics about the data collection in a country

The data collection started 21st September 2016 and was finished 30th September 2016. The information was provided by the Portuguese Archive of Social Information (APIS). APIS is a social science data archive based at the Institute of Social Sciences of the University of Lisbon (ICS-ULisboa).

The data collection process went well without any problems.

3.33.1 Broader ecosystem of DAS operation

Figure 71: Heading concepts values in Portugal

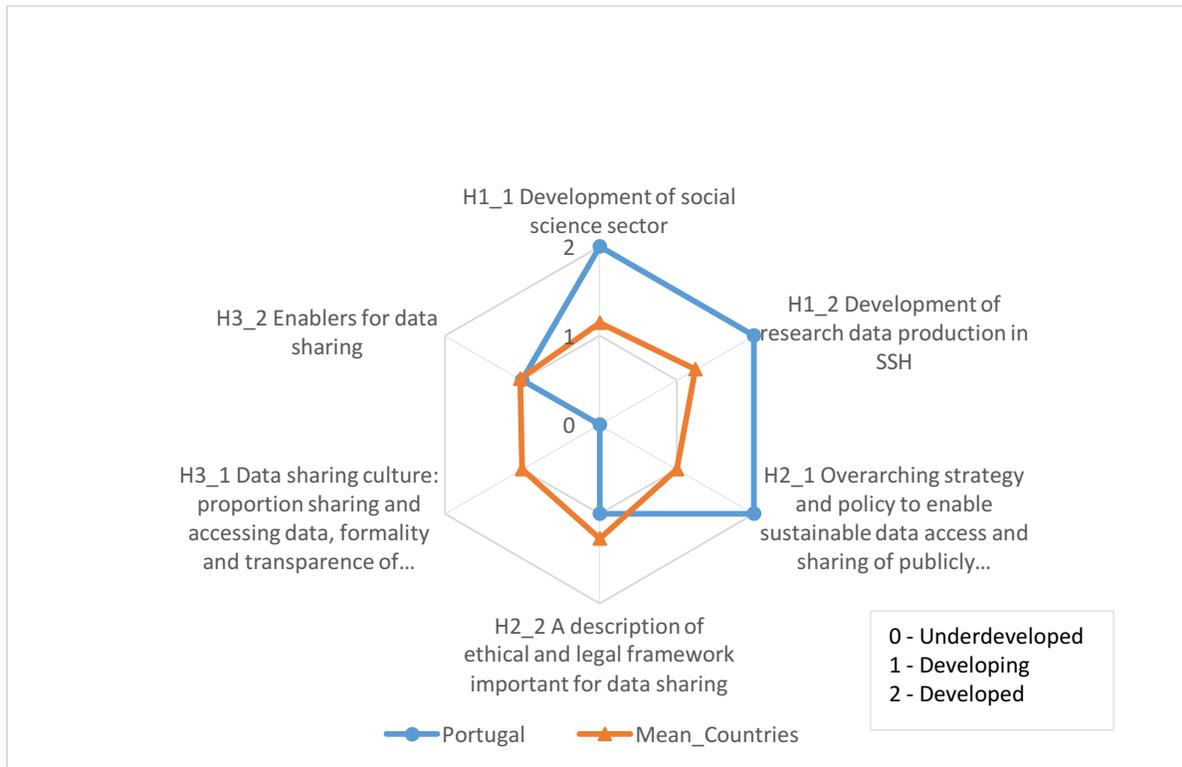


Figure 71 shows that Portugal is develop at the areas ‘Development of social science sector’, and ‘Development of research data production in SSH’. In these areas Portugal is above the average. In the areas ‘A description of ethical and legal framework important for data sharing’ and ‘Data sharing culture proportion sharing and accessing data, formality and transparency of routines’ Portugal is developing and underdeveloped and below the average. Finally, in relation to ‘Enablers for data sharing’ and ‘Overarching strategy and policy to enable sustainable data access and...’, Portugal is – like the average – developing.

Development of social science sector

Overall assessment of SSH development

The financial stability, research capacities and results achieved in the field of social sciences are in the highest quantile (the developed level) and above the average.

In relation to research funding in social science, government and higher education sector is ranked first. The research is also funded abroad (international or cross-border) (ranked second), private non-profit sector (ranked third) and business enterprise sector (ranked fourth).

The Biblioteca do Conhecimento Online – b-on (Online Knowledge Library) makes unlimited and permanent access available, within the research and higher education institutions, to full texts from over 16,750 scientific international publications from 16 publishers, through subscriptions negotiated on a national basis with these publishers. Coming soon b-on will change its search service to the EBSCO EDS Service (<http://www.b-on.pt/en/>).

Development of research data production in SSH

The research data production in SSH is on the developed level and above the average.

The most important studies about politics and social attitudes are the international surveys regularly carried out in Portugal (the European Social Survey (ESS), the Portuguese Election Studies (PES) and, more recently, also the Share).

The average production of research data by the social science institutions is frequent. It is noted that it is a very difficult question to answer without objective indicators. As most of research centres in social sciences field are well rated, it is assessed that the average production is frequent in most of them.

RDM Policy setting

Funders' data management and sharing strategy and/or policy

RDM Policy incentives are on the developing level.

The requirements or recommendations for DMPs in most cases are on the partial level. Portugal is not at the beginning of the process; there is a growing recognition on the importance of the Open Access principles and the government has been making recommendations in this domain. The country has already some Open Access infrastructures, namely the institutional repositories. Concerning the management plans, there is so far any obligation to implement them. But, there is an expectation that next year there will be some recommendations in place.

The way of depositing data in an appropriate disciplinary repository is on partial level. In the already existing repositories, social sciences research data can be deposited with the associated metadata. Some research centres and universities strongly recommend that all research output should be publicly available. However, there are no formal requirements neither sanctions are contemplated for not complying.

The public research funding organizations in Portugal provide, on the partial level, incentives for sharing research data with associated metadata. There are no doubts about the value of research data produced and about the need for long-term preservation; but the investment for the long-term preservation is still scarce. In this relation, cost for managing the data and preparing it for access is on an initial level. Some costs for managing the data can be covered in the scope of project budget but not in a very explicit way.

Legal and ethical framework

The ethical and legal framework is on the developing level and a little below the average.

There is some explicit guidance concerning sharing data but clarification on legal aspects have not been put in place yet.

Data sharing culture

Data sharing culture in on the undeveloped level and below the average.

Data sharing and reuse among social sciences researchers in Portugal is not so common – as self-assessment results indicate, the proportion of researchers sharing data is estimated as low (0-10%). It is stated that the data sharing culture is underdeveloped in the research community and that the situation is the same in all disciplinary fields. One of the reasons in social sciences field lays on the fact that there never was a strong tradition in secondary analysis.

In regards to the proportion of researchers able to access existing third party data they need; it is not possible to provide an estimate. In the same kind of way, it is not possible to provide an estimate for 2011-2016 in relation to routines for data sharing and channels.

Table 26: Attitudes towards data sharing in Portugal

Data sharing has no benefits at all	False
Data sharing creates healthy competition	Neither true, nor false
Data sharing creates negative competition	Probably true
Reuse of existing data can answer new research questions and facilitate advancement of science	Probably true
Data sharing has as a risk that others may misuse and misinterpret data	Probably true
Data sharing involves little effort and minimal costs	False

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from “5”-True to “1” False!

The scores indicate that data sharing in Portugal is beneficial but resource demanding. It probably creates negative competition and has a risk. At the same time, reuse can secure future development in research.

Enablers for data sharing

Enablers for data sharing in Finland are – like the average – on the developing level. The main infrastructure that facilitates data sharing and/or Open Access to research data is RCAAP, an institutional repository funded by the Portuguese national funding agency for science, research and technology FCT (<http://www.rcaap.pt/about.jsp>). There are still others small infrastructures as APIS (Portuguese social information archive) in social sciences or ROSSIO (Portuguese reference infrastructure for Social Science, Arts and Humanities). These infrastructures provide conditions to store and disseminate research data.

In regards to data management and data documentation standards and procedures, it is assessed that some data producers follow these. The use of it is not well developed yet but most of data producers fulfil minimal standards. According to the self-assessment, there are no career rewards related to data sharing.

3.33.2 Capability requirement areas of DAS

Organisational profile

The Portuguese Archive of Social Information (APIS) is a social science data archive based at Institute of Social Sciences, University of Lisbon (ICS-ULisboa). ICS-ULisboa is an associated laboratory dedicated to research and advanced training in social sciences.

In regards to main bodies/departments/units, the archive currently doesn't have such a formal structure.

The primary user community is composed by quantitative social science researchers from the fields of Sociology, Political Sciences and Social Psychology. The secondary user community is composed by students and teachers from the fields of Sociology, Political Sciences and Social Psychology.

The data archive service is 100% public funded, on a project-based funding.

The core services and activities offered by the data archive service are data archiving, data preservation, data dissemination, and data analysis (using NESSTAR).

The overall content of the collection is composed by survey data collected by social scientist with public funding. According to CESSDA topic classification, the main topics of the collection are politics; health; social stratification and groupings; society and culture.

APIS collection meets the following criteria:

- Research project with public funding;
- Scientific value;
- National coverage;
- High expected use considering the designated community;
- Should be related with the topics covered by the archive.

The archive's collection is in Portuguese language. Some metadata is provided in English according to CESSDA mandatory fields.

Organisational infrastructure

Figure 72: Organizational infrastructure in Portugal

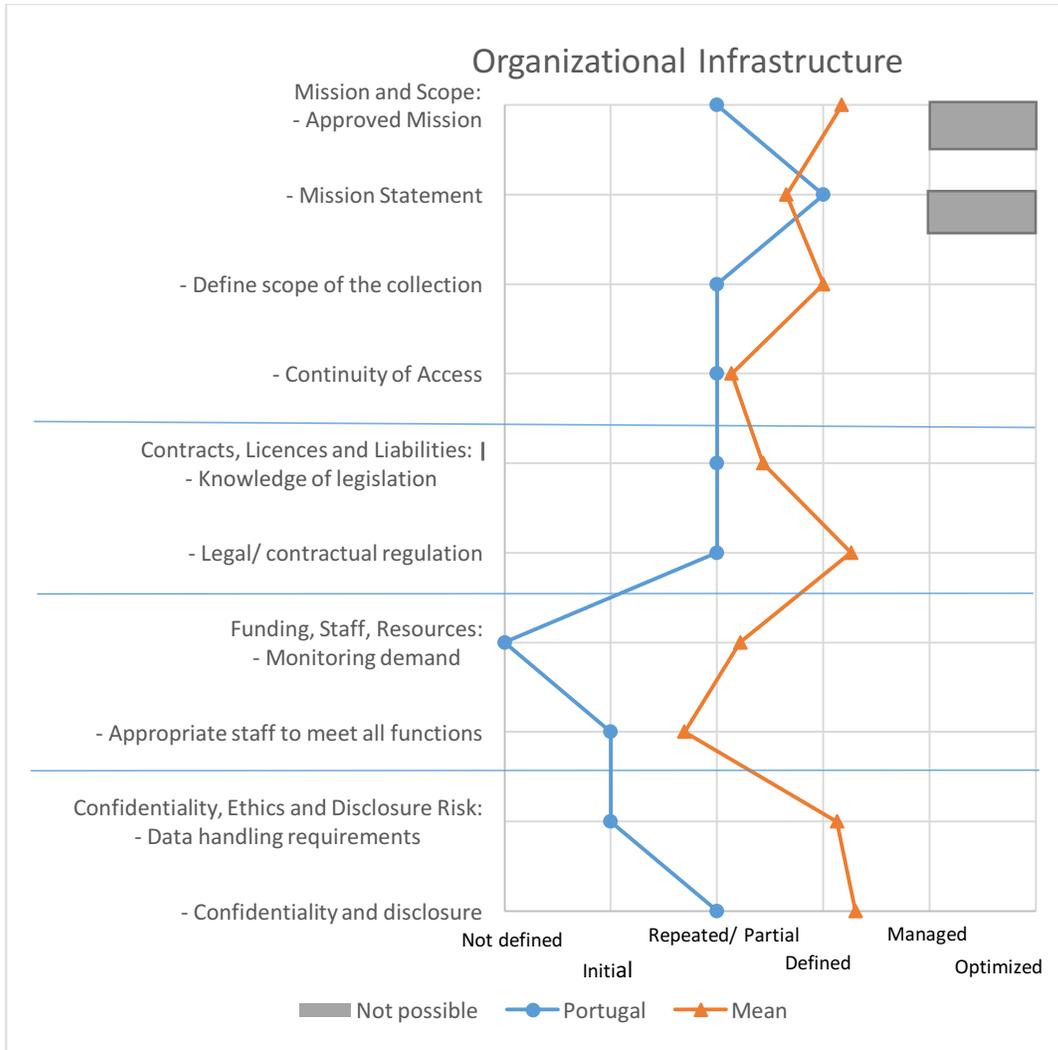


Figure 66 shows that the organizational infrastructure in general is on the repeated/partial level and a little below the average. On other areas, it is on initial level and a little further below the average.

Mission statement

APIS mission is to increase the use of data resulting from social research carried out in Portugal from probabilistic surveys. This is done through rigorous processing, long-term preservation and online dissemination of the data, ensuring that they are reliable and can be openly and friendly used for the purposes of public consultation, secondary analysis and pedagogical use.

The archive has an informal recognition of the research community. Our role is recognised as being part of a research infrastructure included in the national roadmap²⁵⁹.

Continuity of access is on the repeated/partial level and close to the average. The area was assessed at a time when a succession plan was still being written.

Knowledge of legislation is on repeated/partial level. The organisation follows the main laws concerning data protection in Portugal and keeps up to date with the activity of the Portuguese Data Protection Authority (CNPD). The same applies to Intellectual Property Rights, although to a lesser extent.

Legal/contractual regulation is likewise on the repeated/partial level. The use of agreements is fully implemented. The form is available by request and it will become public soon in a new website.

In relation to the appropriate staff (to meet all functions), it is on initial level. The core activities of the archive are ensured. Currently, there are a team of 4 people. There is no appropriate staffing level defined.

The staff consists of:

- 1 Manager
- 1 Data Manager
- 1 International Funding Officer
- 1 IT technician

The manager of the archive is a researcher who allocates a part of his time to APIS; the data manager works at full time and the International Funding Officer and the IT technician give punctual support.

Data handling requirements is likewise on the initial level. The archive complies with social sciences ethical norms and with the Portuguese data protection law.

In relation to confidentiality and disclosure, it is on repeated/partial level. The depositor is required to fully anonymise the data before send it to the archive (instructions are available in the website) The staff checks the data in order to detect any personal data. Disclosure risk is addressed by verifying if top-coding or higher levels of data aggregations are needed.

²⁵⁹ Portuguese Roadmap of Research Infrastructures, available at: https://www.fct.pt/apoios/equipamento/roteiro/2013/docs/Portuguese_Roadmap_of_Research_Infrastructures.pdf (page 32).

Digital object management (data curation) and Technical infrastructure and risk

Figure 73: Digital object management and Technical infrastructure in Portugal

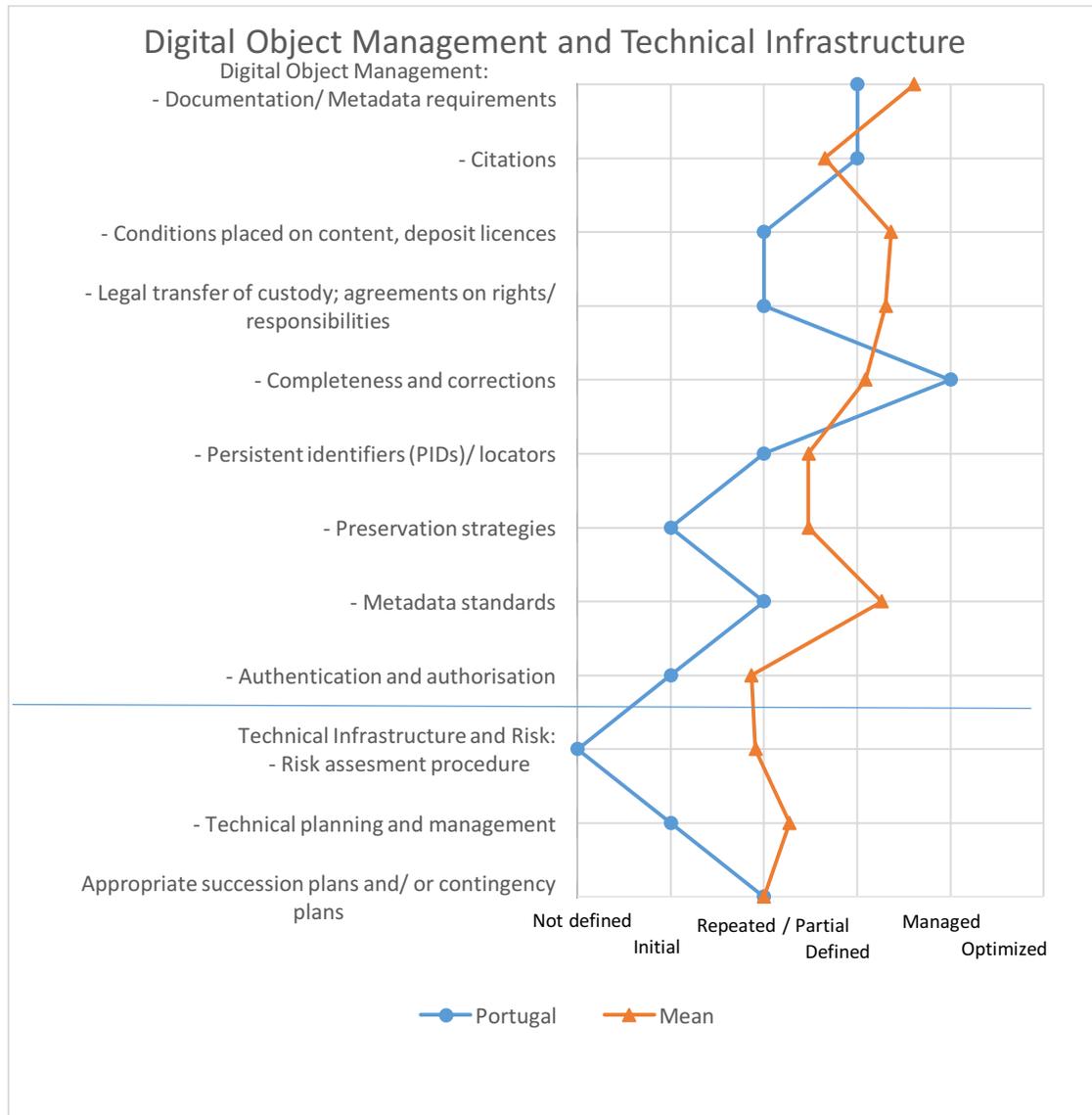


Figure 73 shows that digital object management in general is on the repeated/partial level or a defined level. It differentiates how far below and above Portugal is placed in relation to the average. The technical infrastructure and risk is somewhere between a not defined level and a repeated/partial level.

Digital object management (data curation)

Documentation/Metadata requirements are on the defined level. Required information is defined in the deposit agreement. The agreement includes a short metadata form that aims to collect administrative data concerning the depositor and the data collection (full title, alternative title, producer, funder, authors, abstract, keywords, and related publications). The full description of the data is made by the archive and it is based on the documentation that the depositor provides

(methodological reports, funding application, etc.). The staff also performs quality control checks to data to ensure its completeness and accuracy. All types of error are corrected and subsequent changes of the data are registered in a report.

Conditions placed on content, deposit licenses are on repeated/partial level. The conditions of use defined by the archive are related to data citation; reporting publications stemming from the archive's collection; ethical use of the data and restriction to distribution.

In relation to legal transfer of custody; agreements on rights/responsibilities it is also on the repeated/partial level. There is an agreement which has clear definitions of roles and responsibilities of repository and depositor but adjustments are made on individual basis when needed.

Persistent identifiers (PIDs)/locators are on the repeated/partial level. The archive attributes handle to each dataset using the Scientific Data Hosting Service (SARD) technical facilities. Implementation of DOI is planned for 2017.

In relation to preservation strategies it is on initial level. Preservation policy is still being written.

Metadata standards are on the repeated/partial level. The archive provides data in formats that are widely used by the designated community, such as SPSS and excel. The data is fully documented using the NESSTAR software which is compliant with the DDI metadata standard²⁶⁰.

Technical infrastructure and risk

Risk assessment procedure is on the not defined level. Security plan is still being discussed.

Technical planning and management is on initial level. The technical infrastructure is being re-organised in order to use the national technical infrastructure in a more efficient way. Among other things, this means that it is intended to connect better with important organisations such as RCAAP.

In relation to appropriate succession plans and/or contingency plans it is on the repeated/partial level. At the time the survey was completed, the succession plan was still being written.

3.33.3 Conclusions

The self-assessment was conducted by APIS and the data collection process went very well. The results show that Portugal in general has a low-medium maturity level with a lot of development happening right now and planned for the future.

In relation to the broader eco-system of DAS operation, Portugal is on the developed level in the development of social science sector. There is on regularly basis carried out international surveys in Portugal and the average production is frequent in most of the research centres in social sciences field. In relation to the RDM Policy setting, Portugal is on the developing level. There is a growing recognition on the importance of the Open Access principles and the government has been making recommendations in this domain. The country already has some Open Access infrastructures, namely the institutional repositories. Regarding management plans, some recommendations are expected to be in place next year. There are existing repositories but no

²⁶⁰ A list of preferred formats will be attached to the preservation policy and made available in a new website, available at: <http://www.apis.ics.ulisboa.pt/en/preparing-data-and-documentation/>.

formal requirements for depositing data. The investment for the long-term preservation is scarce but there is a clear understanding of the value of it. The data sharing culture is not so common. An enabler for data sharing is an institutional repository, RCAAP.

In relation to the capability requirement areas of DAS, the organizational profile shows that APIS is a research infrastructure based at Instituto de Ciências Sociais, University of Lisbon. The archive offers data archiving, data preservation, data dissemination, and data analysis as the core services and activities. The organizational infrastructure is primarily accessed as being on the repeated/partial level. In relation to continuity of access a plan has just been available. In relation to legal/contractual regulation, a form will soon be public in a new website. The appropriate staff (to meet all functions) is on the initial level with no appropriate staffing level defined. The digital object management (data curation) is in general on the repeated/partial level. Implementation of DOI is for example planned for 2017 and a preservation policy is being written. Finally, the technical planning and management is assessed somewhere between not defined and repeated/partial level. Security plan is for example being discussed.

3.34 Romania

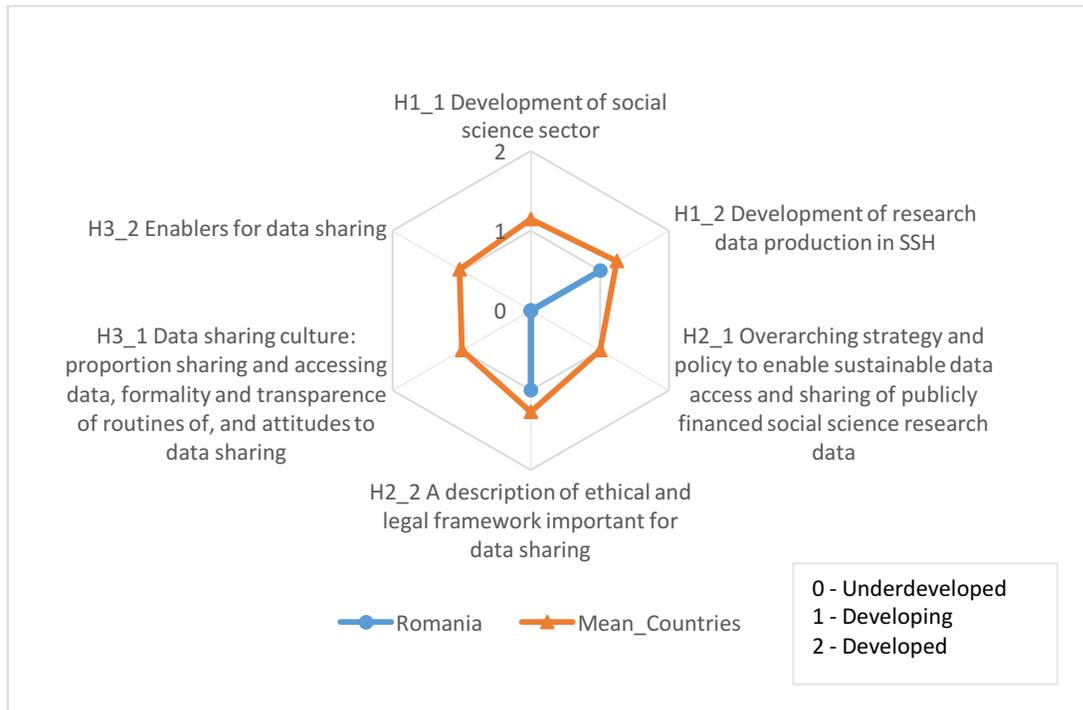
Romania is a country which was very close to joining CESSDA - ERIC in 2013. At that time, the Romanian Ministry for Education and Research backed out in the very last moment, providing no other funding continuity. Besides the current lack of funding, the RODA archive is in a relatively good shape to become the national service provider if Romania decides to join CESSDA.

Specifics about the data collection

Data collection in Romania was done via Adrian Dusa from the Romanian Social Data Archive, who provided most of the information in the web survey between 28.9.2016 and 28.10.2016.

3.34.1 Broader ecosystem of DAS operation

Figure 74: Heading concepts values In Romania



Development of social science sector

The government and higher education and abroad (international or cross-border) are the main sources of funding in of social sciences research in Romania. The Funding of social sciences and productivity of the researchers are at low level.

RDM Policy setting

The RDM policy setting in Romania is non-existent. There seem to be no DMPs requirements in the project funding contracts.

Clarification and support provided on legal and ethical aspects is initial, the IPR is sometimes found in the official discourse, which suggests at least some awareness about the issue, but nothing else.

Data sharing culture

Data sharing and reuse among social sciences researchers in Romania is low – as self-assessment results indicate, the proportion of researchers able to access existing third party data they need is low (0-10%). *“The most part of the data produced in Romania are not available and not shared. In rare cases, they are sometimes shared via personal websites or via informal contacts, as well as sometimes on the data archive's catalogue.”* Thus, informal and not transparent data sharing channels prevail. The exact proportion of social science researchers who have shared their data could not be estimated.

Quotes from the self-assessment survey: *“There is no formal study about the researchers' attitudes regarding data sharing. As a norm, data is very rarely shared (“my data, my property”). Some studies (like the Public Opinion Barometer) do make their data available on their own websites, but that study has ended before 2011. The kind of data everyone is happy to use comes from international projects like ESS or EVS, where data for Romania is freely available (however not in Romania).”*

“Most of the Romanian researchers publish almost exclusively in Romanian (only a handful of researchers publish regularly in top foreign journals), and on the other hand even the top journals have only recently started to make it a requirement to make available the data (in fact they are not require data, in most cases, but the replication file which refer to a certain dataset).”

Enablers for data sharing

It is not common for researchers from Romania to publish in journals that expect data used in the publication to be available for reuse from a trusted digital repository. Quote from the self-assessment survey: *“Most of the Romanian researchers publish almost exclusively in Romanian (only a handful of researchers publish regularly in top foreign journals), and on the other hand even the top journals have only recently started to make it a requirement to make available the data (in fact they are not require data, in most cases, but the replication file which refer to a certain dataset).”*

There are no career rewards to researchers who share data neither are there support services provided to social science researchers, that facilitate data sharing and/or Open Access to research data.

3.34.2 Capability requirement areas of DAS

Organisational profile

Organisation

RODA is the national Romanian institution specialized in archiving electronic data collections obtained by social research. The main goal of the data archive is to preserve these data indefinitely, and to serve as an intermediary between the data owners and data users. RODA is housed within the Faculty of Sociology and Social Work, at the University of Bucharest (<http://www.roda.ro>).

Funding

For the moment, RODA has no permanent funding of any sort or a formal structure.

Content current collection

About 50 studies, as a proof of concept to convince the Ministry of Education and Research of the viability of this institution. Romanian and English versions of archived data are regularly produced.

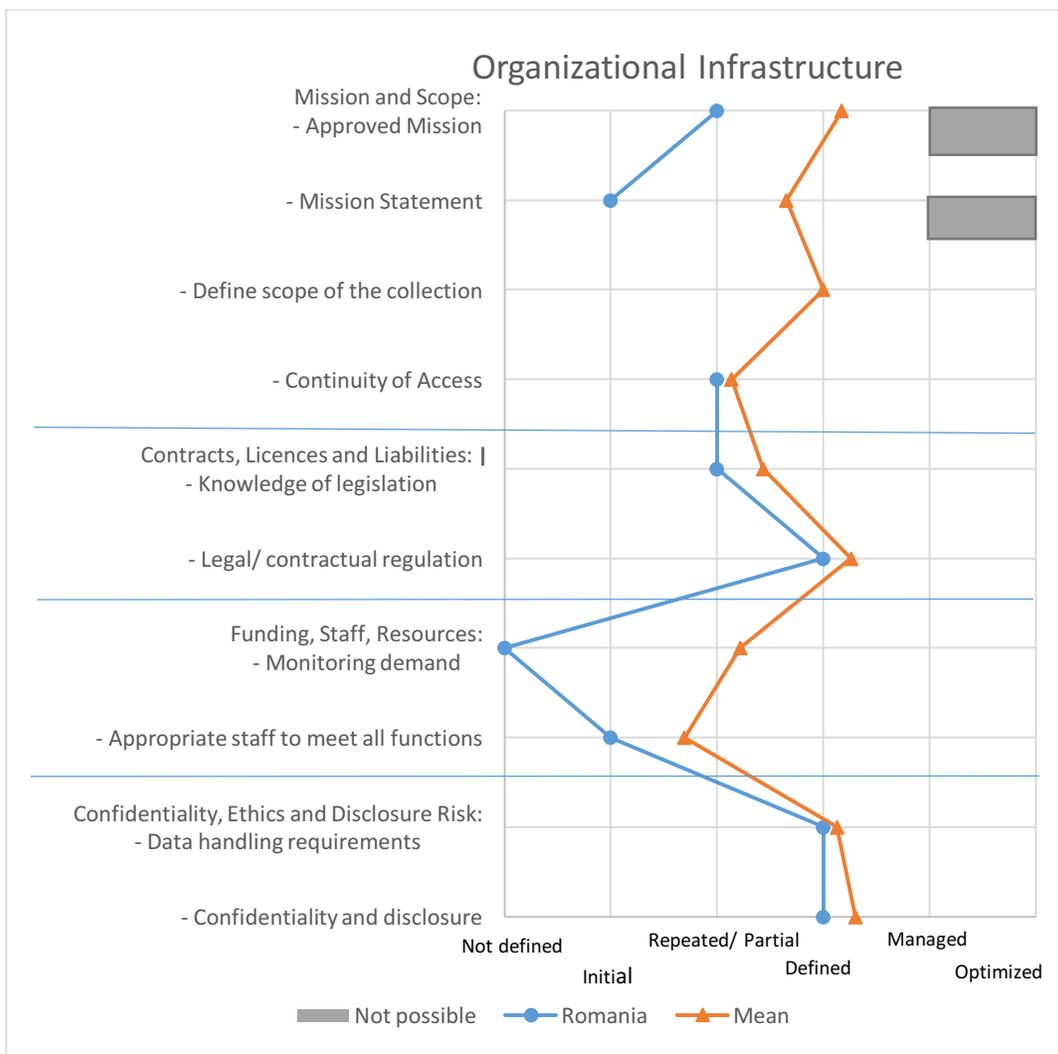
Organisational infrastructure

Mission and scope

RODA is the national Romanian institution specialised in archiving electronic data collections obtained by social research. The main goal of the data archive is to preserve these data indefinitely, and to serve as an intermediary between the data owners and data users.

A partial plan for continued availability is in place, but limited in detail and scope. RODA was supposed to participate in the network of CESSDA Service Providers, but the Romanian Ministry for Education and Research backed out in the very last moment in 2013, providing no other funding continuity.

Figure 75: Organizational infrastructure in Romania



Contracts, licences, liabilities

There is partial knowledge on relevant legislation, or knowledge is not widespread. There is some formalised documentation, but it may be insufficient documentation. Contracts and/or agreements are standardised and implemented according to written procedures; contracts and regulations are made publicly available²⁶¹.

Funding, staff, resources

There is no evidence of monitoring of demand for repository services. As there is no formal requirement from the Ministry, also without any other funding source, RODA cannot monitor such changes.

There is evidence staffing is partially complete, either due to lack of resources or unable to employ qualified staff. The organisation has not defined the appropriate staffing level to support all repository functions and services. Number of staff is dependent on funding, which is missing for the time being.

Confidentiality, ethics and disclosure risk

The organisation complies with norms and legal requirements through systematic written, formal procedures and policies²⁶².

Processes and procedures are in place; standardised information is provided to the depositor prior to the deposit; checks are performed on data after deposit; process and procedure descriptions for handling and altering sensitive data are in place²⁶³.

Summary: The infrastructure of RODA is prepared for operation but currently receives no funding. If Romania decides to join CESSDA, RODA is in a good position to serve as the national service provider.

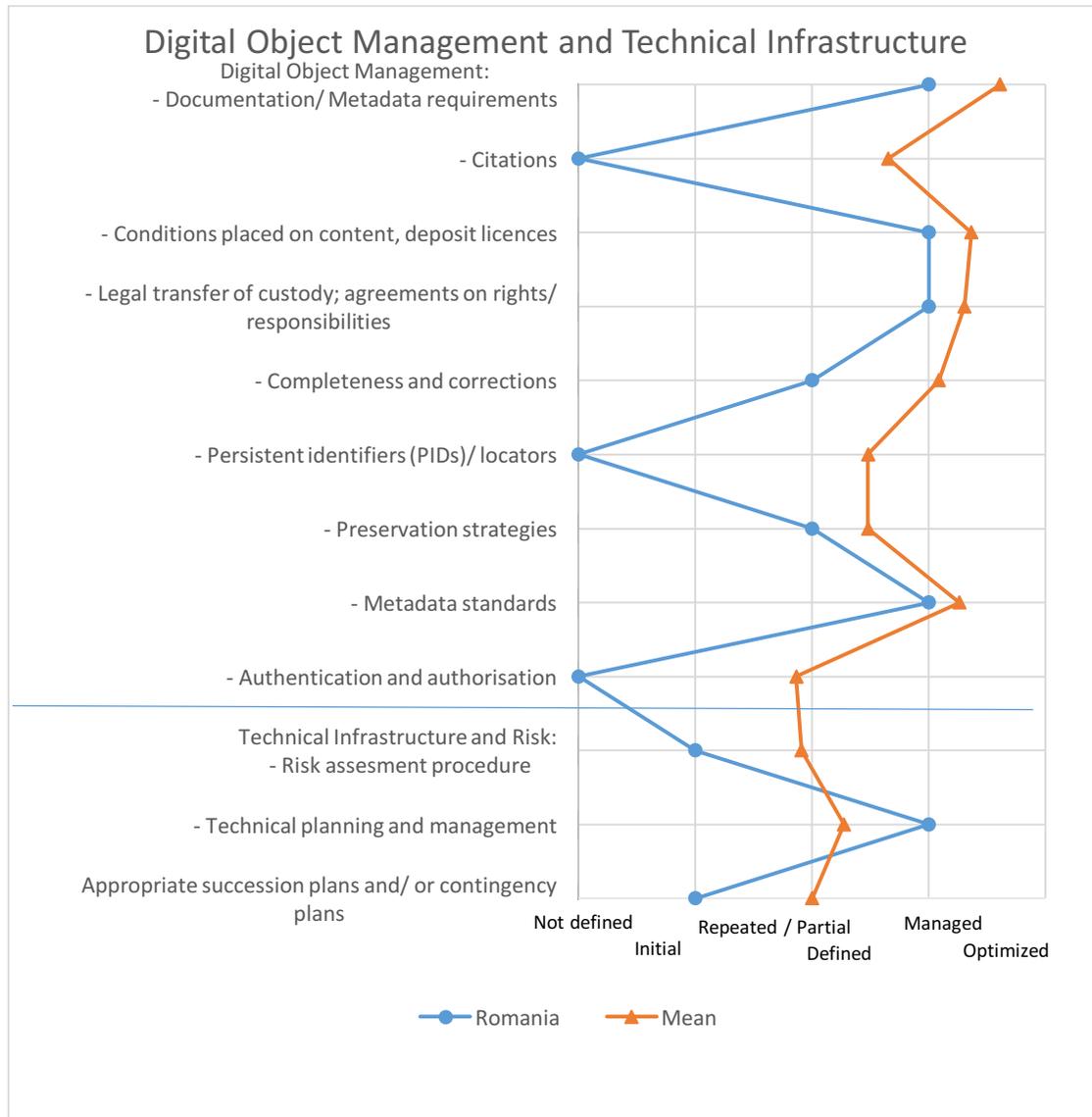
²⁶¹ The English version of the contracts (individual and teaching access) is available at: http://www.roda.ro/static/docs/en/Individual_en.pdf; Romanian version is available at: <http://www.roda.ro/static/docs/en/Didactic.pdf>

²⁶² Deposit form available at: <http://www.roda.ro/static/docs/en/DepositForm.pdf>

²⁶³ Licence agreement available at: <http://www.roda.ro/static/docs/en/LicenceAgreement.pdf>

Digital object management (data curation) and Technical infrastructure and risk

Figure 76: Digital object management and Technical infrastructure in Romania



Documentation/ Metadata requirements

A written formal specification of required information is explicitly defined (e.g. in a collection policy); requirements are compliant with metadata standards that are used and can be understood by Designated Community (e.g. DDI); metadata requirements are accessible and communicated to users/depositors²⁶⁴.

²⁶⁴ Data deposit manual available at: <http://www.roda.ro/en/data-depositing/how-to-submit-data>. Data deposit form available at: <http://www.roda.ro/static/docs/en/DepositForm.pdf>

Citations

Not defined: No citation practices.

Conditions placed on content, deposit licences

All depositors are offered the opportunity to set access conditions on the information that is being deposited; a set of access conditions are formally defined in categories or a template²⁶⁵.

Legal transfer of custody; agreements on rights/ responsibilities

Formal, written agreements and contracts in place; responsibilities and legal transfer of custody clearly defined. Contractual templates are being used consistently; legal and contractual framework is regularly reviewed and updated; all legal and contractual regulations are aligned to higher level policies; roles and responsibilities are identified and maintained²⁶⁶.

Completeness and corrections

There are non-systematised (manual) checks of deposited material in place; processes and procedures are repeated, but they are not formalised or documented. Rectifications are performed repeatedly, either by the repository or by returning data to depositor.

Persistent identifiers (PIDs)/ locators

There is no system for persistent identification at RODA.

Preservation strategies

Contingency action points have matured into a partial strategy; only partly formalised and documented.

Metadata standards

A data/metadata format strategy is explicitly defined, formalised, and communicated to users; data and metadata are provided in formats that are commonly in use and understood by users.

Authentication and authorisation

No authentication approach in place. There is a preliminary attempt by the ministry to create a federated authentication system, but that is for the moment not implemented.

²⁶⁵ Legal agreement available at: <http://www.roda.ro/static/docs/en/LicenceAgreement.pdf>

²⁶⁶ See licence agreement, available at: <http://www.roda.ro/static/docs/en/LicenceAgreement.pdf>

Risk assessment procedure

There is some awareness of the need for risk assessment and some risk assessments are undertaken, but on an ad hoc basis or when requested by the organisation for a specific purpose.

Technical planning and management

The organisation has defined the appropriate technical infrastructure resources to support all functions and services. Quote from the survey “RODA has a very solid system of servers, both in terms of hardware (a recent upgrade worth about 70K euro) and in terms of software, with a recent IT project to upgrade RODA to international standards.”

Appropriate succession plans and/ or contingency plans

There is some awareness of the issue; low institutional commitment to contingency issues; no written, formal processes, procedures, plans or other documents exists.

Summary: Regarding Digital object management and Technical infrastructure RODA has not defined the issues of persistent identifiers and authentication and authorization. It also has no rules with regards to citations. All other aspects of the archive are on the repeated / partial level.

3.34.3 Conclusions

In RODA Romania has a viable potential national service provider. Regarding the future of the archive, the decision of Romania to join CESSDA will be crucial.

3.35 Russia

The social science sector of Russia is well developed. Production of social science research data is frequent and well established. The main source of research funding is the public sector.

No clear policy on Research Data Management exists in the country, nor specific recommendations or requirements regarding Data Management Plans. This is one of the reasons data sharing is still not very common. However, together with requests for secondary data research and demands for more reliable and higher quality data and metadata, it has been increasing for the last decades. Channels for data sharing which lack formality are still the most preferred by researchers.

The perceptions of the attitudes of researchers towards data sharing are mixed, with data misuse and resource and time-consuming activities being the main obstacles to free data sharing.

The Joint Economic and Social Data Archive (JESDA) with roots in the 90s archives and delivers a number of quality and very demanded surveys and research data published in Russia, many of which are directly downloadable on the website of JESDA.

