cessda saw

Research and Innovation Action

CESSDA Strengthening and Widening

Project Number: 674939 Start Date of Project: 01/08/2015 Duration: 24 months

Deliverable 5.2 - PLATFORM FOR KNOWLEDGE EXCHANGE TO DEVELOP HUMAN AND TECHNICAL CAPITAL

Dissemination Level	PU
Due Date of Deliverable	(31/10/16)
Actual Submission Date	(03/01/17)
Work Package	WP 5 Strengthening and widening through global engagement
Task	T5.2
Type	Report
EC Approval Status	16 November 2017
Version	V1.0
Number of Pages	p.1 – p.15

Abstract: D5.2 describes where relevant content from non-CESSDA organizations can be localized and added to the Knowledge-sharing platform (KSP) that is being developed in Task 2.1.

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

History

Version	Date	Reason	Revised by
0.0	11/11/2016	Document created	Birger Jerlehag
0.1	12/11/2016	Text input	Birger Jerlehag
0.2	16/12/2016	1st draft finalised after comments on document	Birger Jerlehag Jonas Recker
0.3	19/12/2016	2nd draft finalised after comments and input from CESSDA MO	Birger Jerlehag Ivana Ilijasic Versic Iris Alfredsson
0.4	22/12/2016	3rd draft after review from DC Chair	Jindrich Krejci
1.0	29/12/2016	Final version after review from Coordinator	Ivana Ilijasic Versic

Author List

Organisation	Name	Contact Information
SND	Birger Jerlehag	Birger.Jerlehag@snd.gu.se
GESIS	Jonas Recker	Jonas.Recker@gesis.org
SND	Iris Alfredsson	iris.alfredsson@snd.gu.se
CESSDA MO	Ivana Ilijasic Versic	ivana.versic@cessda.net

Time Schedule before Delivery

Next Action	Deadline	Care of
First draft	12/11/15	SND
Second draft	15/12/16	SND
Peer review	16/12/2016	SND
Review by the Project Coordinator	19/12/2016	CESSDA AS
Review by the Chair of the Delivery Committee	21/12/2016	CSDA
Approval and Submission by the Project Coordinator to the European Commission	03/01/2017	CESSDA AS

EXECUTIVE SUMMARY

This report is the result of desk research and two reports, D2.1 and D2.2 from WP 2.1 in CESSDA SaW. Task 2.1 of the CESSDA SaW project will create a virtual knowledge-sharing platform (KSP) as a central point of access for the body of knowledge created by CESSDA partners. Task 5.2 will extend the content of the KSP with resources created by other relevant non-CESSDA institutions. This content will be added as external resources accessible via web links. They will be described by a set of metadata fields, harmonized with the metadata schema developed in T2.1, and thus possible to locate in the same way as uploaded resources. This report will help the editorial committee of the KSP to find relevant external content to add to the platform.

Abbreviations and Acronyms

SND	Swedish National Data Service
KSP	Knowledge-Sharing Platform
CV	Controlled Vocabulary
CRISP	Cluster of Research Infrastructures for Synergies in Physics
BioMedBridges	Biomedical sciences research infrastructures
ENVRI	Environmental Research Infrastructures
DASISH	Data Service Infrastructure for the Social Sciences and Humanities
IFDO	International Federation of Data Organizations for Social Science
ICPSR	Inter-University Consortium for Political and Social Research
ANDS	Australian National Data Service
ERIC	European Research Infrastructure Consortium
DPC	Digital Preservation Coalition
NCDD	Netherlands Coalition for Digital Preservation
DCC	Digital Curation Centre
RDA	Research-Data Alliance
IASSIST	International Association for Social Science Information Services and Technology
E-ARK	European Archival Records and Knowledge Preservation
FAIR	Guidelines for data sharing; Findable, Accessible, Interoperable, Reusable
ORCID	Open Researcher and Contributor ID

Table of contents

1. INTRODUCTION

The CESSDA partners has over time accumulated a great deal of knowledge. Much of this knowledge is captured in digital resources such as papers, presentations, reports, guidelines, and training materials. These resources are scattered across the different service providers, sometimes available from the webpage, but in other cases merely stored on internal servers. The still on-going work in Task 2.1 is to create a Knowledge-Sharing Platform as a central point of access for much of the knowledge resources that has been created. This main purpose of this Task is to point out different sources that can expand the Knowledge-Sharing Platform with relevant knowledge resources from non-CESSDA organizations. The report is aimed at the editorial committee that will be responsible for the content of the Knowledge-Sharing Platform, and will aid them in their work.

