

# Guide for Developing National Data Service Plans

CESSDA - Consortium of European Social Sciences Data Archives

**Version :** 1.0

**Date :** November 2016

**Authors :** Christina Bornatici (FORS), Brian Kleiner (FORS)

**Contributing Partners  
and Editors :** Trond Kvamme (NSD), Vigdis Kvalheim (NSD), Aleksandra Bradic  
Martinovic (IEN), Marijana Glavica (FFZG)

**Citation :** Bornatici Christina and Kleiner Brian (2016), *Guide for Developing  
National Data Service Plans*, CESSDA SaW (H2020), v1.0.

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



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 particular purpose. The user thereof uses the information at his/ her sole risk and liability.

# Table of contents

## [Introduction](#)

### [1. General features of the data service](#)

[1.1. Definition of organisation](#)

[1.2. Scope of collection](#)

[1.3. Services and activities](#)

[1.4. Beneficiaries](#)

[1.5. Governance structure](#)

[1.6. Financing schemes](#)

[1.7. Mission](#)

### [2. Description of resources and challenges](#)

[2.1. Host institution of the future data service](#)

[2.2. Human resources and internal structure](#)

[2.3. Partner support and cooperation](#)

[2.4. Specific local resources and challenges](#)

### [3. Going further: Concretely planning for the data service](#)

[3.1. Establishment plan \(before day 1\)](#)

[3.2. Operational plan \(after day 1\)](#)

## [Conclusion and next steps](#)

## [Annex 1 - Available sources of information](#)

## [Annex 2 - The CESSDA Capability Development Model as a planning tool](#)

## [Glossary](#)

# Introduction

This document serves to assist interested actors or “country teams” in developing plans for establishing national data services for the social sciences. By country teams we mean people and/or institutions working towards the establishment of a national data service and taking part in the elaboration of the development plan. This could be national ministries, research councils, universities, or other organisations. The country team may play a key role in the establishment of the data service, but also after it is in operation.

## **The benefits of establishing a national data service**

From a research policy perspective there are benefits of national data services to scientific communities that justify the needed investments. First, through their core activities – data preservation and dissemination – data services make possible long-term access to and wider and more effective use of existing data. This means that publicly-funded data are used more fully beyond their original purposes, and that research funders can concentrate on financing fewer new data collections. Indeed, data gathered for original research can be further exploited by “secondary” users for new insights and scientific contributions. Recent work by Beagrie et al.<sup>1</sup> demonstrates the economic benefits of national data services.

In addition, because data are often not appropriately preserved by researchers, they are lost or become unusable over time. This represents a loss of potential knowledge and cultural heritage. The data service, through appropriate procedures, policies, tools and infrastructures, ensures long-term preservation of data.

Second, data services strengthen research practice and quality by rendering research more transparent and open for replicability. By making data and related documentation available for scrutiny, original research can be tested and evaluated, a key pillar of the scientific method. Also, knowing in advance that the data could be used by others in the future is a way to ensure better quality in methods.

Third, more and more funders and journals are requiring that data be deposited at and available to researchers for secondary use and replication – having archives and repositories as places to store and disseminate data makes possible these requirements.

Finally, the easy availability of secondary data brings value to university teachers and students, who benefit from having data to train and illustrate methods and concepts in their courses and projects.

## **Some examples of data services**

Before beginning with planning, it should be kept in mind that there are different ways to organise a national data service for the social sciences, and that the final decision regarding the makeup of a

---

<sup>1</sup> Beagrie Neil, John Houghton, Anna Palaiologk, and Peter Williams. (2012). *Economic Impact Evaluation of the Economic and Social Data Service*. Available <http://www.esrc.ac.uk/files/research/research-and-impact-evaluation/economic-impact-evaluation-of-the-economic-and-social-data-service/>

data service depends on specific existing conditions and possibilities within a country. Among CESSDA<sup>2</sup> service providers, there are many similarities and differences. Indeed, even if all CESSDA service providers archive and disseminate social sciences data, they differ in their policies, their affiliations with host institutions, their financing schemes, their additional services, activities and types of data, and their resources. Table 1 shows some existing examples that illustrate the diverse possible forms of organisation of national data services<sup>3</sup>.

**Table 1:** Diverse characteristics of five CESSDA member service providers

	<b>Czech Republic - CSDA -</b>	<b>Netherlands - DANS -</b>	<b>Slovenia - ADP -</b>	<b>Switzerland - FORS -</b>	<b>United Kingdom - UKDA -</b>
<b>Host institution</b>	Department of the Institute of Sociology of the Academy of Sciences of the Czech Republic	Institute of Royal Netherlands Academy of Arts and Sciences (KNAW) and Organisation for Scientific Research (NWO)	Organisational unit of the Institute of Social Sciences within the Faculty of Social Sciences at the University of Ljubljana	Independent foundation hosted by the University of Lausanne	Department of the University of Essex
<b>Main funders</b>	The Ministry of Education, Youth and Sports, and the Institute of Sociology of the Czech Academy of Sciences	The Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO)	The Slovenian Research Agency (ARRS)	The State Secretariat for Education, Research, and Innovation, the Swiss National Science Foundation, and the University of Lausanne	The Economic and Social Research Council
<b>Data service activities</b>	Acquisition, preservation and dissemination of data for secondary analysis	Acquisition, preservation and dissemination of data for secondary analysis; data service for data during research and after research, portal with research information, training and consultancy in data management and certification of archives	Acquisition, preservation and dissemination of data for secondary analysis	Acquisition, preservation and dissemination of data for secondary analysis, consulting services for researchers	Acquisition, preservation and dissemination of data for secondary analysis; user support for data management planning and data use; thesaurus development

<sup>2</sup> Consortium of European Social Science Data Archives, <http://cessda.net/>

<sup>3</sup> Please note that the authors have selected these examples, among all CESSDA service providers, only following criteria such as diversity of affiliation, financing scheme, activities and number of staff members. To have a view of all CESSDA service providers, please visit CESSDA website: <http://cessda.net/National-Data-Services/CESSDA-Members>.