Specifics about the data collection in a country

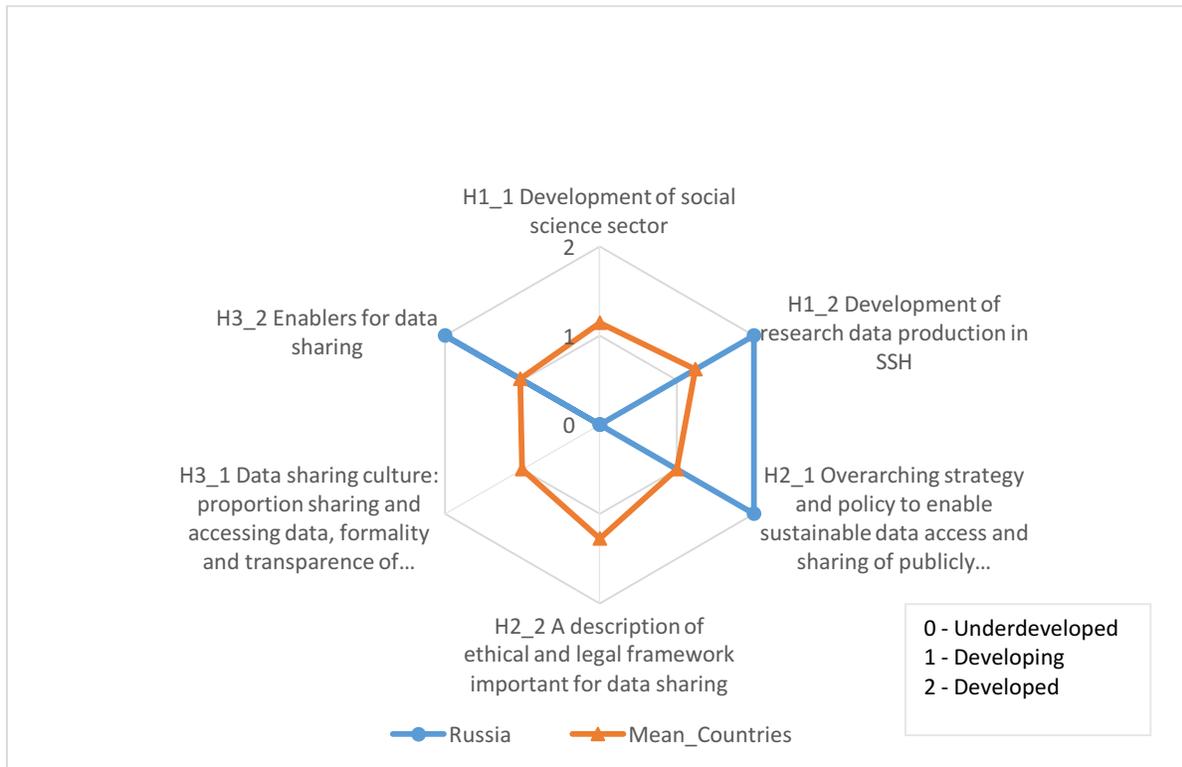
Data collection was carried out during the period between October and December, 2016. The self-assessment survey provided by the CESSDA SaW project was filled in by the director of the Joint Economic and Social Data Archive in Russia, Mrs. Larisa Kosovo in October, 2016.

Two complementary interviews with experts in the field of social science research, Mrs. Ludmila Khakhulina, deputy director of the Levada Analytical Centre in Moscow and Mr. Igor Zadorin from the Research group ZIRCON were conducted in November, 2016. The responses by all three respondents have matched in most of the areas covered by the surveys with few exceptions which are elaborated in this report.

Additional desk research using the sources provided by the contacts was carried out throughout the whole period.

3.35.1 Broader ecosystem of DAS operation

Figure 77: Heading concepts values in Russia



Development of social science sector

Overall assessment of SSH development

Funding of the SSH and productivity of the researchers in Russia are in the lowest quantile. Impact on designated community is small or non-existing. Research in the field of social sciences is mainly publicly funded and certain support for access to commercial databases and software purchase is provided by the government or state funding agencies. As stated by local experts, public policy makers in Russia make decisions based on evidence partially.

Development of research data production in SSH

Research data production in SSH is developed. There are well-established streams of research traditions, national and international, great variety of important types of research data. The average production of research data by the SSH institutions in Russia is frequent. A set of high quality monitoring surveys on the economic situation, well-being and health of the Russian population are conducted periodically, usually monthly or annually. Some of the best known are the Russian Longitudinal Monitoring Survey (RLMS), Monitoring of the economic and social change as well as other periodic surveys on the social structure, human capital and well-being of the population. Some of them are archived and publicly available at the Joint Economic and Social

Data Archive (JESDA) or other academic institutions. According to local experts in the field, leading universities and research institutions have significantly improved the production of SSH data for the last decade.

RDM Policy setting

Since sharing of research data was not very common in the past, requirements or recommendations about preparing Data Management Plans are still uncommon either or, as stated by some experts, non-existent. However, there is growing recognition and awareness about the value of research data produced and about the need for long-term preservation.

Public research funding organizations, regardless of their number, do not explicitly or formally demand data publishing or archiving. Data publishing is mostly a result of own initiatives by social science researchers or the data archive such as JESDA which issues certain requirements and recommendations regarding the data preparation and documentation. Therefore, costs for these activities are usually self-financed by the data producers.

Data sharing culture

Data sharing and reuse among social sciences researchers in Russia is not very common as self-assessment results indicate and as stated by experts who have been interviewed. The proportion of researchers sharing data is estimated as low (0-10%), and the estimate of the proportion of researchers able to access existing third party data they need is very difficult to provide. According to experts, the demand is currently not very high, but is increasing every year which is also reflected by the number of data requests at the Joint Economic and Social Data Archive (JESDA).

There are certain established data sharing channels and routines. Most popular are visiting and acquiring the necessary data from project and personal websites, which lack formality. Additionally, data can be obtained from research magazines if published papers are accompanied by the data used in the study. Archives come in second place, both local and foreign. This is usually the case when more reliable and higher quality data with related metadata is required.

The two experts interviewed and the self-assessment form differently perceive the attitudes of researchers towards data sharing. The results with explanation of the scores are presented below. The overall results indicate that social science researchers in Russia are perceived to acknowledge general benefits of data sharing and healthy competition created by data sharing, as well as importance of data reuse in advancing science. Researchers are perceived not to fear risk that others may misuse and misinterpret data, but consider resources and effort needed to prepare data as an important obstacle to sharing data.

Table 27: Attitudes towards data sharing in Russia

Data sharing has no benefits at all.	Probably false ²⁶⁷
Data sharing creates healthy competition in research.	Probably true ²⁶⁸
Data sharing creates negative competition (for example, being scooped and therefore reduced publication opportunities) for the researcher.	Probably false ²⁶⁹
Reuse of existing data can answer new research questions and facilitate advancement of science.	Probably true ²⁷⁰
Data sharing has as a risk that others may misuse and misinterpret data.	Probably false ²⁷¹
Data sharing involves little effort and minimal cost.	False ²⁷²

Source: self-assessment survey, expert interviews. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from “5”-True to “1” False!

To summarize data sharing culture in Russia, even though data sharing and reuse has been increasing for the last decades, the level of secondary use of research data is still very low. There are established both formal and informal channels for data sharing, however researchers tend to use channels that lack formality, such as websites where data can be easily downloaded. With the increase in demand for more reliable data and especially higher quality metadata, the role of archives like the Joint Economic and Social Data Archive (JESDA) becomes more significant.

The absence of clear requirements or recommendations for Data Management Plans and making data available after completion of research projects is one of the explanations for the gaps in data sharing culture. As experts’ state, if there are no legal requirements, financial support or other types of incentives for RDM, the level of data sharing will remain low.

However, the demand of the research community for more reliable data and higher quality metadata increases. This is manifested by the rise of requests for archived and well-documented data from the DAS, which has clear requirements and guidelines for data documentation.

²⁶⁷ According to the self-assessment form, this statement is true. The perception is that researchers generally do not see benefits in data sharing. Experts interviewed disagree, so the final mean score on the scale from 1 to 5 (where 1 is FALSE and 5 is TRUE) is 2.33.

²⁶⁸ The archive representative holds the position that this is not the view of the majority researchers in the field. However, researchers believe that the research community generally agrees with this statement and therefore the mean score is 3.67.

²⁶⁹ Considered as true by the archive representative, false by the two researchers. Overall score: 2.33.

²⁷⁰ Considered as false by the archive representative, true by the two researchers. Overall score: 3.67.

²⁷¹ Considered as true by the archive representative, false by the two researchers. Overall score: 2.33.

²⁷² The experts’ perceptions correspond to the self-assessment form grade. The overall score is 1 (false). This is one of the main obstacles for researchers to make data available and prepare a high-quality data documentation.

Enablers for data sharing

There are no or few career rewards related to data sharing in the academic community in Russia. Sharing research data does not usually influence career progression within an institution or the research community, however those who have the culture of data sharing and use secondary data are said to have better reputation and better chances of obtaining research funds in the future. There are data support services available to social science researchers that facilitate data sharing and/or Open Access to research data provided by the Joint Economic and Social Data Archive (JESDA). The services include support in data management, data preservation and data access. JESDA has also organized seminars where participants get acquainted with the structure and the content of the archive as well as the usage of the archived data for analytical work.

In summary, the incentives and enablers for data sharing within social science research community in Russia can be benchmarked as developing, but still at a low level.

As data support services to social science research community are available, and there are data producers that follow data management and data documentation standards and procedures that facilitate data reuse, the identifiable gap is related to lack of career-related rewards, financial incentives mainly for the efforts of data preparation and documentation, as well as certain negative attitudes towards data sharing and secondary usage of data.

3.35.2 Capability requirement areas of DAS

Organisational profile

Organisation

The Joint Economic and Social Data Archive (JESDA) was founded on the basis of two structures: The United Archive of Sociological Data (UASD) and the Centre of Statistical Data (CSD). The UASD dates to 2000. Leading research organizations deposited data in the field of social sciences which were published on the internet. The economic part of the JESDA collections appeared and developed within the Higher School of Economics (HSE). The decision to create a united research department was taken by the HSE Academic Council in December 2009.²⁷³

Funding

JESDA is part of the Higher School of Economics (HSE), a state university. The university receives Government Ministry funding as well as some other public and private funding.

²⁷³ Joint Economic and Social Data Archive, About the Archive (December 2016). Retrieved from http://sophist.hse.ru/eng/arch_history.shtml

Core services and activities

The Data Archive has an online catalogue with online access to data and metadata, a set of statistical procedures online and direct downloading of data. The primary and secondary user communities are researchers, students and experts in the respective fields.

Content current collection

The current data collection includes sociological survey data (about 1500 units) and federal statistics on economic issues (about 100 trends). Data is in Russian, partly translated into English.

Organisational infrastructure

Figure 78: Organizational infrastructure in Russia

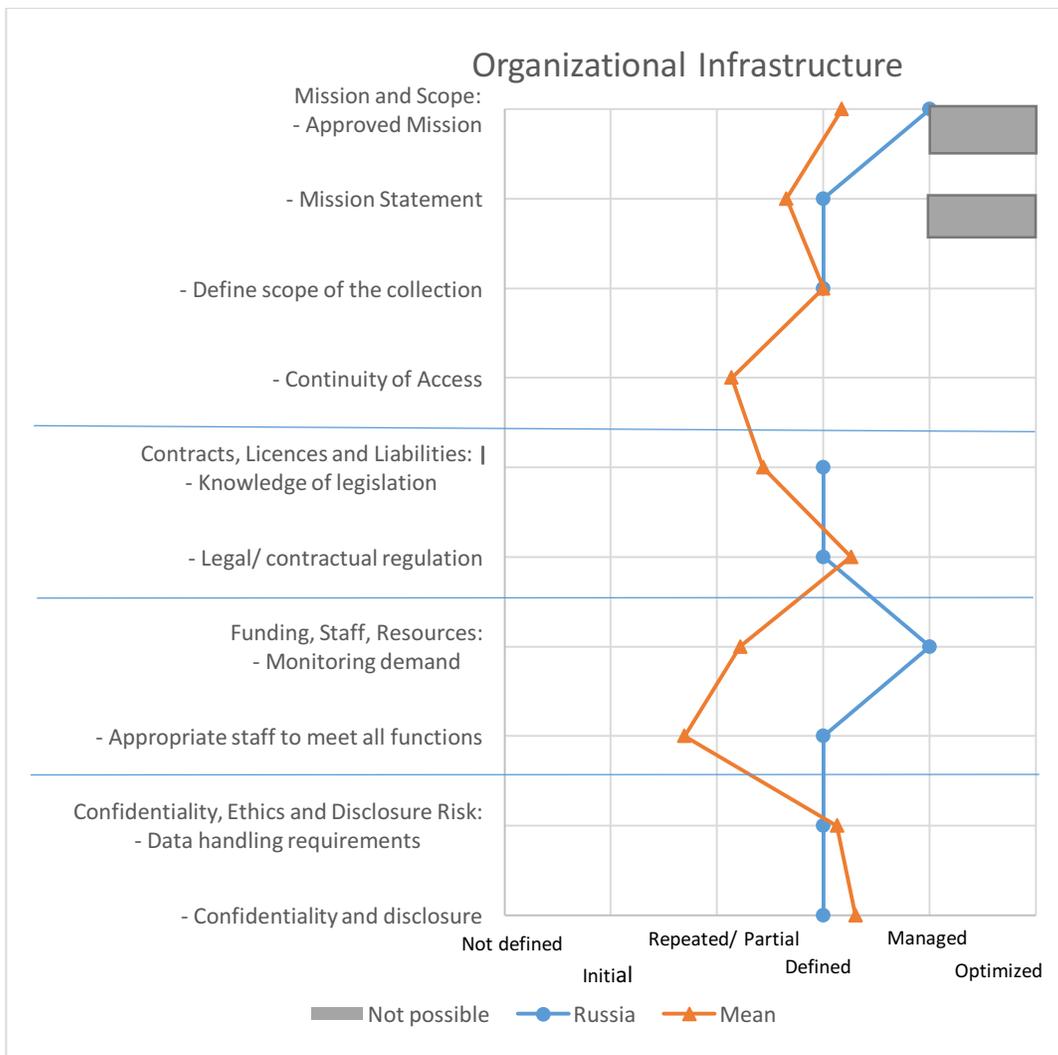
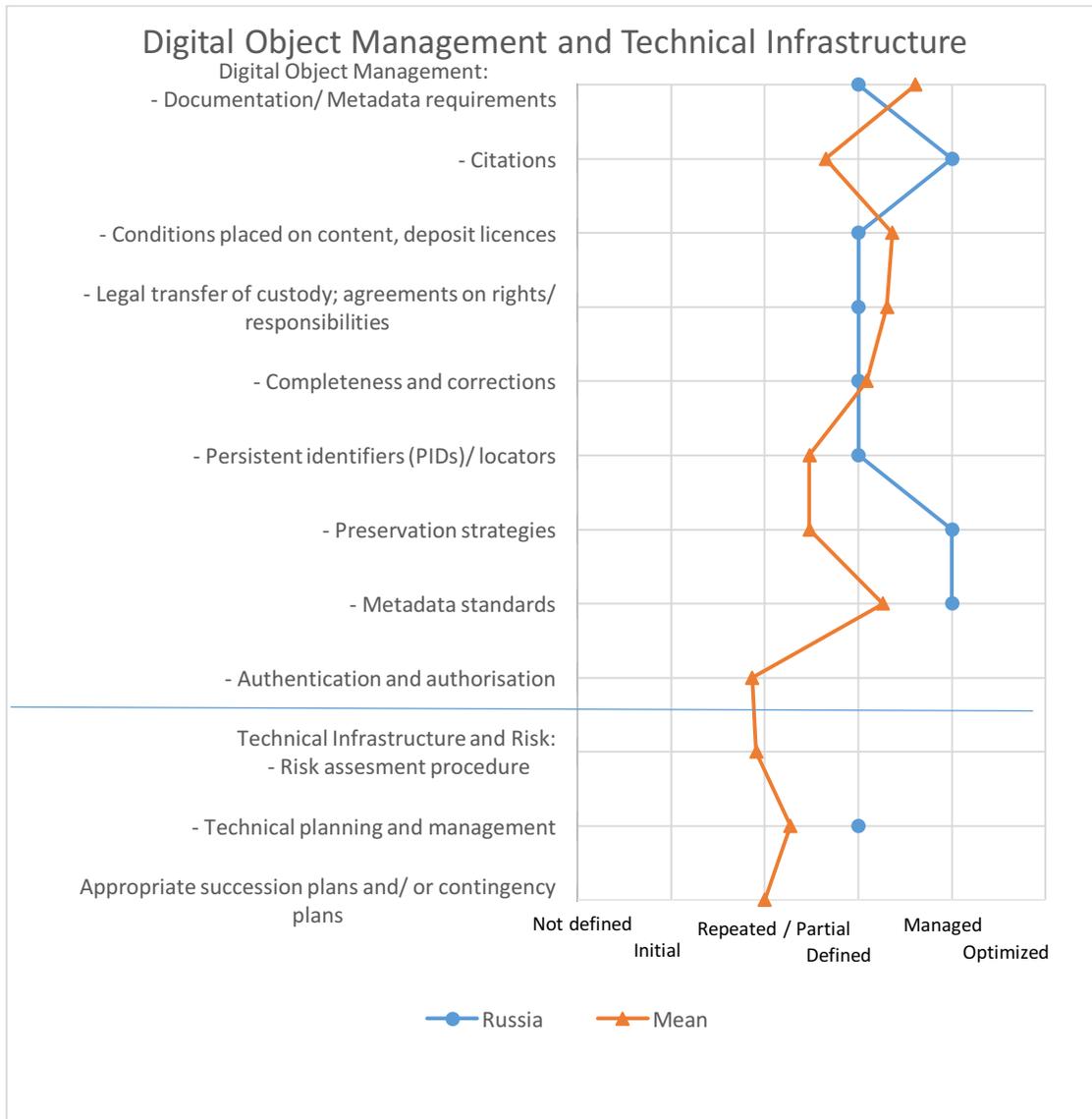


Figure 78 shows that the organizational infrastructure of JESDA is in general at a defined level not significantly deviating from the average. Funding, Staff and Resources are though notably above the mean with Continuity of Access perceived drastically below, not defined.

Digital object management (data curation) and Technical infrastructure and risk

Figure 79: Digital object management and Technical infrastructure in Russia



According to the self-assessment form and as presented in Figure 79, the different aspects of Digital object management and Technical infrastructure are close to the average for all countries included in the study. The few areas far below the average and not defined are Authentication and authorisation; Risk assesment procedure and Appropriate succession plans and/ or contingency plans.

3.35.3 Conclusions

Even though the level of data sharing and secondary data usage is still very low, its significance in Russia has been increasing for the last few decades. The increased number of requests for secondary data usage and for more reliable data and higher quality metadata has impacted the

quality of both data and metadata archived. Few or no rewards are stated to exist for researchers publishing their data upon the completion of research projects, however doing so can increase a researcher's reputation in the research community and his/her chances of obtaining funding in the future. The Joint Economic and Social Data Archive (JESDA) archives a few research data and metadata on socio-economic issues in the Russian Federation with clear requirements and guidelines regarding data preparation and documentation. Further financial incentives for data management and documentation processes as well as education about the benefits of data sharing can further improve the picture of data sharing in Russia.

3.36 Serbia

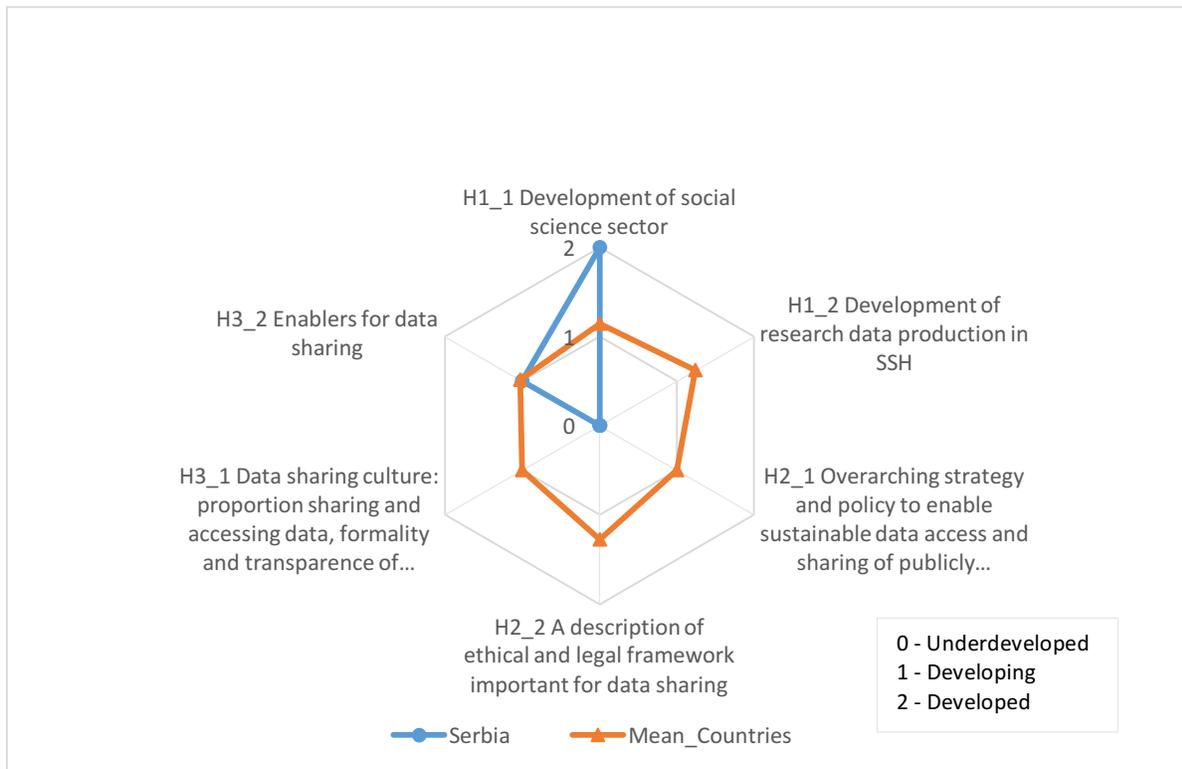
Although the social sciences sector is developed in Serbia and the impact on the designated community is moderate as follows from country macro indicators, research data production, research data management policies, ethical and legal frameworks, data sharing and reuse are all underdeveloped. While the government provides a relatively solid financial support, the results are relatively poorly used in the public decision-making process. There is some development regarding support services for researchers that would facilitate data sharing but for now stronger incentives are absent. Serbia has prototype data archive services already established through the SERSCIDA (see <http://www.rs.serscida.eu/>) and further improvements through SEEDS project are planned in 2017. Serbian partner in SEEDS project (namely [IEN](#)) will prepare 10 datasets for distribution as well as enhance the technical system that was put together within SERSCIDA. Additional to this IEN was involved in RRPP-Rescue project in 2016, where 23 studies from [RRPP project](#) were "rescued" and will be published via [SEEDSbase portal](#), maintained by Swiss data archive - FORS.

Specifics about the data collection in a country

The information for a self-assessment part was entered by Aleksandra Bradić-Martinović and Aleksandar Zdravković, researchers from the Institute of Economic Sciences in Belgrade. The invitation to complete the self-assessment web form was sent on 26/09/2016, and the survey was completed on 12/10/2016. The method used in addition is desk research and the main reference is the assessment of research capacities in the social sciences in Serbia - the report published within the RRPP project framework (RRPP 2010).

3.36.1 Broader ecosystem of DAS operation

Figure 80: Heading concepts values in Serbia



Development of social science sector

Overall Assessment of SSH development

Figure 74 shows that Serbia has a developed social sciences and humanities sector. Funding of the SSH in Serbia are in the highest quantile. The impact on the designated community is strong.

The general intensity of investment in SSH is 0.17 of the GDP. Relative to other disciplines, the intensity of investment in SSH is 21.4% and the number of researchers in SSH per 100.000 capita is 63. The investment in human resources in SSH is 12.803 (per researcher). The most important source of research funding in SSH in Serbia is the government and higher education sector, followed by foreign funding (international or cross-border), the private non-profit sector and last, the business enterprise sector²⁷⁴.

²⁷⁴ "Scientific research in the Republic of Serbia is mainly centralized and administered by the Ministry of Education, Science and Technological Development. Research activity is financed in accordance with the Law on Scientific Research. The Law foresees budget funding of research exclusively through scientific research and developmental projects, while it is expected to carry out funding of other science-related activities through special programs mentioned in the Law. The project financing, resulting from many decades of practice, is based on the public announcement for projects of basic and technological research. The university and independent research organizations realize more than

The government or university provides access to bibliographic and full-text databases and software licences but not to datasets. Furthermore, the country representative reported about the Serbian Library Consortium for Coordinated Acquisition (KoBSON), which is “a new form of organizing the library of Serbia.”²⁷⁵

Development of research data production in SSH

Figure 3.36.1. shows that research data production in Serbia is underdeveloped. Serbia is not involved in international collaborative research or cross-national studies such as CSS, ESS, ISSP and WVS, except for the Comparative Study of Electoral Systems (CSES) in which it was involved in 2012, the European Values Study (EVS) in which it was involved in 2008 and PISA in which it was involved in 2003, 2006, 2009 and 2012. Historically Serbian empirical social science was involved in several collaborative projects in former Yugoslavia, such as Yugoslav Public Opinion Survey lead by the Institut društvenih nauka Univerziteta u Beogradu (Institute for Social Science, University of Belgrade).

In addition, there are some existing studies that systematically assess matters of national importance but the national representative was not specific about that and did not give any example. There is some periodical production as ‘institutions have a tradition of providing some type of research data to a certain extent’. Specifically, “*according to SERSCIDA survey Serbian social science researchers produce 9 data sets per researcher in five years’ period, in average. However, that is only an estimate. The analysis shows that 6% of produced one set of data, 10% produced 2 sets and 13% produced 3 sets. Slightly less, 10%, produced 4 sets, and 15% produced 5 sets. The largest number of*

35% of all the research activities conducted at the level of EU countries. These organizations are holders of the largest number of basic and applied research activities and research activities of importance for the public interest. All scientific and academic institutions deal with permanently low funding for their activities and projects, particularly in the field and empiric research areas. Some research institutes have foreign external financial resources for short term projects, which are phasing out since the interest of donors for Serbia and the whole Balkans as a post-conflict area has significantly decreased. The structure of funds for science financing in Serbia is as follow: 50% budget revenues, 48,3 international projects, 0.8% domestic borrowings and 0.4% international borrowings (SERSCIDA Country Report, 2012). Contribution in budget revenues for social science in Serbia is much lower in comparison to other sciences. One of the main funding scheme in Serbia for social sciences is the Regional Research Promotion Programme (RRPP) Western Balkans, launched in 2007, funded by the Swiss Agency for Development and Cooperation (SDC), aims to establish and strengthen research capacities in the field of social sciences in Albania, Bosnia and Herzegovina, Kosovo, Macedonia, Montenegro and Serbia. The RRPP provides funding to researchers from the region conducting relevant research projects. Until the end of 2016, RRPP covered about 100 research project in the whole region and 26 in Serbia. According to Innovation Union Scoreboard 2011, Serbia’s performance in the research field is below average. It has relative advantages in human resources, open and attractive research systems, ways of financing and economic effects, but it has relative weaknesses in investments from the business sector, links with entrepreneurship, intellectual property, and innovators.”

²⁷⁵ “The initiative for forming the Consortium was launched in November 2001 by the leading research libraries in Serbia, i.e. the National Library of Serbia - Belgrade, Matica Srpska - Novi Sad, University Library "Svetozar Marković" - Belgrade, University Library "Nikola Tesla" - Niš, University Library - Kragujevac, Library of the Academy SANU - Belgrade and Community of university libraries. The main objectives of the association are optimized procurement of foreign scientific information, the transition from paper to electronic editions, improving access to electronic information and promotion of national scientific publishing. This service provides access to: American Chemical Society, American Physical Society, American Psychological Association - APA, ASME Transaction Journals, Cambridge University Press, Emerald Publishing, Institute of Physics Publishing, Oxford Journals, Royal Society of Chemistry, SAGE Publishing, Science Direct, Springer/Kluwer and Wiley Interscience. Through this service Serbian sciences can use several aggregators: Cairn.info, EBSCO, Free Medical, Hein On Line, High Wire, JSTORE, DOAJ, Open Access Archives (just for older issues), Project MUSE and TEEAL. Ministry, through unified public procurement, also provide software for project participants. The problem is that the system of procurement is extremely slow.”

researchers, 24%, produced 6-10 sets, 12% produced 11-20 sets, and 10% more than 21 sets of data. The main problem with data collection is a lack of funding within public projects.”

RDM Policy setting

Research data management policies in Serbia are underdeveloped. Currently Serbia’s National Science Foundation does not require DMP but there is growing recognition and awareness of the need to require DMP. However, for now *“the research institutions do not have any knowledge about DMPs, but researchers are becoming familiar with it through international projects, mainly within H2020”*.

There is also ‘no offering or depositing data in an appropriate disciplinary repository and no long-term curation for valuable research data assets. ‘As the representative noted, “there is no awareness at the level of the Ministry of Education and Science about preserving or sharing research data, although the Ministry is the main public funder of social science research institutions. Institute of Economic Sciences is in the process of establishment of data archive in SSH starting from 2012, supported by EC (FP7 project SERSCIDA, H2020 project CESSDA SaW) and SNSF (SCOPES SEEDS project) but without proper support of Ministry.” In addition, the representative reported that *“based on SERSCIDA Report (2012) there is no research institution in Serbia which has data repository. Up to now, we are not aware that anything has changed on that matter. Some researchers had the opportunity to be engaged in the international projects and get familiar with data preservation. A first organized collection of research data is carried out in the framework of RRPP Data Rescue project (2016).”*

Data sharing culture

The data sharing culture is underdeveloped in Serbia. In fact, data sharing and reuse among social sciences researchers in Serbia is not common – as self-assessment results indicate, the proportion of researchers sharing data is estimated as low (0-10%), and the proportion of researchers able to access existing third party data they need – as low (0-10%). The country representative noted that *“based on SERSCIDA report, the practice of data sharing of Serbian SSH researchers is underdeveloped and regarding that data sharing infrastructure are at the rudimental level there is no reason to believe that something significantly has changed since 2012.”* The country representative was not able to provide an estimate for 2011-2016 regarding established data sharing channels and routines. In fact, up to their knowledge, *“these routines are applied by the SSH researchers, but these are individual cases, so it’s hard to give some reliable assessment of the ranking.”*

The attitudes of researchers toward data sharing, as document based on self-assessment shows, can be characterized as neutral. The country representative says it is true that the “reuse of existing data can answer new research questions and facilitates advancement of science” but that “data sharing has a risk that others may misuse and misinterpret data”. The also mostly agree that “data creates negative competition”. On the other hand, they marked false that statements that “data sharing has no benefits at all” and that it “involves little effort and minimal costs”. Also the statement that “data creates healthy competition” is evaluated close to false.

Attitudes towards data sharing

Table 28: Attitudes towards data sharing in Serbia

Data sharing has no benefits at all	False
Data sharing creates healthy competition	Probably false
Data sharing creates negative competition	Probably true
Reuse of existing data can answer new research questions and facilitate advancement of science	True
Data sharing has as a risk that others may misuse and misinterpret data	True
Data sharing involves little effort and minimal costs	False

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from "5"-True to "1" False!

To summarize data sharing culture in Serbia, data sharing and reuse is not very common, there is no information on its channels, while attitudes can be characterized as neutral to negative.

Enablers for data sharing

Enablers for data sharing are underdeveloped in Serbia. There are no career rewards related to data sharing in the academic community but there are data support services available to social science researchers that would facilitate data sharing and/or Open Access to research data.

There is no awareness regarding the support needed on legal and ethical aspects that facilitate social science data sharing. In addition, the country representative noted that:

"The electronic or digital archiving in the Republic of Serbia is not regulated by special law. However, several legal regulations regulate the work of archives and electronic documents, in the way that they touch upon this field in their provisions. A part of the respective legislation is inherited as a federal regulation from the former Federal Republic of Yugoslavia (FRY), while another part was and remained at the level of Republic regulations. In that sense, there are three legal regulations. First, it is necessary to highlight the Law on Archival Material ("FRY Official Gazette", no. 12/98, 13/98). It is about the previous federal law, which has been applied since 2006, as Republic regulation in Serbia. This Law defines the notion of archival material as original and reproduced written, drawn, printed, photographed, filmed, recorded with a phonograph, recorded legibly with a typewriter, or recorded in some other way - documented material of permanent value and importance for science, culture, and other social needs. As archival material of Federal Republic of Yugoslavia, nowadays in the Republic of Serbia there is the archival material that appeared, amongst other things, in the work of public enterprises, public institutions, and other organizations founded by federal law, and of international gatherings the organizers of which are Federal Republic of Yugoslavia bodies, nowadays bodies of the Republic of Serbia. Also, it defines registry material as overall original and reproduced written, drawn, printed, photographed, filmed, recorded with phonograph, recorded legibly with typewriter or recorded in some other way - documented material and books, files and other records of such material, from which archival material was not selected. According to the law, the bodies and organizations in the work of which the archival material emerges are obliged to mark, date, classify, carry out current archiving and

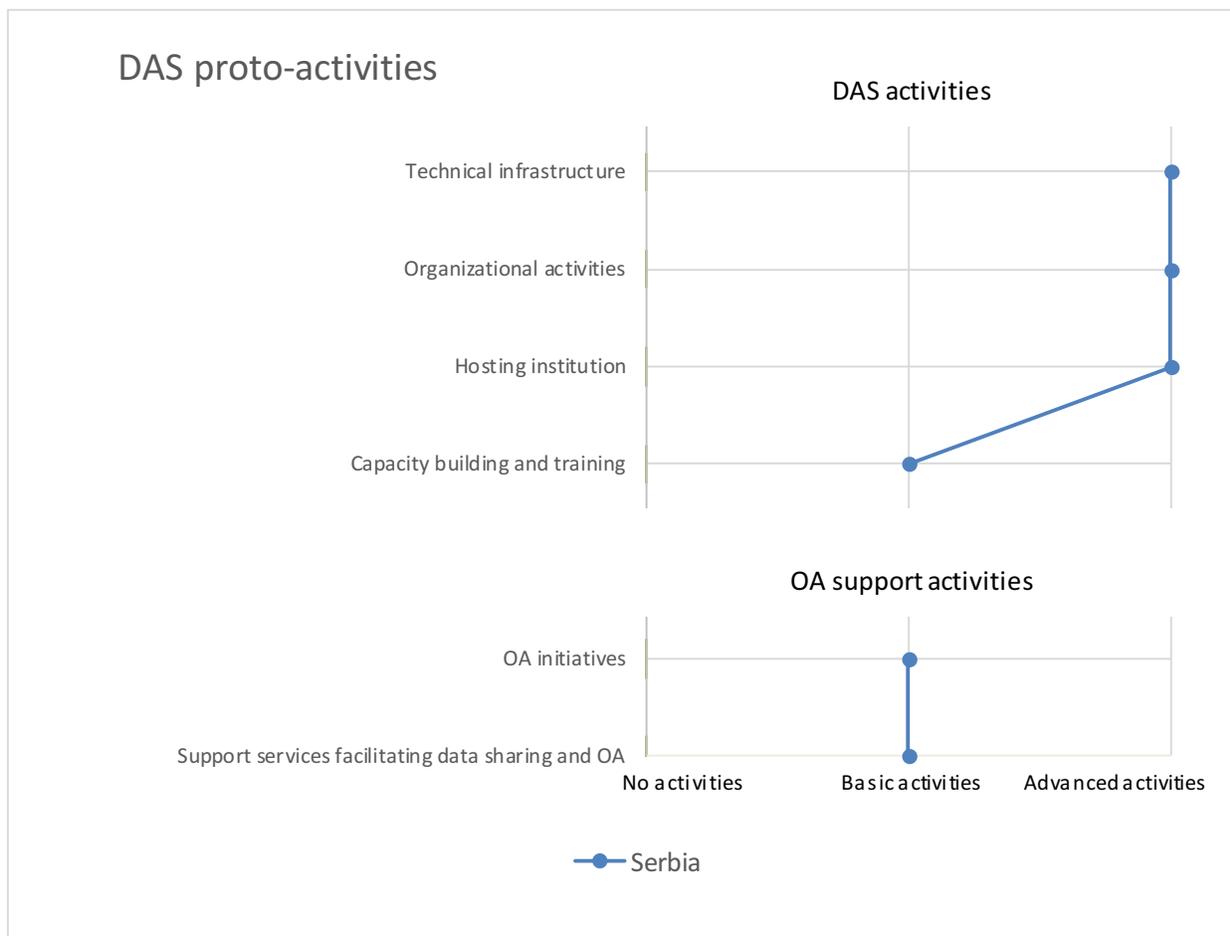
preservation, select archival material, give data, and hand over the registry material to the competent bodies. Since the project subject relates to the authorship works of scientific staff, the Law on Copyrights and Related Rights (“RS Official Gazette”, no. 104/2009) is applied perforce on this field. The Law prescribes that the authorship work is the author’s original intellectual property, which is expressed in certain form, regardless of its artistic, scientific, or other value, its purpose, size, contents and other ways of expression, as well as permissibility of public disclosure of its contents. As the authorship work, there are considered especially written works, such as books, brochures, articles, translations, computer programs in any form of their presentation, including the preparatory material for their production, etc.”

Moreover, social science data producers in Serbia ‘do not follow data management and data documentation standards and procedures that facilitate data reuse’.

In summary, the incentives and enablers for data sharing within social science research community in Serbia can be benchmarked as developing.

3.36.2 DAS proto-activities

Figure 81: DAS and OA activities implementation type in Serbia



DAS activities

In Serbia, no Data Archive Service (DAS) exists yet, but advanced proto-activities related to technical infrastructure, organisation and capacity building are provided on a national and institutional level through SER-DAC (the Serbian Data Center in Social Sciences). Also, there are existing technical infrastructures within the social science field that could possibly be used or applied to a new DAS in social sciences and the activities are quite advanced.

As a result of the SERSCIDA project, SER-DAC, formally established in March 2014 as an organizational unit of the Institute of Economic Sciences Belgrade (IES), has been equipped with basic infrastructure for the data preservation. It is still in the early phase and supported only by international funds through the SCOPES SEEDS project (<http://seedsproject.ch/>) and HORIZON2020 CESSDA SaW project (<http://cessdasaw.eu/>). It can preserve and disseminate all types of data, both quantitative and qualitative, through the FORS data portal SEEDSbase (<https://seedsdata.unil.ch/>). The data scope covers all areas in the social sciences and humanities.

The development of the knowledge and skills, such as data preservation and data dissemination started during the SERSCIDA project and continued through the SEEDS and CESSDA SaW projects.

Among existing national infrastructure associated with recording of scientific work probably COBISS.RS²⁷⁶ may be used for enhancing visibility of archived data (by linking the records of articles in COBISS.RS to DAS).

In Serbia, there are also already activities towards establishing a DAS for the social sciences but only a few and basic. The country representative uploaded a draft version of the Development plan which will be finalised by the end of the SEEDS project (April 2017) which mentioned:

- The involvement, infrastructure and skills of the IES (or Serbian Data Center in Social Sciences - SER-DAC), makes it the more advanced institution and the best option for Serbia to form an institution that could host a DAS for the social sciences, as described in the “enablers of data sharing” section.
- There are several dissemination and outreach activities for key national stakeholders planned within the SEEDS project. For instance, local workshops will be organized for representatives from the research policy makers and funding institutions, statistical offices, national archives, and national and university libraries, and research communities. The purpose of the workshops will be to present the project, develop networks, and to promote the importance of open data, data archiving, and secondary analyses.

Open access (OA) support activities

Some support activities to encourage and facilitate OA exist in Serbia through the participation of the University Computing Center in Belgrade (RCUB) in the OpenAIRE project within European Horizon 2020 program. The aim of the project is for open access to become the dominant open access publishing model. The European Commission insists that all facilities that may be disclosed

²⁷⁶ COBISS (Co-operative Online Bibliographic System and Services) is a platform for shared cataloguing of autonomous library information systems. http://www.cobiss.net/cobiss_platform.htm

as a result of the projects financed under the Horizon 2020 program should be offered in the open-access regime.

To conclude, there are some data support services that facilitate data sharing and OA to research data in social sciences. However, there are only a few activities, not on a regular basis and in only a few institutions.

3.36.3 Conclusions

According to the conducted analysis, it can be concluded:

- Funding of SSH and productivity of the researchers in Serbia are at the high level of development and that impact on designated community is strong.
- Development of research data production in SSH is underdeveloped. Serbia is involved in only a few cross-national studies but there is some national data production, there are some databases that were archived through the RRPP project.
- Research data management policy is not required by the national institutions and data management plans are underdeveloped.
- Ethical and legal framework important for data sharing is underdeveloped. There is a little or no awareness regarding the support needed.
- Data sharing and reuse is not very common in the social sciences research community; there is no information on its channels, while attitudes can be characterized as neutral and therefore underdeveloped.
- Incentives and enablers for data sharing within the social science research community in Serbia are currently not existent but there are data support services available for researchers that would facilitate data sharing. Thus, it can be benchmarked as developing.
- Serbia has basic DAS activities through SERSCIDA, SCOPES-SEEDS and CESSDA SAW project that have created capacities in terms of human resources. Individual people were trained to work; basic protocols and technical solutions are prepared and pilot operations of data services are about to get off. It is expected that also the support environment on the level of liabilities and principles of open access policies will follow.

3.37 Slovak Republic

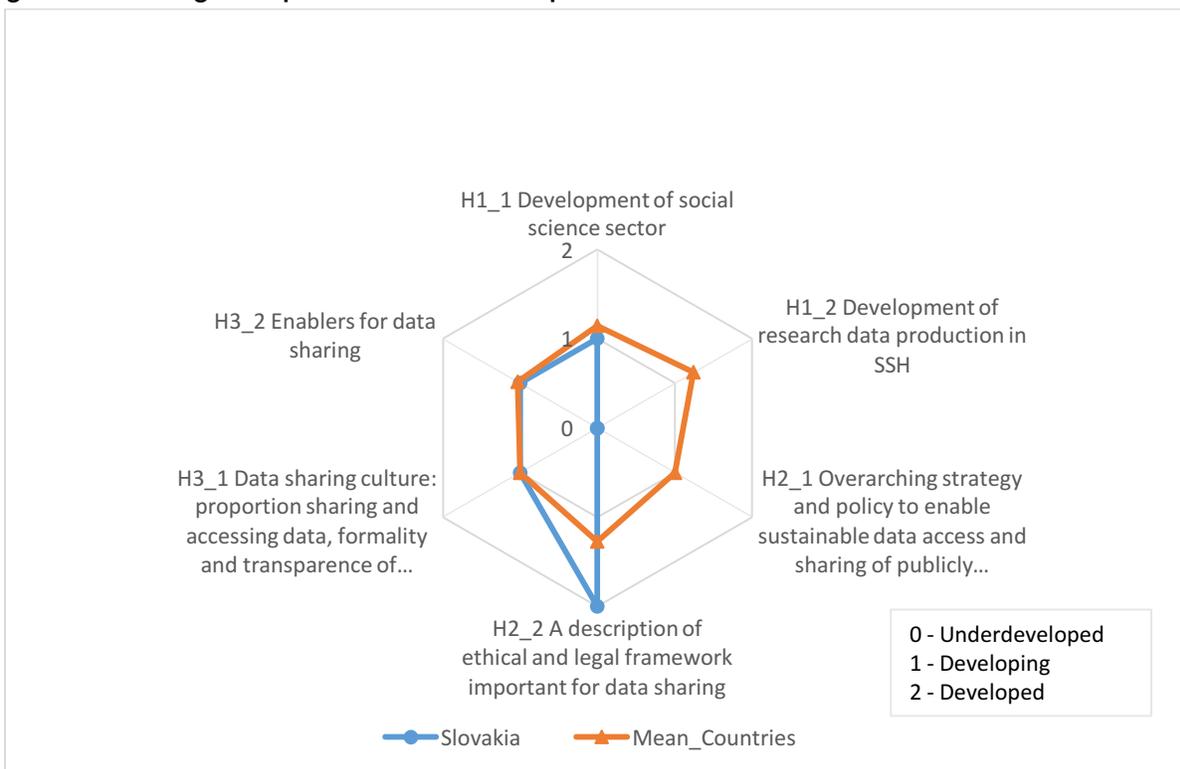
Social sciences and humanities in Slovakia are on the medium level of development and the production of research data is irregular. The country participates in several important international research series but relevant institutions do not promote research data management principles. The data sharing culture is generally underdeveloped. Since the incentives for data sharing are lacking in Slovakia, only one in ten Slovak researchers shares the data he or she produced. Most important institution in the field of research data archiving and open access to data is the Slovak Archive of Social Data (SASD). As an observing member of CESSDA this small archive has a potential, even though its personal and financial resources are very limited. SASD is still unable to provide some important data services and act in accordance with standardized data archiving principles.