2. CONTENT OF THE KNOWLEDGE-SHARING PLATFORM

To determine what kind of resources the users of the KSP would like to see in the portal a survey was conducted in T2.1 (D2.1). It was answered by 47 Data Archive staff from 22 countries including all, at that point of time, CESSDA member countries except Austria. The respondents covered all areas of activity within a Data Archive. The table below lists how many of the respondents have marked a resource as useful for them in their work.

Table 1: Types of resources respondents would like to see in the portal

Type of resource	Percent
Training resources: Guidelines or manuals	95,7
Training resources: Webinars	80,4
Training resources: e-tutorials	80,4
Software tools	78,3
Presentations	76,1
Policy or advocacy documents	76,1
Reports	71,7
Scholarly publications (e.g. articles, collections, monographs)	47,8
User satisfaction surveys	47,8
Blog posts or other social media	34,8
Training resources: other	30,4
Other	8,7

Another important question in the survey was to find out which areas the respondents would like to be covered in the KSP. Each respondent could mark up to five topics.

Table 2: Top five topics that should be covered by the KSP

Topics	Percent
Metadata and standards	67,4
Data Access, Dissemination, Open Data	63
Research Data Management	60,9
Data Protection and Ethics	58,7
Data Processing and Documentation	52,2
Training	50
Archiving (Curation and Preservation)	41,3
Management of Data Archives	41,3
Persistent identifiers	21,7
Pre-ingest, Data Acquisition	17,4
Ingest	13
Other	4,3

To support both the depositors and the future editors/managers of the platform, a 'Collection and Platform Operation Policy' was developed in WP2.1. It describes the scope of the platform, i.e. which content will be collected, and defines selection criteria to support the review and publication process. The selection of external content will be done according to this Policy and in close collaboration with the editorial committee. The external content will in most cases be accessible via weblinks. The owners of the linked content will be informed about the KSP and the rationale behind it.

A metadata schema for the resources uploaded to the KSP is being developed in WP2.1 and will include the following fields to be used for linked external resources.

Metadata fields for T5.2 and for links to external resources:

RESOURCE TYPE:

Might be the same as used in the survey for T2.1 D2.1 CV

AREA/TOPIC:

Might be the same as used in the survey for T2.1 D2.1 CV

TARGET AUDIENCE:

This depends on the collection policy. Could include policy makers, repository staff, educators, researchers …

KEYWORDS:

To allow finer granularity when searching for resources. CV

CREATOR:

Institution and/or person(s) behind the resource.

LICENSE:

under which the resource can be used

LANGUAGE:

CONTACT PERSON:

TITLE OF THE RESOURCE:

DESCRIPTION OF THE RESOURCE (SHORT):

LINK TO THE RESOURCE:

PID:

If the owner has set one

CITATION: (if not stated elsewhere)

These fields will be harmonized with the final metadata schema of the KSP to ensure discoverability on the same level as uploaded resources. These tasks will be done in phase with the implementation process of the KSP.

Table 3: Platform Implementation Process

From	To	Description	Partner(s)
M11	M15	● Description of functional requirements ● Definition of data model ● Definition of metadata schema ● Definition of licensing scheme ● Set up da|ra cooperation for DOI registration	ADP, SND, GESIS, CESSDA AS
M17	M19	● Programming/implementation of platform	GESIS
M20	M22	● Beta-testing	All
M22	M22	● Launch of platform on the CESSDA webpage	GESIS, CESSDA AS

3. LOCALIZATION OF EXTERNAL CONTENT

Task 2.1 of the CESSDA SaW project will create a virtual knowledge-sharing platform (KSP) as a central point of access for knowledge resources created by CESSDA partners. These resources include papers, presentations, reports, guidelines and training materials covering most of the areas where CESSDA SPs are involved. However, this means that the resources to a large extent cover aspects of data management in the rather narrow and specific domain of Social Sciences. CESSDA-SaW is all about strengthening existing SPs and widening of both the membership number and the different types of data being managed by the SPs. This means new tasks in the future, like assisting a fledgeling SP to become a mature institution or how to handle and preserve digital assets different from the traditional ones. T5.2 is about expanding the content of the KSP with resources from other domains, both thematic and geographic, than those coming from the European SSH area that are produced by the CESSDA service providers.
Besides externally funded projects there are several internal "CESSDA Tasks" where the SPs are collaborating in strengthening both the consortia and the individual Service Providers in how to best do their work and develop new services. These tasks span over several subjects such as certification, metadata standards, persistent identification, tools for dissemination and several others. These task groups must have come across a lot of valuable external resources during their work, resources that may well fit in in the KSP. The editorial committee for the KSP will be asked to contact the different Task Groups to harvest such resources.