<b>Other activities</b>	Providing technical and organisational support for large-scale survey research programmes (ISSP, ESS, or the newly established Czech Household Panel Survey), mapping and analysing available sources of data and facilitating their archiving, methodological research and data quality research, data management expertise	Research on sustained access to digital information, services and participation in (inter)national projects and networks	Support and training on research data management, promotion of data usage in teaching	Production of survey data, including national and international surveys, research in empirical social sciences	Research relating to large datasets, data curation and archive certification, disclosure control for personal data, information security and ethical research, expertise on Research Data Management
<b>Type of data</b>	Social sciences data (mainly survey data, small set of qualitative data)	Social and behavioural sciences data, history data, archeology data, language and literature data, life sciences and medicine data	Social sciences data	Social sciences data	Social sciences data, historical data, census
<b>Staff (FTE)</b>	6.6 in total	8.4 for data service activities, 42 in total	6 in total	9 for data service activities, 32.6 in total	59 in total

Of course the selection of the type of organisation will also depend on the specific country context (e.g., politics, traditions), as well as on available resources. More disciplines, activities, and available services will require more resources. It should also be noted that smaller data services (2-3 people) can function well enough with limited resources, as long as the services offered are limited to the basic activities of data archiving and dissemination. On the other hand, data services with limited resources that aim for CESSDA membership might have more difficulties complying with CESSDA service provider requirements.

### How to use this guide

The guide is divided into three main sections. The first two address larger strategic considerations in conceiving and establishing a national data service. The first section aims to assist the country teams in developing a “concept” for a future national data service, that is, the model and key features that are to be put into place. This includes defining: the nature of the organisation; the overall mission; the scope of the collection of data; the offered services and activities; the beneficiaries; and the governance structure. Each subsection (1.1, 1.2, etc.) should be developed as a brief text in the final

development plan. For each subsection, there is an explanation of what information should be included and a list of checkpoints to guide the country teams in their choices and considerations.

The second section of the guide helps country teams to evaluate the concept from section one in relation to realities on the ground, including probable resources and challenges for setting up and maintaining a viable data service over time. This includes subsections on identifying the host institution for the data service, defining the human resources required (number and types of staff, internal structure), and identifying potential partners among the existing network of relevant national and international organisations. The country teams will also have the possibility to develop other aspects that are crucial for them.

The third section of the guide is not compulsory, and is meant to be completed by those countries that are further advanced in their intentions to establish national data services for the social sciences. This section helps country teams to make a concrete step by step plan for new data services, both up to “day one” of operations, and after “day one”.

These three sections will constitute the basis of your national development plan.

---

## **Acknowledgments**

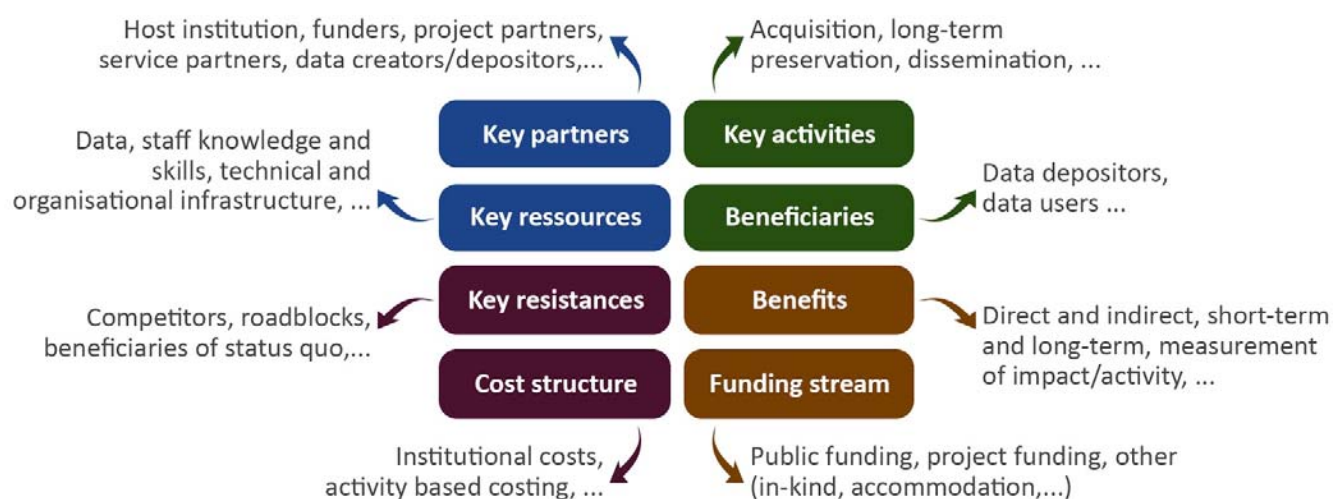
This document includes elements from a variety of existing sources, including from the FP7 project SERSCIDA, CESSDA, and individual CESSDA member service providers.

# 1. General features of the data service

As part of your national development plan, you should first outline the general desired features of the data service. In this first section, the country team should decide on and describe the main objectives and structure of the data service. The resulting concept will define the foundation and help to plan further and more concretely (in sections 2 and 3).

## Brainstorming activity 1

Before filling in the next subsections in depth, we would like to invite you to brainstorm with your team on the following key elements of the (new) organisation and its network. This exercise is intended to help you to develop your ideas related to different relevant aspects, and to think about new and previously unimagined possibilities. It could then inform how you complete the two main sections of the plan for your country. The results of the brainstorming do not need to be part of your national development plan.<sup>4</sup>



*Selected parts of the Archive Development Canvas, developed from Business Model Canvas [www.businessmodelgeneration.com](http://www.businessmodelgeneration.com) for the CESSDA SaW Project by Charles Beagrie © Ltd 2016.*

<sup>4</sup> Further guidance on the complete Archive Development Canvas and additional supporting tools can be found in the CESSDA SaW Cost-Benefit Funding Advocacy Toolkit (see Annex 1 - CESSDA SaW).

## 1.1. Definition of organisation

This subsection of the development plan should introduce the data service with respect to the following:

- Name of the data service
- New independent organisation or housed within a larger organisation
- Main purpose

## 1.2. Scope of collection

This subsection should define the nature and scope of the data that will be included in the new data service collection, with respect to the following:

- Overall fields (social sciences, humanities, medicine)
- Specific disciplines to be prioritised (e.g., sociology, political science, education science, economics, etc.)
- Type of data: quantitative and/or qualitative, contemporary and/or historical, etc.
- Size of data: “big” data or traditional-scale data

## 1.3. Services and activities

This subsection should name and describe the services to be offered by the data service. Services and activities can be scaled and/or developed progressively. Some possibilities follow:

### Main services

- Solicitation and acquisition of data from data producers
- Curation and long-term preservation of data
- Online catalogue of archived data
- Access to and dissemination of data for use in secondary analysis

### Additional services and activities

- Materials and trainings for data depositors (how to prepare data for inclusion in the archive, how to deposit prepared data)
- Materials and trainings for data users (how to find data in the catalogue, how to download data)
- Support, assistance, and advice for data depositors and data users
- Promotional activities and outreach to users
- Data management training for researchers
- Regional collaboration and establishment of networks of service users



### Optional services and activities

- Creation and maintenance of a research inventory of existing research projects in the country (with or without data)
- Access to online tools for data analysis and variable exploration - creating tables, graphs and other analytical results
- Purchased access to foreign databases (e.g., ICPSR, LIS) for national researchers
- Wider promotion of “Open Access” to data

## 1.4. Beneficiaries

This subsection should explain who will be the primary recipients of the various services. The beneficiaries or users can be groups such as:

- University students (doctoral, master, and/or bachelor)
- University researchers
- University teachers
- Pre-university education - secondary school teachers and students
- Researchers from other public institutions (e.g., scientific institutes, government agencies)
- Researchers from private institutions (NGOs, private research companies, newspapers, businesses, associations, etc.)
- Other: policymakers, journalists, general public, etc.

## 1.5. Governance structure

This subsection of the plan should define the main bodies and key stakeholders that will be involved in the governance of the data service, including their roles, responsibilities, and composition.

Examples of main bodies and their stakeholders:

- An executive board
  - Data service director/manager
  - Representatives of host institution(s)
- An oversight board
  - Representatives of ministry
  - Representatives of national statistical institute
  - Representatives of national archives
  - Representatives of research councils or national science foundations
  - Representatives of universities (rector’s office) or university faculties (faculty deans)

- An advisory board
  - Leading researchers of international standing in (social) sciences
  - Representatives of CESSDA member service providers
  - IT specialists

## 1.6. Financing schemes

This subsection should define the specific financing scheme for the establishment process, as well as the general financing scheme to run the data service once it has been established. The financing scheme should involve a description of potential funding sources for different parts of the new data service.

Funding sources could be:

- Hosting institution
- National institutions (e.g. national stakeholders and partners or main beneficiaries)
- International funds and supports (e.g. European Union fundings, etc.)
- Other revenue/funding streams

## 1.7. Mission

Develop a brief mission statement (no more than 1 page) that defines:

- The overall strategic goals
- A larger vision or philosophy
- The expected value and impact on the national scientific community
- The intended future direction of the data service

In addition, the mission statement should summarise the key elements of the earlier subsections, specifically:

- The scope of the data collection
- The main services and activities
- The beneficiaries of the services

→ As it will serve as a communication and promotional tool, it should be simple, clear and straight to the point.

## 2. Description of resources and challenges

This second section of the guide aims to help country teams consider how the concepts and general features of the data service outlined in section one can be implemented in practice, taking into consideration real local and institutional resource conditions. Subsections 2.1-2.3 begin with a set of leading questions. Subsection 2.4 gives the country team the opportunity to go beyond the generalities to address the specific challenges of the local context. However, before working through this section, start with the brainstorming activity below.

### Brainstorming activity 2

In section 1 you defined the general features of a new data service in your country. In this second brainstorming activity, review the points 1.1-1.7 of the previous section while thinking about the strengths and weaknesses of your concepts, concretely in relation to your local situation. You should address potential obstacles to the establishment of a data service and the feasible solutions to overcome them. Some examples:

- 1.3: According to the local resources/situation, in terms of actual and/or potential skills, infrastructure, funding, etc., are the selected services/activities feasible? Accordingly which services/activities should be prioritised? This prioritisation of the services will maybe limit at first the list of beneficiaries. Then who will be your first beneficiaries?
- 1.3: According to the selected activities, what skills and knowledge are necessary? Does the staff, someone in the host institution or in the country already have it? How to develop the knowledge and skills? Is there a local/international partner that could play this role?
- 1.3 - 1.4: How are the selected activities useful for the beneficiaries? Can they be improved to better address their needs? Are these needs somehow already addressed by another institution?
- 1.5: Does the planned composition of the boards include important local stakeholders (in the social science research and academic area, potential funding institutions, etc.)? Is the diversity and the number of involved institutions appropriate? How to contact those institution/people? If you have already contact with institutions that could be interesting partners or board members, are they in favour of the establishment of a data service?
- 1.6: Are the potential funding partners reliable? What would be their funding conditions? Would the funding be secured? For how long? Is the funding based on projects?

The results of this second brainstorming activity will help you to complete the next subsections (2.1-2.3) and to develop subsection 2.4 based on your local situation.

## 2.1. Host institution of the future data service

---

### *Leading questions:*

- Which institutions could host the future data service?
- What are the strengths and weaknesses of each potential hosting institution in relation to the data service concept and the given criteria?
- Could the potential hosts provide a long-term commitment and sustainability?
- Once the host has been determined, what would be the terms of agreement and obligations of the host institution, and would these be acceptable?
- According to the strengths and weaknesses of the selected host institution and the terms of agreement, what are the next steps to take?

### *This subsection of the guide will help the country team to:*

- Select an appropriate host institution
  - Evaluate the strengths and weaknesses of the host institution in relation to the general features outlined in section one
- 

To begin, some country teams may already have identified a probable host institution for a national data service. If this is the case, then the guide can serve to reinforce or else to bring into question this choice. If not, then the suggestions that follow may be particularly useful for selecting an appropriate institution. In the case that a country team plans to establish a new independent organisation, this subsection can be either omitted or used to check that the new organisation has/develops some of the following criteria.

In general, appropriate host institutions should have the capacity to run a successful future data service, even if they currently lack the know-how. Here are some criteria for determining whether a given institution is sufficiently positioned and equipped to become a national data service for the social sciences:

- Research institution in the social sciences, preferably public:** The institution should somehow be involved in data-oriented social sciences research. Public, non-profit institutions are preferred, since these will by nature be in the public interest.
- With existing technical infrastructure:** The institution should be able to rely on good technical infrastructure (proper hardware, software, servers, network and telecommunications), preferably with existing data infrastructure (processes and tools that support data collection, management, storage and dissemination) and with IT expertise and support. A trustworthy technical infrastructure should be in place with suitable measures to ensure the long-term usability of data collections (security and risk management).
- Strong leadership:** Establishing a new data service requires an intensive effort with a strongly committed institution and a leader who will lobby relevant stakeholders.
- Existing staff:** The institution should already have available staff, with an administrative service.

- Interest and motivation:** The institution should be strongly committed to the idea of establishing a national data service, and the individuals involved should be highly motivated.
- Integrated in existing national networks:** The institution should be well integrated into the national research and service network, including links to universities, research institutes, national statistical institutes, and researchers in different disciplines.
- Good relations with ministries/research councils:** There should be existing and positive relations with national ministries or research councils.
- Recognition and important role in the national social sciences community:** The institution should be recognised for its excellence by all stakeholders and the community of researchers.

Also desirable, if possible:

- Familiarity with data archiving issues**
- Existing data management skills**

Of course, no institution can fulfil all of these criteria, and so the goal will be to identify the ones that most closely approximate the ideal.

→ If the host institution has been determined, and a formal preliminary agreement with the host has been reached, then this choice should be integrated into the national development plan, including the name of the institution and a description of the agreed upon terms. A brief justification of the choice of the host institution (based among others on the previous characteristics) would strengthen the plan.

→ If there is no host institution, nor agreement, then this part of the development plan should include the strengths and weaknesses of each potential hosting institution, and according to the assessment which institution would be the first choice.

## 2.2. Human resources and internal structure

---

### *Leading questions:*

- What kinds of skills and capacities are necessary at first or can be developed later on?
- How many staff members are needed in order to successfully carry out the objectives defined in section one?
- How many staff members are really possible given available resources, and what are the implications of this with respect to the services and goals?
- In terms of human resources, what is the bare minimum needed to operate a national data service in the country?
- What kind of internal structure will be there in place?
- What will be the roles and responsibilities of different staff members?

### *This subsection will help the country team to:*

- Determine how many staff members will be needed to adequately meet the main objectives outlined in section one.
  - Define the minimum position types needed, with designated roles and responsibilities.
  - Evaluate and determine the skills and capacities for the positions, identifying areas to be developed.
  - Describe the internal structure of the data service, including any hierarchical relations.
- 

The number of staff members to hire for the data service will depend of course on the available resources and the expected services. Additional resources could also be required to meet the demands of later CESSDA memberships (e.g., Data Seal of Approval). At the beginning, a new data service should at a minimum count on including a manager and a data specialist.

The manager would be responsible for ensuring that the services are available and functioning in relation to the institutional goals. He/she would also be responsible for external relations and collaborations, with respect to stakeholders, client institutions, and international partner data services. This person should ideally have good management skills, a strong national network, and be able to speak passionately and convincingly about open data and research transparency.

A data specialist will be responsible for all aspects of the archiving workflow, including data solicitation, curation, and dissemination. The data specialist should optimally have a background in data sciences or social sciences, but could also have a background in related disciplines, such as history, statistics, or information sciences. The data specialist would be the point of contact for users of the services. If there is no possibility for hiring an IT specialist, then the data specialist should also have strong technical skills.

If resources permit, and the host institution cannot already provide it from existing staff, we recommend hiring an IT specialist who would be responsible for putting into place the necessary servers, hardware, and software, and maintaining the data service website.

If additional resources are available, other staff types could be brought in. For example, a more junior person could be hired for data entry and cleaning activities, thus freeing up the data specialist for other types of work (like outreach, data management training). If not existing already within the host institution, a larger data service would benefit from a financial officer, a human resources officer, and/or a communications officer, as well as clerical/administrative support.

Here is a list of the profiles and activities for the three main staff types:

### **Data service manager**

Profile: Management skills; strong communication skills; social sciences background; proponent of open data and research transparency

Main activities:

- Management of the data service
- Setting of policy and procedural rules
- Relations with national and international stakeholders and partners
- Data service promotion towards the research community (the beneficiaries)

### **Data specialist**

Profile: Data science or archiving background, and/or social sciences/statistical background; data management skills

Main activities:

- Solicitation of data
- Data and documentation ingest
- Data curation and preservation
- Dissemination of the archived data
- Contracts creation for data depositors and data users (maybe with the help of a legal advisor)
- Information and advice to researchers and students

### **IT specialist**

Profile: Systems administration, software programming and database management, software development and maintenance, overall understanding of data exchange protocols in a networked environment and security issues associated with it

Main activities:

- Install and maintain servers
- Install and maintain hardware for staff
- Install and maintain software to archiving and dissemination of data
- Develop and maintain data service website

→ Based on these possibilities, in this subsection of the plan you should describe the staff that you intend to work in the data service, including how many, their roles and responsibilities, and their hierarchical relations (if any).

## 2.3. Partner support and cooperation

---

### *Leading questions:*

- Which institutions could be potential national/international partners?
- What could they provide for the data service (at its establishment, in the longer term)?
- How could they contribute to overcome the current weaknesses or challenges?
- What could the national data service do for the partners?
- What synergies can be put into practice with partner and supporting institutions?
- How will the national data service articulate with other national institutions?
- How will the national data service be oriented internationally?
- How will the national data service collaborate with and benefit from its relationship with CESSDA and CESSDA partners?

### *This subsection provides guidance on:*

- Identifying national stakeholders and relevant partner institutions
  - Identifying international partners and supporting institutions
  - Determining how the data service can benefit from these relationships by finding synergies and areas for cooperation.
- 

### **National stakeholders and partners**

On the national level it is very important for the data archive to be recognised as a valuable national service by all relevant institutions that are part of the science community and higher education system. These are usually institutions that either fund research or research infrastructure or else serve functions that are related to the goals and work of the data service. For instance, national statistical institutes produce micro-data that can be of great value to researchers for secondary analyses. National archives employ many of the same techniques and practices as data services when it comes to the long-term preservation of data. Universities support, train, and produce researchers that are both data providers to archives, but also consumers of archived data.

Moreover, stakeholders, especially the main science policy and financing institutions in a country should be fully informed of the benefits of long-term preservation of research data, as well as the benefits connected with open access to research data. Data services should do everything possible to communicate with these institutions about these benefits and their own related activities. The national development plan should therefore outline how this can be achieved through formal agreements and active communication with relevant national institutions and individuals.

The relevant stakeholders and partner institutions are country-specific, but these could be:

- Ministries that oversee high education, research, and more specifically the social sciences
- Research councils or national science foundations
- National statistical institutes
- National archives
- Universities or university faculties



An effective means of establishing long-term ties with key stakeholders is to put into place an institutional Oversight Board, where relevant elected representatives would be convened once or twice per year to be informed about and to approve the ongoing activities of the data service. Experience shows that such persons, if carefully selected, can become important supporters, allies, and proponents of data services.

In addition, the data service should reach out to and meet the needs of the scientific community. Some data services have established committees of researchers (both data producers and data consumers) that meet occasionally to discuss their practices and needs with respect to services offered by the data service. In this way the data service is more likely to orient its services towards what is of primary importance to its beneficiaries.

→ In this subsection of the national development plan, you should describe how your data service will orient itself at the national level, and what forms of collaboration it may enter into with different partners.

### **International stakeholders and partners**

On the international level, cooperation with existing European data services is crucial because of their extensive experience in archiving procedures and workflows, and research data management. New data services can benefit significantly from this wide body of knowledge and experience.

Such forms of cooperation are facilitated by CESSDA, the Consortium of European Social Science Data Archives. CESSDA and its member service providers organise free workshops, trainings, and web tutorials on various subjects (e.g. data archives and digital preservation, on the Open Archival Information System (OAIS), data security and data protection), in order to build and improve the skills of members (see Annex 1 and Annex 2). Becoming a CESSDA service provider reinforces and signals the quality of the data service's work, as only countries with established data services following international standards and best practices may become member. CESSDA membership also ensures data availability and data sharing throughout the international community.

Also at the European level, data services have in recent years greatly benefited from various large-scale projects funded by the European Commission, such as the FP7 projects Data without Boundaries and SERSCIDA, and the current H2020 projects Big Data Europe, SERISS, and CESSDA SaW. Further, close cooperation of data services with ESFRI research infrastructures as well as other international projects and initiatives dealing with managing and/or publishing research data, such as DARIAH, SHARE, ESS, OpenAIREPlus, and EUDAT, helps to ensure development of an efficient pan-European data sharing and exchange ecosystem.

→ In this part, you should describe how your data service intends to orient itself internationally, and the extent to which it will aim to play a role in these international projects, consortia, infrastructures, and activities.

## 2.4. Specific local resources and challenges

This subsection is not compulsory, but the country teams are highly encouraged to identify one or more specific local challenges and to explore solutions to overcome them. These specific challenges or resources may derive from both brainstorming activities and from answers given in previous subsections of the national development plan. This is the place to assess more in detail the local obstacles and also opportunities. Some examples:

- ❑ Contacts or attempts of contact with the appropriate ministry, the research councils and other partners: contacts already made or planned, the results of these contacts, how to go further, etc.
- ❑ Links between the services and costs: the chosen services lead to different activities, needs for skills and knowledge, budget, and partners (to help execute the services or funding partners), etc.

# 3. Going further: Concretely planning for the data service

This third section of the development plan, which is not compulsory, outlines domains and specific actions that need to be planned for concretely, in order to establish and operate a new data service. For each domain and action, country teams will need to consider the resources required, timeframes, dependencies, and steps.

The plan outline is divided into two stages, one for establishment (before “day 1”), the other for operation and maintenance after the data service begins (after “day1”). This division is, however, not strict, since some activities that are part of the first stage should continue in the second stage (e.g. Develop knowledge and skills of staff through ongoing training), and will also depend on the local situation.

CESSDA and its service providers are available to provide guidance for new data services in implementing their plans (e.g. with policy and procedures development, trainings, etc.).

## 3.1. Establishment plan (before day 1)

### Organisation and internal structure

- Confirm identity of institution(s) to assume the data service (data service, host institution(s), funding institution(s), etc.) (from sections 1 and 2, mainly 1.1, 1.6 and 2.1)
- Define the data service and scope of collection (from section 1, mainly 1.1, 1.2 and 1.7)
- Establish the set of services to be provided (from subsection 1.3)
- Develop a sustainable finance of establishing process and sustainable financing scheme (from subsection 1.6)
- Develop a governance structure (e.g., oversight board, scientific board) (from subsection 1.5)
- Define internal structure of organisation, including an organisational chart (from subsection 2.2)
- Identify work space/facility
- Create a classification scheme of data types to be archived (according to risk level) and corresponding distribution mechanisms
- Define stakeholders and partner institutions (from sections 1 and 2, mainly 1.6, 2.1 and 2.3)
- Establish formal communication channels with and between stakeholders and partner institutions
- Create institutional brand and logo
- Create website
- Begin promotional activities to establish continuous visibility among key stakeholders

## Human resources

- Define roles and responsibilities for likely staff positions (i.e., who does what) (from subsection 2.2)
- Prepare job descriptions, advertise, interview, issue contracts (*near to “day one”, once funding has been secured*)
- Define institutional rules and regulations for staff (e.g., regarding sick leave) (*but only needed if a new institution is envisioned*)
- Develop knowledge and skills of staff through ongoing training

## Technical infrastructure

- Select hardware, including servers, desktop computers, security and backup systems
- Select and prepare software and tools, including statistical analysis programs (like SPSS, STATA), Nesstar
- Select databases and archiving system tools

## Policies, quality control procedures and workflows

- Establish data policies and written protocols for data management and access
- Create a data preservation and succession plan
- Develop policy and plans for data backup and security
- Study legal background and develop contracts (create data deposit agreement as well as end user licenses)
- Ensure standardisation of policy and tools (including compliance with CESSDA and Data Seal of Approval (DSA))

## 3.2. Operational plan (after day 1)

### Implementation and maintenance of technical systems

- Install servers
- Install desktop computers for staff
- Install security and backup systems
- License and install appropriate software (e.g., SPSS, STATA, R, SAS, NVivo)
- License and install tools (e.g., self-deposit software, Nesstar)
- Install databases and catalogues

### **Data acquisition and curation (solicit and process research data)**

- Identify existing research projects with potentially valuable data
- Contact data producers and request deposit of their data
- Process data that arrive through archiving system

### **Activity report, outreach and promotion**

- Create database of users
- Develop basic performance monitoring system (i.e., indicators, measures for monitoring change over time)
- Annual reporting
- Identify target publics for outreach and promotional activities (e.g., research community, policy-makers, media)
- Develop and conduct promotional activities, both short- and long-term
- Develop guides for researchers and training events to promote good practice and data sharing

### **Networking and internal and external communications (e.g., newsletters, email lists)**

- Establish relevant and effective forms of communication with target publics
- Carry out communications on ongoing basis
- Participate in relevant international conferences and workshops
- Join and participate in CESSDA and CESSDA projects and activities

### **Ensuring sustainability and working toward growth**

- Conduct regular meetings with high-level ministry officials
- Keep the public and ministries aware of data service activities and benefits
- Search for new funding opportunities, domestic and foreign
- Submit proposals for grants
- Improve existing services and create new ones
- Work toward fulfilling CESSDA service provider requirements

# Conclusion and next steps

This document has aimed to assist country teams in developing plans for establishing national data services in the social sciences. In order to do so, the country teams can count on the support of CESSDA and its partners. Once the plans are complete, the country teams and/or new data services can also count on the support of the CESSDA main office in Bergen, Norway, especially if the intention is to move toward CESSDA country membership.

But crafting plans for national data services is just a first step. Countries seeking to establish national data services and obtain CESSDA membership must adopt strategic and creative approaches to improve their chances of success. Towards this end, the annex 2 to the Status of CESSDA<sup>5</sup> indicates for the country teams or new national data services what should be achieved in order to become CESSDA service providers. Also, the CESSDA Capability Development Model<sup>6</sup> (see Annex 2) will be the basis upon which an assessment of social sciences service provision is made, and can help in the improvement of the capabilities of existing and future CESSDA service providers.

The CESSDA SaW project and its many national partners encourage country teams to join CESSDA, the growing European network of data services in the social sciences, which can assist in identifying and employing effective solutions and synergies.

---

<sup>5</sup> CESSDA. 2013(2014). "Annex 2: Obligations of Service Providers". In Annexes to the Statutes. Available <http://cessda.net/About-us/Documents>.

<sup>6</sup> The CESSDA SaW Capability Development Model (CESSDA-CDM) is available on the CESSDA website: <http://cessda.net/eng/CESSDA-Services/Projects/CESSDA-SaW/Work-Packages/WP3/CESSDA-CDM>.

# Annex 1 - Available sources of information

Various resources are (freely) available and can assist the country team and/or the new data service in building and improving its knowledge and skills. The resources (tools, training materials, webinars, etc.) have to be gathered in various websites.

**CESSDA** - <http://cessda.net/>

CESSDA Training section (<http://cessda.net/CESSDA-Training>) provides information on the upcoming trainings on various topics. There are also various online materials on the main activities of a data service archive in the social sciences: research data management, data archives and digital preservation, data sharing and security. Customised trainings, given by GESIS, are also available on demand. Moreover, there are working groups (<http://cessda.net/About-us/CESSDA-Working-Groups>) to guide CESSDA members or seeking members regarding especially technical services and trust issues. There are also workshops on these topics. Finally, to complete these resources some tools (i.e. Thesaurus, DDI, Nesstar) are presented (<http://cessda.net/eng/CESSDA-Services/Resources/Tools>).

## CESSDA member service providers

Information on collection and preservation policies, data protection and distribution policy, as well as ideas for mission statements can be found on the CESSDA member service providers websites (<http://cessda.net/eng/National-Data-Services/CESSDA-Members>).

- **UKDA**

Various resources are available on the UKDA website, for example in the Manage data section (<https://www.ukdataservice.ac.uk/manage-data>). UKDA provides also a training manual (<http://www.data-archive.ac.uk/curate/archive-training-manual>) on the following topics: staffing and management, data pre-ingest, data ingest, data access, user support, technical services, communications and publicity, standards and tools. Moreover, trainings (<https://www.ukdataservice.ac.uk/news-and-events/events>) and webinars (<https://www.ukdataservice.ac.uk/news-and-events/webinars>) are organised on various topics.

## International projects

Other resources can be found exploring the links to current and previous CESSDA projects websites (<http://cessda.net/eng/CESSDA-Services/Projects>). Some of them are displayed here:

- **CESSDA SaW (H2020)** – <http://cessdasaw.eu/>

CESSDA SaW webpage contains events announcement (i.e. workshops, webinars) as well as materials from past webinars (<http://cessdasaw.eu/deliverables/>). A knowledge sharing platform and additional supporting tools (e.g., the CESSDA SaW Cost-Benefit Funding Advocacy Toolkit) will also be developed and shared there and in CESSDA website.

- **SERSCIDA (FP7)** – <http://www.serscida.eu/en/>  
One of the main aims of the SERSCIDA project was to support the establishment of social science data archives in the Western Balkan Countries. The following deliverables are particularly interesting:
  - D5.1 – Documents and Materials for Social Science Digital Data Archives** – [http://www.serscida.eu/images/deliverables/D5.1\\_FINAL.pdf](http://www.serscida.eu/images/deliverables/D5.1_FINAL.pdf)  
This report contains examples on the elaboration of: a charter document, data and policies procedures, contracts and licenses, finances and budget, and performance indicators.
  - D5.3 – Report on prototype database** – [http://www.serscida.eu/images/deliverables/D5.3\\_FINAL.pdf](http://www.serscida.eu/images/deliverables/D5.3_FINAL.pdf)  
This report focuses on technical aspects that should be developed in a data service.
- **DASISH (FP7)**
  - Education and training** – <http://training.dasish.eu/training/>  
DASISH provides three training modules on: access policies and licensing, authentication and authorization infrastructure, and persistent identifiers.
  - D4.4 – Comprehensive Policy-Rules for Data Management in social sciences and humanities** – [http://dasish.eu/publications/projectreports/DASISH\\_D\\_4.4-desember2014.pdf](http://dasish.eu/publications/projectreports/DASISH_D_4.4-desember2014.pdf)  
Based on the assessment of the policy rules and procedure implemented by selected data archives and services this report recommends a set of policy-rules covering the full scope of a well-defined preservation policy. Chapter 4.2 provides a checklist with recommended preservation policy elements.
  - D6.5 – Handbook on legal and ethical issues for SSH data in Europe, Part II** – [http://dasish.eu/publications/projectreports/DASISH\\_D6.5\\_februar\\_2015.pdf](http://dasish.eu/publications/projectreports/DASISH_D6.5_februar_2015.pdf)  
This report focuses on legal and ethical issues for the collection, curation, preservation and dissemination of data in the social science and humanities (SSH) area.
  - D6.6 – Report about Preservation Policy-Rules (Preservation Challenges)** – [http://dasish.eu/publications/projectreports/DASISH\\_D6.6\\_september\\_2014.pdf](http://dasish.eu/publications/projectreports/DASISH_D6.6_september_2014.pdf)  
This report looks at ethical and legal issues and challenges that confront researchers, data owners and digital repositories in the emerging European data preservation infrastructure environment.

## International standards

- **Open Archival Information System (OAIS) model**  
OAIS explained within CESSDA website:  
<http://cessda.net/CESSDA-Training/Data-Archives-and-Digital-Preservation>  
Lavoie F. Brian, 2004, The Open Archival Information System Reference Model: Introductory Guide, OCLC Online Computer Library Center, Inc. and Digital Preservation Coalition:  
[https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&ved=0CD8QFjAC&url=http%3A%2F%2Fwww.dpconline.org%2Fdocs%2Fflavoie\\_OAIS.pdf&ei=JQASUsbPH6ah0QWvsICoDw&usq=AFQjCNFVhFOjGWOuI87JQo707-8ftDoDrA&sig2=ioW42lwDyKI2Zd9tWZ4OTw](https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&ved=0CD8QFjAC&url=http%3A%2F%2Fwww.dpconline.org%2Fdocs%2Fflavoie_OAIS.pdf&ei=JQASUsbPH6ah0QWvsICoDw&usq=AFQjCNFVhFOjGWOuI87JQo707-8ftDoDrA&sig2=ioW42lwDyKI2Zd9tWZ4OTw)



- **Standard of trust**

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

Audit and certification of trustworthy digital repositories (ISO 16363):  
<http://www.iso16363.org/>

Criteria for Trustworthy Digital Archives (DIN 31644):

[http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel\\_node.html;isesionid=48D88C87F22A9E4A32B95BF2F47344A5.prod-worker3](http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel_node.html;isesionid=48D88C87F22A9E4A32B95BF2F47344A5.prod-worker3)

- **Policies on open access to research data**

OECD (2007), Principles and Guidelines for Access to Research Data from Public Funding:  
<http://www.oecd.org/sti/sci-tech/38500813.pdf>

European Commission (2012), Commission recommendation on access to and preservation of scientific information:

[https://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/recommendation-access-and-preservation-scientific-information\\_en.pdf](https://ec.europa.eu/research/science-society/document_library/pdf_06/recommendation-access-and-preservation-scientific-information_en.pdf)

European Commission (2012), Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee, and the Committee of the Regions. Towards better access to scientific information: Boosting the benefits of public investments in research:

[https://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/era-communication-towards-better-access-to-scientific-information\\_en.pdf](https://ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-towards-better-access-to-scientific-information_en.pdf)

## Annex 2 - The CESSDA Capability Development Model as a planning tool

The CESSDA Capability Development Model (CESSDA-CDM) was developed in 2016. It 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 want 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.

Unlike existing models, the CESSDA-CDM is primarily aimed at repositories, archives, infrastructures, or other preservation initiatives that are providing, or aim at providing, preservation and data sharing services within the social sciences and humanities. The model provides a structured view of processes across an organisation. It can be used to set process improvement goals and priorities, provide guidance for quality processes and activities, and provide a benchmark for assessing and appraising current practices.

### Model components

The model identifies three subject areas – or Capability Requirement Areas (CRA) that fulfil a high-level objective or principle of a service provider. The CRAs that are identified in the model are the major factors contributing to the development and maturity of a service provider, an organisation or an infrastructure. A CRA may be undertaken by a part of an organisation, the whole of the organisation, or a group of organisations working together, to meet the high-level goals and to support the effective delivery of services to a designated community.

The CESSDA-CDM identifies three Capability Requirement Areas, namely Organisational Infrastructure, Digital Object Management and Technical Infrastructure.

*CRA1 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.

*CRA2 Digital Object Management* 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. The aim of Digital Object Management is to mitigate digital obsolescence, keeping the information accessible to users indefinitely.

*CRA3 Technical Infrastructure* consists of a set of processes 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.

Each of these CRAs has several Capability Process Areas (CPA). Each CPA has a purpose which can be subdivided into a set of objectives. Each objective consists of series of related activities that satisfy the objective of the process. These objectives and activities can be used as guidelines when planning and developing plans for the establishment of national data services.

For a more detailed explanation of the full CESSDA-CDM, see the project web page: <http://dev.cessda.net/CESSDA-Services/Projects/CESSDA-SaW/Work-Packages/WP3/CESSDA-CDM>.

### Links between the guide and the model

This guide has similarities with indicators present in the model, as shown in Table 2. Building your national development plan following this guide will already allow you to complete some elements of the CESSDA-CDM.

**Table 2:** National development plan (NDP) elements and mapping to the CESSDA-CDM

NDP element	CESSDA-CDM element
1.1. Definition of organisation	<b>CPA1.1 Mission and scope</b> ; <i>RA1.1.1.1 Approved Mission, RA1.1.1.2 Mission Statement</i>
1.2. Scope of collection	<b>CPA1.1 Mission and scope</b> ; <i>RA1.1.1.3 Define scope of the collection</i>
1.3. Services and activities	<b>CPA2.1 Data acquisition and ingest;</b> <b>CPA2.2 Data preservation: storage, curation and planning;</b> <b>CPA2.3 Access / Provision</b>
1.4. Beneficiaries	<b>CPA1.1 Mission and scope</b> ; <i>RA1.1.3.1 Define Designated Community</i>
1.5. Governance structure	<b>CPA1.3 Funding, staff, resources</b> ; <i>RA1.3.4.2: Roles and responsibilities within the organisation</i>
1.6. Financing scheme	<b>CPA1.3 Funding, Staff, Resources</b> ; <i>SO1.3.2: Sufficient resources</i>
1.7. Mission	<b>CPA1.1 Mission and scope</b> ; <i>RA1.1.1.1 Approved Mission, RA1.1.1.2 Mission Statement</i>
2.1. Host institution of the future data service	<b>CPA1.3 Funding, staff, resources</b> ; <i>RA1.3.1.1 Institutional Sustainability</i>
2.2. Human resources and internal structure	<b>CPA1.3 Funding, staff, resources</b> ; <i>SO1.3.3: Staff professional development; SO1.3.4: Appropriate expertise</i>
2.3. Partner support and cooperation	<b>CPA1.4 Communication</b>
3.1. Establishment plan	<b>Linked to all model components</b>
3.2. Operational plan	<b>Linked to all model components</b>

# Glossary

**Archive certification** ensures that archives and repositories are trustworthy and operate according to best practices and international standards in the field of digital data preservation. Obtaining certification, for example from the Data Seal of Approval, demonstrates that archives are sufficiently equipped to safeguard data and manage data properly.

**Archiving procedures and workflows** involve documented practices for carrying out different parts of the archiving process, such as ingest, curation, and dissemination. It includes descriptions of who is responsible for particular tasks within the archive, and how the work should be carried out in accordance with existing policies.

**Classification schemes (of data to be archived)** are controlled vocabularies that allow for indexation of data and datasets. Data that are indexed according to classification schemes can then more easily be discovered by users in data catalogues. Such themes are generally thematic or topical.

**Data acquisition** is the process of appraising, selecting and getting data and metadata based on a set of defined criteria; i.e. to identify the data and information that the repository will preserve (ISO 16363). (source: CESSDA SaW Capability Development Model)

**Data curation** refers to the active management of data through its life cycle of interest and usefulness to a designated community. Data curation activities enable data discovery and retrieval, maintain its quality, add value, and provide for re-use over time. As such, it includes all processes in the organization that involves data management. That is, pre-ingest initiatives; ingest functions; archival storage and preservation; and disseminating and providing access to data for its designated community. (source: International Federation of Data Organizations (IFDO), Data preservation, <http://ifdo.org/wordpress/preservation/>)

**Data depositor** is the person or institution that deposits the data in the data service. The data depositor can be a different person or institution than the data producer.

**Data dissemination** consists of distributing or transmitting data to the end-users.

**Data Documentation Initiative (DDI)** is an international standard for describing statistical and social sciences data. Documenting data with DDI facilitates interpretation and understanding, both by humans and computers (<http://www.ddialliance.org/>). (source: CESSDA SaW Capability Development Model)

**Data ingest** consists of accepting data and metadata from data producers/depositors, preparing data and metadata for storage, and ensuring that the information becomes established within the archive. (source: CESSDA SaW Capability Development Model)

**Data Management Plan** is part of grant application or research project delivery that consider essential properties of Research Data Management throughout the project, aiming at Open data as the default. (source: CESSDA SaW Capability Development Model)

**Data preservation** or more specifically, *digital data* preservation, refers to the series of managed activities necessary to ensure continued access to digital materials for as long as necessary. This broad definition of data preservation refers to all of 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 protection** is a set of practices that reduces the likelihood that individual study respondents can be harmed through identification and improper use of personal information. It includes anonymization of data, accreditation and authentication systems to control who has access to the data, constraints on access (e.g., prior approval of the data producer, and user contracts).

**Data security** is the means of ensuring that data is kept safe from corruption or damage and that access is suitably controlled.

**Data user** is the end user of the data disseminated by the data service.

**Long-term access** means that the data should be accessible in 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). (source: CESSDA SaW Capability Development Model)

**Online catalogue** is an online tool containing a list of the data available in the data service and their description (metadata).

**Open access** is the practice of providing unrestricted access to scientific publications, or the movement aiming at this practice. This term is sometimes also used to imply access to research data used for these publications (otherwise addressed by the term “Open Data”). (source: DASISH <http://training.dasish.eu/training/glossary/>)

**Open Archival Information System (OAIS)** reference model is a conceptual framework for an archival system dedicated to preserving and maintaining access to digital information over the long term. ISO 14721:2012 defines the reference model for an open archival information system (OAIS).

(sources: Lavoie, <http://www.oclc.org/research/publications/library/2000/lavoie-oais.html> and ISO14721:2012, [http://www.iso.org/iso/home/store/catalogue\\_ics/catalogue\\_detail\\_ics.htm?csnumber=57284](http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=57284))

**Open data** is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and share alike.

(source: Opendata handbook, <http://opendatahandbook.org/guide/en/what-is-open-data/>)

**Research Data Management (RDM)** 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](http://repository.jisc.ac.uk/6379/16/Training_for_RDM_-_Comparative_european_approaches_May_2016.pdf)). (source: CESSDA SaW Capability Development Model)

**Research inventory** contains information and description on (past and) current research projects in the field of social sciences conducted in the country.

**Secondary analysis** involves the use of existing data, collected for the purposes of a prior study, in order to pursue a research interest which is distinct from that of the original work; this may be a new research question or an alternative perspective on the original question (Hinds, Vogel and Clarke-Steffen 1997, Szabo and Strang 1997).

(source: Heaton, <http://sru.soc.surrey.ac.uk/SRU22.html>)

**Succession plan** is a plan that ensures on-going access to and preservation of the holdings of the organisation in case the data service ceases to operate, or the governing or funding institution substantially changes its scope/obligations. (source: CESSDA SaW Capability Development Model)

**Thesaurus** is a list of subject headings or descriptors usually with a cross-reference system for use in the organisation of a collection of documents for reference and retrieval. (source: Merriam-Webster online dictionary, <http://www.merriam-webster.com/dictionary/thesaurus>)