Specifics about the data collection in a country

This report for Slovakia is based on the data from the self-assessment template filled in by Miloslav Bahna, head of the Slovak Archive of Social Data (SASD), if not stated otherwise. Communication with SASD staff was without any problem. Rest of information was obtained through the desk research.

3.37.1 Broader ecosystem of DAS operation

Figure 82: Heading concepts values in Slovak Republic



Development of social science sector

Overall assessment of SSH development

According to Bahna, the rank of the sources (sectors) of research funding in social science in Slovakia by the amount they provide is following: 1. abroad (international or cross-border) sources, 2. government and higher education sector, and 3. private non-profit sector. The bibliographic and the full-text databases are accessible at public universities and at the Slovak Academy of Sciences. The overall development of SSH in Slovakia was evaluated as one at the medium level. The funding of SSH and the productivity of the researchers are in the mid quantile, while the impact on designated community seems to be rather limited.

Table 29: Assessment of SSH development in Slovak Republic

Indicator	Values for Slovakia	Rating of Slovakia (0 low level, 1 medium level, 2 high level)
GERD in SSH as % GDP	0.15	1
GERD in SSH as % of GERD	16.5	2
Number of researchers in SSH per capita	159	2
GERD in SSH per researcher in SSH	12904	1
Access to databases	-	0
WoS publications	missing	
Average	-	1.2
Overall	-	1

Development of research data production in SSH

Slovakia participates in European Social Survey (waves 2, 3, 4, 5, 6), ISSP (regularly), Comparative Study of Electoral Systems (2010 wave), European Values Study (wave 2, wave 3, wave 4), World Values Survey (participated in 1991), and Programme for International Student Assessment (participated in 2003, 2006, 2009, and 2013). Among existing studies that systematically assess matters of national importance the surveys of ISSP and EVS were emphasized. Between 1995 and 2009 (2010) based on available data the Institute for Public Affairs (the Slovak NGO) published an annual Global Report on the State of Society. The country participates in the international collaborative research. The overall assessment showed that the data production in SSH in Slovakia is in the developing stage, i.e. there are some examples of research excellence and streams of research do stand out.

Table 30: Development of research data production in SSH in Slovak Republic

Indicator	Rating of Slovakia (0 low level, 1 medium level, 2 high level)
International collaborative research	1
Studies of national importance	1
Average production of data	1
Sum	3
Overall	1 (developing)

RDM Policy setting

Public research funding organizations operating in Slovakia provide no incentives for sharing research data and associated metadata. The costs for managing the data sets and preparing them for access are not resourced adequately during the research project lifetime. Mentioned above can be seen as big obstacle for improvement since there are neither career nor financial rewards related to data sharing.

Regarding the clarification and the support provided on the legal and ethical aspects that facilitate social science data sharing, it was mentioned that there are some recommendations and guidance on how to fulfil the legal requirements while sharing the data. In this respect the main focus is on the protection of sensitive personal information. E.g. Slovak statistical office has the guidelines regulating the usage of the detailed/aggregated data for the personal data protection (for the case of LFS).

Data sharing culture

Proportion of social science researchers in Slovakia that have shared the research data they produced in the period between 2011 and 2016 was estimated as rather low (up to 10 per cent).

There are differences in various research areas. There is an awareness about the data sharing in the field of sociology; here it is especially promoted by the Slovak Archive of Social Data well-known to the most of the researchers. However, based on the data collection, it seems that in the other social science disciplines the awareness of data sharing is rather weak. It was estimated (based on the experience of SASD team) that the proportion of the Slovak social science researchers able to access the needed third party data in the period of last five years (between 2011 and 2016), was quite substantial, more than 30 per cent. However, some differences were noted when it comes to the different research areas as third party data seem to be more accessible if produced by a public / governmental institution (e.g. the Statistical office).

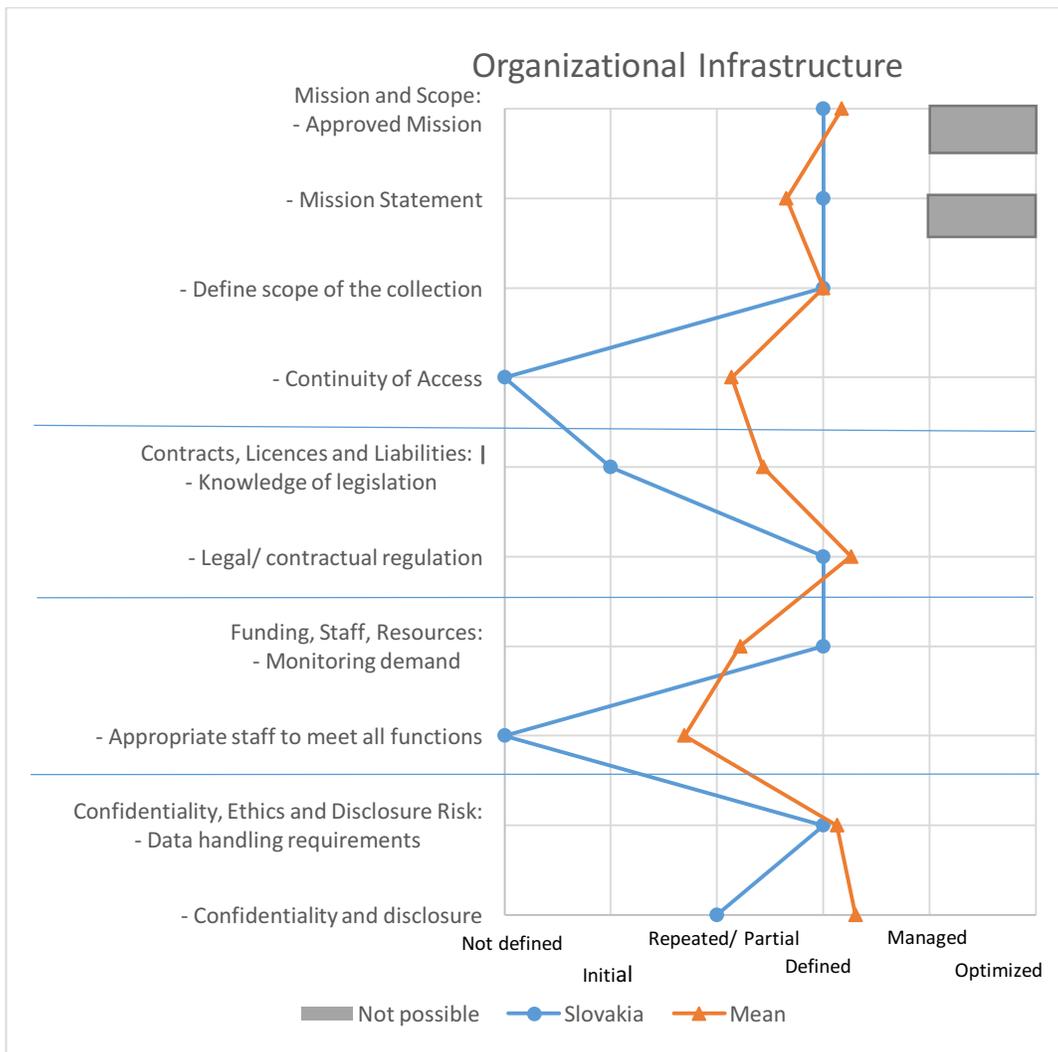
Existing routines for data sharing in Slovakia are based on the preferences of the social science research community. According to the experience of SASD team, most preferred way to obtain the data in Slovakia is via informal contacts (peers and colleagues); the data archive (SASD) was estimated as a second-choice source of data.

Regarding the assessment, it is not very common in Slovakia to publish in journals that expect data used in the publication to be available for reuse from a trusted digital repository. Head of SASD based mentioned above statement on his personal professional experience as an editor in chief of the leading sociological journal in Slovakia (*Sociológia / Slovak Sociological Review*). According to Bahna, national journals do not require research data to be publicly available and the publication activities of Slovak SSH scholars when it comes to the international journals are rather limited.

3.37.2 Capability requirement areas of DAS

Organisational profile and infrastructure

Figure 83: Organizational infrastructure in Slovak Republic



The Slovak Archive of Social Data (SASD) was established in 2004 with the aim of making the documentation, data and information about the results of sociological research accessible for the needs of research and non-commercial use, and to store them in electronic and printed form. The archive is physically and institutionally located at the Institute for Sociology of Slovak Academy of Sciences. The archive's webpage (<http://sasd.sav.sk/en/>) and the data library are located at the web servers of the Slovak Academy of Sciences.

The archive is not a separate department in the Institute but is managed and funded as a part of the library of the Institute for Sociology of SAS. This means only one full time position is shared between library and the Archive. The number of staff appears to be inadequate to support all the functions and services. Many core competencies are missing because of it. There is an ad hoc funding available from the projects at the Institute which finance the development of the data collection by actively searching for older interesting datasets and contacting its owners. Currently another means of funding is the CESSDA SaW project.

SASD complies with norms and legal requirements through systematic written, formal procedures and policies. SASD builds sufficient knowledge and documentation on how relevant aspects of the legislation apply to and affect the holdings and procedures of the organisation. The organisation provides a written, formal and complete mission statement that is available for all relevant users. According to SASD mission statement „*the services of SASD are intended for the non-commercial use of the general public, for the needs of socio-scientific research, for journalists, and for governmental and non-governmental institutes*”. Hence the aim is not only archiving, but also increasing in the quality of public level information about life and changes within the Slovak society. At the same time, the data archive is also intended for the commercial and other research agencies and institutions as a basis for communication with the public through the publication of their research according to the SASD conditions²⁷⁷.

SASD has neither medium-term (3 to 5 year) nor long-term (>5 years) plans in place to ensure the continued availability and accessibility of data in case the organisation ceases to operate.

Core services and activities

Social science scholars and students are the main users of SASD services. International experts on Slovakia and the Central European region form the secondary user community. Key service is archiving empirical data and documentation from accessible sociological research performed in Slovakia and making the data available for secondary analysis. At regular intervals (ideally annually or more frequently) the institution monitors the demand for its repository services. SASD annually reviews the number of downloaded datasets.

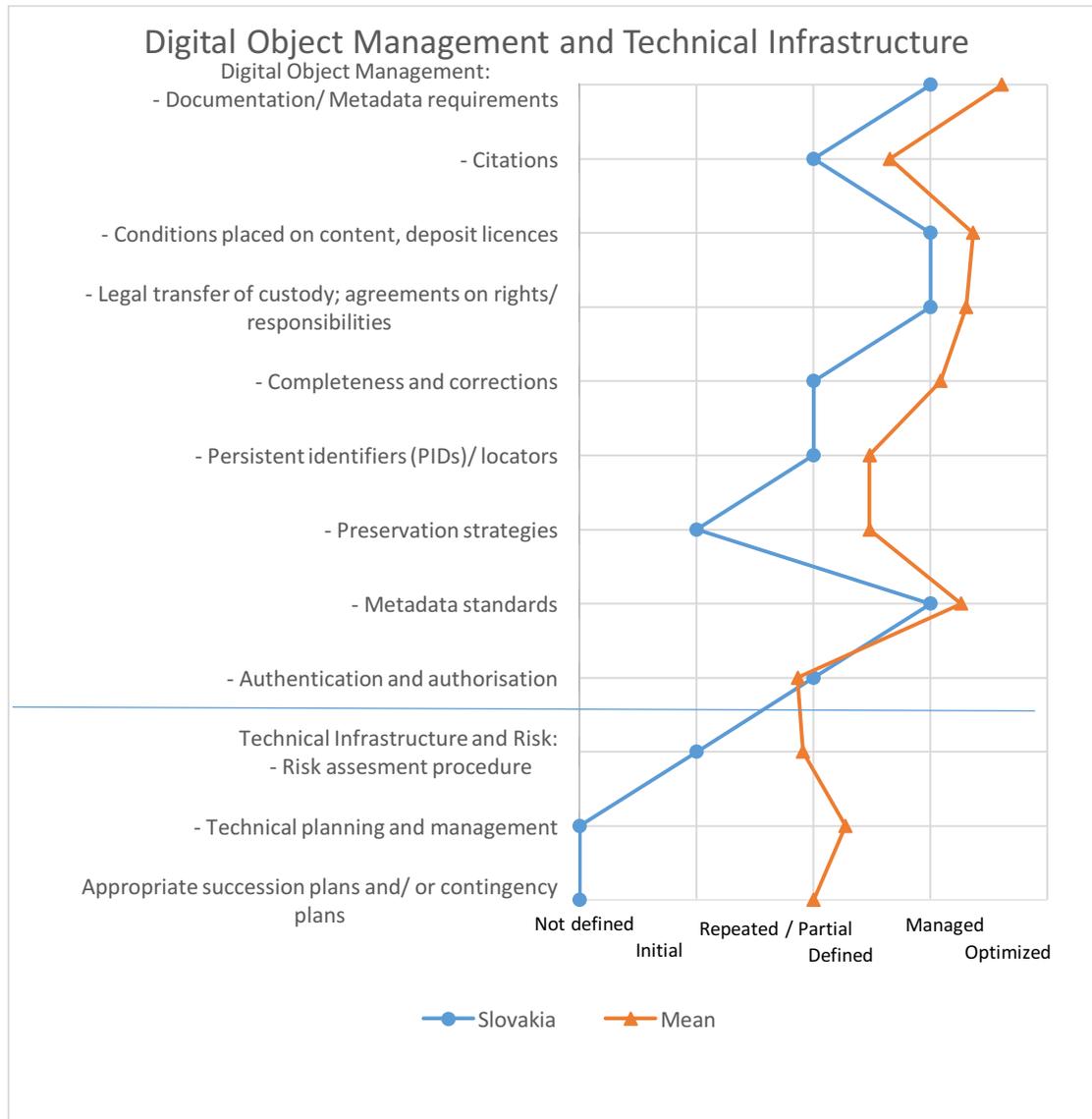
Contracts and/or agreements with both depositors and users are standardised and implemented according to written procedures; contracts and regulations and are made publicly available.

Depositors are repeatedly informed about confidentiality issues, but no formal information template exists. Checks are performed on the deposited data, but there are no written procedures in place (some documentation may exist but it is incomplete). The issue of confidentiality and non-disclosure for collected data is handled on a case by case basis; however no formal guidelines exist.

²⁷⁷ See: http://sasd.sav.sk/en/o_archive.php

Digital object management and Technical infrastructure and risk

Figure 84: Digital object management and technical infrastructure in Slovak Republic



SASD clearly specifies the information (documentation, metadata, provenance) that needs to be associated with the data that is to be deposited²⁷⁸. A written formal specification of required information is explicitly defined (e.g. in a collection policy); requirements are compliant with metadata standards that are used and can be understood by Designated Community (e.g. DDI); metadata requirements are accessible and communicated to users and depositors.

SASD collects the citation information from its users (the provision of this information is required), but it does not distribute it to the depositors. So far, the archive received no request for such information.

²⁷⁸ More elaborate info at http://sasd.sav.sk/en/prispejte_dokumentacia.php

Content of current collection

Scope of collection is explicitly defined in the official mission statement of SASD. Statements on scope are connected to policies; policies are, in their turn, connected to specific processes and procedures. The main language of the collection is Slovak, while approximately one third of the collection has English language documentation.

The core of the data collection consists of data produced by the Institute for Sociology, mostly national data from studies within international survey programmes like ISSP, ESS, and EVS. There are also few datasets from other data producers. Almost all the datasets are data from the survey based research.

3.37.3 Conclusions

Empirical social sciences in Slovakia still produce limited number of data; though there is some data relevant for secondary analysis (mostly the data from international research series with Slovak participation). Data sharing culture is underdeveloped at the moment. There are no incentives for data sharing in Slovakia and only one in ten of Slovak researchers share produced data. Leading institution in area of data archiving and open access to data is SASD. After twelve years of development SASD is well established data archive. The Slovak archive is now the Observing member of CESSDA. In order to become a „mature” archive SASD needs an appropriate staff to fulfil all necessary functions.

3.38 Slovenia

Research data production in SSH in Slovenia is on quite high level despite limited resources. Slovenia participates in most of the important international research series fielded in Europe.

Culture of data sharing is still underdeveloped (estimated percentage of researchers sharing data is low) but near future is promising in this regard thanks to new strategies of Slovenian government and new conditions of European grants (H2020).

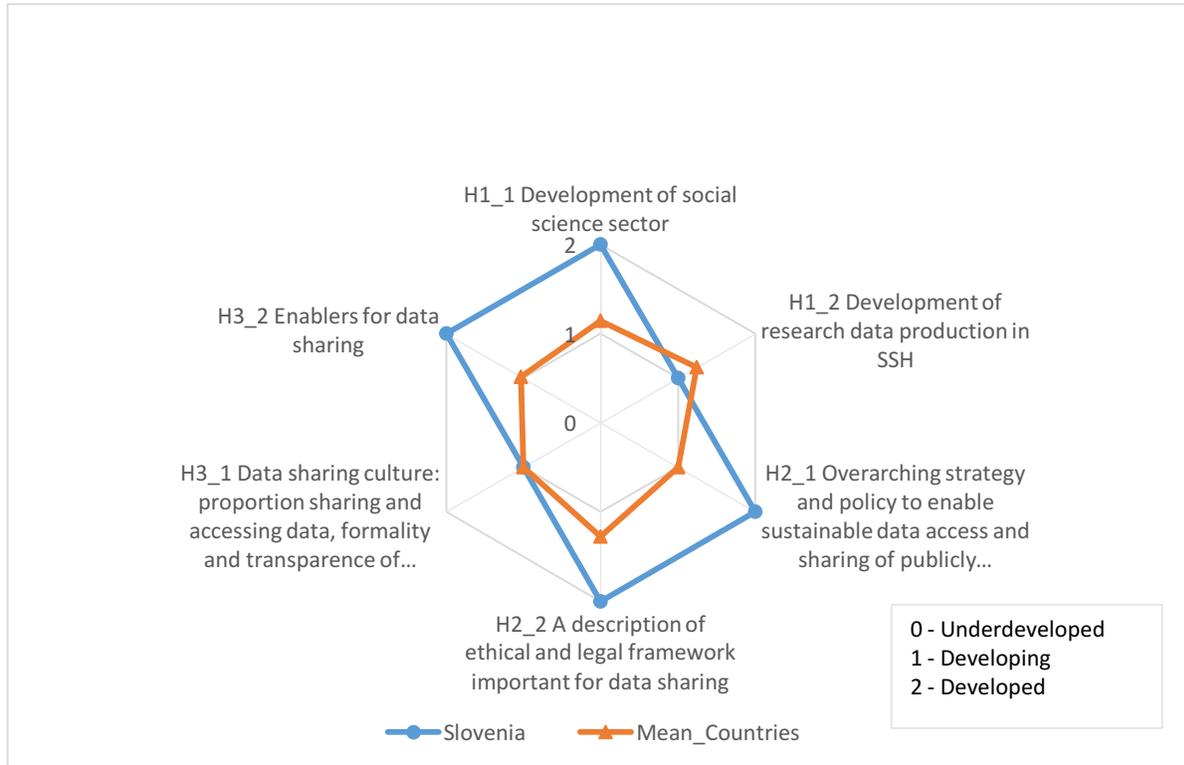
Leading institution in the area of research data archiving is Social Science Data Archives (ADP – Arhiv družboslovnih podatkov), which could be characterized as a rather small but „mature“ repository using highly elaborated standards of data archiving. Apart from providing the infrastructure services in SSH, the Slovenian data archive plays a very important role in promoting the data sharing culture on the national level.

Specifics about the data collection in a country

The data collection for this report was conducted using a Web-form during October 2016. If not stated otherwise, the data for Slovenia comes from the self-assessment template filled in by the team of the ADP, namely by the head of ADP Janez Štebe, and Sonja Bezjak and Irena Vipavc Brvar.

3.38.1 Broader ecosystem of DAS operation

Figure 85: Heading concepts values in Slovenia



Development of social science sector

Overall assessment of SSH development

According to the collected data, the rank of the sources (sectors) of research funding in social science in Slovenia by the amount they provide is as following (starting from the biggest source): 1. Government and Higher education sector, 2. Sources from the abroad (international or cross-border), 3 Business enterprise sector, 4. Private non-profit sector.

Bibliographic and full-text databases (some OECD datasets) and national software licences appear to be accessible in Slovenia. However, the whole coverage is not yet achieved in all the mentioned fields. For instance, the Consortium of Libraries negotiates and selects the scientific literature services. Commercial data sets and software licenses are also selectively and sometimes partially covered, depending on institutional policies and resources available.

The overall development of SSH in Slovenia was rated as a high level one; although the funding of SSH and productivity of the researchers are in the mid quintile and the impact on designated community is estimated as rather limited.

Table 31: Assessment of SSH development in Slovenia

Indicator	Values for Slovenia	Rating of Slovenia (0 low level, 1 medium level, 2 high level)
GERD in SSH as % GDP	0.17	2
GERD in SSH as % of GERD	6.6	0
Number of researchers in SSH per capita	91	1
GERD in SSH per researcher in SSH	32461	2
Access to databases	access on certain places ²⁷⁹	2
Average	-	1.4
Overall	-	2

*Development of research data production in SSH***Table 32: Development of research data production in SSH in Slovenia**

Indicator	Situation in Slovenia	Rating of Slovenia (0 low level, 1 medium level, 2 high level)
International collaborative research	ISSP, EVS, CSES, ESS, PISA	1
Studies of national importance	Slovene public opinion survey, political barometer.	0
Average production of data	periodical	1
Sum	-	3
Overall	-	1 (developing)

Slovenia is a regular participant of European Social Survey (participated in all the waves), ISSP (regular participation since 1991) and Comparative Study of Electoral Systems (all the waves). It

²⁷⁹ Government or universities in Slovenia provide access to commercial bibliographic and full-text databases (such as WoS, Scopus, EBSCO or JSTOR, etc.), datasets and national software licences.

also participated in European Values Study (wave 2, wave 3, and wave 4), Programme for International Student Assessment (participated in 2006, 2009, and 2013) (see Table 3.38.2).

Among existing studies that systematically assess matters of national importance for Slovenia ADP mentioned Slovene public opinion survey (starting in 1968). Some telephone polls devoted to current problems and political barometer were also pointed out. Generally, production of SSH data in Slovenia was characterized as „periodical“.

Overall research data production in SSH in Slovenia was characterized as one in the „developing stage“ with some research excellence examples and particular streams of research standing out, either qualitative or quantitative, and also some examples of advanced international collaborative research.

RDM Policy setting

Regarding requirements or recommendations about preparing Data Management Plans (DMPs) as an integral part of on-going project activities we can say generally that there is the expectation or recommendation to have DMP in place on the political level. However, it's not officially implemented as part of funders' policies; and it is generally not part of the culture yet. There is the expectation or recommendation to offer or deposit data in an appropriate disciplinary repository or equivalent data archive service as well. RDM requirement is included in National strategy of open access to scientific publications and research data in Slovenia 2015-2020 (approved by Slovenian Government in 2015) and in the Action plan expected to be confirmed in 2017²⁸⁰.

The National strategy mentioned in that an appropriate funding is in fact essential to ensure the sustainability of the infrastructure to provide an open access to scientific information²⁸¹. Cost for managing the data and preparing it for an access can be implicitly covered in research budgets. The Slovenian research agency (ARRS) mentioned 'open access' in its funding contracts and it interprets an overhead of research project budget as a budget that could be also used for RDM. In the Action plan draft it is said that this will change to more explicit requirements and corresponding budget if planned and grounded in the project description.

Situation in Slovenia regarding clarification and support provided on legal and ethical aspects aimed on facilitation of the social science data sharing is only partially satisfying. Although Social Science Data Archives provide recommendations and guidance, the level of awareness among researchers and research institutions is still low.

²⁸⁰ “The aforementioned two types of research data have to be deposited, preferably in a research data repository. Research data repositories are online archives for research data. They can be subject-based/thematic, institutional or centralized. As a priority, the research data has to be deposited at the authorized national data centres from the list of the Slovenian Research Agency.” Quote from National strategy of open access to scientific publications and research data. Available at:

http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Zakonodaja/Strategije/National_strate

²⁸¹ “The national open access infrastructure for scientific information in the form of publications and research data is composed of people, organisations, equipment (machines 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.” Quote from National strategy of open access to scientific publications and research data. Available at:

http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Zakonodaja/Strategije/National_strate

Data sharing culture

Proportion of social science researchers in Slovenia that have shared the research data they produced in the period between 2011 and 2016 was estimated as quite low (0-10%). ADP informed about the survey they conducted in 2016 at the Faculty of Social Sciences, University of Ljubljana, which is one of the biggest producers of social science data in Slovenia. According to the results of this survey, less than 10 % of research projects data (funded by public funds) were deposited to ADP or any other repository (Internal report). This was explained by the assumption that many data producers themselves see their data as ephemeral and not interesting for further secondary analysis.

Table 33: Attitudes towards data sharing in Slovenia

Data sharing has no benefits at all	Probably true
Data sharing creates healthy competition	False
Data sharing creates negative competition	Probably true
Reuse of existing data can answer new research questions and facilitate advancement of science	Neither true, nor false
Data sharing has as a risk that others may misuse and misinterpret data	Probably true
Data sharing involves little effort and minimal costs	Neither true, nor false

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from "5"-True to "1" False!

According to ADP's estimate in the period between 2011 and 2016 only 10-30% of science researchers in Slovenia were able to access existing third party data when they needed them. All publicly funded research projects of Slovenian Research Agency (ARRS) are listed in the SICRIS²⁸² database, but there is no recorded evidence concerning the usage of the third-party research data in those projects. Considering above mentioned shortcomings is not possible to estimate the proportion of SSH researchers who are able to access existing third party data.

Probably the most frequently used way to obtain data for secondary analysis in Slovenia is via informal contacts (peers and colleagues), while the data archive or repository and project or personal websites were listed as less frequent channels.

There are, however, certain career rewards related to data sharing. In Slovenia, the research data deposited in Social Science Data Archives (ADP) are peer reviewed and can be listed as a scientific publication.²⁸³ If certain level of quality (methodology, relevance, etc.) is recognized, then researcher can get scientific credits (points). Researchers who use research data deposited in ADP

²⁸² Database available at: <http://www.sicris.si/>

²⁸³ The document is available in Slovenian language only: Pravilnik o postopkih (so)financiranja, ocenjevanja in spremljanju izvajanju raziskovalne dejavnosti (neuradno prečiščeno besedilo št. 4) - v veljavi do 29.07.2016. See Attachment: Priloga 1: Bibliografska merila znanstvene in strokovne uspešnosti. (Section 2. H). Accessed by: <http://www.arrs.gov.si/sl/akti/prav-sof-ocen-sprem-razisk-dej-jan15.asp> (28.2.2017)

are obliged to cite this data and to report to ADP about their use. Despite these obligations the researchers often do not report and on the other hand scientific journals often do not check if the data used is properly cited in the articles (in the text and in the reference list); and data citation index is not developed yet.

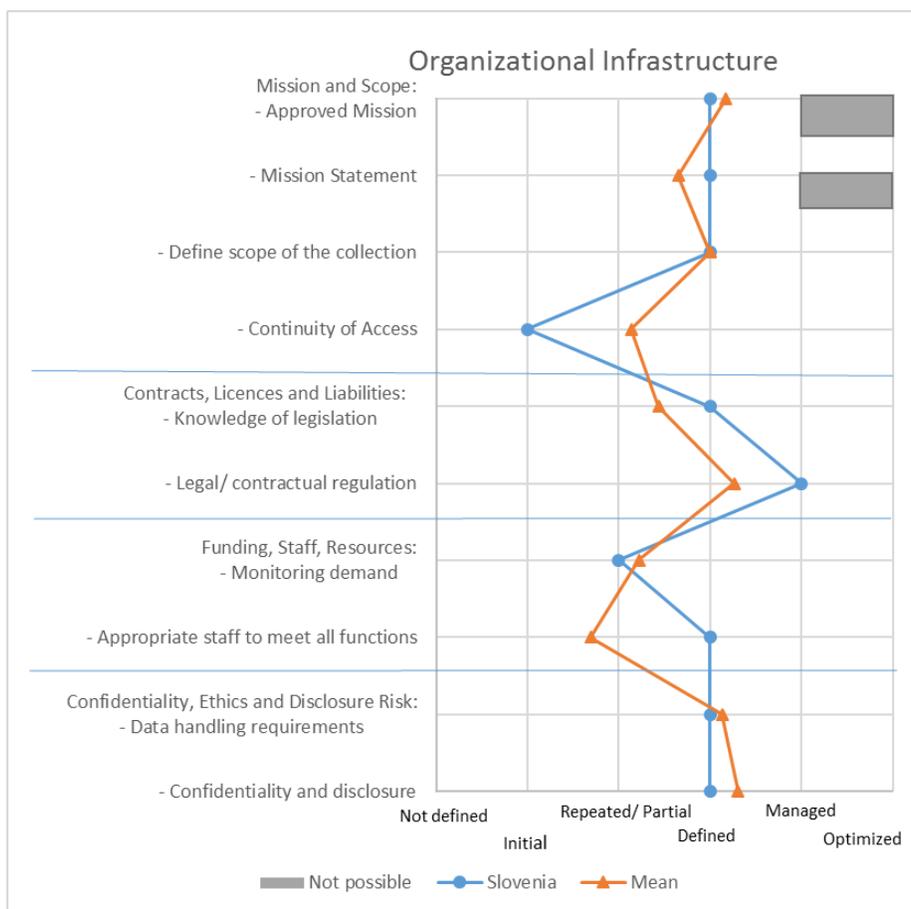
ADP offers certain support services to social science researchers in Slovenia. These services facilitating data sharing and Open Access to research data includes training, workshops, webinars, online reference materials, support to data management planning, support to long-term preservation of data, access to data, and data citation standards.

There are some data producers who follow data management and data documentation standards and procedures in Slovenia. With H2020 and new Action plan requirements the improvement of the depositing practices are expected.

3.38.2 Capability requirement areas of DAS

Organisational profile and infrastructure

Figure 86: Organizational infrastructure in Slovenia



The ADP were established in 1997 as an organizational unit within the Institute of Social Sciences at the Faculty of Social Sciences, University of Ljubljana. ADP (<http://www.adp.fdv.uni-lj.si/eng>). Due to the small size of the organisation, responsibilities are shared among several people in the organisation. Some of the work is also outsourced. The basic structure and roles of ADP is defined;

the staff is appropriate to support all the functions and services. ADP has quite simple structure since it is a small unit without division into departments.

Data archive services are funded only from the public money. The most important source is National Research Agency (Infrastructure centre); the funding is based on 6 years contract in scope of infrastructure program ["The Research Infrastructural Centres Network of University of Ljubljana."](#) Additional funding is available via international or national projects (approximately 25% of the total budget of ADP).

Core services and activities offered by the ADP are as following:

- Archiving of research data in accordance with OAIS process (promote / actively gathers, archive and distribute data);
- Actively promoting RDM (mainly through workshops);
- Promoting Open access to data (members of national committees);
- Assisting other archives in the process of development and establishment.

ADP provides written, formal and complete mission statement that is available for all relevant users. According to this mission statement ADP is a Slovenian national social science research data infrastructure for social sciences, whose mission is to offer data services in supporting research, education and general public benefit. It is involved in international infrastructures and cooperates in developing sustainable policies promoting the open-access. The goal of digital preservation policy in ADP is sustainably designed quality assurance and promotion of ingest, storage and access services to target users of valuable research data from the field of social sciences in Slovenia and abroad.

There is some evidence to support the notion that ADP and its activities receive approval from stakeholders. Collected data suggest that the organisation is on an ad hoc basis "encouraged" to continue its activities, but no formal mandate or formalised approval exist at the moment.

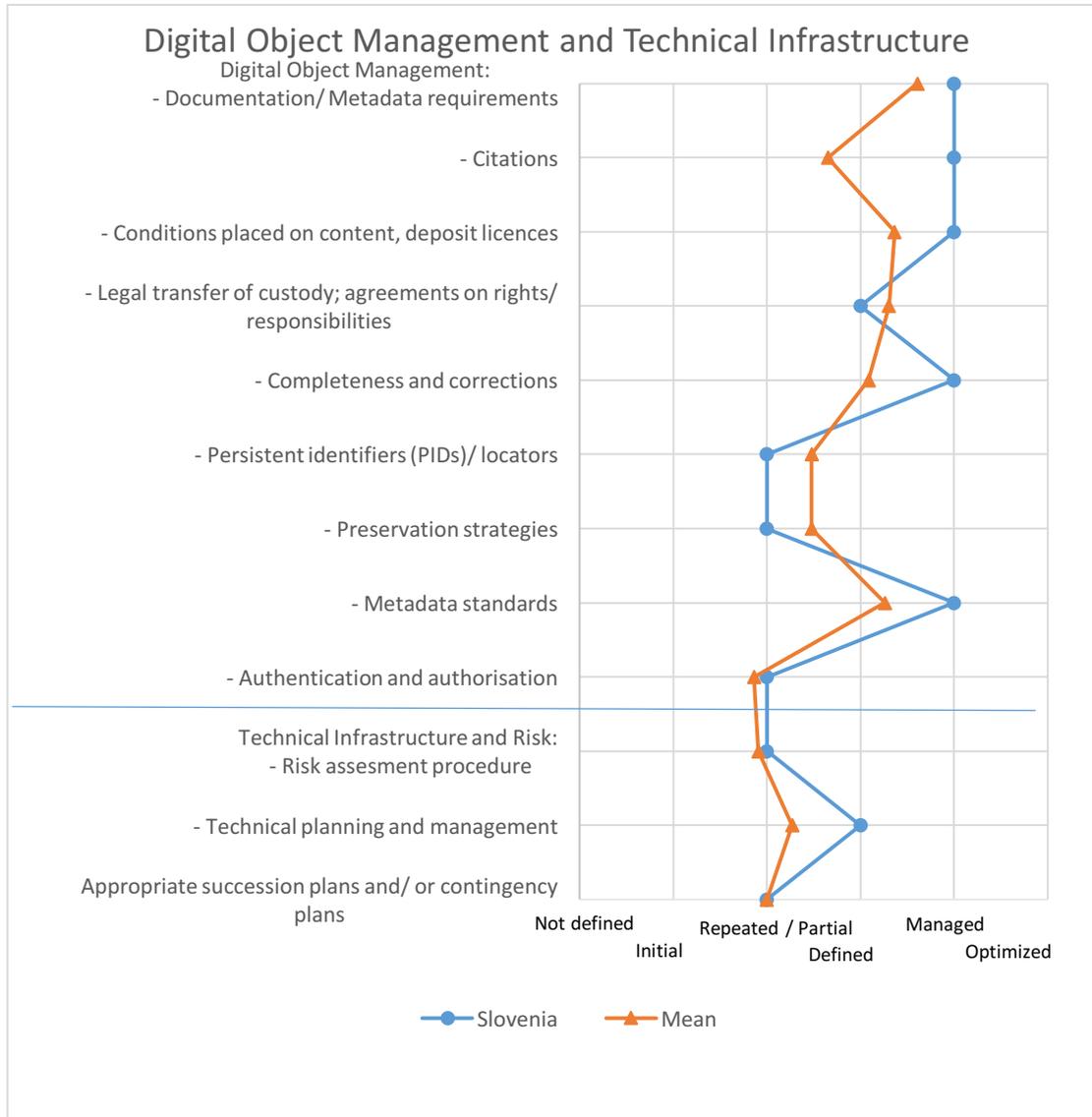
The sustainability of services of the national social science research datacentre of social sciences is an obligation, made by the Republic of Slovenia, in the framework of membership in the international infrastructure unit CESSDA ERIC. Long-term role of ADP on the national level is shown by the consistent support from the ministry responsible for scientific-research work and supporting the establishment of ADP. From 2004 onwards, funding is enabled through the infrastructure programme "Network of Research and Infrastructural Centres", which is currently financed for the period from 2015 to 2020.

Digital object management (data curation) and Technical infrastructure and risk

ADP has sufficient knowledge and documentation on how relevant legislation affects the holdings and procedures of the organisation in area of research data. In its operation ADP closely follows national legislation from the relevant areas, especially: Personal Data Protection Act, which defines the responsibilities of ADP as a micro-data distributor, and the Law on Copyright and Related Rights. In collaboration with the Statistical Office of the Republic of Slovenia of providing access to official anonymised microdata, ADP also follows National Statistics Act. Service access procedures are based on legal or contractual regulations. Legal and/or contractual framework is

regularly reviewed and updated; all regulations are aligned to higher level policies; roles and responsibilities are identified and maintained²⁸⁴.

Figure 87: Digital object management and Technical infrastructure in Slovenia



In order to register for the data access, a user needs to fill in a registration form available on-line²⁸⁵. In this registration form the user states his/her personal information, status (student, researcher, public official, other etc.) and should specify the purpose for which the data are to be used (educational, scientific, public, or commercial). In order to proceed with the registration, it is necessary to comply with the terms of the data usage, which amongst others, demand the respect for professional codes of ethics and oblige to cite the author/s of the data and the archive. ADP

²⁸⁴ More info on License Agreement between Archive (ADP) and Data Producer is available at: http://www.adp.fdv.uni-lj.si/eng/za_dajalce/izjava_o_izrocitvi/

²⁸⁵ See: Registration to access data - <http://www.adp.fdv.uni-lj.si/registracija>

complies with applicable research discipline norms and relevant legal data handling requirements through systematic written, formal procedures and policies.

The data depositors at ADP should meet the requirements of confidentiality and non-disclosure for collected individual data. ADP checks for fulfilment of these requirements; adequate processes and procedures for ensuring mentioned principles are in place. In the pre-ingest phase all possible questions with the data producer are resolved, including ethical questions, confidentiality and data anonymization issues, authorship of data, and the right of the archive (or its successor) to take over the data and to assure the process of digital preservation. All the rules on data access are also defined at that point.

ADP occasionally reviews changes in demand for the repository services but only reacts to significant changes. There are surveys among users and feedback from workshops is collected regularly. The requirements for data deposit are specified clearly (information on which kind of documentation, metadata, and provenance is needed). Mentioned requirements are aligned with the archive 's policies and other processes, as well as relevant documents. There are regular reviews and assessments (of success) of the information requirements²⁸⁶.

ADP offers and provides regularly reviewed and updated functions and mechanisms for proper citation of the data. Citation is also added to each study description²⁸⁷. Instructions are also provided at trainings. Automatic export to EndNote XML, RIS and Zotero is planned for the near future.

ADP has regularly reviewed and updated the mechanisms and functions that allow the depositor to set the access conditions for the deposited data. Depositors are instructed at special trainings. Some information is also available on the website of the archive²⁸⁸. Depositor is free to choose the licence in Licence agreement form²⁸⁹.

The agreements defining the roles of depositor and archive and the legal responsibility of both sides are in place. The contractual and legal regulations at ADP also make sure that the deposited materials do not infringe on any intellectual property rights or personal rights. The contractual templates are being used consistently. Legal and contractual framework is regularly reviewed and updated. All legal and contractual regulations are aligned to higher level policies.

ADP has set necessary processes for checking the completeness and the correctness of deposited materials. All relevant modifications are measured and registered; processes and procedures are reviewed and updated regularly²⁹⁰.

²⁸⁶ More on data archiving/workflow/standards can be found at http://www.adp.fdv.uni-lj.si/za_uporabnike/o_arhiviranje/ (SI only) or: Irena Vipavc Brvar (2016): Hands-on: curating the RRPP quantitative data. SEEDS Workshop II, Ljubljana, available at: http://seedsproject.ch/wp-content/uploads/2016/02/SEEDS_WS2_VIPAVC_DIP_Data.pdf; Irena Vipavc Brvar (2016): AIPs and data management: best practice in curation. SEEDS Workshop II, Ljubljana, available at: http://seedsproject.ch/wp-content/uploads/2016/02/SEEDS_WS2_VIPAVC_DIP_Metadata.pdf; or: Irena Vipavc Brvar (2015): Ingest - Acquisition and deposit. SEEDS Workshop I. Belgrade, available at: http://seedsproject.ch/wp-content/uploads/2016/02/SEEDS_WS1_VIPAVC_Acquisition0.pptx

²⁸⁷ An example of the instruction for data citation available at: <http://www.adp.fdv.uni-lj.si/opisi/sjm13/> (choose EN language and find the "How to cite a study" part).

²⁸⁸ See: http://www.adp.fdv.uni-lj.si/eng/za_dajalce/izjava_o_izrocitvi

²⁸⁹ More information on licence agreement available at: http://www.adp.fdv.uni-lj.si/media/img/datoteke/license_agreement_ADV_V2_R0.pdf

²⁹⁰ For more elaborated information on workflow in ADP see its website http://www.adp.fdv.uni-lj.si/za_uporabnike/o_arhiviranje/. See also the following publications: Štebe, Janez, Vipavc Brvar, Irena (2011). Analiza

Mechanisms and systems for identification and location of the data are partly in place in ADP. There is a certain directory structure or hierarchy to make the location of data easier, but it does not comply with formalised DOI systems. Certain mechanisms are being repeatedly used, but there is a lack of formalisation and written procedures²⁹¹.

The preservation strategy describing how the organisation would act upon identified risks, and how it would address topics like the degradation of storage media, the obsolescence of media drives, and safeguards against accidental or intentional digital corruption of preserved materials is only partly formalised and documented²⁹².

ADP provides data and metadata that are in accordance with standards used and understood by the providers and users. Format usage and enquiries are measured and assessed; and the format strategy is regularly reviewed and updated. More information on data types and related material formats can be found at the ADP's website²⁹³.

User authentication and control is limited mainly due to the non-existence of such support for Nesstar, which ADP uses for the distribution of the data. However, the Slovenian archive also manages its own database of users and they plan to use AAI in the relation to new system (FEDORA). Users need to register to access data²⁹⁴.

ADP has defined the appropriate technical infrastructure resources to support all the functions and services. Risk assessments are undertaken and significant changes are made to the technical infrastructure. The backup scheme is currently waiting for the update. The technical solution is now based on the parallel server. Most of the important data are already moved to the servers of Slovene academic network (ARNES) and the National university library system (NUK). ADP is aware of the importance of the contingency issues and they plan to prepare some systematic approach to it; at the moment, there are no written statements, procedures or processes. ADP has signed the contract with NUK concerning the data archiving; although NUK is not supposed to distribute the data in case ADP ceases to exist.

stanja in perspektiva digitalne hrambe v Arhivu družboslovnih podatkov (ADP). Knjižnica, letnik 55, številka 1, str. 57-85. URN:NBN:SI:DOC-I29E9VOL from <http://www.dlib.si> (Analysis of situation and perspective of digital preservation at ADP) and Irena Vipavc Brvar (2016): Hands-on: curating the RRPP quantitative data. SEEDS Workshop II, Ljubljana, available at: http://seedsproject.ch/wp-content/uploads/2016/02/SEEDS_WS2_VIPAVC_DIP_Data.pdf and Irena Vipavc Brvar (2016): AIPs and data management: best practice in curation. SEEDS Workshop II, Ljubljana, available at: http://seedsproject.ch/wp-content/uploads/2016/02/SEEDS_WS2_VIPAVC_DIP_Metadata.pdf

²⁹¹ More info on data identification in: Irena Vipavc Brvar (2016): AIPs and data management: best practice in curation. SEEDS Workshop II, Ljubljana. Available at: http://seedsproject.ch/wp-content/uploads/2016/02/SEEDS_WS2_VIPAVC_DIP_Metadata.pdf (p.8, system in development). There is an agreement with National University Library (NUK) to use URN, presented in: Irena Vipavc Brvar (2016): PID Non-user stories: ADP. CESSDA PID workshop. Cologne. Available at:

http://cessda.net/eng/content/download/760/6722/file/3-1_Pres_Non-UserStory_ADp.pdf

²⁹² More information (in Slovene): Janez, Vipavc Brvar, Irena (2011). Analiza stanja in perspektiva digitalne hrambe v Arhivu družboslovnih podatkov (ADP). Knjižnica, letnik 55, številka 1, str. 57-85. URN:NBN:SI:DOC-I29E9VOL. Available at: <http://www.dlib.si> (Analysis of situation and perspective of digital preservation at ADP)

²⁹³ A detailed description in Irena Vipavc Brvar (2016): AIPs and data management: best practice in curation. SEEDS Workshop II, Ljubljana. Available at: http://seedsproject.ch/wp-content/uploads/2016/02/SEEDS_WS2_VIPAVC_DIP_Metadata.pdf

²⁹⁴ See: <http://www.adp.fdv.uni-lj.si/registracija/>

Content of current collection

The linguistic composition of the archive's collections suggests that most of the datasets available at ADP are in Slovene language (with some exception of international or foreign studies they distribute as a primary distributor). Metadata are available in both, Slovene and English. Current data collection of Slovenian archive consists mostly of quantitative data (survey data), with some exception (few qualitative studies, interviews). Regarding the data format, most of data set are received in SPSS format. The documentation is available in Acrobat (pdf) format. Data sets are mostly from sociology, political science and other social sciences²⁹⁵. Most of the users of data archived at ADP are social science researchers, as well as master and doctoral students (mostly from the social science faculties in the country), and librarians (who ask ADP to help with RDM or with depositing materials to trusted repositories).

3.38.3 Conclusions

The production of social science data in Slovenia seems to be quite intensive. The country participates in most of the important international research series fielded in Europe. Research data production in SSH in Slovenia is under development; and, despite the limited sources, at the moment there is a number of research excellence examples.

Leading institution in social data archiving is Social Science Data Archives (Arhiv družboslovnih podatkov – ADP), which could be characterized as a rather small but „mature” repository with 20 years of experience and many good archiving practices established (in accordance with OAIS process). The role of ADP as an important infrastructure on the national level is proved by the consistent support from the ARRS in scope of a special infrastructure program (which is currently financed until 2020). The sustainability of services of the national research datacentre of social sciences is an obligation, made by the Republic of Slovenia, in the framework of membership in the international infrastructure unit CESSDA ERIC.

Despite the mentioned above intensive production of social science data in Slovenia, the practices of data sharing are still not at highest level at the moment. Apart from providing the infrastructure services in SSH, the Slovenian archive seems to play an important role in promoting the data sharing culture on the national level.

²⁹⁵ A detailed description of topics is available at: http://www.adp.fdv.uni-lj.si/opisi/vsebinska_podrocja/

3.39 Spain

As no information has been gained through the survey, all conclusions can only refer to the desktop research. It seems that there is not yet established a common infrastructure for research data in Spain, but their activities and projects aim to improve the situation.

Specifics about the data collection in a country

In the internal contact list the Centro de Investigaciones Sociológicas (CIS) was mentioned as a potential Service Provider but no further contact details or persons were listed. The director of the CIS Data Bank Department was contacted twice to support the survey or help to find persons who could do it instead. We received no answer to our mails. Thus, we do not have any information gathers by the survey.

There was no reaction after sending the draft report.

3.39.1 Broader ecosystem of DAS operation

Development of social science sector

Spain has participated in cross-national studies, like CCS, ESS, ISSP, CSES, EVS, WVS and PISA.

RDM Policy setting

The Spanish Foundation for Science and Technology (FECYT) and the Network of Spanish University Libraries (REBIUN) have established a platform for national scientific repositories.

RECOLECTA²⁹⁶ is the result of the collaboration since 2007 between the Spanish Foundation for Science and Technology (FECYT) and the Network of Spanish University Libraries (REBIUN) run by the Conference of Vice-Chancellors of Spanish Universities (CRUE). Their work is aimed at creating a nationwide infrastructure of Open Access scientific repositories.

As part of the RECOLECTA project a report on “The preservation and reuse of scientific data in Spain. Report of the good practices working group” was published by the end of 2012.

“In Spain, the recently approved “Science, Technology and Innovation Law”¹¹ added stimulus to the creation of infrastructures to support scientific information, with a section of the Law dedicated specifically to the deposit of scientific articles funded by the General State Budgets in institutional or thematic repositories.

This report stems from the Recolecta project and emphasises certain important considerations which need to be taken into account in the design and implementation of a research data-management policy, with particular emphasis on the situation in Spain compared to other countries. In this report, we define the variety of types of research data together with those involved in its management (institutional and thematic repositories, funding agencies, existing data centres, researchers, librarians and data-management experts, etc.). We also reflect on financial issues stemming from the creation of an interoperable, data-management infrastructure. Finally, this report aims to contribute to future

²⁹⁶ See: <https://www.recolecta.fecyt.es/?language=en>

initiatives which will be required for the management of research data under the scope of the new Science, Technology and Innovation Law.” (Report, p 7)

There is an International Registry on Research Data (ODiSEA) and a project called DATASEA and DATASEA Extended. It is a Spanish initiative to improve the know-how and best practices in research data storage and curation²⁹⁷.

3.39.2 Data archive service (DAS) proto-activities

DAS activities

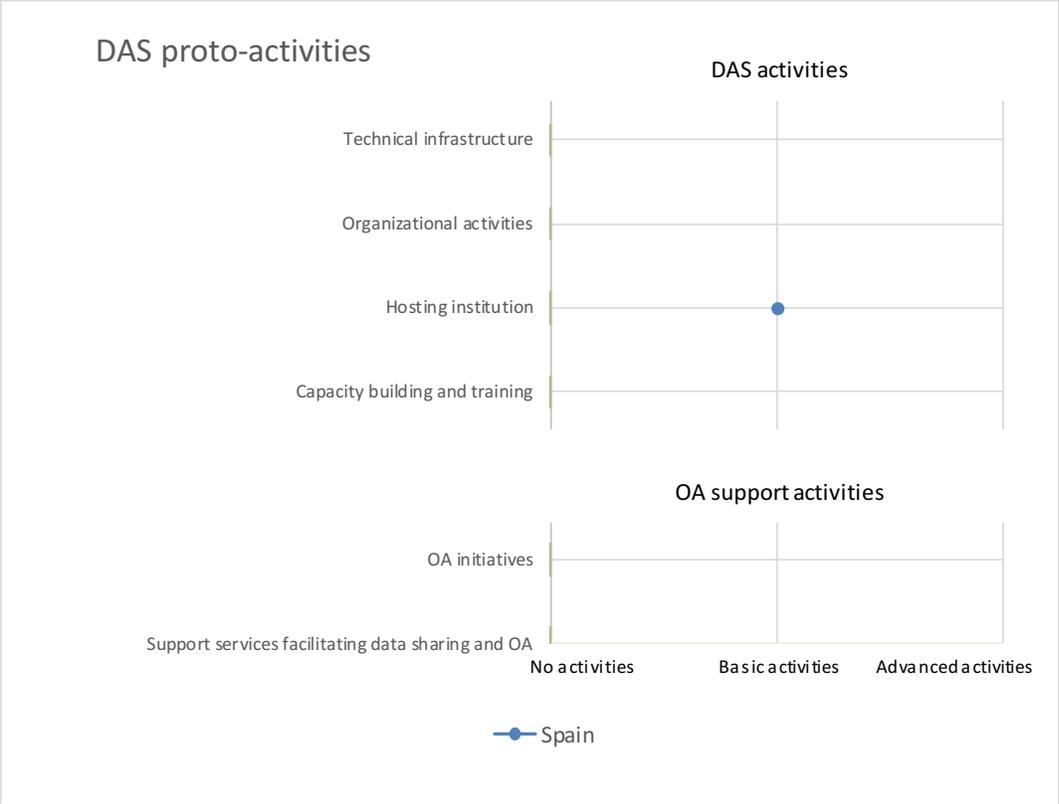
Currently, there is no DAS in Spain. CIS (Centro de Investigaciones Sociológicas), an independent administrative body, with its own legal status and funding, dependent on the Ministerio de la Presidencia could possibly host and be used for or applied to a new DAS.

CIS main tasks are to analyse and develop scientific knowledge on Spanish society by taking surveys and doing qualitative research studies. These studies are performed either at the initiative of the entity itself, or through agreements with public or not-for-profit private institutions. The majority of the CIS research activity focuses on carrying out public opinion surveys. These surveys include electoral studies, its monthly public opinion barometers, monographic studies on different aspects of Spanish society and the surveys resulting from CIS involvement in international projects. Through agreements, the CIS also fosters the organisation of courses and seminars (or participation in such activities) aimed at training, the analysis or dissemination of research results, within fields associated with sociology and political science.

The Centre has its own training schemes and research promotion programmes, together with a publications department, which plays an essential role in research dissemination in the social sciences in Spain. The CIS has a staff of approximately 100 people -civil servants and employees-, including technical, administrative and support personnel.

²⁹⁷ See: <http://www.datasea.es/dt/index.php/el-proyecto/descripcion>

Figure 88: DAS and OA activities implementation type in Spain



3.39.3 Conclusions

See summary above.

3.40 Sweden

All in all, Sweden has a medium-high maturity level. The country is in general above the average. Sweden scores medium-high in relation to the broader eco-system of DAS-operation and high in relation to the capability requirement areas of DAS.

There are a lot of strengths in relation to the broader eco-system, such as well-established traditions in data production and studies assessing matters of national importance. There is also potential for developing the RDM policy setting and the data sharing culture.

The conclusion is that Sweden is doing well and will continue to develop in the future.

Specifics about the data collection in a country

The data collection started 21st September 2016 and was finished 30th September 2016. The information was provided by Swedish national data service (SND). SND is a national unit at the University of Gothenburg.

The data collection process went well without any problems.

Broader ecosystem of DAS operation

Figure 89: Heading concepts values in Sweden

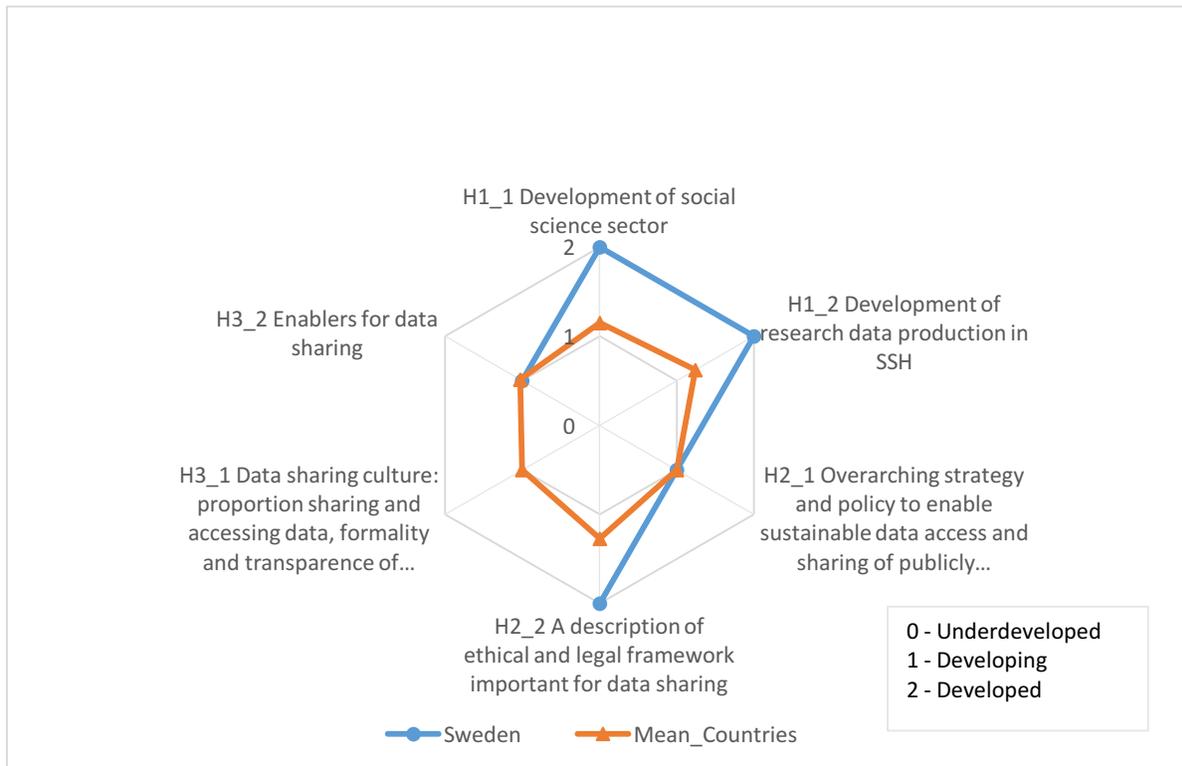


Figure 82 shows that Sweden is on the developed level at the areas 'Development of social science sector', 'Development of research data production in SSH', and 'A description of ethical and legal framework important for data sharing'. Here the country is above the average. Sweden is on the developing level in the areas 'Overarching strategy and policy to enable sustainable data access

and sharing of publicly financed...' and 'Enablers for data sharing' and therefore like the average. There is missing an estimate for 'Data sharing culture'.

Development of social science sector

Overall assessment of SSH development

The financial stability, research capacities and results achieved in the field of social sciences are in the highest quantile (on the developed level) and above the average.

In relation to research funding in social science, government and higher education sector is ranked first. The research is also funded by private non-profit sector (ranked second), abroad (international or cross-border) (ranked third) and business enterprise sector (ranked fourth).

- An overview of the Swedish research system in international comparison²⁹⁸
- Funding of research in Swedish higher education 1995-2006²⁹⁹
- (Current cost) for R&D in the higher education sector by source of funds 2013³⁰⁰

There are provided access to bibliographic and full-text databases, datasets, and software licenses.

Development of research data production in SSH

The research data production in SSH is on the developed level and above the average.

Following studies do systematically assess matters of national importance produced by social science researchers in Sweden:

- The Swedish National Election Studies³⁰¹
- The SOM Surveys³⁰²

The average production of research data by the social science institutions is frequent; the institutions have well established tradition in data production.

RDM Policy setting

Funders' data management and sharing strategy and/or policy

The overarching strategy and policy to enable sustainable data access and sharing of publicly financed social science research data is – like the average – on the developing level.

²⁹⁸ See: <https://publikationer.vr.se/produkt/forskningsbarometern-2016-en-overblick-av-det-svenska-forskningssystemet-i-internationell-jamforelse/>

²⁹⁹ See: <http://www.vr.se/download/18.67a0e785117e88b91b480001151/Rapport1.2008.pdf%3Efinansiering%20avForskninginom%20densvenskah%C3%B6gskolan1995%E2%80%932006%3C/a%3E%20%20%20%20%3Ca%20href=Revenu>

³⁰⁰ See: <http://www.scb.se/en /Finding-statistics/Statistics-by-subject-area/Education-and-research/Research/Research-and-development-in-the-higher-education-sector/Aktuell-Pong/8803/136742/>

³⁰¹ See: <http://valforskning.pol.gu.se/english/?!languageId=100001&disableRedirect=true&returnUrl=http%3A%2F%2Fvalforskning.pol.gu.se%2F>

³⁰² See: http://som.gu.se/som_institute/-surveys/?!languageId=100001&disableRedirect=true&returnUrl=http%3A%2F%2Fsom.gu.se%2Fundersokningar%2F

The requirements or recommendations for DMPs in most cases are on initial level. There is a growing recognition and awareness of need to require DMP.

The way of depositing data in an appropriate disciplinary repository is on the partial level, while the long-term curation for valuable research data assets is initial. Finally, the cost for managing the data and preparing it for access are at an initial level.

Legal and ethical framework

The ethical and legal framework is on the developed level and above the average.

There are recommendations and guidance provided on how to respect the legal requirements while sharing data. More concrete, where are explicit statements about data sharing and support is given in form of CODEX, rules and guidelines for research (<http://codex.vr.se/en/index.shtml>) and ethical vetting (<http://www.epn.se/en/start/>).

Data sharing culture

There is missing an estimate for 'Data sharing culture'. There are no estimates for the number of scholars in the Sweden that share their data. The estimate is probably low. In the same kind of way it is impossible to provide estimates for proportion of researchers able to access existing third party data they need. It is not possible to rank routines for data sharing and so on.

Enablers for data sharing

Enablers for data sharing are – like the average – on the developing level.

There is web guidance, workshops, online reference materials, contact and info point, metadata creation and publishing tools, support to data management planning, support to long-term preservation of data, and access to data in the Swedish national data service. According to self-assessment, there are no career rewards related to data sharing.

Finally, some data producers follow data management and data documentation standards and procedures.

3.40.1 Capability requirement areas of DAS

Organisational profile

Swedish national data service (SND) is a national unit at the University of Gothenburg.

SND is one single unit, publicly funded by the Swedish Research Council and University of Gothenburg.

The primary user communities are social sciences, humanities and health & medicine.

The core services and activities offered by the data archive service are: A) Access to existing data within and outside of Sweden, B) Support and guidance to researchers, C) Legal advice concerning research data, and D) Persistent identification and citing of research data.

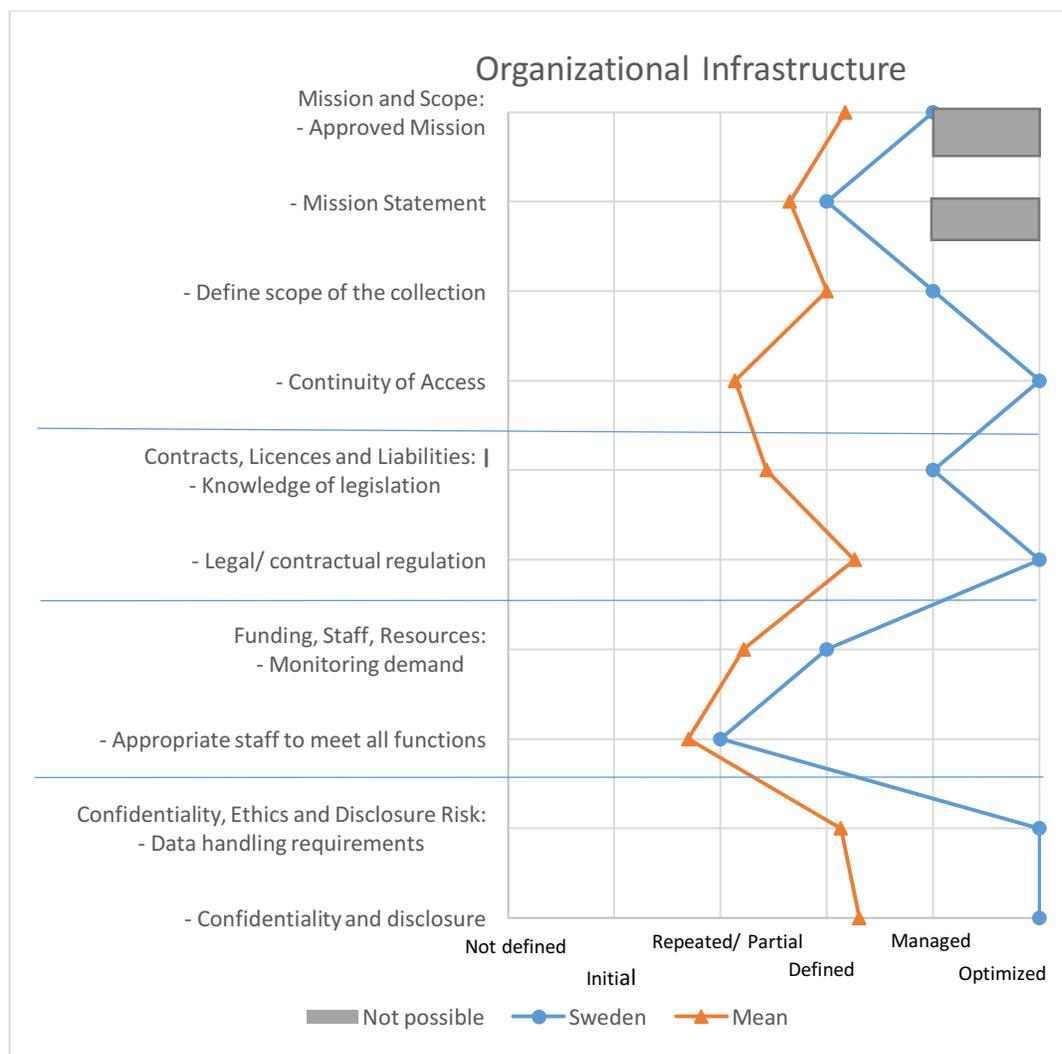
The overall contents of the current collection are A) Social science numeric survey data, B) Archaeological geospatial data, C) Qualitative data, mainly in humanities, and D) Numerical health data both survey and register data.

The linguistic composition of the archive's collection is in Swedish, while metadata is translated to English.

Organisational infrastructure

Figure 3.40.2 shows that the organizational infrastructure in general is on the managed or an optimized level. In all aspects, Sweden is placed above the average. It can be mentioned that Sweden has a succession plan which states that there should be a solution how to make the data available if the organization ceases to exist.

Figure 90: Organizational infrastructure in Sweden

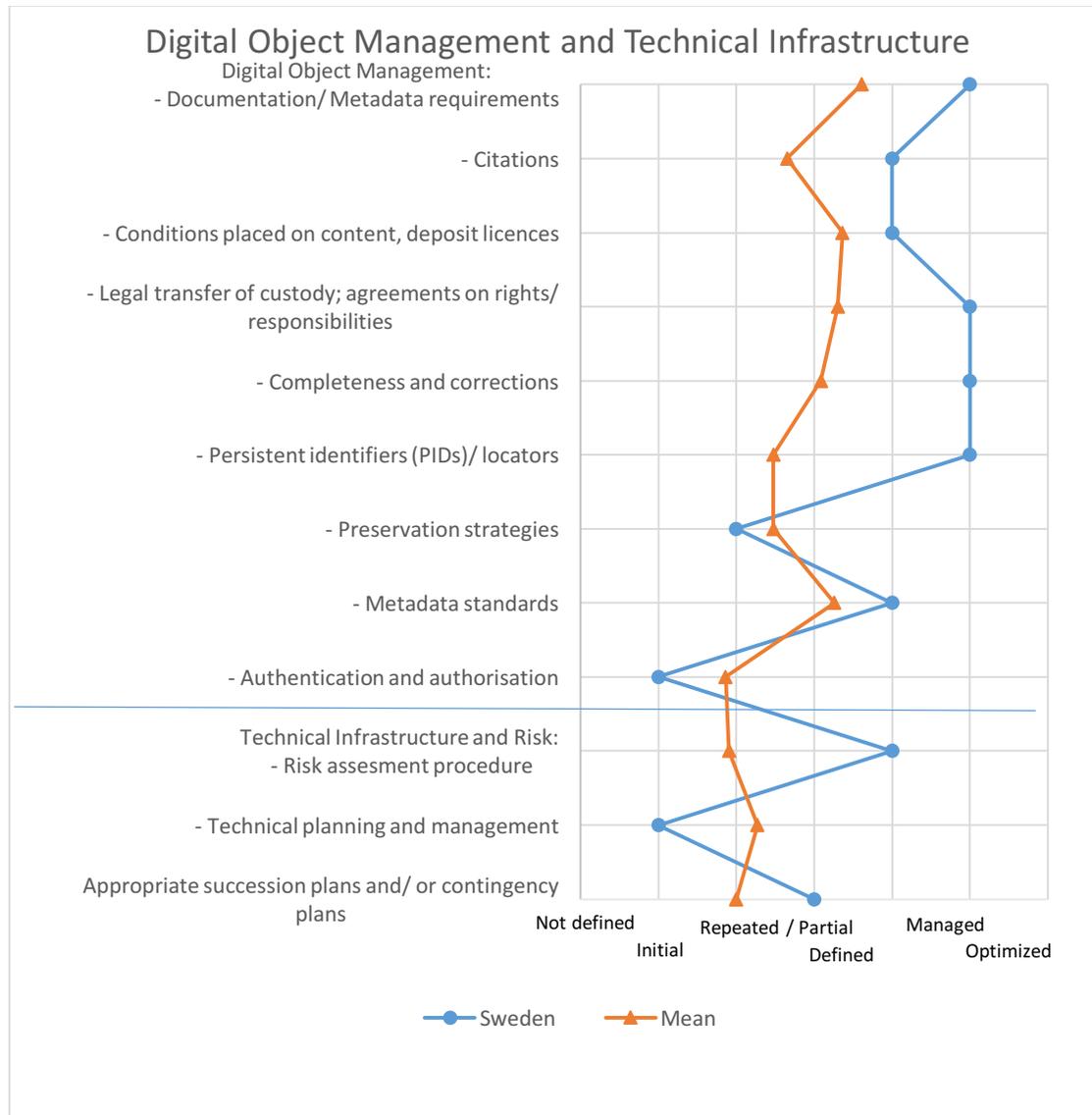


Mission statement

Statement from the Swedish Research Council: “Makes data available for research in the humanities, social sciences and medicine. Also, has the responsibility to provide documentation and to develop standards for documentation and storage of research data”³⁰³.

Digital object management (data curation) and Technical infrastructure and risk

Figure 91: Digital object management and Technical infrastructure in Sweden



³⁰³ Swedish Research Council, available at (only in Swedish): <http://vr.se/forskningsinfrastruktur/infrastrukturforforskning.4.12276aba1326e7bd62a800017663.html>

Figure 91 shows that digital management (data curation) in general is on the managed or optimized level and above the average. The technical infrastructure and risk is divided among three levels; both the initial, the repeated/partial, and the managed level.

Documentation/Metadata requirements

SND provides information and guidance online about the process of depositing data at SND. Data depositors provide metadata through a Deposit Form and by attaching related material. Submitted data without appropriate contextual information is blocked at the ingest stage and the depositor is asked to supply the necessary information³⁰⁴.

Citations

SND ensures that all its data are usable and are identified uniquely. SND provides data with Digital Object Identifiers (DOIs) as well as downloadable data citations for use in publications based on analyses of the data³⁰⁵.

Conditions placed on content, deposit licences

Depositors set access conditions in the deposit agreement³⁰⁶.

Legal transfer of custody, agreements on rights/responsibilities

Conditions for handling and making data accessible are stated in the deposit agreement³⁰⁷.

Completeness and corrections

Each study is processed by a data manager. All changes made by SND to a SIP to create an AIP to allow digital preservation are documented by SND data managers. Complete information about each data and metadata file as well as change logs and additional administrative metadata are stored in SND's internal management system (SIMS), and certain information is also published with the metadata. The workflow at SND follows the OAIS Reference Model and the DDI research data life cycle model. The internal document "Arbetsprocesser vid SND" (Work processes at SND) contains an overview of the work processes including acquisition, data processing, long time preservation and dissemination. The intranet based handbook "SND Handboken" contains detailed descriptions of the work flows. Upon signing the deposit agreement, the data producer is supplied with information about how SND manages the deposited research data.

³⁰⁴ Form available at: <https://snd.gu.se/en/describe-and-deposit-data/form>; Documentation available at: <https://snd.gu.se/en/deposit-data/documentation>

³⁰⁵ Tab datasets, available at: <https://snd.gu.se/en/catalogue/study/SND0870>

³⁰⁶ Agreement available at: <https://snd.gu.se/en/deposit-data/deposit-agreement>

³⁰⁷ See agreement above.

Persistent identifiers (PIDs)/locators

SND ensures that all its data are usable and are identified uniquely. SND provides all datasets with Digital Object Identifiers (DOIs).

Preservation strategies

SND has a draft to a preservation plan/policy from 2009, and it is currently under revision. SND uses migration strategies for providing long-term preservation of the SND data holdings. The data formats that SND uses for storing the data are chosen with long-term preservation in mind, avoiding proprietary, closed or rarely used file formats. Extensive metadata are collected and stored to ensure the usability of data. Developments and progress in technologies are followed closely by involvement in national and international organizations working with research data sharing and preservation.

Metadata standards

SND uses DDI Lifecycle as its main metadata standard.

Authentication and authorisation

Work on authentication and authorization has started.

Risk assessment procedure

Data at SND are stored and distributed according to the Regulation for IT-Security at University of Gothenburg³⁰⁸.

Technical planning and management

An overview of the technological infrastructure is underway.

Appropriate succession plans and/or contingency plans are on the defined level. Repository have implemented functions and mechanisms to be adequately prepared for major institutional changes; processes and procedures are formalised and defined.

Appropriate succession plans and/or contingency plans

SND has a succession plan (only in Swedish).

³⁰⁸ Regulation for IT security at the University of Gothenburg, available at:
http://medarbetarportalen.gu.se/digitalAssets/1531/1531415_regulations-for-it-security-revision2015_rev_pl.pdf

3.40.2 Conclusions

The self-assessment was conducted by the Swedish national data service (SND) and the data collection went very well. The results show that Sweden in general has a medium-high maturity level.

In relation to the broader eco-system of DAS operation, Sweden is on the developed level in the development of the social science sector. There are studies systematically assessing matters of national importance and well established traditions in data production. In relation to the RDM Policy setting, it is also assessed on the developing level. It is for example argued that there is a growing recognition and awareness of the need to require DMP. The data sharing culture isn't assessed on a level and it is noted that the number of scholars sharing data probably is low. It is not possible to give estimates for the proportion of researchers able to access existing third party data or to rank routines for data sharing. Enablers for data sharing are on the developing level with a lot of different guidance provided by SND.

In relation to the capability requirement areas of DAS, the organizational profile shows that SND offers core services and activities like access to data, support and guidance, legal advice, and persistent identification and citing. Both the organizational infrastructure and digital object management (data curation) is in general on the managed or optimized level and above the average. Finally, the technical infrastructure and risk is divided among three levels; the initial, the repeated/partial, and the managed level.

3.41 Switzerland

The analysis shows that the development of the social science sector is on the average comparable to other countries, yet research data production is comparatively more developed and with established research traditions. Among enablers and incentives for data sharing, results range from underdeveloped to developed, with underlying evidence for planning toward improvement in near future. Research data management policies are developed but ethical and legal frameworks are still developing. Data sharing and reuse is underdeveloped, while incentives and enablers for data sharing are developing.

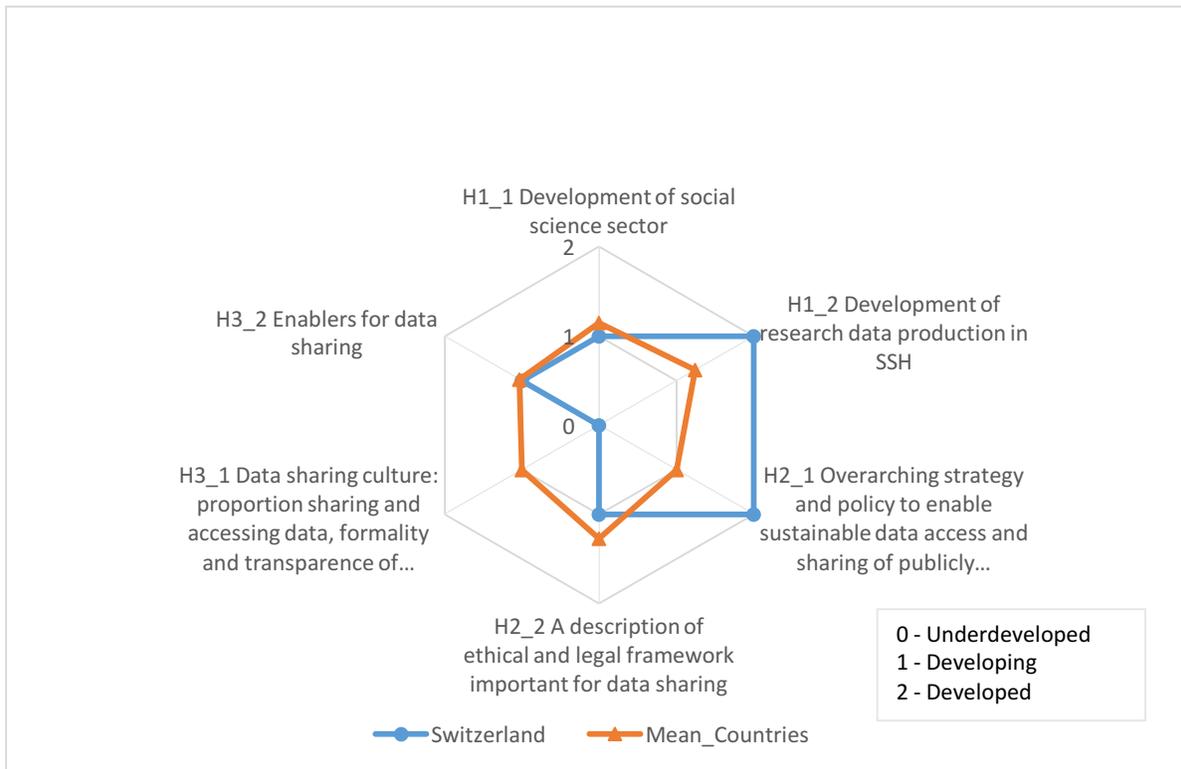
The data archive unit within FORS is well established, with a defined mission and scope. There is room for improvement in obtaining more firm external guarantees for long-term service, as the area of continuity of access is self-assessed as only initial. Legal regulations are defined while knowledge of legislation again is still an area for further improvement. The funding, staff and resources are partial, while confidentiality, ethics and disclosure risks are defined. Conditions placed on content, deposit licences, legal transfer of custody, agreements on rights, preservation strategies and technical planning and management are managed. Documentation requirements, citations, metadata standards, authentication and authorisation and risk assessment procedures are defined. Completeness and corrections, persistent identifiers and appropriate succession plans are only partial.

Specifics about the data collection in Switzerland

The information was entered primarily by Christina Bornatici, a researcher from FORS, and was reviewed and validated by Brian Kleiner, head of data services. The method used was self-evaluation.

3.41.1 Broader ecosystem of DAS operation

Figure 92: Heading concepts values in Switzerland



Development of social science sector

Overall Assessment of SSH development

The social science sector and humanities (SSH) sector in Switzerland is developing and its funding and the productivity of the researchers are in the mid quantile. The impact on the designated community is limited.

In 2012, the expenses allocated to Research and Development attained 2.97% of the Swiss GDP, according to the Federal Statistical Office³⁰⁹. It collects also information on the number of researchers per capita: in 2012, there were 7.24 full-time equivalent researchers in the universities, private sector and government per thousand economically active persons³¹⁰. Regarding the ranking of the sources of research funding, the State Secretariat for Education,

³⁰⁹ See http://www.oecd-ilibrary.org/science-and-technology/main-science-and-technology-indicators/volume-2016/issue-1_msti-v2016-1-en;jsessionid=2ne2m0htks26m.x-oecd-live-02 (in German). For English see the OECD (2016) publication: http://www.oecd-ilibrary.org/science-and-technology/main-science-and-technology-indicators/volume-2016/issue-1_msti-v2016-1-en;jsessionid=2ne2m0htks26m.x-oecd-live-02.

³¹⁰ See <https://www.bfs.admin.ch/bfs/fr/home/statistiques/developpement-durable/cockpit/formation-recherche-innovation/nombre-chercheuses-chercheurs.html>.

Research and Innovation (SERI), the most important source is the government and public funding, followed by foreign funding (international or cross-border), the higher education sector, the business enterprise/private sector and the private non-profit sector³¹¹.

The Switzerland government or universities provide access to commercial bibliographic and full-text databases, datasets and software licences³¹².

Development of research data production in SSH

In Switzerland research data production is developed. There are well-established streams of research traditions, national and international, with a great variety of important types of research data.

Switzerland is involved in international collaborative research and cross-national studies, such as the European Social Survey (ESS), International Social Survey Program (ISSP), Survey of Health, Ageing and Retirement in Europe (SHARE), European Values Survey (EVS), Comparative Candidate Survey (CCS), Comparative Study of Electoral Systems (CSES), and Program for International Student Assessment (PISA). However, it is not involved in the World Values Survey (WVS). In addition, there are several other social science studies that systematically assess matters of national importance, among them:

- The Swiss Electoral Studies (Selects) is a research project that study voting behavior in Switzerland in depth in the latest four national elections (since 1995). <http://forscenter.ch/en/our-surveys/selects/>
- Voto, Vox and VoxIt surveys have been conducted after each federal vote and offered insight into the votes of Swiss citizens (since 1977). <http://forscenter.ch/en/data-and-research-information-services/2221-2/special-projects/vox-voxit/#voxit>
- The Swiss Household Panel (SHP) observes since 1998 social change, in particular the dynamics of changing living conditions and representations in the population of Switzerland. <http://forscenter.ch/en/our-surveys/swiss-household-panel/>
- MOSAiCH is a biennial cross-sectional survey that focuses on the Swiss population's values and attitudes toward a wide range of social issues. <http://forscenter.ch/en/our-surveys/international-surveys/mosaich-issp-2/>
- Transitions from Education to Employment (TREE) surveys post-compulsory educational and labour market pathways of school leavers in Switzerland. http://www.tree.unibe.ch/index_eng.html
- The Swiss Survey on Children and Youth (COCON) is an interdisciplinary project, which examines the social conditions, life experiences and psychosocial development of children and adolescents in Switzerland from a life-course perspective. <http://www.cocon.uzh.ch/en/>

³¹¹ State Secretariat for Education, Research and Innovation (SERI): *The Swiss research and innovation system*, Figure A 2.3, p. 39. Available at: <https://www.sbf.admin.ch/sbfi/en/home/topics/research-and-innovation-in-switzerland/forschung-und-innovation-in-der-schweiz-2016/das-schweizer-forschungs--und-innovationssystem--teil-a.html>.

³¹² Only commercial software licences, not national.

- Swiss federal surveys of adolescents (ch-x) run over a two-year period and include some 50'000 young adults. Topics are different in each wave of the survey (in 2016-2017: mobility of young adults). <http://www.chx.ch/fr>
- Sicherheit (Security) is an annual study that evaluates long-term trends and tendencies in public opinion on foreign, security, and defence policy issues in Switzerland. <http://www.css.ethz.ch/en/publications/sicherheit.html>
- Freiwilligen-Monitor (Volunteer Work Bulletin) is a reliable source of information about volunteer work in Switzerland. <http://sgg-ssup.ch/en/volunteer-work-bulletin.html>
- Swiss Job Market Monitor (SJMM www.stellenmarktmonitor.uzh.ch) is devoted to a systematic monitoring and analysis of the Swiss labour market. For this purpose, SJMM has compiled a dataset of job openings in the Swiss economy dating back to 1950. The dataset is based on representative samples of job advertisements drawn from all advertising channels that are relevant at the given time.

Social science institutions in Switzerland have a well-established tradition in data production, and the production of high-quality research data is frequent.

RDM Policy setting

Research data management (RDM) policies are developed. Nevertheless, currently the Swiss National Science Foundation (SNSF) does not require RDM, and writing a Data Management Plan (DMP) is not required when applying for funding or subsidies. However, one of the SNSF measures of the Multi-Year Programme 2017–2020³¹³, consists of long-term curation, i.e. *"introducing new measures for improving research data management and help to ensure good scientific practice"* (p. 13). Moreover, the public research funding organizations do not offer any incentives for sharing research data with associated metadata, such as providing adequate resources for managing the data and preparing it for access.

Switzerland has defined formal requirements about offering and depositing data in an appropriate disciplinary repository; however, there is little or no monitoring. Specifically, article 47b of Funding Regulations of the SNSF says that *"the data collected with the aid of an SNSF grant must also be made available to other researchers for further research and integrated into recognised scientific data pools."*³¹⁴

There are partial requirements regarding long-term curation for valuable data assets. As the Swiss representative wrote, *"it is expected or recommended to assess the value of research data and resources providers declared their motivations for continuing to invest in sustaining the assets. The FAIR data principles (findable, accessible, interoperable and reusable) are highlighted in these recommendations."* Moreover, *"the Swiss National Science Foundation is currently developing its Open Data policy, and will likely put more emphasis on long-term curation in different ways."*

In Switzerland, there are only initial incentives for sharing research data with associated metadata. The Swiss representative wrote that *"the cost for managing the data and preparing it for access can be implicitly covered in the overall research project budget"*.

Finally, the clarification and support provided on legal and ethical aspects that facilitate social science data sharing are in the initial phase, i.e. there is growing awareness about the problem and about the need to provide clarification on legal aspects but only scarce or no organised support is

³¹³ See: http://www.snf.ch/SiteCollectionDocuments/mehrjahresprogramm_2017_2020_e.pdf

³¹⁴ See: http://www.snf.ch/SiteCollectionDocuments/allg_reglement_16_e.pdf (p.15)

given. The Swiss representative wrote that “the legal and ethical framework depends on disciplines and institutions”.

Data sharing culture

The data sharing culture in Switzerland is underdeveloped. Data sharing and reuse among social sciences researchers is not very common – as self-assessment results indicate, the proportion of researchers sharing data is estimated as low (0-10%)³¹⁵, and the proportion of researchers able to access existing third party data they need – as medium (10-30%)³¹⁶.

There are no available statistics regarding data sharing channels and routines in Switzerland.

As the self-assessment results shows, the attitudes tend to be mixed. Even though social science researchers in Switzerland in general seem to be worried about data misuse and misinterpretation, and consider data sharing costly and time consuming, they acknowledge that there are some benefits to data sharing. It should be noted that this is the perception of the data archive service of FORS on the quantitative researchers’ attitudes. The remaining three statements were not evaluated.

Table 33: Attitudes towards data sharing in Switzerland

Data sharing has no benefits at all	False
Data sharing creates healthy competition	Estimate not available for 2011-2016
Data sharing creates negative competition	Estimate not available for 2011-2016
Reuse of existing data can answer new research questions and facilitate advancement of science	Estimate not available for 2011-2016
Data sharing has as a risk that others may misuse and misinterpret data	True
Data sharing involves little effort and minimal costs	False

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from “5”-True to “1” False!

To summarize the data sharing culture in Switzerland, data sharing and reuse is not very common. Many researchers are not familiar with the available channels for data sharing, and some tend to be sceptical about sharing their own data, although most are in favour of data sharing generally

³¹⁵ Approximately, 2000 new project descriptions were entered in the FORS repository between 2011 and 2015. Over the same period, only 65 projects with datasets were submitted, archived and disseminated within FORS. This makes only 3.25%.

³¹⁶ This is difficult to estimate as there are no published statistics.

Enablers for data sharing

Enablers for data sharing in Switzerland are developing. There are no career rewards related to data sharing in the academic community, to quote the country representative -- “This is indeed the major issue. Data sharing is not rewarded.”

On the other hand, there are data support services available to social science researchers that facilitate data sharing and/or open access to research data. Specifically, FORS, the Swiss Centre of Expertise in the Social Science (<http://forscenter.ch/en/>), provides this service.

Some data producers follow data management and data documentation standards and procedures, i.e. data producers that deposit their research data in FORSbase (FORS data archiving and dissemination tool) follow DDI standards for their metadata, which is dictated by our system. The practices of data producers that deposit their data elsewhere are not known.

In summary, the incentives and enablers for data sharing within the social science research community in Switzerland can be benchmarked as developing.

3.41.2 Capability requirement areas of DAS

Organisational profile

Organisation

FORS (<http://forscenter.ch/>) is a national centre of expertise in the social sciences, which is housed within the University of Lausanne. Its primary activities consist of: 1) the preservation and dissemination of data for use in secondary analysis (the data service); 2) the production of survey data, including national and international surveys; 3) research in empirical social sciences, with focus on survey methodology; 4) consulting services for researchers in Switzerland and abroad.

The data service consists of 1) the data archive, 2) the research inventory, and 3) the data promotion group. The archive consists of 6 staff members and provides the usual archive services. The inventory is managed by one person who maintains the database of research project descriptions. The promotion group has 4 staff members and is responsible for outreach and publicity for data and targeted activities³¹⁷.

The primary user community are researchers, PhD candidates, and students in social sciences, in Switzerland and abroad, while the secondary user community are researchers affiliated with government departments, offices of national statistics, and public and commercial survey institutes and polling organisations.

Funding

The data archive within FORS is entirely financed by the State Secretariat for Education, Research, and Innovation (SERI). The surveys conducted by FORS are funded by the Swiss National Science Foundation (SNSF). FORS receives to a lesser extent funding and support from the University of Lausanne. Third party projects represent additional funding sources.

³¹⁷ Organisational Chart (FORS), available at : <http://forscenter.ch/en/about-us-2/foundation/organigramme/>

The University of Lausanne, the host institution for FORS, provides substantial financial and material contribution (administration, HR applications and salaries, infrastructure, some computer/network maintenance etc.).

Core services and activities

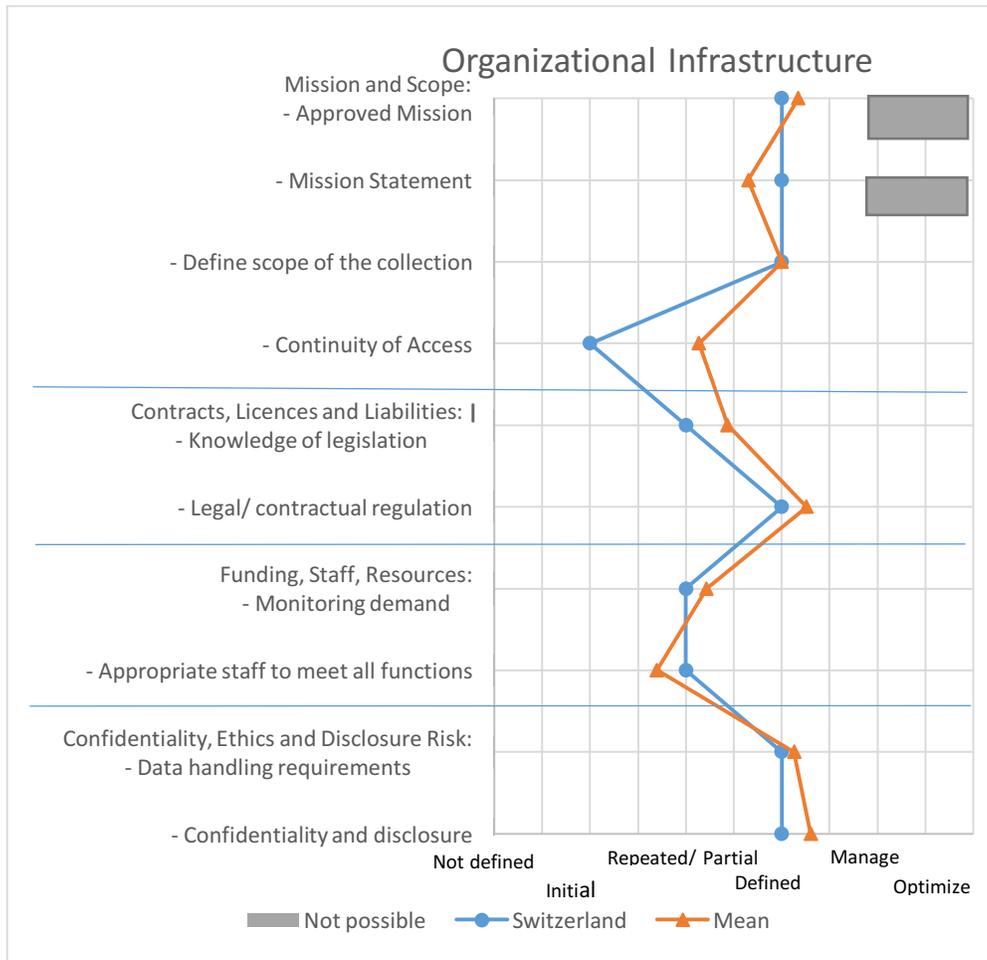
The core services and activities are research inventory, acquisition, preservation and dissemination of data for secondary analysis, consulting services for researchers in Switzerland and abroad, and data service activities (data management guidance and training). Other activities within FORS are production of survey data, including national and international surveys, research in empirical social sciences with focus on survey methodology, collaboration with researchers and research institutes in the social sciences in Switzerland and internationally.

Content current collection

Social science data are archived and disseminated since 1992 (previously within SIDOS, and since 2008 within FORS). There are over 500 datasets, including international and national surveys as well as others surveys in various disciplines of the social sciences (i.e., political science, education, sociology, health, demography). Most of the data are quantitative, with a small but growing number of qualitative datasets (e.g., text transcriptions). The data deposited and metadata are available in German, French and/or English, depending on the projects.

Organisational infrastructure

Figure 93: Organizational infrastructure in Switzerland



Mission and scope

The organisation has repeated/partial approval for its mission statement, i.e. there is some evidence to support the notion that the organisation and its activities receive approval from stakeholders; the organisation is on an ad hoc basis “encouraged” to continue its activities, but no formal mandate or formalised approval exist.

An explicit written, formal and complete statement on the role, mandate, purpose and mission of the organisation is defined and available for all relevant users at³¹⁸.

The scope of the collection is explicitly defined and promulgated. However, there are no contingency plans - only informal intent/agreements. Currently there is no succession plan in case FORS ceases to exist. Various options are being examined and a solution should be given in 2017.

³¹⁸ Mission statement, available at: http://forscenter.ch/wp-content/uploads/2015/05/Mission-statement_EN_v21.pdf

Contract, licences and liabilities

The organisation has only repeated/partial knowledge on relevant legislation. Specifically, FORS consulted a legal expert with respect to the national legislation (most notably data protection laws), and has been assured that its policies and practices are consistent with the national laws. However, the knowledge about this is not widespread across the organisation and there is limited documentation. Contracts and/or agreements are defined, i.e. they are standardised and implemented according to written procedures and made publicly available. The standard deposit³¹⁹ and user contracts³²⁰ are available on the FORS website and on FORSbase, the FORS data archiving and dissemination tool.

Funding, staff, resources

The organisation partially monitors demand for its repository services and occasionally reviews changes in demand for the repository services and only reacts to significant changes. For instance, they conduct surveys of users and researchers every two years.

The staffing is only partially complete, as the organisation has not defined the appropriate staffing level to support all repository functions and services.

Confidentiality, ethics and disclosure risk

The organisation complies with norms and legal requirements defined through systematic written, formal procedures and policies.

Processes and procedures are defined as in place, and standardised information is provided to the depositor prior to the deposit. Checks are performed on data after deposit. Process and procedure descriptions for handling and altering sensitive data are in place. Specifically, information is provided in documents on data protection³²¹, the deposit contract³²² and the policy on archiving qualitative data³²³.

Digital object management (data curation) and technical infrastructure and risk

Documentation/metadata requirements

A written formal specification of required information is explicitly defined, and requirements are compliant with metadata standards that are used and can be understood by the designated

³¹⁹ See https://forsbase.unil.ch/media/general_documentation/en/deposit_contract_FORS_en.pdf

³²⁰ [User contract, available at: https://forsbase.unil.ch/media/general_documentation/en/download_contract_en.pdf](https://forsbase.unil.ch/media/general_documentation/en/download_contract_en.pdf)

³²¹ Documents on data protection, available at: <http://forscenter.ch/en/data-and-research-information-services/2221-2/deposit-data/data-protection/>

³²² Deposit contract, available at: https://forsbase.unil.ch/media/general_documentation/en/deposit_contract_FORS_en.pdf

³²³ FORS policy and procedures on qualitative data, available at: http://forscenter.ch/wp-content/uploads/2015/05/FORS-policy-and-procedures-on-qualitative-data_20_04_2016.pdf

community (e.g. DDI). Metadata requirements are accessible and communicated to users/depositors.

On the FORS website, there are the following documents: guidelines for deposit³²⁴, quantitative data preparation³²⁵ and qualitative data preparation³²⁶.

Citations

Citation practices are required and offered to all depositors but also defined and formalised through templates or other written documents. Processes and procedures are documented. Specifically, a standardised citation is automatically created by FORSbase, the data deposit system, and offered to the depositors. The depositors can modify it. Finally, the approved citation is automatically written into the end user contract. The citation is publicly available in the data catalogue as metadata for all datasets. Conditions placed on content, deposit licences

The organisation manages regular reviews and updates of set of conditions aligned with high level policies. Specifically, “the depositors can set access restrictions upon the use of data and special permission. The depositors can keep control over access to its data by requiring that any data provision is subject to its authorization. If the depositors chose this option, they will receive a copy of each data request by e-mail, together with the details of the person who is requesting the data, as well as a description of the intended analyses. Data will be released only after receiving the depositor's approval.”

Legal transfer of custody, agreements on rights/responsibilities

There is monitoring of the usage of agreements and contracts that are reviewed and updated regularly. Actions are taken where contracts/agreements appear not to be working effectively or are not in accordance with a higher level of policies. Examples for the end-user contract³²⁷ and deposit contract³²⁸ are given.

³²⁴ Guidelines for deposit:

<http://forscenter.ch/en/data-and-research-information-services/2221-2/deposit-data/guidelines-for-deposit/>

³²⁵ Quantitative data preparation: <http://forscenter.ch/en/data-and-research-information-services/2221-2/deposit-data/guidelines-for-deposit/quantitative-data-preparation/>

http://forscenter.ch/wp-content/uploads/2013/11/Guide_to_depositing_quantitative_data_at_FORS1.pdf

³²⁶ Qualitative data preparation:

<http://forscenter.ch/en/data-and-research-information-services/2221-2/deposit-data/guidelines-for-deposit/qualitative-data-preparation-4/>

http://forscenter.ch/wp-content/uploads/2013/11/Guide-to-depositing-qualitative-data-at-DARIS_EN_26-04_20161.pdf

³²⁷ End-user contract, available at:

https://forsbase.unil.ch/media/general_documentation/en/download_contract_en.pdf

³²⁸ Deposit contract, available at:

https://forsbase.unil.ch/media/general_documentation/en/deposit_contract_FORS_en.pdf

Completeness and corrections

There are non-systematised (manual) checks of deposited material in place; processes and procedures are repeated but they are not formalised or documented. Rectifications are performed repeatedly, either by the repository or by returning data to the depositor.

Persistent identifiers (PIDs)/locators

There are mechanisms and systems for identification and location that are partly in place but do not comply with formalised DOI systems. The mechanisms are being repeatedly used, but there is lack of formalisation and written procedures. In 2017, DOIs will be implemented for all datasets.

Preservation strategies

The organisation periodically reviews and updates the strategy³²⁹.

Metadata standards

A data/metadata format strategy is explicitly defined and formalised and communicated to the users. Data and metadata are provided in formats that are commonly in use and understood by the users.

Authentication and authorisation

The repository uses digital identities, identity management, authentication and authorization to control access to data. The AAI is formalised, systematised and documented. FORS rely in large part on a manual authentication. Only people who have a domain name from a Swiss or an international university get automatic authentication. People with other domain names must apply for access.

Risk assessment procedure

There is a systematic analysis of security when there are changes to the technical infrastructure, but for the moment the information does not exist in one document but in various documents dealing with security and technical infrastructure.

Technical planning and management

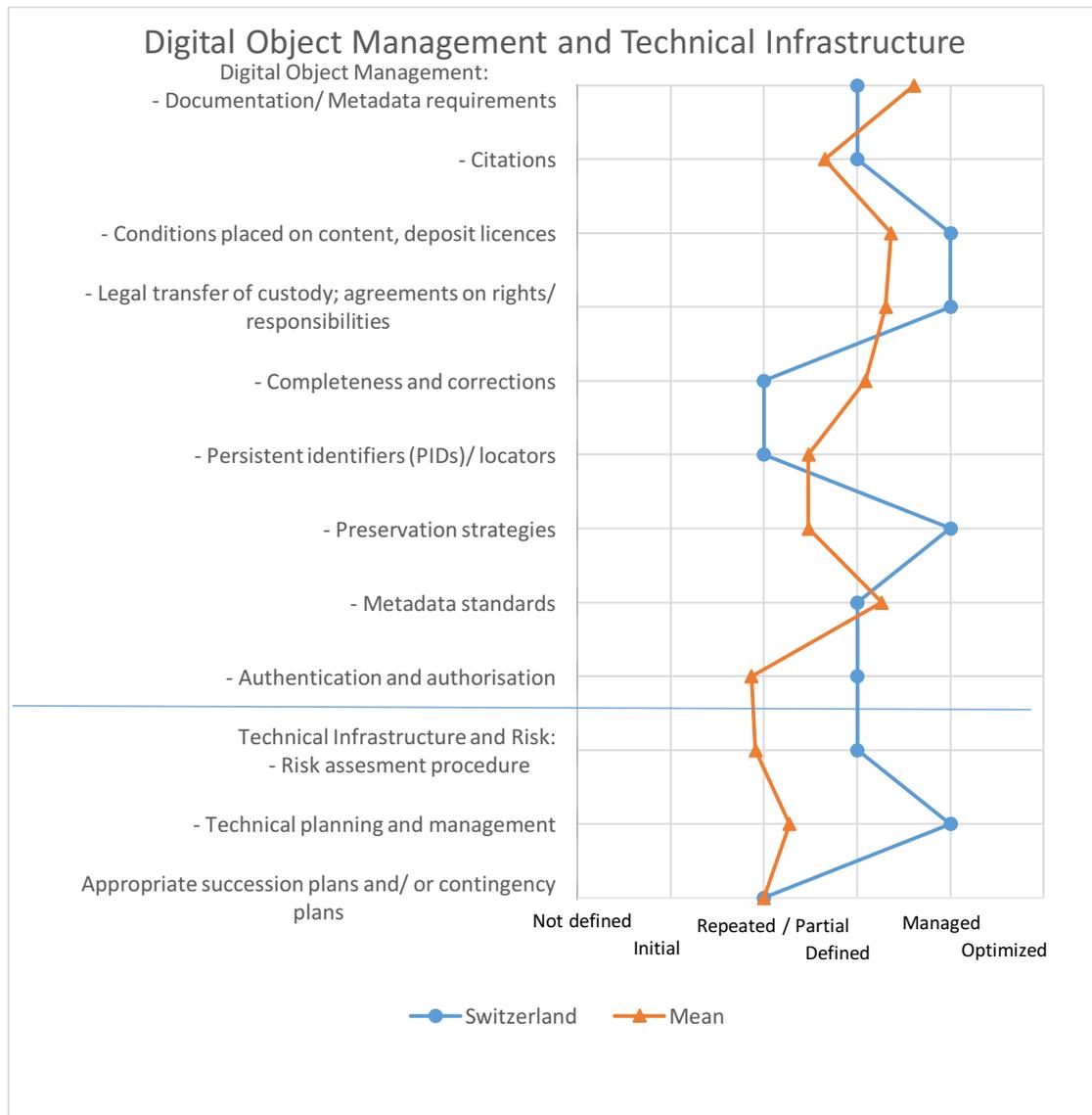
Technical infrastructure resources appropriate to support functions and service needs are monitored and analysed at regular intervals. The monitoring is documented.

³²⁹ Preservation policy document available at: <http://forscenter.ch/wp-content/uploads/2015/05/Preservation-Policy1.pdf>.

Appropriate succession plans and/or contingency plans

There is some awareness of the need to plan and prepare for contingency issues. However, there are no written statements, procedures or processes.

Figure 94: Digital object management and technical infrastructure in Switzerland



3.41.3 Conclusions

Regarding the broader ecosystem of data archive services operations, we can conclude that Switzerland is already at a high level of development. Although research data management policies are not in place within national funding institutions, formal requirements about offering and depositing data are defined. There is a growing awareness regarding ethical and legal frameworks important for data sharing which are developing, but no organised support is given. Data sharing and reuse are still not the norm, and researchers have mixed attitudes towards data sharing and re-use. Currently there are also no explicit incentives and enablers for data sharing and reuse.

However, there are some well-established data management support services available for researchers.

Regarding the capability requirement areas of the data archive service within FORS, the organisational infrastructure has a defined mission and scope, except for the continuity of access, which is in the initial phase. Partial are also the funding, staff and resources, both in terms of monitoring demand and appropriate staff. Also, completeness and corrections, persistent identifiers and appropriate succession plans are only partial. Except for those, the majority of continuous DAS service functions are on a high 'defined' level, while for remaining areas there are technical (new software) and organisational (external agreements, human resources investment) developments that promise to fulfil the current gaps in the near future.

3.42 Turkey

No report. Even after repeated contacts representatives of scientific community didn't respond to the invitation.

3.43 Ukraine

The social science sector of Ukraine is moderately developed. The main source of research funding is the public sector. However, social science research data is not well-established. No clear policy on Research Data Management exists in the country. A data sharing is not very common as well. The perceptions of the attitudes of researchers towards data sharing are mixed. Meanwhile, requests for secondary data research have been increasing over the last years.

Specifics about the data collection in a country

Data collection was carried out in late 2016. Unfortunately, there is no data from the self-assessment survey provided by the CESSDA SaW project.

A major source of information was *Peer Review of the Ukrainian Research and Innovation System*³³⁰. *Background Report*³³¹ written by Klaus Schuch, Gorazd Weiss, Philipp Brugner, Katharina Buesel (Centre for Social Innovation (ZSI), Vienna, Austria) in May of 2016. A complimentary data was provided by Andrii Gorbachyk, Dean of Faculty of Sociology, Taras Shevchenko National University of Kyiv (TSNU).

³³⁰ Klaus Schuch, Gorazd Weiss, Philipp Brugner, Katharina Buesel (2016) Peer Review of the Ukrainian Research and Innovation System. Centre for Social Innovation (ZSI), Vienna, Austria.

³³¹ Klaus Schuch, Gorazd Weiss, Philipp Brugner, Katharina Buesel (2016) Background Report. Peer Review of the Ukrainian Research and Innovation System. Centre for Social Innovation (ZSI), Vienna, Austria.

3.43.1 Broader ecosystem of DAS operation

Development of social science sector

Overall assessment of SSH development

While the economic and geopolitical framework of the country transformed enormously during the last 25 years, its social sciences were not enough responsive in adapting to this new situation. Funding of the SSH and productivity of the researchers are low. Research in the field of social sciences is mainly publicly funded. Still an evidence-based public policy making is in an embryonic phase of development in Ukraine.

According to *Peer Review of the Ukrainian Research and Innovation System* public funding of social research as well as general research and development funding mechanism needs to introduce independent competitive project funding, which almost absent now. Establishment of the National Research Foundation (NRF), as stipulated by the new Law on Scientific and Technical Activity, is the first step in that direction. In the area of research infrastructures, a picture is highly problematic as the research infrastructure facilities for Ukrainian researchers are overall outdated and financial resources to renew research facilities have been very low.

Today Ukrainian government sets the macro level priorities of spending funds from national budget for research. The Ministry of Education and Science (MES, <http://www.mon.gov.ua/>) is responsible for secondary education, special secondary education, higher education and science. Ministry distributes among universities state budget funding dedicated for research. The MES also organizes competition to allocate grants for research in the universities. Another important player is the National Academy of Sciences of Ukraine (NASU)³³². It has own budget, which is independent of the MES. Institutes of the NASU conduct research in many fields, including social sciences.

There are also private research centers, various NGO's with some research facilities and capacities. These organizations are conducting political, economic, sociological, marketing surveys. Their funding is from different sources (state budget, local government, international foundations, international organizations, private companies etc.). Still a share of these organizations in social research is relatively small.

Development of research data production in SSH

Research data production in SSH is developing in Ukraine. Institute Sociology of the NASU, which was founded in November 1990, is major empirical data producer in social sciences. A set of monitoring surveys on the social and economic situation, social and political attitudes of the Ukrainian population are conducted periodically. The Institute of Sociology of the NANU was conducting national survey on the social and economic attitudes every year since 1994.

Ukrainian research institutions also participated and produced data for major international social research projects as the European Social Survey, World Values and European Values studies, International Social Survey Program, Health behavior in school-aged children and others.

³³² See: <http://www.nas.gov.ua/EN/Pages/default.aspx>

RDM Policy setting

Since sharing of research data was not very common in the past, requirements or recommendations about preparing Data Management Plans are non-existent. However, there is growing recognition and awareness about the value of research data produced and about the need for long-term preservation. Public research funding organizations do not explicitly or formally demand data publishing or archiving. There are no requirements and recommendations regarding the data preparation and documentation.

Data sharing culture

Data sharing and reuse among social sciences researchers is not common enough in Ukraine. There are few data sharing channels and routines. Social researchers can use the data of various international projects freely if this data is published via international open access electronic archives (like a data of the European Social Survey, ESS).

In case of national studies based on empirical data social researchers need to ask local Ukrainian research teams for an individual permission to use the available data. Also, it is possible to acquire the necessary data from project and personal websites. In this case, there are no formal rules and transparency.

Also, sometimes there are bilateral institutional agreements on the data use. For example, undergraduate and postgraduate students of the TSNUK are able to use national data of the ESS, data of monitoring surveys from the Institute of Sociology of the NANU actively according special agreement between Faculty of Sociology of the TSNUK and Institute of Sociology of the NANU. In a similar way, a data of some surveys of the KIIS is also available to the TSNUK students. Data sharing and reuse of data (secondary analysis of raw data) is quite rare in Ukraine. There are no permanent public or private services for this. It has negative influence on the culture of data analysis and data use.

To summarize data sharing culture in Ukraine, the level of secondary use of research data is low. There are only few formal channels for data sharing and however researchers tend to use channels that lack formality. The absence of clear requirements for making data available after completion of research projects is one of the explanations for the problems in data sharing culture. On the other hand, the demand within the research community for reliable data and higher quality data increases.

Enablers for data sharing

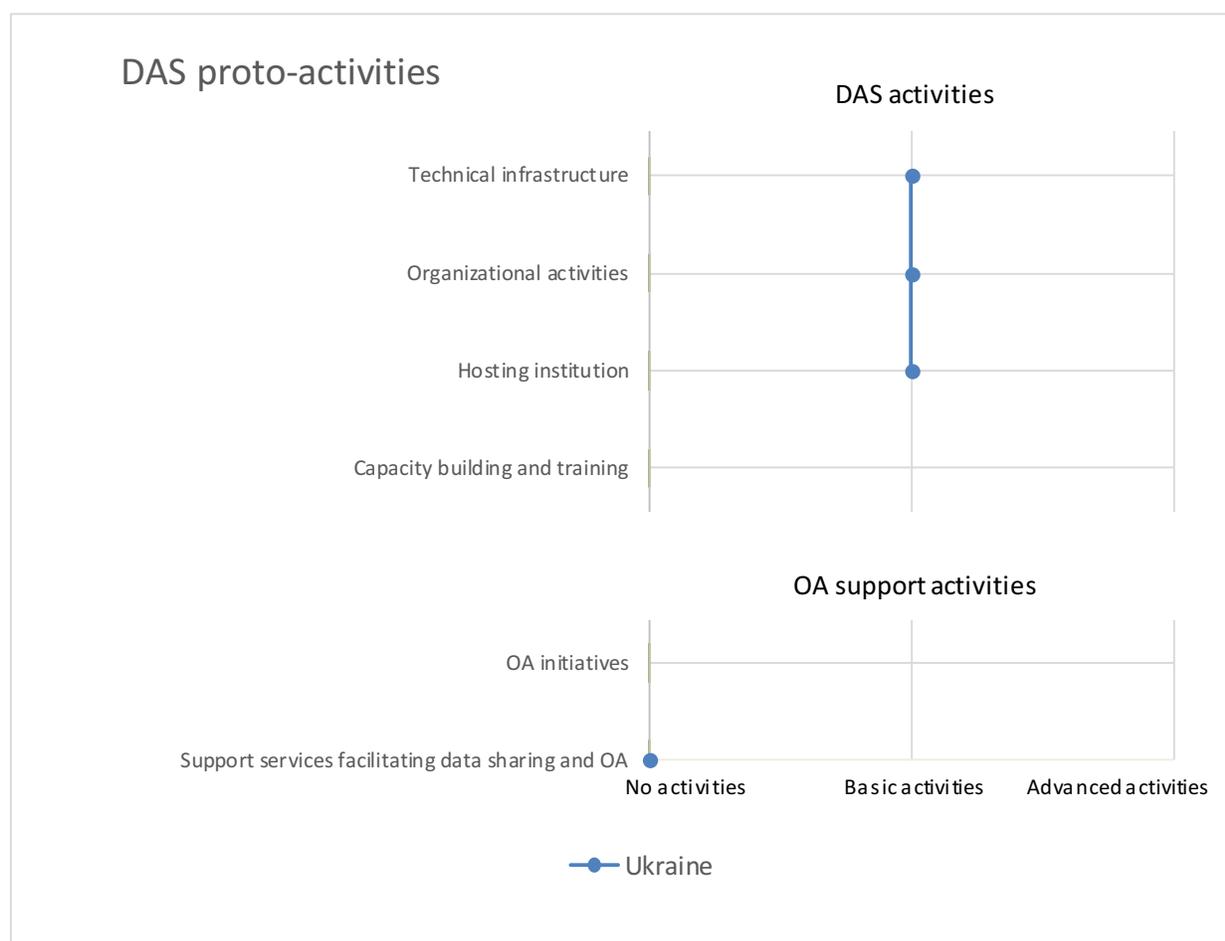
There are no clear career rewards or financial incentives related to data sharing in the academic community in Ukraine. Sharing research data does not influence career development within an institution or the research community. There are no data support services available to social researchers that facilitate data sharing. However, three Ukrainian research institutions, i.e. the Taras Shevchenko National University of Kyiv (TSNUK, <http://www.soc.univ.kiev.ua/en>), Kyiv International Institute of Sociology (KIIS, private research institute <http://www.kiis.com.ua/?lang=eng>) and Institute of Sociology of the NANU (<http://www.i-soc.com.ua/institute/>) made efforts to establish national data service based on a survey data. Still this data service is only underdevelopment and not functioning properly. As well as there are no data documentation standards and procedures that facilitate data reuse.

Data support services are highly traditional and main data institution here are academic libraries. Many of them have electronic catalogues available via internet. Still social science researchers do not have access even to major full-text databases as Web of Science/ Clarivate Analytics, Scopus, JSTOR and so on. There is only support for access to national databases via traditional libraries.

In summary, the incentives and enablers for data sharing within social science research community in Ukraine is on a low level.

3.43.2 Data archive service (DAS) proto-activities

Figure 95: DAS and OA activities implementation type in Ukraine



DAS activities

There is no public institution for the social research data management and archiving in Ukraine. As already was noted the Taras Shevchenko National University of Kyiv, Kyiv International Institute of Sociology and the Institute of Sociology of the NANU attempted to establish a data archive for

the empirical sociological research. A test version of this service with the data files from the KIIS³³³ will be available in near future. A national survey data on the social and economic attitudes from the Institute of Sociology of the NANU is available for researchers by an individual request as well. The future DAS plans to distribute its data in Ukrainian and Russian languages. It is expected that the data user communities of this service will include primarily researchers and students.

There is no special funding for this trilateral initiative to launch some empirical data services. Moreover, research grants from the MES or the NASU have no funding for data management, sharing and preservation. In general, research grants include funding only of salaries of researchers and very little is given for equipment, for experimental work or fieldwork.

3.43.3 Conclusions

The level of data sharing and secondary data usage is low in Ukraine. However, its significance and demand for reliable data is increasing. Recently three major Ukrainian social research institutions are making first steps towards survey data archiving and the long-term data preservation. Further financial incentives for data management and documentation as well as broader education about the benefits of data sharing are highly needed in Ukraine.

The Peer Review of the Ukrainian Research and Innovation System recommended that Ukrainian researchers, both from academia and business, need to have access to modern research infrastructures to have the possibility to conduct cutting-edge research, and the MES should also seek to establish beneficial agreements with European research infrastructures that allow Ukrainian researchers to have access to these research infrastructures and to conduct research there together with their fellow peers from other countries.

³³³ See <http://www.kiis.com.ua/?lang=eng&cat=db>

3.44 United Kingdom

This report summarises the results of a CESSDA-SaW survey on the strengths and potentials of data archiving services related to social science research in the United Kingdom. The responses made clear that the United Kingdom hosts a wealth of data from social science research and that the ingestion, storage, maintenance, and access to the data, is well established, organised and rooted in a strong research tradition. Overall a very high maturity level is demonstrated and in most areas, the United Kingdom scores (far) above the European average that was calculated in this survey. The foundations are laid for a sound technical, organisational and legal infrastructure that allows for excellent research in the realm of the social sciences. There remains work to be done in the areas of data sharing and creating a data sharing culture. For example, consistent and clear guidelines on data sharing amongst all relevant organisations and funding bodies have not yet been realised. Researchers are also not always offered the appropriate incentives for sharing their data. Despite the high maturity level faced in the United Kingdom, the respondents noted challenges around maintaining sufficient resources and being able to train and appoint the proper staff.

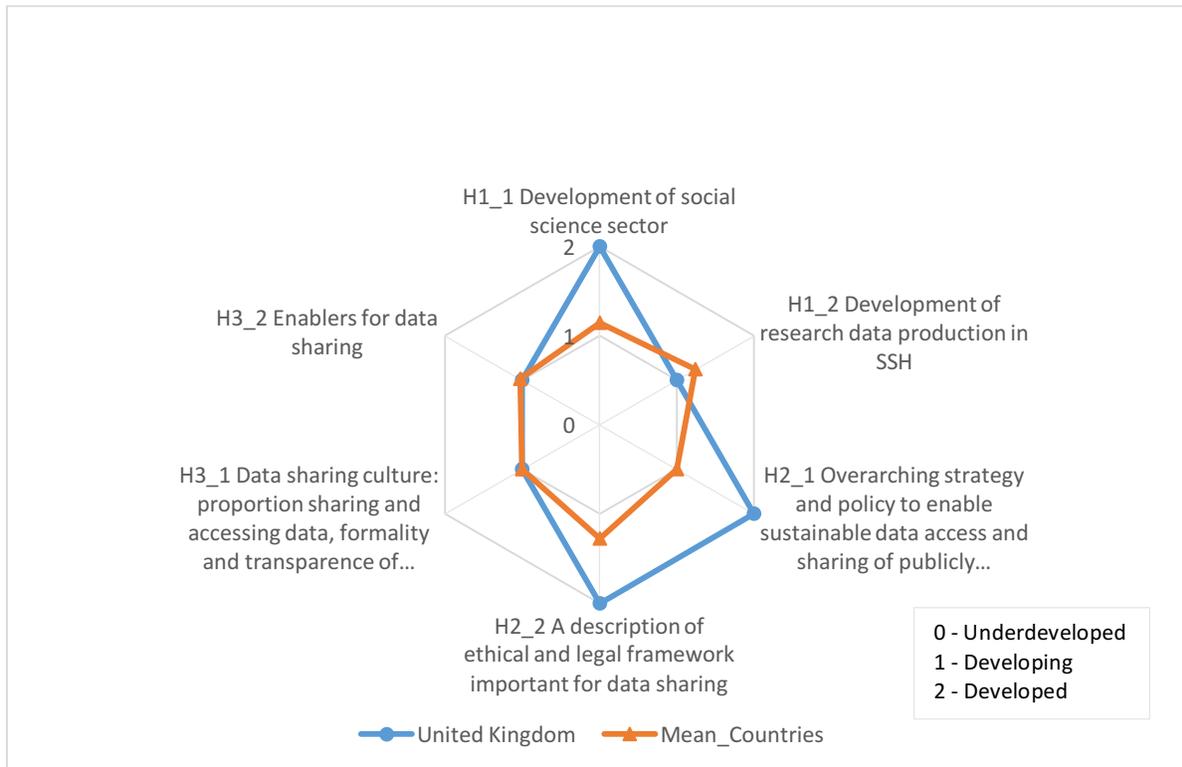
Specifics of data collection in country

Matthew Woollard, director of the UK Data Archive (<http://data-archive.ac.uk>) and director of the UK Data Service (<https://www.ukdataservice.ac.uk>), and his colleagues Veerle van den Eynden and Hervé L'Hours, responded to the survey in October 2016. The collection of information went smoothly and was provided well-structured and with the proper references for (online) details. In the United Kingdom, the issue of data sharing, open data and open science is well on the agenda. A very relevant report from the UK Data Service on open research³³⁴ was released only a few weeks after this survey was completed, references to this report are provided but it was not possible to use the content of this report. As authors of the report were also involved in this CESSDA-SaW survey, it can be assumed that much of the observations noted in the report is also represented in this survey. Despite this minor issue, the provided information allowed for a good analysis of the current situation concerning the UK Data Archive and the social science research data landscape in the United Kingdom.

³³⁴ Van den Eynden, Veerle et al. (2016). *Towards Open Research: practices, experiences, barriers and opportunities*. Wellcome Trust. Available at: <https://dx.doi.org/10.6084/m9.figshare.4055448>.

3.44.1 Broader ecosystem of DAS operation

Figure 96: Heading concepts values in the UK



In Figure 96, the broader ecosystem of DAS operation in the United Kingdom is well developed. Areas open for further development concern the development of research data production in Social Sciences and Humanities (SSH), the enablers for data sharing and the several surveyed aspects of the data sharing culture.

Development of social science sector

Overall Assessment of SSH development

The financial stability, research capacities and results achieved in the field of the social sciences in the United Kingdom are at the highest recorded level and above the average of the surveyed European countries. Funding of SSH and productivity of the researchers are in the highest quantile; impact on the designated community is strong.

Research funding is not only provided by the government and higher education sector, but also by the private non-profit sector, the business enterprise sector and from abroad, ranking in this order. It is also very beneficial that the government and universities provide access to commercial and bibliographic full text databases, access to datasets and provide and support software licences. However, there are differences between organisations concerning the types and levels of subscription to commercial services.

In the United Kingdom, the GERD (Gross domestic Expenditure on Research and Development) in SSH, as % GDP, rates 0.21.

Development of research data production in SSH

In the United Kingdom, the level of research data development is developing and slightly below the mean of all the countries that have been surveyed. This situation is (possibly) due to the fact that the United Kingdom in principal does not participate in all the surveyed international initiatives, listed below. Obviously, this slightly reduces the “data production” score of the United Kingdom.

International social science data initiatives in which the United Kingdom is involved:

- European Social Survey (ESS)
- Comparative Study of Electoral Systems (CSES)
- European Values Study (EVS)
- World Values Survey (WVS)
- Programme for International Student Assessment (PISA)

The United Kingdom is not involved in the following initiatives:

- Comparative Candidate Survey (CCS)
- International Social Survey Program (ISSP)
- Generations and gender programme (GGP)

At the national level, however, there are many surveys carried out by government departments and research organisations and these data are available through the UK Data Service³³⁵.

There is a good tradition in the generation of survey data by many organisations. The Economic and Social Research Council (ESRC) funds much social science data production in the United Kingdom on a continuous basis, both large-scale longitudinal data production³³⁶ and small-scale research. Also, government departments have social science research professionals who carry out regular data production, grouped into the Government Social Research Profession³³⁷.

RDM Policy setting

The overarching strategy and policy to enable sustainable data access and sharing of publicly financed social science research data is at a well-developed level and above the average of the surveyed European countries.

The **general situation** with regards to requirements and/or recommendations about preparing Data Management Plans (DMP's), aiming at open data as the default among public funders of social science research in the United Kingdom, is well established. A DMP is usually a requirement, clear guidance is issued, support and tools are provided, the content of DMP and exemptions from full open access are defined³³⁸. However, in general, it is not clear to what extent the DMP's are used

³³⁵ Available at: <https://www.ukdataservice.ac.uk/get-data/key-data>

³³⁶ Available at: <http://www.esrc.ac.uk/research/our-research/>

³³⁷ Available at: <https://www.gov.uk/government/organisations/civil-service-government-social-research-profession>

³³⁸ DMP requirements specified in the Economic and Social Research Council (ESRC) data policy: <https://www.ukdataservice.ac.uk/manage-data/plan/dmp-esrc>

as part of the funding approval process. There is little evidence that DMP's are being judged in the review process or that they are being used during research to implement good data management practices.

Quality requirements set by the public research funding organisations are in place for primary data and for metadata. Not fulfilling these requirements may result in sanctions. Also for this aspect of the general data policy it is not always clear how these requirements are used and/ or assessed by the funders and in review procedures.

Formal requirements, based on collective, contractual arrangements of roles and responsibilities between different stakeholders, are in place. Clear definitions are established by the funders of public research on how data should be documented, what data is preserved, and for how long. Funders require that periodic monitoring is performed to assure that the resources are used in such a manner that the amount of benefit is maximised.

How well defined the data sharing requirements often are, is a subject to a discussion. At the UK Data Archive, itself it is somewhere between the qualifiers "formal" and "managed". For example, there are generally clear definitions of what data is preserved, but whether the value judgements for selecting materials for curation are applied because this is the best use of resources is not always clear. In general, for the United Kingdom, some data are curated because "*they've always been curated*", and some data which has a greater potential reuse value may not be curated.

Public research funding organizations in the United Kingdom provide **incentives** for sharing research data and associated metadata. Costs for managing the data and preparing it for access are often resourced adequately during the lifetime of research projects. There is explicit recognition that additional cost for preparing the data for access are legitimate project cost that can partially cover the RDM cost up to a certain limit. The situation differs amongst the funding bodies, scaling from "partly" to "defined". In some cases, costs are fully covered "directly" in an award, but, for example the Engineering and Physical Sciences Research Council (EPSRC) would expect that indirect costs are used for some of these activities (and therefore might be used in a university's FEC calculations). It is best to label the policies around incentives and financial support for data sharing as "partial" as costs are not automatically covered (by institution or by funder) based on DMP.

Legal and ethical framework for data sharing

In the United Kingdom, organised services to support and encourage legally and ethically sound data sharing practice are widely available. For data producers, data users (researchers) and interested others, the UK Data Service provides detailed guidance on the legal and ethical aspects of sharing social sciences data on their website³³⁹, and runs a programme of regular training events on these topics. Also, the main public funder ESRC provides clear information in their Research Data Policy and their Framework for Research Ethics on its expectations for data sharing within a legal and ethical context. When social science data are deposited with the UK Data Service, data are checked for any disclosure information, assigned access criteria in line with the consent (or other) agreements made with research participants. For data use, the UK Data Service provides

The ESRC funding guide: <http://www.esrc.ac.uk/files/funding/guidance-for-applicants/research-funding-guide/>
DMP guidance provided by the UK Data Service: <https://www.ukdataservice.ac.uk/manage-data/plan/dmp-esrc>
DMP guidance for peer reviewers (ESRC): <http://www.esrc.ac.uk/files/about-us/policies-and-standards/data-management-plan-guidance-for-peer-reviewers/>

³³⁹ Available at: <https://www.ukdataservice.ac.uk/manage-data/legal-ethical>

different levels of access to data, to facilitate sharing of a wide range of data, including those that may contain disclosure information: open data, safeguarded data, controlled data (these levels are similar to those in the CESSDA Data Access Policy).

Data sharing culture

Data sharing and the reuse of data generated by social science researchers is facilitated, promoted and becoming more and more the standard. This aspect of the data cycle is still in a developing stage, however, and at an average level (estimated > 30%).

It is estimated in self-assessment survey that in the United Kingdom more than 30% of the social science researchers have shared the research data they produced in the period between 2011 and 2015. The data sharing culture is still developing as in most of the European countries. The estimate is backed up by following calculations and reports.

Counting from 2013 on, more than 30% of the research projects funded by ESRC have shared their data via the UK Data Archive (proportion calculated as number of datasets received by UK Data Service versus number of research grants funded by ESRC³⁴⁰. More than half (55%) of 250 researchers have made data available in the last 5 years (2010 - 2016) (Ven den Eynden et al. 2016)³⁴¹. From the Wellcome Trust-funded³⁴² researchers, 30% of social science researchers have made data available in last the 5 years (Ven den Eynden et al. 2016). Obviously, this is related to the fact that these funders strongly advocate further re-use of data and provide the guidance to do so. There are also disciplinary variations, for example many prominent economy journals nowadays require deposit of all raw data and code as supplementary materials for replication and validation purposes, and this increases the data sharing percentages significantly.

Concerning the use of shared social science data, it is estimated in self-assessment survey that more than 30% of the academic social science researchers have accessed data from the UK Data Service. In principle, all social science researchers working within academia have access to all data made available through the UK Data Service.

A real ranking of the preferred routines of data sharing for the social science research community in the United Kingdom was difficult to provide³⁴³.

It is the respondents' impression that the risk that others may misuse and misinterpret your data is still a thought holding back social science researchers from sharing their data. Also, the fear of time consuming effort and high costs is an element inhibiting the sharing of data. The Open Research report (Ven den Eynden et al. 2016), based on > 300 survey responses, provides a more detailed insight on these considerations and thoughts amongst the social science research communities.

³⁴⁰ See Figure 3 in the online publication Van den Eynden, Veerle and Louise Corti (2016, May.) Advancing research data publishing practices for the social sciences: from archive activity to empowering researchers. Available at: <http://link.springer.com/article/10.1007/s00799-016-0177-3>

³⁴¹ See the Open Research report Van den Eynden, Veerle et al. (2016) Towards Open Research: practices, experiences, barriers and opportunities. Wellcome Trust. Available at: <https://dx.doi.org/10.6084/m9.figshare.4055448>

³⁴² The Wellcome Trust (<https://wellcome.ac.uk>) is a major funder of international research and supports researchers exploring ideas in science, population health, medical innovation, the humanities and social sciences and public engagement.

³⁴³ The respondents had difficulties with the question on ranking the routines for data sharing to reflect the preferences of the social science research community in the UK. Do routines reflect the "source" from which researchers may access data or the "method" by which researchers make data available?

Attitudes towards data sharing

The table below provides several statements on data sharing and the scoring, on a five-point scale from “5”-True to “1”-False, that best matches the overall attitudes of social science researchers in the UK. The scoring is based on experiences of the UK Data Archive and previously published reports in the period from 2011 - 2016.

Table 34: Attitudes towards data sharing in the UK

Data sharing has no benefits at all	Neither true, nor false
Data sharing creates healthy competition	Probably false
Data sharing creates negative competition	Probably true
Reuse of existing data can answer new research questions and facilitate advancement of science	Probably true
Data sharing has as a risk that others may misuse and misinterpret data	True
Data sharing involves little effort and minimal costs	False

Source: self-assessment survey. Question asked: Please, score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011-2016, on a five-point scale from “5”-True to “1” False!

Enablers for data sharing

Enablers for data sharing are at an average level in the United Kingdom and still developing.

When it comes to the relationship between data sharing and the benefit to the career of researchers in the social sciences it appears that in practice the sharing of your research data does not benefit your career (see also the previously mentioned Open Research report). Most likely this is a “does not benefit yet” as most research funders and research organisations do emphasise that these types of activities may be used to enhance career progression. The possibility of rewards may not have had a serious opportunity to filter down too many academic researchers.

Several organisations in the United Kingdom provide social science researchers with the means to manage, share and publish their data as open online data.

ESRC funded:

- ADRN - a group of universities and national statistics agencies spread throughout the UK, who work together to help researchers get access to de-identified, linked administrative data.
- The UK Data Archive is an internationally acknowledged centre of expertise in acquiring, curating and providing access to social science and humanities data. The UK Data Archive is the lead organisation of the UK Data Service., which provides access to the UK's nationally and internationally significant social science data assets for research, teaching, skills development and policy-making

Other:

- The Digital Curation Centre (DCC), providing advice on digital curation activities, as well as Data Management Planning
- Many Higher Education Institutions provide data management and sharing guidance, (staff) support and a data repository service to their researchers.

Most social science data producers also follow the data management and data documentation standards and procedures that are provided to facilitate data reuse. See examples at the UK Data Service data management guidance³⁴⁴. Another example of one of the many universities that provide data management guidance to their researchers, is the University of Edinburgh³⁴⁵.

Concerning the completion and standardisation of metadata; that is usually part of the guidance and documentation on RDM. It is the respondents' impression that data documentation standards are not widely used in the United Kingdom.

3.44.2 Capability requirement areas of DAS

Organisational profile

Organisation

The UK Data Archive is situated at the University of Essex (Wivenhoe Park, Colchester, Essex CO4 3SQ, United Kingdom).

The UK Data Service, the United Kingdom Service Provider to CESSDA, is a national distributed data service funded by the ESRC to provide research access to the UK's largest collection of social, economic and population data. The lead partner is the UK Data Archive, based at the University of Essex. As part of its contribution the UK Data Archive provides an information secure (ISO27001 compliant³⁴⁶) Trustworthy Data Repository (Data Seal of Approval compliant, <http://www.datasealofapproval.org>). The other partners in the UK Data Service are: Jisc, the Universities of Manchester, Southampton, Leeds, and Edinburgh, and University College London (UCL). The UK Data Archive is a department in the University of Essex, and carries out activities beyond the UK Data Service. Over 80 staff members are employed by the UK Data Archive.

Separate services that are being provided such as:

- Access Services
- Administrative Data Service
- Application Development and Maintenance
- Big Data Network
- Collections Development
- Communications and Impact

³⁴⁴ See <https://www.ukdataservice.ac.uk/manage-data>

³⁴⁵ See <http://www.ed.ac.uk/information-services/research-support/data-management>

³⁴⁶ See <http://www.iso.org/iso/iso27001>

- Digital preservation Systems and Security
- Enhancing and Enriching Historical Census Microdata
- Ingest Services
- Management Information Systems
- Producer Support
- Resource and Management Services
- User Training and Support

The designated user community of the Data Archive is made up of social science and related data users within Higher Education (HE) and Further Education (FE) in the United Kingdom, though best efforts are made for all users (for the purpose of this definition, quantitative and qualitative data collections created by or for historians are considered to be “social scientific”). All users are expected to have a basic understanding of social science methods and techniques relevant to the data collections being accessed.

Funding

The UK Data Service is funded by the Economic and Social Research Council (ESRC). The UK Data Archive, which is the lead partner in the UK Data Service, is a department of the University of Essex and is thus partially funded by the University of Essex. The UK Data Archive also receives project funding from other organisations, but this is generally limited to 10% of the annual income.

Core services and activities

The UK Data Service provides unified access to the UK’s largest collection of social, economic and population data. Their collections include most major United Kingdom government-sponsored surveys, cross-national surveys, longitudinal studies, UK census data, international aggregate, business data, and qualitative data. Data users can browse UK Data Service collections online and register to analyse and download them. Open Data collections are available for anyone to use. The UK Data Service is designed to meet the data needs of researchers, students and teachers from all sectors, including academia, central and local government, charities and foundations, independent research centres, think tanks, business consultants and analysts, communities and the commercial sector, providing:

- Access to high-quality local, regional, national and international social and economic data
- Support for policy-relevant research in the higher education, public and commercial sectors
- Guidance and training for the development of skills in data use
- The development of best practice in digital preservation and data sharing standards
- The sharing of expertise with international data providers to remove barriers to accessing data
- Pre-Ingest and ingest services
- Administration
- Management
- Archival Storage

Curation Services at the following levels, depending on the nature of the collection, is provided:

- B. Basic curation – e.g. brief checking, addition of basic metadata or documentation
- C. Enhanced curation – e.g. creation of new formats, enhancement of documentation
- D. Data-Level Curation as C above but with additional editing of deposited data for accuracy

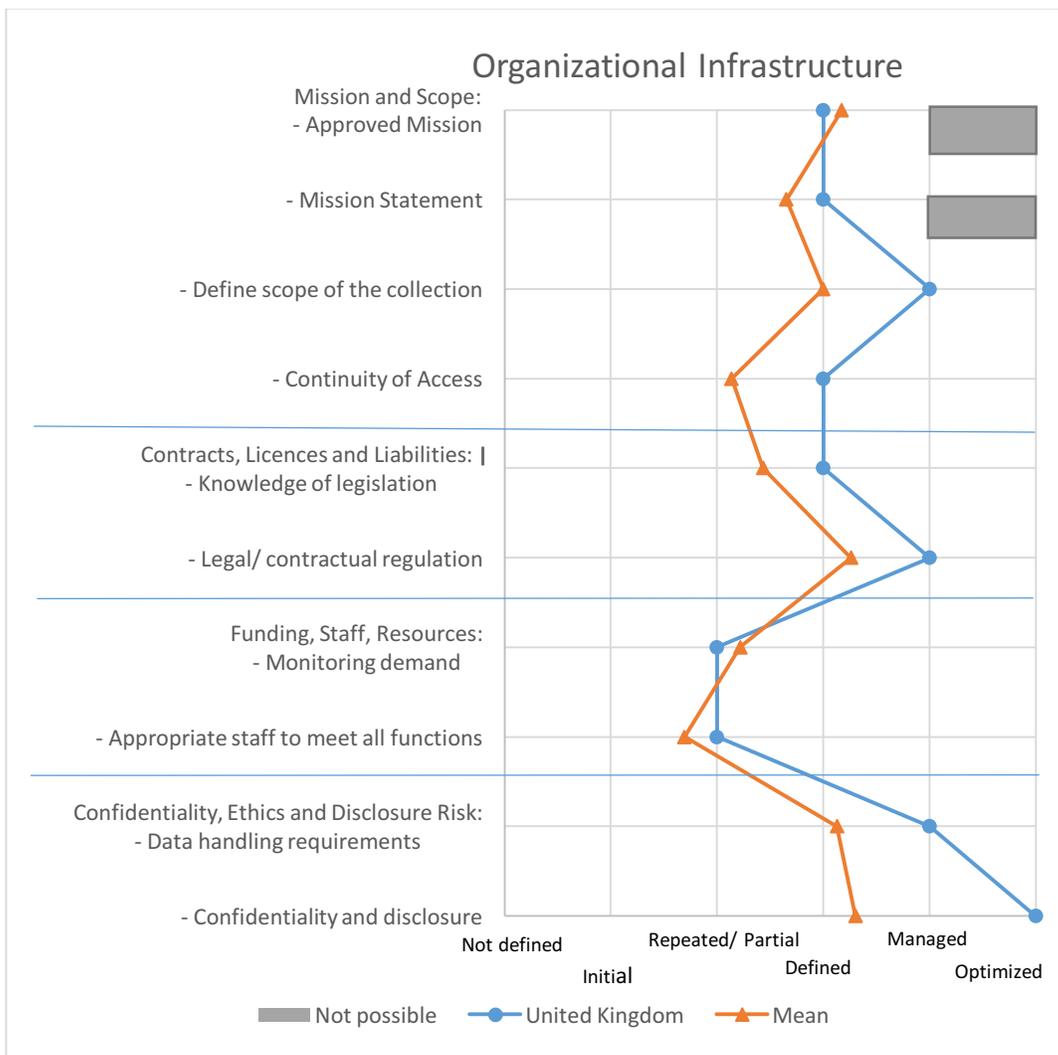
Content current collection

Social and economic data, major UK government-sponsored surveys, cross-national surveys, longitudinal studies, UK census data, international aggregate, business data, and qualitative data.

The language is primarily English for deposits. English for all metadata and documentation generated as part of the curation activities.

Organisational infrastructure

Figure 97: Organizational infrastructure in the UK



Mission statement

The formal mission of the UK Data Archive is to provide users with seamless and flexible access to a wide range of data resources to facilitate high quality social and economic research and education³⁴⁷.

The Strategic Plan of UK Data Archive describes the institute in the Mission Statement³⁴⁸ as a “leading centre of expertise in data archiving in the United Kingdom. It holds the largest collection of digital data in the social sciences and humanities in the UK and makes them available to the higher and further education sector and beyond for research, teaching and learning. Its mission is to maintain its leading role by continuously improving the quality and breadth of its data products and services in response to user needs and technological changes. To this end it collaborates with national and international organisations to develop new data standards and tools. It also aims to build bridges between data creators and data users across different sectors and disciplines”. The organisation thus provides a written, formal and complete mission statement that is available for all relevant users.

Approved mission

The Mission Statement forms part of the Strategic Plan which is approved at the highest levels.

Define scope of the collection

The UK Data Archive has defined an explicit scope for the repository and their collection, which guides the selection and appraisal of data. The definition of scope is monitored for compliance with policies, processes and procedures. Actions are taken where processes appear not to be aligned to scope, policy or mission statement and regular reviews of scope and mission of repository are performed.

Continuity of Access

Formal, written succession/contingency plans are in place to ensure the continued availability and accessibility of data in case the UK Data Archive ceases to operate. This includes formal written agreements with external organisational frameworks.

Knowledge of Legislation

All staff has access to sufficient and documented knowledge on all relevant legislative aspects, both at the national as well as the international level.

³⁴⁷ See <https://www.ukdataservice.ac.uk/about-us/purpose>

³⁴⁸ UK-DATA strategic plan. Available at: <http://www.data-archive.ac.uk/media/196518/ukda-strategicplan20102015full.pdf>

Legal/contractual regulation

The legal and/or contractual framework is regularly reviewed and updated; all legal and/or contractual regulations with the service users are aligned to higher level policies; roles and responsibilities are identified and maintained³⁴⁹.

Compliance to procedures and policies are monitored and assessed. Non-compliance incidents are recorded. Information security procedures are in place, sufficient to be accredited to ISO 27001. However, the UK Data Archive doesn't review policies on a regular basis "for compliance to research discipline norms" because they review them "for compliance against information security standards and national legislation for privacy protection."

Appropriate staff to meet all functions

The UK Data Archive provides its service based on the appropriate number of staff to support all basic repository functions and services. However, appropriate staff and resources to address the ever-growing needs of the user community remain an issue.

Confidentiality and Disclosure

Concerning data related to human subjects the UK Data Archive requires that data meet the requirements of confidentiality and non-disclosure. The usage and success of confidentiality and disclosure mechanisms are continuously assessed, reviewed and updated; monitoring of wider legal framework (e.g. national and EU regulations); regular and formalised contact with relevant stakeholders; automated checks and anonymisation mechanisms may be in place³⁵⁰.

³⁴⁹ See <https://www.ukdataservice.ac.uk/get-data/how-to-access/conditions>

³⁵⁰ See also: <https://www.ukdataservice.ac.uk/manage-data/legal-ethical> and <https://www.ukdataservice.ac.uk/deposit-data>.

Digital object management (data curation) and Technical infrastructure and risk

Figure 98: Digital object management and Technical infrastructure in the UK

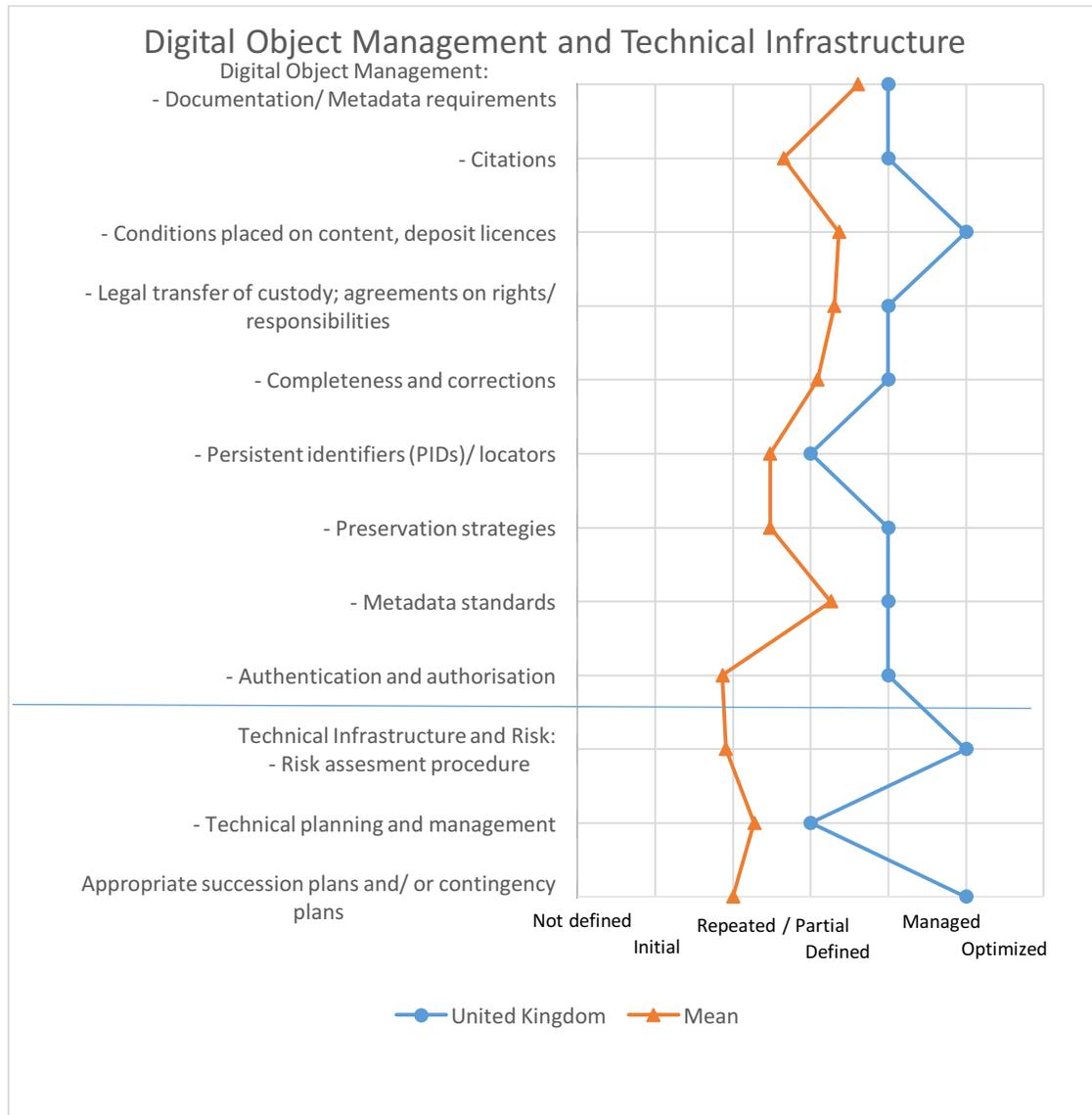


Figure 98 demonstrates that digital object management is in general at the managed level and all elements are above the average. The Technical Infrastructure and Risk is at a level between defined and managed and is also far above average.

Documentation and metadata requirements

The UK Data Archive clearly specifies the information (documentation, metadata, provenance) that needs to be associated with the data that is to be deposited. The services are at the managed level. Documentation and metadata requirements are aligned with policies and other processes

and procedure documents. There are regular reviews and assessments (of success) of the information requirements³⁵¹.

Citations

The UK Data Archive offers and provides functions and mechanisms for proper data citations, citation mechanisms are regularly reviewed and updated³⁵².

Conditions on content and deposit licences

Mechanisms to allow for the data depositors to assign the proper data access conditions to the data are in place. Regular review and updates of conditions are established; formalised feedback mechanisms and cooperation with user groups and other relevant stakeholders (e.g. how to deal with funder policies that have open access requirements)³⁵³.

Legal transfer / agreements on rights and responsibilities

Agreements that confirm the legal transfer (or other consensual agreements) of the information that is being deposited are in place, and the agreement includes clear definitions of roles and responsibilities of repository and depositor. The usage of agreements and contracts is being monitored and actions are taken where contracts/agreements appear not to be working effectively or are not in accordance with higher level policies³⁵⁴.

Completeness and corrections

Processes and mechanisms are in place to verify the deposited material for completeness and correctness. It is ensured that variables and values are accurate according to the documentation supplied, are sufficiently labelled for reuse and that variable names in a dataset match variable names in a codebook. Checks and modifications are measured and registered; processes and procedures are reviewed and updated regularly. Automated processes are implemented where appropriate³⁵⁵.

³⁵¹ See: <https://www.ukdataservice.ac.uk/deposit-data>

³⁵² See: <https://www.ukdataservice.ac.uk/use-data/citing-data>

³⁵³ See: <https://www.ukdataservice.ac.uk/manage-data/legal-ethical/access-control>

³⁵⁴ See: <https://www.ukdataservice.ac.uk/manage-data/rights>

³⁵⁵ See: <https://www.ukdataservice.ac.uk/manage-data/format/quality>

Persistent identifiers

There are mechanisms and systems in place at the UK Data Archive to persistently identify and locate data and metadata (either by following external systems like DOI, or by internal PID systems); all processes and procedures are documented and formalised³⁵⁶.

Persistent identifications of individuals (ORCID etc.) is evolving, the same concerns organisations (re3data is a partial solution). Identification of linked subsets of new and novel forms of data present challenges in contrast to the traditionally versioned data/metadata/documentation “study” object model.

Preservation strategies

The UK Data Archive has a preservation strategy in place that is periodically reviewed and updated. The strategy includes issues such as the degradation of storage media, the obsolescence of media drives and protection against accidental or intentional digital corruption of preserved material.

Metadata standards

The UK Data Archive provides data and metadata that are in accordance with standards that are used and understood by the Designated Community. Format usage and enquiries are measured and assessed; format strategy is regularly reviewed and updated. Acceptable file formats for delivery are defined, this is done through continuous user training and support interaction with users, rather than through a file format questionnaire or other formal mechanism.

Authentication and authorisation

The organisation uses AAI (Authentication and Authorization Infrastructure) to control access to data, all AAI mechanisms and functions are measured, assessed and regularly reviewed and updated (Shibboleth for depositors and end users, Microsoft Active Directory for internal users).

Technical Infrastructure and Risk

Risk assessment procedure

At regular intervals, the risk assessment and analysis processes concerning the risks of changes to the technical infrastructure, are reviewed. Any modifications to the processes are documented and communicated internally. This level of analysis is required for ISO 27001 Certification.

³⁵⁶See: <https://www.ukdataservice.ac.uk/use-data/citing-data>

Technical planning and management

Technically the UK Data Archive has defined the appropriate technical infrastructure resources to support all functions and services and addresses the 'right staff, right skills' challenge.

Succession / contingency plans

To safeguard the UK Data Archive continuation and the access to and preservation of its holdings, the procedures and processes are continuously assessed; the wider community is monitored (technology watch); regular and formalised contact with relevant stakeholders is maintained; assessment of success of training programs and staff knowledge is performed; procedures are assessed and measured towards cost-benefit assessments and repository performance; and the UK Data Service benefits from repository functions at the UK Data Archive.

3.44.3 Conclusions

The CESSDA-SaW self-assessment for the United Kingdom was performed by the UK Data Service. The Data Archive provided well-structured information about their own situation as well as a clear picture of the social science research data landscape in the United Kingdom.

What the information clearly canvasses, is the high maturity level. This is based on a strong research tradition where, obviously, data have always played an important role, but also because developments and investments are continuing at the UK Data Archive and other organisations and institutes alike. Therefore, in relation to the broader eco-system of DAS operation, the United Kingdom is on the developed level in the development of the social science sector.

The United Kingdom does not participate in all major international projects that were listed in the survey on which this report is based. At the national level, there are many data-initiatives running, but overall the development of research data production is still in a developing stage. RDM Policy setting is assessed on the developed level. Guidelines and support for DMP is well established, although how DMP's are utilized in an optimum way, i.e. by the funding agencies, is not always clear. The legal and ethical framework for data sharing is in place, provided through the UK Data Service.

The data sharing culture in the United Kingdom is on a developing level, both in terms of the enablers as well as developing a data sharing culture. As in many disciplines of science there is still some reluctance amongst researchers to share data, be it for practical (how exactly) or financial (extra costs) reasons. Platforms to share data and the relevant guidelines are in place. However, there is certainly room for more and better developments.

In relation to the capability requirement areas of DAS, the organizational profile of the UK Data Archive shows that the organization is sufficiently equipped with staff, offering their services through several thematic sections. The user community of the UK Data Archive is made up of social science and related data users within Higher Education and Further Education in the UK and internationally, though best efforts are made for all users.

Both the organizational infrastructure and digital object management (data curation) and technical infrastructure are on the managed level, or just below. For most assessed elements, the level of development is certainly above average. The Mission Statement of the organisation is part of its Strategic Plan which is approved at the highest levels. A point of concern according to the respondents is the issue of sufficient funding and the appropriate staff to fulfil the ever-growing needs of the user community.

Appendix 1: List of tables and figures

List of tables:

- Table 1: Target countries for CESSDA SaW
- Table 2: Attitudes towards data sharing in Austria
- Table 3: Participation of Bosnia and Herzegovina in a selection of international surveys
- Table 4: Attitudes towards data sharing in Bosnia and Herzegovina
- Table 5: Attitudes towards data sharing in Croatia
- Table 6: Assessment of SSH development in Cyprus
- Table 7: Development of research data production in SSH in Cyprus
- Table 8: Attitudes towards data sharing in Cyprus
- Table 9: Attitudes towards data sharing in Czech Republic
- Table 10: Attitudes towards data sharing in Denmark
- Table 11: Attitudes towards data sharing in Estonia
- Table 12: Attitudes towards data sharing in Finland
- Table 13: Participation of Germany in a selection of international surveys
- Table 14: Attitudes toward data sharing in Germany
- Table 15: Attitudes toward data sharing in Greece
- Table 16: Attitudes toward data sharing in Hungary
- Table 17: Attitudes towards data sharing in Israel
- Table 18: Attitudes towards data sharing in Italy
- Table 19: Attitudes towards data sharing in Latvia
- Table 20: Attitudes towards data sharing in Lithuania
- Table 21: Attitudes towards data sharing in FYRM
- Table 22: Attitudes towards data sharing in The Netherlands
- Table 23: Attitudes towards data sharing in Norway
- Table 24: Assessment of SSH development in Poland
- Table 25: Development of research data production in SSH in Poland
- Table 26: Attitudes towards data sharing in Portugal
- Table 27: Attitudes towards data sharing in Russia
- Table 28: Attitudes towards data sharing in Serbia
- Table 29: Assessment of SSH development in Slovak Republic
- Table 30: Development of research data production in SSH in Slovak Republic
- Table 31: Assessment of SSH development in Slovenia
- Table 32: Development of research data production in SSH in Slovenia
- Table 33: Attitudes towards data sharing in Slovenia
- Table 33: Attitudes towards data sharing in Switzerland
- Table 34: Attitudes towards data sharing in the UK

List of figures:

- Figure 1: Heading concepts values
- Figure 2: DAS and OA activities implementation type
- Figure 3: DAS and OA activities implementation type in Austria
- Figure 4: DAS and OA activities implementation type in Belarus
- Figure 5: Heading concepts values in Belgium
- Figure 6: Organizational infrastructure in Belgium
- Figure 7: Digital object management and Technical infrastructure in Belgium
- Figure 8: Heading concepts values (BiH)
- Figure 9: DAS and OA activities implementation type in Bosnia and Herzegovina

Figure 10: Heading concepts values in Bulgaria
Figure 11: DAS and OA activities implementation type Bulgaria
Figure 12: Heading concepts values in Croatia
Figure 13: DAS and OA activities implementation type in Croatia
Figure 14: Heading concepts values in Cyprus
Figure 15: DAS and OA activities implementation type in Cyprus
Figure 16: Heading concepts values in Czech Republic
Figure 17: Organizational infrastructure in Czech Republic
Figure 18: Digital object management and Technical infrastructure in Czech Republic
Figure 19: Heading concepts values in Denmark
Figure 20: Organizational infrastructure in Denmark
Figure 21: Digital object management and Technical infrastructure in Denmark
Figure 22: Heading concepts values in Estonia
Figure 23: Organizational infrastructure in Estonia
Figure 24: Digital object management and Technical infrastructure in Estonia
Figure 25: Heading concepts values in Finland
Figure 26 Organizational infrastructure in Finland
Figure 27: Digital object management and Technical infrastructure in Finland
Figure 28: Heading concepts values in France
Figure 29: Organizational infrastructure in France
Figure 30: Digital object management and Technical infrastructure in France
Figure 31: Heading concepts values in Germany
Figure 32: Organisational infrastructure in Germany
Figure 33: Digital object management and Technical infrastructure in Germany
Figure 34: Heading concepts values in Greece
Figure 35: Organizational infrastructure in Greece
Figure 36: Digital object management and Technical infrastructure in Greece
Figure 37: Heading concepts values in Hungary
Figure 38: Organizational infrastructure in Hungary
Figure 39: Digital object management and Technical infrastructure in Hungary
Figure 40: Irish research landscape, picture provided by HRB
Figure 41: Heading concepts values in Ireland
Figure 42: Organizational infrastructure in Ireland
Figure 43: Digital object management and Technical infrastructure in Ireland
Figure 44: Heading concepts values in Israel
Figure 45: DAS and OA activities implementation type in Israel
Figure 46: Heading concepts values in Italy
Figure 47: Organisational infrastructure in Italy
Figure 48: Digital object management and Technical infrastructure in Italy
Figure 49: Heading concepts values in Kosovo
Figure 50: DAS and OA activities implementation type in Kosovo
Figure 51: Heading concepts values in Latvia
Figure 52: DAS and OA activities implementation type in Latvia
Figure 53: Heading concepts values in Lithuania
Figure 54: Organizational infrastructure in Lithuania
Figure 55: Digital object management and Technical infrastructure in Lithuania
Figure 56: DAS and OA activities implementation type in Luxembourg
Figure 57: Heading concepts values in FYRM
Figure 58: DAS and OA activities implementation type in FYRM
Figure 59: DAS and OA activities implementation type in Moldova
Figure 60: Heading concepts values in Montenegro
Figure 61: DAS and OA activities implementation type in Montenegro
Figure 62: Heading concepts values in The Netherlands
Figure 63: Organizational infrastructure in The Netherlands
Figure 64: Digital object management and Technical infrastructure in The Netherlands
Figure 65: Heading concepts values in Norway

Figure 66: Organizational infrastructure in Norway
Figure 67: Digital object management and Technical infrastructure in Norway
Figure 68: Heading concepts values in Poland
Figure 69: Organizational infrastructure in Poland
Figure 70: Digital object management and Technical infrastructure in Poland
Figure 71: Heading concepts values in Portugal
Figure 72: Organizational infrastructure in Portugal
Figure 73: Digital object management and Technical infrastructure in Portugal
Figure 74: Heading concepts values In Romania
Figure 75: Organizational infrastructure in Romania
Figure 76: Digital object management and Technical infrastructure in Romania
Figure 77: Heading concepts values in Russia
Figure 78: Organizational infrastructure in Russia
Figure 79: Digital object management and Technical infrastructure in Russia
Figure 80: Heading concepts values in Serbia
Figure 81: DAS and OA activities implementation type in Serbia
Figure 82: Heading concepts values in Slovak Republic
Figure 83: Organizational infrastructure in Slovak Republic
Figure 84: Digital object management and technical infrastructure in Slovak Republic
Figure 85: Heading concepts values in Slovenia
Figure 86: Organizational infrastructure in Slovenia
Figure 87: Digital object management and Technical infrastructure in Slovenia
Figure 88: DAS and OA activities implementation type in Spain
Figure 89: Heading concepts values in Sweden
Figure 90: Organizational infrastructure in Sweden
Figure 91: Digital object management and Technical infrastructure in Sweden
Figure 92: Heading concepts values in Switzerland
Figure 93: Organizational infrastructure in Switzerland
Figure 94: Digital object management and technical infrastructure in Switzerland
Figure 95: DAS and OA activities implementation type in Ukraine
Figure 96: Heading concepts values in the UK
Figure 97: Organizational infrastructure in the UK
Figure 98: Digital object management and Technical infrastructure in the UK

Appendix 2: Guidelines and communication protocol for interviewers

A. General introduction to data collection methods, roles and principles

In this section we describe data collection methods, general principles to be followed, and suggested roles and their responsibilities in the data collection process. These cover the principles and instructions to be followed for data collection with all three methods – desk research, self-assessment, and semi-structured expert interviews with stakeholders (as well as for combining data sources and data collection methods, when needed).

Roles and responsibilities in data collection

Role of the Task leader and Task force in data collection

Task 3.2 leader coordinates and monitors the data collection efforts. Task leader is the contact point for resolving difficult cases related to data collection for a country.

Task force consists of representatives for sub-tasks working on different sections of web form to enable consultation related to specific items/indicators.

Task leader together with task force consult and support partner responsible for collection of data in a country in case there are some difficulties, and together with partner responsible for data collection make the final decisions with regard to use of resources for data collection.

Partner responsible for data collection in a country

She or he is a task 3.2 participant that has responsibility to coordinate the data collection efforts via different sources - desk research, self-assessment template and semi-structured interviews.

In cooperation with the task leader, and if, possible, with the contact person for the country/ service provider/DAS, partner responsible for the data collection in the country:

(a) plans data collection (see section E1) by selecting the most appropriate method(s) for data collection - desk research, self-assessment or semi-structured interview - for each item/indicator (and motivates it, if needed) within the suggested resources for the country (see [document on division of responsibilities](#));

(b) selects documents for desk research (mainly for type III and IV countries, for specification see [document on division of responsibilities](#)).

(c) selects experts for semi-structured interviews (mainly for type III and IV countries, see [document on division of responsibilities](#)),

Partner responsible for the data collection also

(d) initiates self-assessment by sending the invitation letter and supporting documents and personalised link(s) to the web form (see sections F and B for details)

(e) monitors and coordinates the issues related to the self-assessment and answers questions related to the self-assessment questions and entering self-assessment information in the web form;

(f) reviews the self-assessment information entered in the web form, and, if necessary, does the follow-up interviews (mainly type I and II countries, see [document on division of responsibilities](#));

(g) processes and aggregates the information when/if necessary (for example, assigns or adjusts the assessment scores on basis of semi-structured interview and desk research results in the Partners' Content Area of the web form; mainly for type III and IV countries, see [document on division of responsibilities](#));

(h) gathers and processes feedback from the country representatives regarding the final scores in the web form (based on self assessment, desk research and interviews and entered in the web form); if needed, adjusts the scores after the feedback;

(i) manages referenced resources in the Resource management section of the web form.

Contact person for the country/service provider/DAS

Contact person for the country/service provider/DAS should have the role of facilitator within the country/institution.

Contact person should:

- (a) assist partner responsible for data collection in the country with information on relevant contacts (within institution, regarding self-assessment template as well as experts for semi-structured interviews) and documents,
- (b) coordinate that the most appropriate people in the institution(s) answer the self-assessment questions in the web form (self-assessment can be done by more than one respondent in the organisation or organisations);
- (c) assist the partner responsible for the country with desk research (especially, if the documents are not in English or the partner responsible for the country does not have the linguistic competence to navigate them);
- (d) assist partner responsible for the country in case clarifications and/or follow-up interviews needed in the review process of self-assessment scores in the web form;
- (e) where possible and necessary, the contact person could also assist in modifying/adapting the semi-structured expert interview guidelines for the situation in the country and/or linguistically;
- (f) provide (or coordinate provision of) feedback regarding the final scores (based on self-assessment, desk research and interviews) entered in the web form;
- (g) assist partner responsible for the country in providing the referenced resources for the Resource management section of the web form.

If a contact person for the country/service provider/DAS willing to fill the role cannot be found (most probably type III and IV countries, for specification see [document on division of responsibilities](#)), the partner responsible for data collection in the country should assume most of the responsibilities described here.

Interviewer

Interviewer is a person conducting a semi-structured interview with an identified key expert, following the interview guidelines. This is usually the **partner responsible for the country**, but can be the **contact person for the country/institution/DAS**, if agreed with partner responsible for the country. Interviewer could be asked to

- (a) contact the selected interview person
- (b) conduct the interview,
- (c) prepare a short summary of answers to the interview questions in one document, and use notes following the questions in a web form to substantiate the coding decisions.
- (c) in some cases should also be responsible to send the interview summary to the interviewee, and communicate the feedback to the contact person for the country or partner responsible for data collection in the country (if agreed while discussing participation in the interview).

Interviewee/interview person

Interviewee/interview person is an expert participating in semi structured interview.

Respondent

Respondent is a person assigning self-assessment scores for a country in the web form (or a part of it). Respondent should be able to provide links to the relevant documents, if asked; provide clarifications and answer the follow-up interview questions, if necessary.

Planning data collection

There are three main data collection methods used. Desk research includes both review of official sources (like policy documents, statistical data) as well as literature review (articles, reports etc. on previous data collection efforts). The field research consists of self-assessment and semi-structured interviews. The self-

assessment in a form of survey is primarily aimed at SP/DAS representatives as respondents (main method of data collection for type I and II countries, see [document on division of responsibilities](#)). Semi-structured expert interviews will be used to reach a wider group of stakeholders for assessing the country-specific scientific, cultural, policy, legal and funding setting.

Data from all three data collection methods as well as references to the relevant resources, are entered in the web form. In countries where no SP/DAS exists, partner responsible for a country assigns scores in the web form based on information from desk research and semi-structured interviews.

As not all data collection methods can or should be applied in all situations (because of the specific situation in the country or limited resources), data collection should follow some general principles, regarding the division of responsibilities, allocation of resources and collection of raw data versus aggregated information.

1. Where possible, the responsibility for data collection should go to the partner responsible for the country (and not to the contact person for country/SP/DAS). However, for both semi-structured interviews as well as for desk research it is reasonable to ask the contact person to assist in finding relevant contacts, documents, and other relevant information. For self-assessment, the contact person should be the respondent or coordinator for other respondents in the institution filling in the template.

2. We should, where possible, choose the solution requiring fewer resources (time, effort etc.), if it does not decrease the value of information considerably (for example, choose desk research instead of semi-structured expert interviews, if possible and relevant; use self-assessment instead of semi-structured interviews; decide to focus on one or two semi-structured interview blocks instead of all; prepare structured interview summaries instead of full transcripts, etc.).

3. In conclusion, as this is not an academic research project, and in most cases the countries assessed and having a SP/DAS have considerable knowledge of their situation already, the rule of a thumb should be:

start with the known desk research sources (especially for Part I);

then use the self-assessment (especially Part II that is based on the CMD-model);

finally, semi-structured interviews should be used

only for indicators where this information is essential (to get the research-community perspective, for example; thus shortening and focusing the interviews considerably)

where nothing else is available, and

for follow-up interviews if needed in addition to self-assessment.

These principles have been partly implemented by drafting the data collection instruments (for example, the web form distinguishes between the self-assessment questions to be answered by SP/DAS respondent and the Partners' Content Area, where scores based on the results of desk research or semi-structured interview can be assigned). There is, however, still some flexibility that should be managed by the partner responsible for a country within the limits of resources.

For more detailed description on how to document data collection planning, see section E.

Data aggregation from multiple sources

If multiple methods of data collection (e.g., desk research, and/or self-assessment, and/or semi-structured interviews) are used, assigning the final score that is entered in the web form, should be done by partner responsible for a country (in cooperation with, or taking into account feedback from contact person).

The approach should be as follows:

- web form should be seen as a placeholder for numeric score, wherever supporting evidence comes from;
- in case different data sources or data collection methods are used:
- the web form should already contain the final, aggregated score acquired by different data collection methods;
- aggregation should be done by the partner responsible for the country in cooperation with the contact person

- final score should be entered in the form by the partner responsible for the country;
- any unstructured documentary or verbal supporting evidence should be summarized in a for that purpose reserved placeholder - a free text field in the web form (in case of aggregation, filled in the web form by the partner responsible for the country together with the explanation of the aggregation method);
- other supporting material - full text documents, interview summaries etc. - would be kept and organised following **Common recommendations and scenarios template for Data management planning** (see more in section E) and managed in the Resource management section of the web form.

B. Protocol for selection of interviewees and communication

This section addresses the matter of choosing and contacting the appropriate respondents for each of 3.2 data collection instruments: the self-assessment survey and the semi-structured interview.

Self-assessment survey

Selection protocol

Within CESSDA members (type I countries), the self-assessment survey will be filled in by the service providers (SPs).

In non-CESSDA member countries, where there are already established data services (type II countries) or grassroots efforts to establish data services (type III countries), i.e. relevant partners of CESSDA SaW, SERSCIDA and SEEDS projects, representatives of each SP/DAS will also fill in the self-assessment survey.

The selection of the person(s) answering the survey is decided internally in each service and will be explained in a [SP contact letter](#).

Communication protocol

For CESSDA SaW 3.2 partners, no special communication protocol is needed. The 3.2 partners are in charge of finding appropriate respondent(s) in their own institution. If needed, they can use (part of) the email sent to the SPs (cf. next paragraph).

For type I, II and III countries that are not CESSDA SaW 3.2 partners (about 20 countries), the protocol of communication is different. The initial contact (email) should include a brief description of the CESSDA SaW project and CESSDA (only for types II and III), as well as task 3.2 goals to achieve with the data collection, and the optimal profile or characteristics of respondents. The 3.2 task leader will draft the template text of this first email. The email should be sent to the director of the SP or the contact person for CESSDA SaW if any is mentioned in the [CESSDA SaW contact list](#).

Follow-up contacts can be less formal, by email or by phone. Each contact should be documented in the [CESSDA SaW contact list](#).

Semi-structured interview

Selection protocol

The semi-structured interview will be conducted with informants who have a particularly good view of the social science research community in their respective countries. As these semi-structured interviews will be conducted either to complete the information gathered with the desk research and the self-assessment survey (in types I, II and III countries) or to gather information about type IV countries, two distinct strategies of interviewee selection apply.

In the first case (types I, II and III countries), it might be that the desk research and self-assessments will not be able to address all points in the collection instrument, thus necessitating semi-structured interviews. In these cases, 3.2 partners may ask the SPs to indicate the most appropriate organisation or person to contact for an interview.

In the second case (type IV countries, and if needed type III countries), the experts may be identified during the desk research. Also, 3.2 partners can ask different organisations (OpenAIRE national open access desks, faculties/research institutes in the social sciences, funding institutions, ministry/state agencies in charge of social sciences) to give the contact information of 1 or 2 experts from within the social science research community.

The expert, someone recognised for her/his important role in the national social science landscape could be a:

- Senior researcher in the social sciences working in a public research institution in the social sciences;
- Dean or her/his representative of a faculty of social sciences;
- Representative of the main funding institution in the social sciences and/or the open data movement (e.g., research councils or national science foundations);
- Representative of the ministry or state agency in charge of research in the social sciences.

Communication protocol

Contacts can be made by email/phone by CESSDA SaW 3.2 partners. The initial contact (by email) should include a brief description of the CESSDA SaW project, of CESSDA, and task 3.2 goals to achieve with the data collection and, if no name has been given by the SPs or desk research, the ideal profile or characteristics of the interviewee (i.e. expert in the social sciences... as described above). A template text of this first email is in [section F](#). The email should be sent by the 3.2 partner responsible for data collection in the country. Follow-up contacts can be less formal. Each contact should be documented in the [CESSDA SaW contact list](#).

C. Guidelines, suggestions and remarks regarding use of data collection methods

This section provides some suggestions on how to conduct the data collection process with each of the methods.

Desk research includes both a review of official sources (like policy documents, statistical data) as well as a literature review (articles, reports and similar), and should be limited to the country in question, if not indicated otherwise.

Data from desk research is entered in the web form and, if necessary, combined with self-assessment and semi-structured interviews for the final score by the partner responsible for the country.

Desk research procedure

For desk research, the work flow should be as follows:

- where it is indicated in the web form that desk research should/could be used (Partners' Content Area of the web form or instructions for self-assessment, accessible with button "Show partners' content in the web form"), the partner responsible for the country in cooperation with the contact person for the country decides which documents are relevant;
- partner responsible for the country (or if agreed, the contact person for the country) assigns a score based on the contents of the document(s);
- the score resulting from desk research, if needed, can later be combined (see section A, Data aggregation from multiple sources) with data from self-assessment or semi-structured interviews;
- references to the relevant document(s) and aggregation method should be provided in Resource management section of the web form;
- the final scores entered in the web form, if possible, should be assigned in agreement with the contact person for the country (SP/DAS representative or similar).

Some indicators are compiled collectively for all countries by one partner. In this case the partner responsible for data collection in the country might be asked to assist. Scores will appear in items in the Partners' Content Area in the web form.

Semi-structured interview process

For semi-structured interviews with experts/stakeholders:

- partner responsible for the country together with the contact person for country/SP/DAS (if available), identifies and selects relevant stakeholders (according to the Protocol for selection of interviewees), and makes the first contact according to the communication protocol (see sections B and F for more detail);
- following the principle of resource allocation, after evaluating potential use of other data sources and data collection methods (see section A, Planning data collection for more detail), the semi-structured interview guide for the stakeholder is produced by partner responsible for the country together with the contact person for country/SP/DAS;
- the process of tailoring semi-structured interview guidelines should be as follows:
- partner responsible for the country together with the contact person, following respondent selection guidelines, identify relevant interviewees (policy makers, researchers, OpenAire and other OA people, etc.; see also section on selection of interviewees and communication protocol);
- potential interviewees are contacted, and an appropriate mode of interview data collection event (personal, Skype, telephone, e-mail interview etc.) is agreed upon;
- the interview guidelines, stating main themes and question blocks, relevant issues, and questions are created by the partner responsible for the country, adapting interview guidelines from section D (Interview Guide for semi-structured interviews with experts), by for example, taking out questions that are not relevant to the interviewees expertise, or to the country or to the limitations of resources - using the semi structured interview only when the two other methods do not work (thus limiting number of covered blocks of information, number of questions).
- interviews can be conducted in person, via Skype or similar, or in some cases - even as written interviews via e-mail,

- when needed, could be recorded; audio recording is recommended mostly in face-to-face interviews with nobody taking notes; the recording should be seen as a source for preparing the written interview summary;
- the preferred interview language is English, short interview summary of answers to the interview questions in one document should also be in English; (see section D);
- the summary should be sent to the interviewee for review and accept (if agreed), and, if needed, clarification.

Interview results (summaries) should be used to assign values; the scores (in combination with scores from other sources, if used) are entered in the web form by the partner responsible for the country, with short explanation about aggregation method in a notes space following the questions in the web form;

*Example. The indicator 3.2.1 in the web form, Partners' Content Area, measures incentives for data sharing by desk research in WoS and Scopus databases and in semi-structured expert interviews with the question **How common it is for researchers in your country in your research area to publish in journals that expect data used in the publication to be available for reuse from a trusted digital repository, and why?** Answers to the question in a summarised form should be found in the interview summary. Based on the summary, the score "Very common", "Rather common", "Not that common" or "Not common at all" should be assigned. Interview results should be combined with desk research results to be collected collectively for all countries. The interview summary should be referred to in the field for comments/notes, and uploaded in the Resource management section. The desk research score should also appear in the notes as well as aggregation method, the mean value between the desk research result and interview score. The final score is reviewed by the contact person of the country and adjusted by the partner responsible for the country, if needed.*

Self-assessment procedure

The self-assessment is initiated by the person responsible for data collection the country, and done and/or coordinated (in case there are several people filling it in the institution) by the contact person.

The process should, in general, go like this:

- a link to a web form where self-assessment results should be entered, should be sent to contact person for the country/SP/DAS together with a letter describing the data collection process as well as the approximate deadline (see respondent selection and communication protocol, section B and template contact letter, section F);
- the results of the relevant parts of self-assessment are filled in the web form by the contact person and/or other respondents in the country/SP/DAS, based on experience and practices in the institution/country, and, documents and reports describing these;
- responsible partner for the country monitors the self-assessment process, answers questions and reviews the self-assessment data entered in the web form;
- first step review – check, if all the necessary information is there and if it is consistent; if necessary, clarifications and follow-up interviews are done, and self-assessment scores changed in the web form;
- second step review – in the process of aggregating self-assessment information with data from other sources, some scores might be changed;
- the final scores entered in the web form should be assigned in agreement with the contact person for the country (SP/DAS representative or similar; if possible).

The time and resources to be put into self-assessment should be negotiated by the partner responsible for data collection and contact person, and, if needed, the task leader (see section A).

Navigation of Limesurvey web form

Limesurvey web form serves as a database for data collected via self-assessment, desk research and semi-structured interviews - scores, references, supporting evidence any additional notes and comments etc.

The central part of the web form is the self-assessment area that is accessible for country representatives having a link specific to the country (sent to the contact person for the country).

There are some differences in the scope of the self-assessment area between the countries. For example, type III and IV countries self-assessment would generally cover I1.-I4. sections of the web form, while for type I and II countries also II1.-III1. sections.

Navigation between question blocks / sections in the self-assessment area is possible

- by using button “**Question index**”, opening a drop-down menu on the upper right corner, or
- by using buttons “**Previous**” and “**Next**” at the bottom of the page to get to the previous or next question block/ section.

Answers can be saved by clicking the button “**Resume later**” in the top right corner or at the bottom of the page, and retrieved when restarting work by clicking the button “**Load unfinished survey**”.

Partners’ Content Area is mainly for the use of the partner responsible for data collection in the country. It can be accessed by grey-coloured button on the top right corner, “**Show partners content**”, and exited by clicking the same button in red, “**Hide partners content**”.

In addition to self-assessment items, it includes data items/indicators that are collected by desk research and/or semi-structured interviews. Desk research data are sometimes collected separately for each country and sometimes collectively for all countries. In such cases, the scores are visible here.

There are also cases when self-assessment data should be combined with semi-structured interview or desk research data. A final score is entered in Partners’ Content Area by the person responsible for data collection. The score should be decided in agreement with the contact person for the country, and it should follow the aggregation process described in the notes/comments area (see more details on aggregation in Section A, Data aggregation from multiple sources). For this task, contact persons can access Partners’ Content Area by clicking the “**Show partners content**” button.

In the Partners’ Content area of the web form, scores are entered by the contact person for the country and/or respondents (if none, by the person responsible for data collection).

The **Resources management area** is seen and used by partners responsible for data collection in the country to describe and manage other supporting material - full text documents, interview summaries etc. (for more detail, see section E, Common recommendations and scenarios template for Data management planning).

D. Interview Guide for semi-structured interviews with experts: content areas, indicator/related question blocks and issues to be addressed

Introduction

Focus: relevant areas of expertise

Remarks for interviewer: At the beginning of the interview, the interviewer should inform the interviewee that the answers to these questions should cover whole SSH sector in his/her country and that he/she should answer the questions from his/her knowledge or experience.

Themes:

interviewee's background/ areas of expertise, academic/other working position;

academic and related achievements;

research experience;

knowledge/ experience of RIs

Block A: Development of social sciences sector in the country

Focus:

Funding capacities, human resources and infrastructure, international collaboration and national studies as driver of data production in the country.

Themes:

Estimation of funding characteristics and size of the SSH researchers' community

major public and private funders for research in SSH

state (funding agencies) support regarding access to commercial databases with scientific papers or data

Development of research data production in SSH in the country

existence of national studies as an incentive for data production

volume and frequency of data production in SSH in the country

assessment of quality of produced data in SSH in the country

Questions:

Estimation of size and other characteristics of the SSH researchers' community

How would you **rank the sources (sectors) of research funding in SSH in your country?** (Government sector, Business enterprise sector, Abroad, Private non-profit sector, Higher education sector). Could you assess the share of each source of funding?

Remark: Private non-profit sector includes civil non-market organization like trade unions, consumers' associations, charity institutions etc. Further discussion about models of funding and expectations in the future should be encouraged.

Does your government or state funding agency **support access to commercial databases with scientific papers (i.e. EBSCO, JSTOR) or datasets (i.e. Bankscope) and software purchase?**

Remark: The emphasis here is on the support to complete scientific community rather than institutions. Also, the term "software" here is related to specialized applicative software for scientific analysis.

Do the **public policy makers in your country make decisions based on evidence?**

Remark: Interviewee should be encouraged to give additional remarks about attitudes of government authorities toward the usefulness of the academic research for the policy decision making.

Development of research data production in SSH in the country

Are there **existing studies of national importance produces by SSH research** in your country? Can you elaborate on this (number and frequency of studies)?

Remark: The phrase "studies of national importance" relates to data-producing studies that systematically assess conditions and public opinion on the matters of national importance, like public opinion surveys or election surveys.

How would you characterize **the average production of research data by the SSH institutions in your country?** (Rare - ad hoc, Periodically or Frequently)

Remark: It should be clarified to the interviewee that research data here encompass all SSH data produced, including NGOs (not only by academic institutions).

Block B: RDM policy and support setting

Focus:

Policies for data documentation and management facilitating data sharing and ethical and legal framework.

Themes:

Explore overarching strategy and policy to enable sustainable data access and sharing of publicly financed SSH data

general situation regarding requirements or recommendations concerning the preparation of a Data Management Plan (DMP)

research funding organization's requirements or recommendations about quality-assured social science research data with associated metadata

RDM policy requirements: sustainability and long term curation

incentives for data sharing: covering costs in the country for managing the data resourced

Ethical and Legal framework important for data sharing

general situation in a country with regards to clarification and support provided on legal and ethical aspects that facilitate social science data sharing

Questions:

Explore overarching strategy and policy to enable sustainable data access and sharing of publicly financed SSH data

How would you characterize the general situation with regards to **requirements or recommendations regarding the preparation of Data Management Plans (DMP)** as an integral part of funding applications and of on-going project activity? **Is it** aiming at Open data as the default among public funders of SSH research in your country?

To what extent have public **research funding organizations operating in your country issued requirements or recommendations about quality-assured SSH research data with associated metadata?** Do they require the researcher to offer or deposit the data in an appropriate disciplinary repository?

Remark: For the countries of the 3rd or 4th group, where expected answers are low or initial, it may be better to have questions 1) and 2) separated into two parts, to avoid possible issue of double-barrelled answers

To what extent have **public research funding organizations** operating in SSH in your country **issued requirements or recommendations** or **show awareness about the long-term curation for valuable research data assets**, evaluated and selected regarding reuse potential.

Remark: If the answer cannot be specified in detail, the interviewer should proceed with the following question: Is there an intention or awareness to address research data management and sharing in any form and on any level of public funder policy or strategy or principles declaration? Please describe and refer to specific examples!

Overall, do the **public research funding organizations** operating in your country **cover costs for managing and preparing data, as incentives for sharing research data with associated metadata?**

Ethical and Legal framework important for data sharing

How would you characterize the **general situation in your country** with regard to **clarification and support provided on legal and ethical aspects** that facilitate SSH data sharing (IPR, data protection...).

Block C: Data sharing culture

Focus:

Behaviour/ practices, attitudes and perceived barriers and incentives for data sharing, RDM support and practices

Themes

Data sharing and access to data

Attitudes towards sharing and reuse of data

Enablers for data sharing

Data support services and RDM practices

Questions

Data sharing and access to data

How common is it for SSH researchers in your country to **share the research data they produce**? Can you assess the **proportion of SSH researchers sharing the research data they produce**? Are there any **differences between research areas**?

How many researchers in your country **would be interested in re-use of third party data**? Which **proportion of SSH researchers can access the third-party data they need**? Are there any **differences between research areas**?

In your country, **which of the four routines for data sharing** - data archive or repository, supplementary data in a journal (alongside paper), project or personal websites, informal contacts (peers and colleagues) - **is the most common / frequent in the SSH, and why**? Can you rank them? **Are there any other ways to share research data in your research community/area**?

In your country, which of the **four routines for accessing third party data for reuse** - data archive or repository, supplementary data in a journal (alongside paper), project or personal websites, informal

contacts (peers and colleagues) - is the most common / frequent in the SSH, and why? **Are there any other ways to access third party research data in your research community/area?**

Attitudes towards sharing and reuse of data

How well do the following statements describe the attitudes to data sharing in SSH research community in your country? What are the prevalent attitudes in your research community regarding the benefits for data sharing?

Is the general attitude that **data sharing promotes healthy competition, or rather negative competition?**

What about **the data sharing and its impact on advancement of science?**

What about **data sharing and risk that others may misuse and misinterpret the data?**

What about the **effort required for data sharing, and its costs for the researchers?**

Enablers for data sharing

How common it is for researchers in your country to publish in journals that expect data used in the publication to be available for reuse from a trusted digital repository?

In your country and within the SSH research discipline(s), could you say whether there are **careers rewards related to data sharing?** If yes

What influence, if any, may sharing research data have on career progression within institution or community or due to government rules?

What about **the success rate in obtaining research funding?**

What about **better standing within the research community and other.**

Data support services and RDM practices

What cooperative and shared **data support services that facilitate data sharing and OA** to research data are available in the in your country in different stages of research project?

Is there support (e.g. **trainings, workshops, webinars, online reference materials, help desk or other consultancy, etc.**) given to researchers from libraries or other institutions regarding, for example, **data management plans, data preservation, and data access?** If yes, then can you describe them?

In your country, how do researchers in your research area **manage and document the data to facilitate data reuse?**

What **standards do they use, and what procedures in data management they follow** to facilitate data reuse?

Block D: Data infrastructure (where no formal DAS exists)

Focus:

Assessment of data archive proto-activities and open access support activities in countries where no formal DAS exists yet.

Themes

Institutionalisation of DAS proto-activities

availability of technical infrastructure

existence of organisational activities
availability of capacity building and training
Open access support activities
availability of open access (OA) projects or initiatives
cooperation with OpenAIRE National Open Access Desks (NOAD)

Questions

Institutionalisation of DAS proto-activities

Are there **existing technical infrastructures** (such as **repositories, online tools, databases, or online catalogues**, etc.) in your country **that could possibly be used for or applied to a new data archive service (DAS)**? Also, is it planned to use them for a future DAS?

Are there **any activities in your country towards establishing a DAS** for the social sciences?

Are **there institutions that could host a DAS in your country**?

Are there existing **initiatives in your country to develop the knowledge and skills of people who might at some point be employed at a DAS**? This could include: research and data management (mainly data management plan), data preservation, and data access.

Open access support activities

Are there **open access projects or initiatives in your country, either funded by the government or by grassroots**?

How could a future DAS for social sciences cooperate **with OpenAIRE National Open Access Desks (NOAD)**?

E. Recommendations for planning and documenting data collection and data management

These recommendations sum up data collection and data management approach for planning and reporting aims. This will help coordinating the activities on a country level between different roles and help establishing the monitoring of data collection, while also documenting the information generating process and outcome.

The planning and documenting data collection and data management consists of three steps:

1. Data collection planning, where the partner responsible for country should plan the methods to be used and experts to be contacted, and can seek advice from the Task leader, if needed. For data collection planning we use the [CESSDA SaW contact list](#), mainly columns "Purpose of the contact" and "Notes".
2. Documenting of the completion of data collection plan is meant to be filled in for each person or documentary resource, who or which will contribute information. This is part of Resources management area of the web form, and includes references to data sources (in some cases for desk research - full text documents or, for semi-structured interviews - interview summaries).
3. Recording of the notes/comments on indicator level includes detailed documentation of resources (see point 2.) at the level of questions/indicators notes. Each comment, that will summarise the decision about the indicator values, will include a reference to any of the resource (personal or desk research), mentioned in 3. This is implemented at the web form section level, where the resources from the second part can be selected, followed by the free text comment at the level of questions/indicators notes.

1. Data collection planning and resource management

Timeline and scope of data collection should be determined and agreed upon at the beginning of data collection, most preferably between partner responsible for the country and the contact person in the country. The Task leader will be informed about the draft plan from the CESSDA SaW contact list information, provided by partner, and can assist in case of difficulties (e.g. can allocate additional human resources, or support decisions about alternative methods usage). Partner responsible for the country should establish a plan about whom to select and then contact for self-assessment or interview in a the [CESSDA SaW contact list](#) as described which main source will be used for desk research. Document resources that are of a common interest should be documented in the common [Existing studies review](#). This includes documents that can be referred for final all countries report or that can serve as example for other partners. Actual use of all desktop resources will be documented in the steps 2 and 3 of documenting and recording of evidence.

Information about persons participating needs to be included in the [CESSDA SaW contact list](#) template for each person (be it contact person, respondent or interviewee). If something changes in planning, and new interviewees or additional contact persons are planned, this should be added to a list of contacts.

Each main contact person that we attempt to contact should be listed at the Main contact persons' section of a CESSDA SaW contact list. For planning purposes, each contact should first be assigned to the 'Planned' in a 'Status of contact' of the Follow-up section of the contact list, with a 'Date planned' when the contact will be attempted. Specific CESSDA SaW 'Task 3.2' reference should be made at the 'Purpose of contact', followed by the role (Contact person, Interviewee, Respondent). Notes could be used for any additional information regarding a contact, e.g. which section a Respondent or Interviewee would be assigned, a timeline of contact attempts if more than one, until the final status of a contact is reached (either 'Answered', 'Transferred' or 'No answer'), visible in a Status of contact.

If contact-specific consultation is needed with the Task leader and Task force members, this should be described in the Notes column, followed with the email to the Task leader as a contact, to initiate the discussion. Agreed upon conclusions from consultation can be added to the timeline of contact in the Notes.

Be aware of the limited resources and plan the data collection accordingly, following the principles stated in other sections of the Guidance! The person responsible for the collection of data in a country can consult with the task leader (together with task force) with regard to use of resources for data collection.

2. Documenting data collection and recording of answers

For each person and method, a combination of accessing and adding information to the web form (MS6 Instrument), a record should be established about the person entering the information (e.g. Partner responsible for country in case of Desk research and Interviewing, else it is Contact person or person assigned by him/her to fill part of self-assessment template). In addition, the role and method should be visible, of the information that will be entered, and in case of interview or desk research, who provided information or which main resource was used. For each method/resource (written or in person) combination a separate record needs to be established.

This information should allow the estimate of the workload that each partner spend in data collection activities.

(1) Person entering the information (Name)

(2) **Short description of the reference**

(3) Method (Options: Desk research/ Semi-structured interview) (only one can be selected; for each method and source the whole template set should be repeated as a separate set of information)

(4) A) Interview: **Contacted person / interviewee name:**

(Details for each contact should be documented in the [CESSDA SaW contact list](#).)

If Desk research B) **Please upload files used as a reference for data input:**

If Interview C) Link to, or name of interview summary file

Note: Original recordings of interviews text or original documents used are to be stored in a Lime survey system for review process. The files will be uploaded in a step (4).

(Repeatable for each method / contact or main source)

This is part of Resources management area of the web form.

3. Recording of the notes/comments following each of the indicators

For each note (or section) there should be a description about who entered a note and in what capacity/role. This is needed for communication reasons, as different persons could write notes. It is also needed for documentation of the aggregation process (e.g. reasoning behind assigning codes in a self-assessment template) and to collect feedback from participants.

(1) Name (Partner responsible for the data collection in the country; contact person for the country/ service provider/DAS)

(2) Method (Desk research/ Self-assessment / Self-assessment monitoring / Self-assessment follow-up / Semi-structured interview)

If Interview or Desk research: reference to the source used from 2.

(3) Date

(4) Short summary of desk research results or interview answers, discussion about triangulation of results if different methods used, and explanation about how indicator values were derived.

Information should be entered in the free text field related to the indicator / section, and should include this information, if applicable (but it does not necessarily have to be structured in this way).

F. The initial contact e-mail template

We can use official CESSDA look for a letter available [here](#).

1 A motivation for collaboration

1.1 Self-assessment survey // Communication protocol

Instructions: The protocol of communication is for type I, II and III countries that are not CESSDA SaW 3.2 partners:

The initial contact (email) should include a brief description of the CESSDA SaW project and CESSDA (only for types II and III), as well as task 3.2 goals to achieve with the data collection, and the optimal profile or characteristics of respondents.

The email should be sent to the director of the SP or the contact person for CESSDA SaW if any is mentioned in the [CESSDA SaW contact list](#).

1.1.1 Service Providers: directors

Dear colleague

We kindly invite you as a director of a European research data service (Service provider) to take part in our survey. This study is carried out by CESSDA AS and its partners in the CESSDA SaW project, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement 674939. The aim of CESSDA SAW project is to strengthen and widen CESSDA as the European infrastructure for social science data archives. Brief descriptions of the CESSDA, CESSDA SaW project, and the purpose of the survey are provided further below.

In case you don't have time or are not in the position to respond, please, let us know as soon as possible who we may contact in your organisation regarding further details of the data collection.

Participation is voluntary and you may withdraw at any time. By participating you agree that the information provided will be made accessible through publications and reports of the CESSDA SaW project and beyond.

1.1.2 CESSDA SaW contact person

Dear colleague

We kindly invite you as a contact person for a research data service (Service provider) to take part in our survey. This study is carried out by CESSDA AS and its partners in the CESSDA SaW project, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement 674939. The aim of CESSDA SAW project is to strengthen and widen CESSDA as the European infrastructure for social science data archives. Brief descriptions of the CESSDA, CESSDA SaW project, and the purpose of the survey are provided further below.

As a contact person, we would like to ask you to:

- (a) Assist with gathering information on relevant contacts and documents that we need for desk research;
- (b) Coordinate that appropriate people in the institution(s) **answer the questions in the online self-assessment survey (you may see print version of the survey questions [attached](#))**;
- (c) Assist with selecting and contacting the relevant person(s) for follow-up interviews.

As partners, responsible for the country we will assist and monitor the process of answering the questions, and if needed for clarification, ask follow-up interviewing questions.

Participation is voluntary and you may withdraw at any time. By participating you agree that the information provided will be made accessible through publications and reports of the CESSDA SaW project and beyond.

1.2 Semi-structured interview// Communication protocol

*Instructions: As these semi-structured interviews, will be conducted either to complement the information gathered with the desk research and the self-assessment survey (in types I, II and III countries) or to gather information about type IV countries, **two distinct strategies** of interviewee selection apply.*

1. types I, II and III countries: 3.2 partners may ask the SPs to indicate the most appropriate organisation or person to contact for an interview

2. type IV countries (and if needed type III countries): the experts may be identified during the desk research. Also, 3.2 partners can ask different organisations (OpenAIRE national open access desks, faculties/research institutes in the social sciences, funding institutions, ministry/state agencies in charge of social sciences) to give the contact information of 1 or 2 experts from within the social science research community.

The expert, someone recognised for her/his important role in the national social science landscape could be a:

- Senior researcher in the social sciences working in a public research institution in the social sciences;
- Dean or her/his representative of a faculty of social sciences;
- Representative of the main funding institution in the social sciences and/or the open data movement (i.e. research councils or national science foundations);
- Representative of the ministry or state agency in charge of research in the social sciences.

Contacts can be made by email/phone by CESSDA SaW 3.2 partners. The initial contact (by email) should include a brief description of the CESSDA SaW project, of CESSDA, and task 3.2 goals to achieve with the data collection and, if no name has been given by the SPs or desk research, the ideal profile or characteristics of the interviewee (i.e. expert in the social sciences... as described above). The 3.2 task leader will draft the template text of this first email, but the email should be sent by the 3.2 partner in charge of the country.

1.2.1 Type I, II and III countries

Dear Mrs or Mr

We recognize that you play an important role in the social science landscape in your country on the national level. Therefore we kindly ask you to take part in an interview, which aims to complete the information we have already gathered with desk research and a self-assessment survey, completed by the national data archive service contact. This study is carried out by CESSDA AS and its partners in the CESSDA SaW project, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement 674939. The aim of CESSDA SAW project is to strengthen and widen CESSDA as the European infrastructure for social science data archives. Brief descriptions of the CESSDA, CESSDA SaW project, and the purpose of the survey are provided further below.

Participation is voluntary and you may withdraw at any time. By participating you agree that the information provided will be made accessible through publications and reports of the CESSDA SaW project and beyond.

If you don't have time or believe that we should contact another expert or your colleague, please let us know as soon as possible and suggest who we may contact instead.

Please confirm your participation by responding to this email, after which we can arrange the interview. We would appreciate your response by date.

1.2.2 Type IV: See above 1.2.1 Type I, II and III countries and adapt

2 A brief description of CESSDA (only for types II and III)

CESSDA (Consortium of European Social Science Data Archives - see more on <http://cessda.net/>) is a pan-European Research Infrastructure that coordinates activities of social science data archives across Europe. The aim is to promote the results of social science research, re-use of research data, and thereby to support national and international research and cooperation. CESSDA achieves its goals collectively by its members facilitating researchers' access to important data resources of relevance to the European social science research agenda regardless of the location of either researcher or data.

Currently CESSDA membership consists of 15 national state members and 1 observer. The aim is to achieve full European coverage, to strengthen the network and to ensure sustainability of its data for the widened network.

3 A brief description of the CESSDA SaW project

The CESSDA SaW (Strengthening and widening) project examines the barriers to and the potential value and benefits from membership in CESSDA in each European country.

One of the approaches used by the CESSDA SaW project is a bottom-up analysis of current prospects in European Research Area (ERA) countries of becoming full CESSDA members. It does so by inquiring into the situation in the country regarding support toward organising a sustainable research data infrastructure service on a country level. This is an aim of our survey/interview.

The findings will serve as a starting point for further development planning, and the results will be used to inform external stakeholders on necessary steps to be taken to achieve a national data service sustainability, and to promote CESSDA membership.

Based on analysis from the bottom-up phase, a top-down approach will follow that will provide advice and support for initialising and further developing National Data Archive Service organisations in the ERA.

5 Conclusions

So, let us once again invite you to take part in the survey/interview, which will help us to gain a better overview of the situation in different countries, and will help with organising further steps to achieve the goals of CESSDA as a true pan-European Research Infrastructure.

For any further information please contact me as CESSDA SaW partner responsible for country data collection!

CESSDA SaW partner name

CESSDA SaW Partner Institution + phone, email...

Appendix 3: Questionnaire text extract from the Web form

CESSDA SaW Survey

CESSDA SaW (Strengthening and Widening) Project is a 2-year project funded by European Union's Horizon 2020 research and innovation programme under grant agreement number 674939.

The **CESSDA SaW (Strengthening and widening) project** examines the barriers to and the potential value and benefits from membership in CESSDA in each European country.

The findings will serve as a starting point for further development planning, and the results will be used to inform external stakeholders on necessary steps to be taken to achieve a national data service sustainability, and to promote CESSDA membership.

We kindly invite you to take part in our survey. Participation is voluntary and you may withdraw at any time. By participating you agree that the information provided will be made accessible through publications and reports of the CESSDA SaW project and beyond.

Navigation of Limesurvey web form

Navigation between question blocks / sections in the self-assessment area is possible:

by using button "**Question index**", opening a drop-down menu on the upper right corner, or

by using buttons "**Previous**" and "**Next**" at the bottom of the page to get to the previous or next question block/ section.

Answers can be saved by clicking the button "**Resume later**" in the top right corner, and retrieved when restarting work by clicking the same link (URL) previously provided.

Partners' Content Area is mainly for the use of the partner responsible for data collection in the country. It can be accessed by **grey-coloured button** on the top right corner, "**Show partners content**", and exited by clicking the same button in red, "**Hide partners content**".

Recording of the notes/comments following each of the indicators

For each Note, there should be a description about who entered a note and in what capacity/role. This is needed for communication reasons, as different persons could write notes.

The following convention is recommended:

- | | | | | | | | | |
|-----|--------------------------------|-------|-----------|----|-----|--------|--|-------|
| (1) | | | | | | | | Name |
| (2) | Purpose | (e.g. | reference | to | the | source | | used) |
| (3) | | | | | | | | Date |
| (4) | Note content in a summary form | | | | | | | |

Information should be entered in the free text field related to the indicator / section, and should include this information, if applicable (but it does not necessarily have to be structured in this way).

1.1. Broader ecosystem of DAS operation: Structural conditions

The aim of this section is to assess overall development of the social sciences and research data production as an important determinant of the quantity and quality of the research data produced and potential of data reuse. Focus is mostly on the issues of financial stability, research capacities and results achieved.

Name of the person entering information:

1.1. Development of social science sector

Financial stability, research capacities and results achieved in the field of social sciences in a country.

1.1.5. How would you rank the sources (sectors) of research funding in social science in your country by the amount they provide?

All your answers must be different and you must rank in order.

Please number each box in order of preference from 1 to 4

Government and Higher education sector

Business enterprise sector

Abroad (international or cross-border)

Private non-profit sector

Eurostat data for total GERD structure of funding are collected centrally for all countries and are available [here]. Please consult other data sources as well, such as reports for previous projects or official EU and government reports.

Notes/comments:

1.1.6. Does the government or university provide access to commercial bibliographic and full-text databases (i.e. WoS, Scopus, EBSCO, JSTOR, ...), datasets (i.e. Bankscope) and/or national software licences?

Please choose **all** that apply:

bibliographic and full-text databases

datasets

software licences

none of the above

1.1.7. Number of Web of Science (WoS) Core Collection publications in SSH per researcher.

Only numbers may be entered in this field

Data are collected centrally for all countries and are available [here].

1.1.8. In your opinion, do the public policy-makers in your country make decisions based on evidence?

Please choose **only one** of the following:

absolutely yes

yes in most cases

only partially
no in most cases
absolutely no

Make a comment on your choice here:

Please give your opinion based on experience and policy making practice. If possible, opinion should be underpinned with some relevant documents.

Enter documents URL-s and/or make a comment on your choice here:

All in all, what would be your overall assessment of financial stability, research capacities and results achieved in the social sciences for this country?

Please choose **only one** of the following:

LOW LEVEL - Funding of social sciences and productivity of the researchers are among the lowest quantiles. Impact on designated community is small or non-existing.

MEDIUM LEVEL - Funding of social sciences and productivity of the researchers are among the mid quantiles. Impact on designated community is limited.

HIGH LEVEL - Funding of social sciences and productivity of the researchers are among the highest quantiles. Impact on designated community is strong.

Make a comment on your choice here:

1.2. Research data production in SSH

Prevalence of high quality research data with high potential for reuse.

1.2.2. Are there any existing studies that systematically assess matters of national importance (public opinion survey, election survey, etc.) produced by social science researchers in your country?

Please choose **only one** of the following:

yes
no
other

Make a comment on your choice here:

Please list and briefly describe existing studies in the comments section. Give your answer based on your own experience and published reports if available (please provide references to these reports).

1.2.3. In your opinion, how would you characterize the average production of research data by the social science institutions in your country?

Please choose **only one** of the following:

Rare production (data are produced ad hoc)

Periodical production (institutions have tradition in producing some type of research data to a certain extent)

Frequent production (institutions have well established tradition in data production)

Make a comment on your choice here:

Please give your answer based on experience and published reports (if available).

I.2. Broader ecosystem of Data archive services (DAS) operation: RDM Policy setting

The aim of this section is to explore overarching strategy and policy to enable sustainable data access and sharing of publicly financed social science research data. This can range from declared awareness about principles and soft recommendations to explicit requirements, the fulfilment of which is actively supported and rewarded. Data archive services (DAS) can have a recognized and important support role in such an environment, which you can refer to in comments to answers and questions.

Name of the person entering information:

2.1. Funders' data management and sharing strategy and/or policy

Overarching strategy and policy to enable sustainable data access and sharing of publicly funded social science research data.

2.1.1. How would you characterize the general situation with regards to requirements or recommendations about preparing Data Management Plans (DMPs) as an integral part of on-going project activity, aiming at Open data as the default among public funders of social science research in your country?

The requirements or recommendations for DMPs in most cases (funding authorities, project calls) are:

Please choose **only one** of the following:

None

Initial: There is growing recognition and awareness of need to require DMP

Partial: There is the expectation or recommendation to have DMP in place

Defined: Formal requirement, little monitoring and support

Managed: DMP is a requirement, clear guidance is issued, support and tools are provided, the content of DMP and exemptions from full open access are defined

Cannot answer: Why?

Make a comment on your choice here:

2.1.2. Overall, in your experience, how thoroughgoing it is that the public research funding organizations operating in your country have issued requirements or recommendations about quality-assured social science research data with associated metadata?

Offering or depositing data in an appropriate disciplinary repository.

Please choose **only one** of the following:

None

Initial: There is growing recognition and awareness of the need to have disciplinary specific place of deposit and support services

Partial: There is the expectation or recommendation to offer or deposit data in an appropriate disciplinary repository or equivalent data archive service

Defined: Formal requirement, little or no monitoring

Managed: Formal requirement, sanctions for not complying with regulation are in place (such as reduced payment, etc.), full support and guidance is provided

Can't answer: Why?

Make a comment on your choice here:

See recommendation from the EAGDA Report, "clear expectations for study leaders on the use of established repositories with archiving facilities, in order to assure quality and ensure discoverability of data" - <https://wellcome.ac.uk/sites/default/files/governance-of-data-access-eagda-jun15.pdf>

2.1.3. How thoroughgoing it is that the public research funding organizations operating in social sciences in your country have issued the requirements or recommendations or show awareness about the following aspects of social science research data with associated metadata?

Long-term curation for valuable research data assets, evaluated and selected in terms of reuse potential.

Please choose **only one** of the following:

None

Initial: There is growing recognition and awareness about the value of research data produced and about the need for long-term preservation; scarce or no investment and support for long-term curation provided.

Partial: It is expected or recommended to assess the value of research data and resources providers declared their motivations for continuing to invest in sustaining the assets. The FAIR data principle (findable, accessible, interoperable and reusable) are highlighted in these recommendations.

Formal: Requirement to assess research data appropriate for data curation. It is understood that the best use of resources involves making choices based on value judgements and selecting material for curation. Investment and support for long-term curation is in place, based on contractual arrangements.

Managed: Formal requirement, based on contractual collective arrangements of roles and responsibilities among different stakeholders, clear definitions of what data is preserved, how it is documented, and for how long. Funder requires that periodic monitoring is performed if the best use of resources is made for expected amount of benefit.

Can't answer: Why?

Make a comment on your choice here:

2.1.4. Overall, in your experience, do the public research funding organizations operating in your country provide the following incentives for sharing research data with associated metadata?

Cost for managing the data and preparing it for access are resourced adequately during research project lifetime.

Please choose **only one** of the following:

No

Initial: Cost for managing the data and preparing it for access can be implicitly covered in the overall research project budget

Partial: There is explicit recognition that additional cost for preparing the data for access are legitimate project cost that can partially cover the RDM cost up to a certain limit

Defined: Costs for RDM are fully covered and adequate, based on DMP plan in project documentation

Can't answer

Make a comment on your choice here:

2.2. Legal and ethical framework

A description of the ethical and legal framework important for data sharing.

2.2.1. How would you characterize the general situation in your country with regards to clarification and support provided on legal and ethical aspects that facilitate social science data sharing (IPR, data protection..)?

Please choose **only one** of the following:

No awareness

Initial: There is growing awareness about the problem and about the need to provide clarification on legal aspects, scarce or no organised support is given

Partial: There are recommendations and guidance provided on how to respect the legal requirements while sharing data

Defined: Explicit statements about data sharing - ethical and legal aspects are embedded in ethical codes and/or legal documents that govern research and data management activities, little organised support besides guidance and recommendations is given

Managed: Organised services are widely available to support and encourage legally and ethically sound data sharing practice

Please, give a descriptive qualification of your answer as appropriate. What kind of support is given? *

I.3. Broader ecosystem of DAS operation: Data sharing culture/practice

This section focuses on general data sharing practices in social sciences and those that influence the availability of data for reuse as well as real and perceived barriers to data sharing and incentives for sharing.

Name of the person entering information:

3.1. Data sharing culture

The following questions cover the prevalence of data sharing and reuse, and existing routines and attitudes towards data sharing in the social science community in your country. Please answer the questions based on reports published between 2011 and 2016 (if available), and recent experiences in your institution.

3.1.1. What proportion of social science researchers in your country have shared the research data they produced in the period between 2011 and 2016?

Please give an estimate of the proportion of researchers, based on experience in your institution and published reports on data sharing (if available).

Please choose **only one** of the following:

low (0-10%)

medium (10-30%)

high (>30%)

Unable to provide estimate for 2011-2016

Describe, explain or note differences in different research areas, add a reference for the estimate or make another comment on your choice here:

3.1.2. What proportion of social science researchers in your country have been able to access existing third party data they need in the period between 2011 and 2016?

Please give an estimate of the proportion of researchers, based on experience in your institution and published reports on data access and reuse (if available).

Please choose **only one** of the following:

low (0-10%)

medium (10-30%)

high (>30%)

Unable to provide estimate

Describe, explain or note differences in different research areas, add a reference for the estimate, or make another comment on your choice here:

3.1.3. Can you rank the following routines for data sharing to reflect the preferences of the social science research community in your country between 2011 and 2016?

Please base the ranking on experiences in your institution and published reports on data sharing, access and reuse (if available).

All your answers must be different and you must rank in order.

Please select at most 5 answers

Please number each box in order of preference from 1 to 6

Data archive or repository

Supplementary data in a journal (alongside paper)

Via project or personal websites

Via informal contacts (peers and colleagues)

Other (please, describe in the field below)

Unable to provide estimate for 2011-2016

If any of several of the routines not relevant, please, leave them in the left-sided box

Describe, explain or note differences in different research areas, add a reference for the estimate, or make another comment on your choice here:

3.1.4. Please score the statements below to best match the overall attitudes of social science researchers in your country, based on experience in your institution and previously published reports, in the period from 2011 to 2016, on a five-point scale from 5 - true, to 1 - false.

Please choose the appropriate response for each item:

	TRUE	FALSE	Unable to provide estimate for 2011-2016
Data sharing has no benefits at all.			
Data sharing creates healthy competition in research.			
Data sharing creates negative competition (for example, being scooped and therefore reduced publication opportunities) for the researcher.			
Reuse of existing data can answer new research questions and facilitate advancement of science.			
Data sharing has as a risk that others may misuse and misinterpret data.			
Data sharing involves little effort and minimal costs.			

Describe, explain or note differences in different research areas, add a reference for the estimate, or make another comment on your choice here:

3.2. Enablers for data sharing

This section focuses on questions about incentives and rewards related to data sharing in the social science community in your country, as well as skills and practices of researchers and available support services and tools for data sharing and reuse.

3.2.1. How common it is for researchers in your country in your research area to publish in journals that expect data used in the publication to be available for reuse from a trusted digital repository, and why?

Please choose **only one** of the following:

Very common

Rather common

Not that common

Not common at all

Unable to provide estimate

Make a comment on your choice here:

3.2.2. Would you say that there are career rewards related to data sharing, if you consider social science researchers in your country?

This can include any kind of career rewards, e.g. influence on career progression within institution or community or due to government rules, higher success rate in obtaining research funding, better standing within the research community and other.

Please choose **only one** of the following:

Yes, a lot

Yes, to some extent

No

Please list and describe these rewards, or make a comment on your choice here:

3.2.3. According to your experience, are there data support services provided to social science researchers in your country, that facilitate data sharing and/or Open Access to research data (regarding for example, data management plans, data preservation, and data access)?

Examples of services: web guidance, trainings, workshops, webinars, online reference materials, helpdesk or contact and info point, metadata creation and publishing tools (Nesstar, NADA, DataVerse...), linkages between papers and data, support to data management planning, support to long-term preservation of data, access to data

Please choose **only one** of the following:

Yes

No

Please list the data support services available to social science researchers, and institution(s) providing the services here: *

Only answer this question if the following conditions are met:[IY3Y2Y3](#) == 'Y'

3.2.4. Would you estimate that social science data producers in your country follow data management and data documentation standards and procedures that facilitate data reuse?

Please choose **only one** of the following:

Most data producers follow data management and data documentation standards and procedures.

Some data producers follow data management and data documentation standards and procedures.

None

Please list and describe best practice, use of metadata standards among data producers etc. if possible, with links to sources, or make another comment on your choice here: *

I.4. Broader ecosystem of DAS operation: Data infrastructure (where no formal DAS exists)

The goal of this section is to assess data archive proto-activities and open access support activities in countries where no formal Data Archive Service (DAS) exists yet.

Your contribution will help us to identify pioneers in your country, the key players involved in DAS related activities, and to describe their current expertise, the level of technical infrastructure development, as well as their overall activities and potential regarding open science issues in general.

4.1. DAS proto-activities

4.1.1. In your country, are there existing technical infrastructures (national and/or institutional infrastructures in the social sciences such as repositories, online tools, databases, or online catalogues, etc.) that could possibly be used for or applied to a new DAS?

Please choose **only one** of the following:

- Yes
- No
- Do not know

Please, describe them and state if it is planned to use them for a future DAS: *

Do these kinds of infrastructure exist already in other scientific disciplines in your country?

Please choose **only one** of the following:

- Yes
- No

Please describe the existing infrastructure in other disciplines in your country: *

4.1.2. Are there any activities in your country towards establishing a DAS for the social sciences?

Please choose **only one** of the following:

- Yes
- No
- Do not know

Please specify what these are (e.g., policy development, governance, human resources): *

4.1.3. In your opinion, are there institutions that could host a DAS for the social sciences in your country?

Please choose **only one** of the following:

- Yes
- No
- Do not know

Please, name the institution(s): *

4.1.4. Are there any existing initiatives in your country to develop the knowledge and skills of people who might be employed at a DAS at some point?

This could include: research and data management (mainly data management plan), data preservation, and data access.

Please choose **only one** of the following:

- Yes

No

Do not know

Please specify what these are: *

4.2. Open access support activities

4.2.1. Are there open access (OA) projects or initiatives (e.g. OA promotion) in your country, funded either by the government or by grassroots?

Please choose **only one** of the following:

Yes

No

Do not know

Please describe them: *

II. Capability requirement areas

The purpose of this section is to carry out the self-assessment of the active data archive service DAS activities following the specification of the CESSDA SaW Capability Development Model (CESSDA-CDM), which is explained here:

<http://dev.cessda.net/CESSDA-Services/Projects/CESSDA-SaW/Work-Packages/WP3/CESSDA-CDM>.

We will provide you with a set of statements that describe processes or a general set of activities that can be performed in an archive or a repository.

The processes and activities are divided into three main categories:

Organisational infrastructure,

Digital object management (or data curation) and

Technical infrastructure

II.1. Capability requirement areas: Organisational infrastructure

In this section, we present a set of key process areas that support the development and maintenance of a viable preservation organisation. Please indicate the development stage of your organisation on each of the following statements:

Name of the person entering information:

1.1. Mission and scope

1.1.1.1. Does your organisation have a formal approval for its mission statement?

Please choose **only one** of the following:

Not defined: The organisation has not received any approval.

Initial: Some informal support from external stakeholders, no formal mandate, approval or contractual obligations exist.

Repeated/partial: There are some evidence to support the notion that the organisation and its activities receives approval from stakeholders; the organisation is on an ad hoc basis “encouraged” to continue its activities, but no formal mandate or formalised approval exist.

Defined: The organisation has a received formal approval through clearly defined mandate.

eg. mandated by funders, statement signed off by governing board, etc.

Mission statement: statement that reflects the organization's mission and commitment to the preservation of, long-term retention of, management of, and access to digital information.

1.1.1.2 Are there explicit written statements on the role, mandate, purpose and mission of the organisation; are the statements available for anyone to read?

Please choose **only one** of the following:

Not defined: A mission statement does not exist.

Initial: There is some evidence that the organisation has recognised its mission (implicitly) but there are no formal, written documents to substantiate it.

Repeated/partial: There are some relevant output to support and substantiate the mission of the organisation, but they are only partially complete and/or inconsistent.

Defined: The organisation provides a written, formal and complete mission statement that is available for all relevant users.

Please quote the mission statement here: *

Only answer this question if the following conditions are met:
((IY1Y1Y2.NAOK >= "1"))

1.1.1.3. Does your organisation have defined (as part of the mission statement) an explicit scope for the repository and their collection, which guides the selection and appraisal of data?

Only answer this question if the following conditions are met:
((IY1Y1Y2.NAOK >= "2"))

Please choose **only one** of the following:

Not defined: No defined scope

Initial: Awareness of the need to define scope of the repository; there are some ideas of scope, but nothing is formalised or explicitly communicated.

Repeated/partial: Scope of collection is implicitly defined by selection/appraisal actions and routines; some types of data are repeatedly appraised, but without formal or explicit scope in place.

Defined: Scope of collection / repository explicitly defined and promulgated, e.g. by a collection policy or in an official mission statement. Statements on scope are either integrated into, or connected to, policies; policies are connected to specific processes and procedures.

Managed: The definition of scope is monitored for compliance with policies, processes and procedures; actions are taken where processes appear not to be aligned to scope, policy or mission statement; regular reviews of scope and mission of repository.

Optimized: Regular reviews and updates of scope based on monitoring of, and communication with, Designated Community and other relevant stakeholders; delimitation of scope based on cost-benefit analyses and measurements of repository performance.

1.1.1.4. Does your organisation have a medium-term (3 to 5 year) and long-term (>5 years) plans in place to ensure the continued availability and accessibility of data in case the organisation ceases to operate?

Please choose **only one** of the following:

Not defined: There is no evidence of a continuity of access, or is not applicable.

Initial: No contingency plans or only informal intent/agreements.

Repeated/partial: A partial plans in place, but limited in detail and scope

Defined: Formal, written succession/contingency plans; formal written agreements with external organisational frameworks.

Managed: The organisation monitors its organisational environment to determine when to execute plans.

Optimised: There are regular reviews and updates to the plans and agreements to reflect changes in the organisation's environment. E.g. due to funding issues or other "crisis"

Enter documents URL-s and/or make a comment on your choice here:

1.2. Contracts, licences, liabilities

1.2.1.1. Does your organisation have sufficient knowledge and documentation on how (relevant aspects of) the (national or international) legislation applies to and affects the holdings and procedures of the organisation?

Please choose **only one** of the following:

Not defined: There is no evidence of knowledge or documentation about national legislation and how it applies to the organisation or infrastructure.

Initial: Very limited knowledge; or very limited documentation.

Repeated/partial: There is partial knowledge on relevant legislation, or knowledge is not wide-spread. There is some formalised documentation, but it may be insufficient documentation.

Defined: Sufficient and documented knowledge on all relevant legislative aspects is available to all staff.

Managed: There are regular analysis of the national legislation, application and effects. Non-observance in the application of the legislation are documented.

Optimized: Procedures and mechanisms are in place to update this knowledge and are used.

Enter documents URL-s and/or make a comment on your choice here:

1.2.1.4. Are service access procedures based on legal or contractual regulations that are settled in agreements between the service users and the repository; and the contractual and legal regulations, ensures that the parties do not infringe any intellectual property rights (IPR) of any other person(s) or institution(s)?

Please choose **only one** of the following:

Not defined: No contracts, regulations or agreements in place.

Initial: Legal and/or contractual issues are solved on an ad hoc basis; case-by-case.

Repeated/partial: Contracts and/or agreements are set up on a regular basis; contracts are not formalised or standardised; no procedures for implementation.

Defined: Contracts and/or agreements are standardised and implemented according to written procedures; contracts and regulations are made publicly available.

Managed: Legal and/or contractual framework is regularly reviewed and updated; all legal and/or contractual regulations are aligned to higher level policies; roles and responsibilities are identified and maintained.

Optimized: The usage and success of access licenses and the access conditions framework are continuously assessed; monitoring of wider legal framework (e.g. national and EU regulations); regular and formalised contact with relevant stakeholders.

Enter documents URL-s and/or make a comment on your choice here:

1.3. Funding, staff, resources

1.3.2.2. Does your organisation monitor demand for its repository services, growth and funding?

Please choose **only one** of the following:

Not defined: There is no evidence of monitoring of demand for repository services.

Initial: The organisation only reacts to significant changes in demand for the repository services when the changes occur. It does not attempt to adjust its funding.

Repeated/partial: The organisation occasionally reviews changes in demand for the repository services and only reacts to significant changes.

Defined: At regular intervals (ideally annually or more frequently) the institution monitors the demand for its repository services.

Managed: The organisation actively monitors demand and response for its repository services and is aware of trends.

Optimized: At regular intervals, the organisation utilises the monitoring information to adapt its level or other aspects of its repository service provision and ensures there is adequate funding.

Enter documents URL-s and/or make a comment on your choice here:

1.3.4.1. Does your organisation have the appropriate number of staff to support all repository functions and services?

Please choose **only one** of the following:

Not defined: There is no evidence that the organisation has defined a level for the appropriate number of staff to support all functions and services. The number of staff appears to be inadequate to support all functions and services. Many core competencies are missing.

Initial: There is evidence staffing is partially complete, either due to lack of resources or unable to employ qualified staff. The organisation has not defined the appropriate staffing level to support all repository functions and services.

Repeated/partial: There is evidence that there is likely to be an appropriate number of staff to support all repository functions and services, however the organisation has not defined what.

Defined: The defined level for the appropriate number staff to support all functions and services is met.

Managed: Staffing levels appropriate to support functions and service needs are monitored and assessed at regular intervals. The monitoring is documented.

Optimized: At regular intervals, the appropriate staffing level is reviewed with respect to the function and service needs and the defined appropriate level is adjusted accordingly.

Enter documents URL-s and/or make a comment on your choice here:

1.5. Confidentiality, ethics and disclosure risk

1.5.1.1. Does your organisation comply with applicable research discipline norms and legal data handling requirements (within the scope of social science data)?

Please choose **only one** of the following:

Not defined: Not relevant or no requirement for compliance.

Initial: Informal and inconsistent adherence to legal and research discipline norms for handling confidential data.

Repeated/partial: The organisation complies with norms consistently through repeated action, however there are no written procedures or procedures for exceptions and errors.

Defined: The organisation complies with norms and legal requirements through systematic written, formal procedures and policies.

Managed: Compliance to procedures and policies are monitored and assessed. Non-compliance incidents are recorded.

Optimized: Reviews of the procedures and policies for compliance to research discipline norms for confidential data handling are performed at regular intervals.

Enter documents URL-s and/or make a comment on your choice here:

1.5.1.3. Does your repository require that data depositors ensure that data meet requirements of confidentiality and non-disclosure for data collected from human subjects?

Please choose **only one** of the following:

Not defined: Not applicable; not relevant; or there is no such activity.

Initial: There is some awareness of the issue, but decisions and procedures are ad hoc and performed on a case-by-case basis; there are no written processes and procedures in place for dealing with confidentiality, disclosure and data protection issues; no predefined criteria or non-disclosure agreements/statements are available for depositors/users.

Repeated/partial: Depositors are repeatedly being informed of confidentiality issues, but no formal information template exists; checks are being performed on the deposited data, but there are no written procedures in place (some documentation may exist but it is incomplete).

Defined: Processes and procedures are in place; standardised information is provided to the depositor prior to the deposit; checks are performed on data after deposit; process and procedure descriptions for handling and altering sensitive data are in place.

Managed: Processes and procedures are integrated into high level policies; there are regular reviews and updates of processes and procedures; there are mechanisms and procedures in place for staff training on confidentiality, disclosure risks and anonymisation.

Optimized: The usage and success of confidentiality and disclosure mechanisms are continuously assessed, reviewed and updated; monitoring of wider legal framework (e.g. national and EU regulations); regular and formalised contact with relevant stakeholders; automated checks and anonymisation mechanisms may be in place.

In some cases, the repository may alter sensitive data to create anonymised data that can be distributed to its user community

Please, describe/list, if possible, with links to sources documents: *

II.2. Capability requirement areas: Digital object management (data curation)

In this section, we present a set of key process areas required to maintain and provide access to digital information in an authentic form, for as long as required and across changing technical environments. Digital Object Management (DOM) is closely related to the term "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. The aim of DOM and digital curation is to mitigate digital obsolescence, keeping the information accessible to users indefinitely.

Based on the maturity level descriptions for each of the following statements, please indicate the development stage of your organisation.

Name of the person entering information:

2.1.1.2. The organisation clearly specifies the information (documentation, metadata, provenance) that needs to be associated with the data that is to be deposited.

Please choose **only one** of the following:

Not defined: Not specified; no awareness.

Initial: There is some awareness of the documentation and metadata that is needed for deposit, but it is not formalised; information communicated to users/depositors on an ad hoc basis.

Repeated/partial: Documentation and metadata requirements implicitly defined by acquisition and deposit activities and routines; no formal or explicit requirements in policy or other written documents.

Defined: A written formal specification of required information is explicitly defined (e.g. in a collection policy); requirements are compliant with metadata standards that are used and can be understood by Designated Community (e.g. DDI); metadata requirements are accessible and communicated to users/depositors.

Managed: Documentation and metadata requirements are aligned with policies and other processes and procedure documents. There are regular reviews and assessments (of success) of the information requirements.

Optimized: Regular reviews and updates of requirements based on technology watch, monitoring of, and communication with Designated Community and other relevant stakeholders.

Enter documents URL-s and/or make a comment on your choice here:

2.1.2.4. The organisation offers and provides functions and mechanisms for proper data citations.

Please choose **only one** of the following:

Not defined: No citation practices.

Initial: Citations are offered when requested by the depositor; ad hoc and case-by-case approach; no practices or strategies are written down.

Repeated/partial: Citation practices are being repeated and offered regularly; lacks formalisation and systemisation.

Defined: Citations are required and offered to all depositors; formalised through templates or other written documents; processes and procedures are documented.

Managed: Citation mechanisms are regularly reviewed and updated.

Optimized: Mechanisms for requesting and providing citation are monitored and measured; there are systemised reviews and updates of citation mechanisms based on technology watch and formal, regular communication with Designated Community and other relevant stakeholders.

Enter documents URL-s and/or make a comment on your choice here:

2.1.2.5. The organisation has mechanisms and functions in place that allow the depositor to place access conditions on the information that is being deposited.

Please choose **only one** of the following:

Not defined: No awareness; no mechanisms/procedures in place.

Initial: Some awareness of the issue; access conditions are set up and agreed upon when required by the depositor, but most deposits are done without any conditions on use.

Repeated/partial: Most depositors are offered an opportunity to define conditions of use on the information they deposit, but conditions are not formally defined; no template or set of predefined categories.

Defined: All depositors are offered the opportunity to set access conditions on the information that is being deposited; a set of access conditions are formally defined in categories or a template.

Managed: Regular review and updates of set of conditions; aligned with high level policies.

Optimized: Regular review and updates of set of conditions; formalised feedback mechanisms and cooperation with user groups and other relevant stakeholders (e.g. how to deal with funder policies that have open access requirements).

Enter documents URL-s and/or make a comment on your choice here:

2.1.2.6. The organisation has in place agreements that confirm the legal transfer (or other consensual agreements) of the information that is being deposited, and the agreement includes clear definitions of roles and responsibilities of repository and depositor (i.e. the repository legally takes control of the deposited material so that they can make the necessary changes to the data to prepare it for long-term storage and to distribute it to their consumers). The contractual and legal regulations also make sure that the deposited material does not infringe on any intellectual property rights (IPR) of any other person(s) or institution(s).

Please choose **only one** of the following:

Not defined: No agreements on legal transfer of custody.

Initial: Awareness of legal custody, but no formal or written agreements in place; some agreements are made ad hoc, on an individual basis; some/most deposits are made without any kind of agreements on legal transfer of custody.

Repeated/partial: Most of deposits are made with agreements in place, but adjustments are made on individual basis when needed. Agreements are not formalised into 'templates' that are being used regularly and systematically.

Defined: Formal, written agreements and contracts in place; responsibilities and legal transfer of custody clearly defined. Contractual templates are being used consistently; legal and contractual framework is regularly reviewed and updated; all legal and contractual regulations are aligned to higher level policies; roles and responsibilities are identified and maintained.

Managed: Monitoring of the usage of agreements and contracts; actions are taken where contracts/agreements appear not to be working effectively or are not in accordance with higher level policies; reviewed and updated regularly.

Optimized: The usage and success of transfers and agreements are continuously assessed; monitoring of wider legal framework (e.g. national and EU regulations); regular and formalised contact with relevant stakeholders.

Enter documents URL-s and/or make a comment on your choice here:

2.1.3.1. The organisation has processes and mechanisms in place that verify the deposited material for completeness and correctness. That is, the organisation inspects data files to ensure that variables and values are accurate according to the documentation supplied; that variables and values are sufficiently labelled for reuse; and that variable names in a dataset match variable names in a codebook.

Please choose **only one** of the following:

Not defined: No awareness; no inspections of data. Data deposited without data inspection.

Initial: Some awareness of data inspection; there are non-systematised (manual) checks of deposited material (e.g. checked for basic metadata completion; visual checks of data, etc.), but the procedures are ad hoc. There are no written procedures or process documents and there is a lack of mechanisms for detecting technical errors in deposit/transmission. There are no formalised processes and procedures for rectifying data and/or metadata.

Repeated/partial: There are non-systematised (manual) checks of deposited material in place; processes and procedures are repeated, but they are not formalised or documented. Rectifications are performed repeatedly, either by the repository or by returning data to depositor.

Defined: Systemised checks of all data and metadata are in place; procedures and processes defined and formalised in written documents. Functions and mechanisms are in place for rectifying data and metadata, including processes for contacting and/or returning data to depositor when necessary. Some processes may be automated.

Managed: All completeness and correctness checks and modifications are measured and registered; processes and procedures are reviewed and updated regularly. Automated processes are implemented where appropriate.

Optimized: All completeness and correctness checks and modifications are measured, registered and assessed regularly; processes and procedures are reviewed and updated regularly based on technology watch and formalised communication with Designated Community.

Enter documents URL-s and/or make a comment on your choice here:

II.2. Capability requirement areas: Digital Object Management - part 2

Name of the person entering information:

2.2.1.2. The organisation has mechanisms in place that generates persistent, unique identifiers for all its data holdings.

Please choose **only one** of the following:

Not defined: No system for persistent identification.

Initial: There is some awareness of the need for persistent identifiers and locators, but actions are sporadic and ad hoc; there are no formalised systems, processes or procedures in place.

Repeated/partial: Mechanisms and systems for identification and location are partly in place (e.g. there may be a certain directory structure or hierarchy to make the locating of data easier), but does not comply with formalised DOI systems; mechanisms are being repeatedly used, but there is lack of formalisation and written procedures.

Defined: There are mechanisms and systems in place to persistently identify and locate data and metadata (either by following external systems like DOI, or by internal PID systems); all processes and procedures are documented and formalised.

Managed: PID systems and locators are regularly reviewed and updated; mechanisms are aligned to higher level preservation goals and plans.

Optimized: All mechanisms and functions are monitored and measured; there are systemised reviews and updates of PID systems based on technology watch.

Enter documents URL-s and/or make a comment on your choice here:

2.2.3.3. The organisation has in place documented preservation strategies that are relevant to its holdings. The preservation strategy describes how the organisation will act upon identified risks, and will typically address the degradation of storage media, the obsolescence of media drives, and safeguards against accidental or intentional digital corruption of preserved material.

Please choose **only one** of the following:

Not defined: No preservation strategies in place.

Initial: Informal, ad hoc 'contingency action points' are in place, but a full comprehensive strategy is lacking; action points are not formalised or connected to a policy.

Repeated/partial: Contingency action points have matured into a partial strategy; only partly formalised and documented.

Defined: Fully formalised and documented preservation strategy in place; connected to preservation policies and repository strategies, and to processes and procedures.

Managed: Strategy is periodically reviewed and updated.

Optimized: Usage of strategy is measured and assessed; processes, functions and mechanisms are under constant improvement and continuously integrated into the strategy and the higher-level policies.

Enter documents URL-s and/or make a comment on your choice here:

2.3.1.5. The organisation provides data and metadata that are in accordance with standards that are used and understood by the Designated Community.

Please choose **only one** of the following:

Not defined: No strategy for formats and standards.

Initial: There is some awareness of the formats and standards that are used by Designated Community; the repository provides data and metadata in user-specified formats on an ad hoc basis, when enquired.

Repeated/partial: Some standards and formats are repeatedly in use; lacks a defined and explicit strategy; there are no lists of available formats.

Defined: A data/metadata format strategy is explicitly defined and formalised and communicated to users; data and metadata are provided in formats that are commonly in use and understood by users.

Managed: format usage and enquiries are measured and assessed; format strategy is regularly reviewed and updated.

Optimized: format strategy is regularly reviewed and updated based on technology watch and formalised communication with users / Designated Community; format strategy is integrated into higher level data policies.

Enter documents URL-s and/or make a comment on your choice here:

2.3.2.3. The organisation uses AAI (Authentication and Authorization Infrastructure) or other direct or federated user authentication approaches to control access to data

Please choose **only one** of the following:

Not defined: No authentication approach in place.

Initial: There is some awareness of the need to implement authentication approaches; access control is performed on an ad hoc, case-by-case basis.

Repeated/partial: An authentication infrastructure emerges by repeated use of authentication approaches; lacks standardisation and formalisation.

Defined: The repository uses digital identities, identity management, authentication, authorization to control access to data; AAI is formalised, systematised and documented.

Managed: All AAI mechanisms and functions are measured, assessed and regularly reviewed and updated.

Optimized: All AAI mechanisms and functions are monitored and measured; there are systemised reviews and updates of AAI systems based on technology watch and formal, regular communication with Designated Community and other relevant stakeholders.

Enter documents URL-s and/or make a comment on your choice here:

II.3. Technical infrastructure and risk

Name of the person entering information:

3.1 The organisation undertakes risk assessments and overall risk analysis when there are changes to the technical infrastructure which may affect the security or resilience of any service, component or procedure.

Please choose **only one** of the following:

Not defined: There is no evidence that risk assessments are undertaken.

Initial: There is some awareness of the need for risk assessment and some risk assessments are undertaken, but on an ad hoc basis or when requested by the organisation for a specific purpose.

Repeated/partial: Risk assessments are undertaken when significant changes are made to the technical infrastructure, but these are not analysed.

Defined: There is a documented risk assessment methodology (and tool) that is used to make a systematic analysis of security and infrastructure resilience risk factors when there are changes to the technical infrastructure.

Managed: The use of the risk assessment methodology and analysis is monitored. Any significant risks are managed immediately as part of risk mitigation procedure, with issues, incidents and discrepancies are documented.

Optimized: At regular intervals, the risk assessment and analysis processes are reviewed. Any modifications to the processes are documented and communicated internally.

Enter documents URL-s and/or make a comment on your choice here:

3.2. The organisation has sufficient and appropriate technical infrastructure resources (e.g. suitable software, technical services, and appropriate management plans) to ensure that all functions and services of the repository are supported.

Please choose **only one** of the following:

Not defined: There is no evidence or awareness that the organisation has appropriate technical infrastructure resources to support all functions and services. Lack of skills and technology is inadequate to support all functions and services.

Initial: There is evidence that the technological infrastructure is partially complete, either due to lack of resources or unable to employ qualified staff. The organisation has not defined the appropriate technical resources or level to support all repository functions and services.

Repeated/partial: There is evidence that there is likely to be an appropriate level of technical infrastructure resources to support all repository functions and services, however the organisation has not defined what this level should be.

Defined: The organisation has defined the appropriate technical infrastructure resources to support all functions and services.

Managed: Technical infrastructure resources appropriate to support functions and service needs are monitored and analysed at regular intervals. The monitoring is documented.

Optimized: At regular intervals, the appropriate technical infrastructure resources are reviewed with respect to the function and service needs and the description is adjusted accordingly.

Enter documents URL-s and/or make a comment on your choice here:

3.5.1. To ensure on-going access to and preservation of its holdings, the organisation has appropriate succession plans and/or contingency plans in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope/obligations.

Please choose **only one** of the following:

Not defined: No succession plans, contingency plans, and/or escrow arrangements.

Initial: There is some awareness of the issue; low institutional commitment to contingency issues; no written, formal processes, procedures, plans or other documents exists.

Repeated/partial: Some awareness of the need to plan and prepare for contingency issues; no written statements, procedures or processes.

Defined: Repository have implemented functions and mechanisms to be adequately prepared for major institutional changes; processes and procedures are formalised and defined.

Managed: Functions and mechanisms are regularly reviewed and updated; all plans are integrated into higher level policies; all relevant staff are trained in contingency issues; roles and responsibilities are identified and maintained; some relevant parts of the ISO 17799 are employed.

Optimized: Procedures and processes are continuously assessed; monitoring of wider community (technology watch); regular and formalised contact with relevant stakeholders; assessment of success of training programs and staff knowledge; procedures are assessed and measured towards benefit-cost assessments and repository performance; repository maintains ISO 17799 certification.

Enter documents URL-s and/or make a comment on your choice here:

III.1. Background information: Organisational profile

Name of the person entering information:

1.1. Please describe core services and activities that are offered by the data archive service:

1.2. Please describe how the data archive service is funded:

Whether it is fully or partly publicly or privately funded, and/or share of project-based funding vs. long-term contractual funding [example: 100 % public funding, through Research Council and Government Ministry funding]

1.2. Physical and administrative location of the data archive service.

Also: describe relationship to other suppliers/partners that are of importance to the services (if applicable).

e.g. part of university/housed within a larger organisation, etc.

1.4. Please describe the main bodies / departments / units that the data archive service consists of:

Including roles and responsibilities of different bodies / stakeholders. Number of departments, number of personnel (of a DAS department).

1.5. Please describe the linguistic composition of the archive's collections:

i.e. the main language(s) of the data and metadata of the collection, often English, but may be in any language, or a combination of languages.

1.6. Please provide a brief description of the overall contents of the current collection:

i.e. subjects, types of data, etc.

1.7. Please define primary and secondary user communities [if applicable]:

Thank you for your time, we will keep in touch in case there are some things to be clarified!

You can continue to [CESSDA SaW project website](#) or just close this page.

Submit your survey.

Thank you for completing this survey.

Appendix 4: Classification protocols for assigning development level

Development of SSH sector

1.1. Development of social science sector	Low level	Medium level	High level
Definitions	0	1	2
Overall assessment of financial stability, research capacities and results achieved in the social sciences in a country (funding, human resources and infrastructure conditions; impact and prestige in society)	Funding of SSH and productivity of the researchers are in the lowest quantile; impact on designated community is small or non-existing	Funding of SSH and productivity of the researchers are in the mid quantile; impact on designated community is limited	Funding of SSH and productivity of the researchers are in the highest quantile; impact on designated community is strong

First group of indicators from 1.1.1 to 1.1.4 are data on GERD and number of researchers. According to the centrally collected data, some referent quantile' thresholds are obtained in order to sort countries into categories and assess indicator values. This is presented in the following table

Indicator	GERD in SSH as % GDP	GERD in SSH as % of GERD	Number of researchers in SSH per capita	GERD in SSH per researcher in SSH
low (0 point)	$x \leq 0.09$	$x \leq 8$	$x \leq 60$	$x \leq 10000$
medium (1 point)	$0.09 < x \leq 0.15$	$8 < x \leq 16$	$60 < x \leq 95$	$10000 < x \leq 20000$
high (2 points)	$x > 0.15$	$x > 16$	$x > 95$	$x > 20000$

Indicator 1.1.5 on structure of the funding according to the sources is only descriptive, so it cannot be used to compute some particular values.

Indicator 1.1.6 on support to SSH in form of software licenses and access to databases gives three option to check (software licenses, datasets, and full-text databases). If all options are checked, score of the indicator is 2, if two are checked score is 1, otherwise it is 0.

After scores of the indicators are assessed, **average** score of all indicators should be computed and overall level of SSH development should be assessed. Preliminary threshold for the level categorization are given as:

if overall indicator less than 1 – low level

if overall indicator is equal to or higher than 1, but lower than 1.4 – medium level

if overall indicator is equal to or higher than 1.4 – high level

The next table illustrates three examples of countries being classified to low, medium and high level of SSH development according to the average scores of indicators:

Indicator	Macedonia	Croatia	Netherlands
GERD in SSH as % GDP	0	1	2
GERD in SSH as % of GERD	2	1	1
Number of researchers in SSH per capita	1	1	1
GERD in SSH per researcher in SSH	0	1	2
Access to databases	1	1	2
Average	0.8	1	1.6
Overall	0	1	2

Development of research data production in SSH

1.2. Development of research data production in SSH	Underdeveloped	Developing	Developed
Definitions	0	1	2
Prevalence of high quality research data with high potential for reuse	Rare or no data producing research projects, dispersed and low quality existing data, absence of studies of national importance	Some examples of research excellence, streams of research stand out, either qualitative or quantitative, some examples of international collaborative research	Well established streams of research traditions, national and international, great variety of important types of research data

In order to assess development of research data production in SSH, the following system of scoring is assigned to the three indicators, which belong to this subsection, as presented in the following table:

Indicator	International collaborative research	Studies of national importance	Average production of data
Underdeveloped (0 point)	five or less	if answer is "no",	rare
Developing (1 point)	6	if answer is "other"	periodical
Developed (2 points)	seven or more	if answer is "yes"	frequent

After scores of the indicators are assessed, **summary** score of all indicators should be computed and overall level of SSH development should be assessed. Preliminary threshold for the level categorization are given as:

- if overall indicator is 1 or 2 – underdeveloped
- if overall indicator is 3 or 4 – developing
- if overall indicator is 5-6 – developed

The next table illustrates three examples of the same countries as in previous examples.

Indicator	Macedonia	Croatia	Netherlands
International collaborative research	0	1	2
Studies of national importance	0	0	2
Average production of data	1	1	2
Sum	1	2	6
Overall	0	0	2

RDM Policy setting

The classification of **Research data management (RDM) requirements and incentives** fall into the following three categories:

0 – Non-existent: Not aware of the need, not seen as a priority.

1 – Emerging: Declared awareness about importance and intentions of formulation of policy principles and strategy supporting data sharing motivation

2 – Developed: Partially or fully operationalized strategy and policy developed and implemented in calls on key aspects enabling data sharing.

The classification is based on the following four indicators:

I2_1_1 'Research data management (RDM) policy requirements: Data management plan' /

1 "None"

2 "Initial: There is growing recognition and awareness of need to require DMP"

3 "Partial: There is the expectation or recommendation to have DMP in place"

4 "Defined: Formal requirement, little monitoring and support"

5 "Managed: DMP is a requirement, clear guidance is issued, support and tools are provided, the content of DMP and exemptions from full open access are defined "

I2_1_2 'RDM policy requirements: Appropriate place of data deposit defined ' /

1 "None"

2 "Initial: There is growing recognition and awareness of the need to have disciplinary specific place of deposit and support services "

3 "Partial: There is the expectation or recommendation to offer or deposit data in an appropriate disciplinary repository or equivalent data archive service "

4 "Defined: Formal requirement, little or no monitoring"

5 "Managed: Formal requirement, sanctions for not complying with regulation are in place (such as reduced payment, etc.), full support and guidance is provided"

I2_1_3 'RDM policy requirements: Sustainability and long-term curation' /

1 "None"

2 "Initial: There is growing recognition and awareness about the value of research data produced and about the need for long long-term preservation; little or no investment and support for long-term curation provided."

3 "Partial: It is expected or recommended to assess the value of research data and resources providers declared their motivations for continuing to invest in sustaining the assets. The FAIR data principle (findable, accessible, interoperable and reusable) are highlighted in these recommendations."

4 "Formal: Requirement to assess research data appropriate for data curation. It is understood that the best use of resources involves making choices based on value judgments and selecting material for curation. Investment and support for long-term curation is in place, based on contractual arrangements."

5 "Managed: Formal requirement, based on contractual collective arrangements of roles and responsibilities among different stakeholders, clear definitions of what data is preserved, how it is documented, and for how long. Funder requires that periodic monitoring is performed if the best use of resources is made for expected amount of benefit."

I2_1_4 'Incentives for data sharing: Cost for managing the data resources' /

1 "No"

2 "Initial: Cost for managing the data and preparing it for access can be implicitly covered in the overall research project budget"

3 "Partial: There is explicit recognition that additional cost for preparing the data for access are legitimate project cost that can partially cover the RDM cost up to a certain limit "

4 "Defined: Costs for RDM are fully covered and adequate, based on DMP plan in project documentation"

Median values of all four indicators are classified as follows:

(1=0 - Non-existent) (1.5 thru 2=1 - Emerging) (2.5 thru 5=2 - Developed)

A description of ethical and legal framework important for data sharing

Underdeveloped	Developing	Developed
Little or no awareness among different stakeholders, legal and ethical concerns are mainly used as an excuse for not sharing data at all	Growing awareness about the problems and some general guidance is provided	Specific guidance, advice and support services are available to prevent and overcome legal and ethical obstacles to data sharing and to actively support optimal mode of access to data

The classification is based on one indicator:

12.2.1. Clarification and support on legal and ethical aspects provided

1. "No awareness"

2. "Initial: There is growing awareness about the problem and about the need to provide clarification on legal aspects, little or no organised support is given"

3. "Partial: There are recommendations and guidance provided on how to respect the legal requirements while sharing data"

4. "Defined: Explicit statements about data sharing - ethical and legal aspects are embedded in ethical codes and legal documents that govern research and data management activities, little organised support besides guidance and recommendations is given"

5. "Managed: Organised services are widely available to support and encourage legally and ethically sound data sharing practice"

Values are classified into the heading categories as follows:

(1=0 - Underdeveloped) (1.5 thru 2=1 - Developing) (2.5 thru 5=2 - Developed)

Data sharing culture

Detailed matrix for the procedure for the aggregation of the values of indicators 3.1.1 to 3.1.4 to a high-level score for concept 3.1.

Underdeveloped	Developing	Developed
A Intensity of sharing and reuse		
Sharing rare Reuse (access) rare 3.1.1. AND 3.1.2. = 1 (low)	Sharing not that common Reuse (access) not that common 3.1.1. OR 3.1.2. = AT LEAST 2 (medium)	Sharing very common Reuse (access) very common 3.1.1. AND 3.1.2. = 2 or 3 (high)
B Formality and transparency of sharing routines		
No sharing routines 3.1.3. none of the routines (1) to (3) ranked AND / OR routine (5) ranked as first Data archive or repository (1) Supplementary data in a journal (alongside paper) (2) Via project or personal websites (3) Via informal contacts (peers and colleagues) (4) Other (5)	Informal and not transparent sharing routines 3.1.3: routines (3) AND (4) ranked among first three	Formal and transparent sharing routines 3.1.3: routines (1) AND (2) ranked among first three
B Attitudes towards sharing (5-4=TRUE; 3=NEUTRAL; 1-2=FALSE)		
Negative or indifferent attitudes For indicator 3.1.4, <ul style="list-style-type: none"> · At least, two out of statements (1), (2) and (3) are TRUE, AND · At least two out of statements (5), (6) and (7) are FALSE or NEUTRAL. Data sharing has no benefits at all (1) Data sharing creates negative competition for the researcher (2) Data sharing has as a risk that others may misuse and misinterpret data. (3) Data sharing creates healthy competition in research (5) Reuse of existing data can answer new research questions and facilitate advancement of science (6)	Indifferent attitudes For indicator 3.1.4t least, 4 out of six attitude statements are NEUTRAL	Positive attitudes For indicator 3.1.4, <ul style="list-style-type: none"> · At least, two out of statements (5), (6) and (7) are TRUE, AND. · At least two out of statements (1), (2) and (3) are FALSE or NEUTRAL.

Data sharing is easy and requires little effort and costs from researchers (7) <i>The algorithm is to be used for derivation of a higher score only.</i>		
Underdeveloped = 2 out of 3 conditions A, B, C on this level	Developing = 2 out of 3 conditions A, B, C on this level	Developed = 2 out of 3 conditions A, B, C on this level

Enablers for data sharing

Detailed matrix for procedure how the values of indicators 3.2.1 to 3.2.4 are aggregated to a high-level score for concept 3.2.

Underdeveloped: no enablers	Developing: some enablers	Developed: many enablers
3.2.2. = No (There are no career rewards related to data sharing - e.g. influence on career progression within institution or community or due to government rules, higher success rate in obtaining research funding, better standing within the research community)	3.2.2. = Yes, to some extent	3.2.2. = Yes, a lot
3.2.3= No (There are no data support services for researchers in social sciences that facilitate data sharing and OA to research data in social sciences)	3.2.3=Yes AND no more than 2 institutions and 2 services (example services are: data librarian in institution, web guidance, trainings, workshops, webinars, online reference materials, helpdesk or contact and info point, metadata creation and publishing tools (Nesstar, NADA, DataVerse...), linkages between papers and data; support to data management planning, support to long term preservation of data, access to data) named	3.2.3= Yes AND more than 2 institutions and 2 services (example services are: data librarian in institution, web guidance, trainings, workshops, webinars, online reference materials, helpdesk or contact and info point, metadata creation and publishing tools (Nesstar, NADA, DataVerse...), linkages between papers and data; support to data management planning, support to long term preservation of data, access to data) named

3.2.4= No evidence (There is no evidence of data management and data documentation standards and procedures that social science data producers follow and that facilitate data reuse)	3.2.4= There is some evidence	3.2.4= There is a lot of evidence
Underdeveloped: at least, 2 out of 4(3) indicators score on the described level, no indicator on “developed” level	Developing: max 2 indicators on “underdeveloped” level, max 2 on “developed” level.	Developed: at least 2 out of 4 (3) indicators score on the described level, no indicator scores on “underdeveloped” level

DAS proto-activities

Overall assessment 1: *Data Archive Service (DAS) activities implementation type*

0 = No DAS activities	1 = Basic DAS activities	2 = Advanced DAS activities
<i>No DAS exists and no activities – related to technical infrastructure, organisation or capacity building – are provided on a national and/or institutional level.</i>	<i>No DAS exists but basic activities –related to technical infrastructure, organisation or capacity building – are provided on a national and/or institutional level.</i>	<i>No DAS exists but advanced activities –related to technical infrastructure, organisation or capacity building – are provided on a national and/or institutional level.</i>

> Indicators to use: 4.1.1, 4.1.2, 4.1.3, 4.1.4

Assess whether the country belongs rather to the group 0, 1 or 2 according to the answers given and/or the information found during the desk research for indicators 4.1.1, 4.1.2, 4.1.3, 4.1.4. You might also base your assessment (and/or explanations, recommendations) on other indicators responses (in this section or in the whole survey/instrument).

Overall assessment 2: *Open access (OA) support activities implementation type*

0 = No OA support activities	1 = Basic OA support activities	2 = Advanced OA support activities

<p>Support activities to encourage and facilitate OA are rare or not existing in the social science research community.</p>	<p>Some support activities to encourage and facilitate OA exist, but not on a regular basis and only in few institutions.</p>	<p>Support activities to encourage and facilitate OA are well established (common and in most institutions), known and used by the social science research community.</p>
--	---	---

--> Indicators to use: 4.2.1, 3.2.3

Assess whether the country belongs rather to the group 0, 1 or 2 according to the answers given and/or the information found during the desk research for indicators 4.2.1, 3.2.3. You might also base your assessment (and/or explanations, recommendations) on other indicators responses (in this section or in the whole survey/instrument).

Appendix 5: Glossary

Authentication and Authorization Infrastructure (AAI): An AAI is an infrastructure to verify a user's identity (authentication) and to verify that a user has the rights to access the service the user has requested (authorisation) [DASISH].

Authenticity: The degree to which a person (or system) regards an object as what it is purported to be. Authenticity is judged based on evidence [OAIS].

Data access: The OAIS entity that contains the services and functions which make the archival information holdings and related services visible to Consumers.

Data Documentation Initiative (DDI): The DDI is an international standard for describing statistical and social science data. Documenting data with DDI facilitates interpretation and understanding, both by humans and computers [<http://www.ddialliance.org/>].

Data Management Plan (DMP): Data Management Plan is part of grant application or research project delivery that consider essential properties of RDM throughout the project, aiming at Open data as the default.

Data preservation: or more specifically, digital data preservation refers to the series of managed activities necessary to ensure continued access to digital materials for if necessary. This broad definition of data preservation refers to all the actions required to maintain access to digital materials beyond the limits of media failure or technological change. Long-term preservation can be defined as the ability to provide continued access to digital materials, or at least to the information contained in them, indefinitely. (Source: IFDO, Data preservation, <http://ifdo.org/wordpress/preservation/>)

Data Seal of Approval (DSA): Self-assessment for trusted digital repositories. The DSA is granted to repositories that are committed to archiving and providing access to research data in a sustainable way.

Designated Community: An identified group of potential Consumers who should be able to understand a 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 [OAIS].

Digital Object Identifier (DOI): A DOI name is an identifier (not a location) of an entity on digital networks. It provides a system for persistent and actionable identification and interoperable exchange of managed information on digital networks [www.doi.org].

Dublin Core: The Dublin Core Schema is a metadata standard and contains a small set of vocabulary terms that can be used to describe web resources (video, images, web pages, etc.), as well as physical resources.

GERD - Gross Domestic Expenditure on Research and Development is total intramural expenditure on research and development performed on the national territory during a given period [<https://stats.oecd.org/glossary/detail.asp?ID=1162>].

Ingest: The preservation functional entity that contains the services and functions that accept data and metadata from data producers/depositors, prepares data and metadata for storage, and ensures that the information becomes established within the archive.

Integrity: the quality of being complete and unaltered in all essential respects. Can be a component of Authenticity, when combined with identifier [APARSEN]. Data integrity is the opposite of data corruption, which is a form of data loss. The overall intent of any data integrity technique is the same: ensure data is recorded exactly as intended, and upon later retrieval, ensure the data is the same as it was when it was originally recorded. Data integrity aims to prevent unintentional changes to information. [Core Concepts of Information Systems Auditing]

Intellectual Property Right (IPR): IPR is grouping or class of several different legal regimes that generally concerns creations of the human mind. Copyright is one of the legal regimes that fall under the umbrella of intellectual property. [http://corecopyright.org/2009/12/03/copyright_ip/]

Long Term: A period of time long enough for there to be concern about the impacts of changing technologies, including support for new media and data formats, and of a changing Designated Community, on the information being held in an archive. This period extends into the indefinite future [OAIS].

Maturity Level: a Maturity Level estimates the level of maturity for an Activity. Maturity Levels are defined in a scale ranging from 0 to 5: N/A (0), Initial (1), Repeatable (2), Defined (3), Managed (4), and Optimised (5).

Open Archival Information System (OAIS): An Archive, consisting of an organization, which may be part of a larger organization, of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community. The term **OAIS** also refers, by extension, to the ISO OAIS Reference Model for an OAIS. This reference model is defined by recommendation CCSDS 650.0-B-1 of the Consultative Committee for Space Data Systems; this text is identical to [ISO 14721:2003](#) which is superseded by [ISO 14721:2012](#). It meets a set of responsibilities, as defined in section 4, that allows an OAIS Archive to be distinguished from other uses of the term 'Archive'. The term 'Open' in OAIS is used to imply that this Recommendation and future related Recommendations and standards are developed in open forums, and it does not imply that access to the Archive is unrestricted [OAIS].

Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH): The OAI-PMH is a low-barrier mechanism for repository interoperability. It is a protocol developed for harvesting (or collecting) metadata descriptions of records in an archive so that services can be built using metadata from many archives. An implementation of OAI-PMH must support representing metadata in Dublin Core, but may also support additional representations.

Open data as the default: (See <https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex>)

Persistent Identifier (PID): A maintainable identifier that allows one to refer to and have reliable access to a resource or object over long periods. A PID must always be resolvable through a resolution system [APARSEN].

Research Data Management (RDM): 'Research Data Management comprises the different components of the research data lifecycle, from data creation to data preserving, sharing and re-use.' 'RDM is an integral part of the wider research process, contributing the standards and principles of research, and applicable not just to the research data lifecycle, but throughout the lifecycle of research projects as a whole.' (See http://repository.jisc.ac.uk/6379/16/Training_for_RDM_-_Comparative_european_approaches_May_2016.pdf)