There are a lot of proficient Social Science data archives and data libraries beside of CESSDA. These share the same rationale as the CESSDA service providers do, use the same metadata

standards and techniques and serve the same communities as CESSDA do. Some of them produce a great deal of the educational material that is used in the data archiving world.

There are several other research domains besides the Social Sciences and all of them are faced with the same chores when it comes to collect, preserve and disseminate research data. The four EC-funded cluster projects (**CRISP** – physics, **BioMedBridges** – life sciences, **ENVRI** – environmental sciences, and **DASISH** – social sciences and humanities) and their successors have created a lot of knowledge resources that are of interest to CESSDA and the SPs when it comes to managing and disseminating research data to a designated community and how to build and manage research infrastructures.

Several knowledge centres have been created over the years and they house large collections of informational and educational material spanning over all aspects of data management from the policy level to the practical hands on tutorials. As they already are knowledge-sharing platforms of a kind the task here will be how make their resources discoverable in the KSP.

Data Repositories are often 'one of a kind' nationally so colleagues are often found in other countries and continents. This have led to the formation of several interest groups where practitioners from different fields meet and share their know-how and best practises. Presentations and other outcome from these gatherings contains good, often real-life, examples of how to solve problems or simplify tasks and chores in the daily work.

Broadening the view and looking outside the academic institutions there are a multitude of other institutions, mostly on the national level, working with the same goals in sight. They are mostly found within the LAM sector, (Libraries Archives and Museums) where they work with both creating and preserving digital heritage assets and making them accessible to the public. They themself and their collaborative organizations produce resources that are of interest for the KSP.

Running and managing a data repository is dependant of having access to relevant technical equipment such as reliable networks, trustworthy backup systems and of course computers. The hardware parts are the same as in any other business but there the similarities end. The lack of useful and efficient software tools in the data curation field has lead to a multitude of homegrown solutions of which a few has reached a wider audience and evolved into widely used applications. To ensure continued development of and open source access to the tools they are often placed in a not-for-profit organization owned by the user community and financed by membership fees. These companies are a good place to look to find knowledge resources about the more technical know-how needed to run a repository.

4. SUMMARY

Finding relevant external content will start with an assessment of resources published by institutions with the mission of simplifying access to and usage of (research) data. These fall mainly into one of five categories;

- Data Repositories
- Collaborative organizations
- Interest groups
- National heritage institutions
- Business organizations

Data Repositories

An obvious starting point will be at the non-CESSDA members of the International Federation of Data Organizations for Social Science (IFDO) which is a worldwide organization for data services. Major actors like ICPSR and ANDS are very active in creating knowledge resources for the repository domain. Other sources are other ERIC:s with data management responsibilities.

Collaborative organizations

Lack of knowledge like technical skills and knowhow in data management and data archiving led to the formation of knowledge centres financed by research institutions, funders and other stakeholders. Their mission was to create and produce educational material, hold workshops and promote data management. Some of the major ones are the Digital Preservation Coalition (DPC), Coalition for Digital Preservation (NCDD) and the Digital Curation Centre (DCC).

Interest groups (based on individual membership)

Most interesting for the CESSDA community are the International Association for Social Science Information Services and Technology (IASSIST) and the Research Data Alliance (RDA). Especially RDA is an important organisation to follow closely as output and recommendations from the workgroups are adopted by many institutions and turn into factual standards and best practises.

National heritage institutions

Many National Archives and National Libraries are faced with similar tasks as Data Repositories. They collaborate in many ways and some of their projects like E-ARK have produced both tools and know-how that are of interest for the CESSDA community.

Business organizations

There are many, mostly not-for-profit, companies specialising on specific parts of the chain of tasks that will make research data usable. These tasks are mostly on the more technical side of data management. DataCite for instance specialises on Persistent Identifiers for research data to make them citable in the same way publications are, ORCID creates unique identities for individuals where they can store their CVs, while others like DuraSpace creates technical solutions and platforms for handling and storing digital resources.

APPENDIX 1

Links

BioMedBridges	http://www.biomedbridges.eu
ENVRI	http://envri.eu
DASISH	http://dasish.eu
IFDO	http://ifdo.org/wordpress/
ICPSR	https://www.icpsr.umich.edu/icpsrweb/
ANDS	http://www.ands.org.au
DPC	http://www.dpconline.org
NCDD	http://www.ncdd.nl/en/
DCC	http://www.dcc.ac.uk
IASSIST	http://www.iassistdata.org
RDA	https://www.rd-alliance.org
E-ARK	http://www.eark-project.com
DataCite	https://www.datacite.org
ORCID	http://orcid.org
DuraSpace	http://www.duraspace.org

APPENDIX 2

List of tables: