# Cost-Benefit Advocacy Toolkit
# User Guide


**Illustrations by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark**



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

# Contents


Background 3
CESSDA SaW Task 4.6 3
Intended users for the toolkit 3
How to use the toolkit 3
Concepts 4
Design criteria 4
Effort grading levels 4
Toolkit Landing Pages and DOIs 4
Toolkit components 4
Factsheets 6
Benefits Factsheet 7
Costs Factsheet 7
Return on Investment Factsheet 8
Worksheets 8
The Archive Development Canvas 8
Benefits Summary for a Data Archive 10
Case Studies 10
ADP (Slovenia) – Case study on user satisfaction surveys 11
FSD (Finland) – Case study on using cost and benefit tools 11
LiDA (Lithuania) - Case study on toolkit advocacy focus groups 11
UKDS (UK) - Case study on use of the Economic and Social Data Service (ESDS) economic impact study 11
Selected External Tools 12
KRDS Benefits Analysis Tools 12
Curation Costs Exchange 12
KRDS Activity Model and Costs Framework 13
DANS Cost Model and Balanced Score Card 13
ESDS Impact Study 14
ADP User Satisfaction Survey 14
User Surveys 15
Value and Impact Studies 17

# Background

## CESSDA SaW Task 4.6: Understanding the economic impact of social science data archives

CESSDA is the Consortium of European Social Science Data Archives. The CESSDA SaW project "Strengthening and widening the European infrastructure for social science data archives" is funded by the European Commission as part of its Horizon2020 programme.

The project proposal for CESSDA SaW included a task (task 4.6) focused on understanding the economic impact of social science data archives. Its aim is to develop a benefit/cost advocacy programme and supporting toolkit for data archives. This task was led by Charles Beagrie Ltd, with support from the Slovenian Social Science Data Archive (ADP), the Finnish Social Science Data Archive (FSD), the Lithuanian Social Science Data Archive (LiDA), the University of Tartu in Estonia (UTARTU) and the UK Data Archive (UKDS).

The cost-benefit advocacy toolkit was able to draw on a range of pre-existing work by the consortium partners. However, it needed to develop the methodology and a toolset of documents in order for this to be applicable to a range of European countries, and in new and emerging as well as established social science data archives.To help shape the toolkit, an online user requirements survey, and a series of focus groups and workshops were held. The toolkit was developed in incremental stages up to its release in April 2017.

The full results of the user requirements survey, focus groups and workshops, and the detailed description of how the toolkit was developed are reported in the Deliverable D4.9 Cost-Benefit Advocacy Toolkit Report.

## Intended users for the toolkit

The toolkit has been developed for its primary audience of staff in existing or proposed national social science archives in Europe. It will help them to understand the approaches and tools available and support their advocacy to funders and policymakers. However we expect the toolkit may be of interest to other audiences including staff in institutional data repositories, students, researchers, funders, policymakers and communities in other data intensive disciplines, even if the toolkit is not specifically tailored and adapted for them.

## How to use the toolkit

This user guide forms part of the toolkit and provides an overview of the toolkit components and the additional external tools that have been selected, and supplemental guidance in their use. As a cost-benefit advocacy toolkit, its major use will be supporting funding and business cases but elements are likely to be relevant in advocacy to other groups including depositors and users of the archive, and in supporting the broader operational tasks of the archive. Creative Commons by Attribution licensing (CC-BY) is intended to allow you to easily re-use locally any material from the factsheets and case studies in the toolkit.

The toolkit comprises of: three factsheets (Benefits, Costs, and Return on Investment); four case studies from Social Science Data Archives (ADP in Slovenia, FSD in Finland, LiDA in Lithuania, and UKDS in the UK); and two worksheets (the Archive Development Canvas, and the Benefits Summary for a Data Archive). In addition, the toolkit describes and links to a number of pre-existing external tools and relevant studies.

# Concepts

## Design criteria

Based on initial ideas and desk research for the toolkit and feedback from the user requirements survey and focus groups, a number of design criteria were established for the toolkit:

- Ease of use for individual staff and smaller archives;
- Short documents;
- Good infographics;
- High-quality synthesis of existing reports and studies;
- Making existing tools (e.g. worksheets) easier to use/tailored for social science data archives;
- Creative Commons CC-BY licensing wherever possible for ease of re-use.

We have applied these criteria wherever possible, including the use of effort grading levels.

## Effort grading levels

Part of making the toolkit components easy to use has been applying indicators of the effort that may be required to apply them. Depending on the maturity of the service and its existing level of resources, different components of the toolkit or selected external tools will be more appropriate than others. To help in the selection, three broad categories of effort have been defined and each of the toolkit components and external tools has been graded according to the effort needed. All the core toolkit components developed by CESSDA SaW require perhaps only a few hours to a day of effort to learn and apply. However, many of the linked external tools will take more effort ranging from days to months to apply.

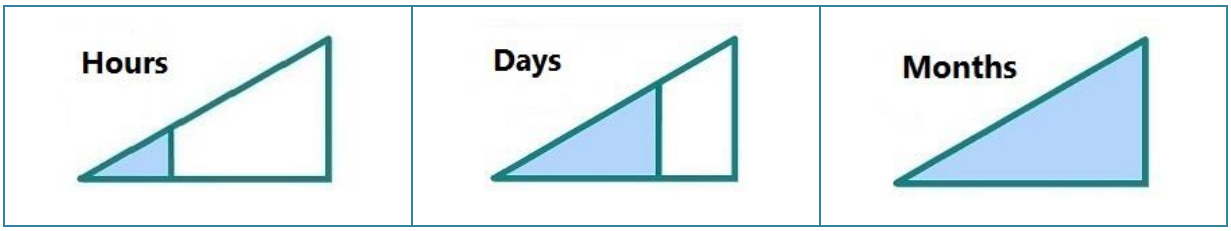



***Illustration of the effort grading levels applied to toolkit components***

## Toolkit Landing Pages and DOIs

We have assigned separate Digital Object Identifiers (DOIs) to all the individual components of the toolkit and for the toolkit as a whole. These provide persistent identifiers that will resolve to the appropriate landing pages over time.

If you wish to cite or provide a link to the toolkit as a whole the DOI http://dx.doi.org/10.18448/16.0013 will resolve to the CESSDA SaW Cost-Benefit Advocacy Toolkit landing page providing an overview of the toolkit and links to all the components.

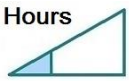
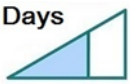
DOIs for individual components of the toolkit are embedded in the links in the appropriate sections of the user guide that follow.

## Toolkit components

The CESSDA SaW developed components form the core of the toolkit and consist of factsheets, worksheets, case studies and this user guide. These are all intended to be easy to use (effort of a few hours up to a day to read/ apply initially) and suitable for a wide range of individuals and data archives to understand and use.

| Effort | Core toolkit components |
|---|---|
|  | Factsheets:<br>Benefits<br>Costs<br>Return on Investment (ROI) |
| | Worksheets:<br>Archive Development Canvas<br>Benefits Summary for a Data Archive |
| | Case studies:<br>ADP (Slovenia) User Satisfaction Surveys<br>FSD (Finland) Using Benefit and Cost Tools<br>LiDA (Lithuania) Toolkit Focus Groups<br>UKDS (UK) Use of the ESDS Impact Study |
| | User Guide |
| | Deliverable D4.9 Cost-Benefit Advocacy Toolkit Report |

We have also selected a number of key pre-existing external tools that are described and linked in components of the toolkit. These external tools have not been produced by the project and will have their own independent copyright and licensing terms. Most will require a significantly higher level of effort to apply, ranging from a few days to potentially months of activity.

| Effort | External tools |
|---|---|
|  | KRDS Benefits Analysis Tools |
|  | Curation Costs Exchange |
| | ESDS Impact Study |
| | ADP User Satisfaction Survey |
|  | KRDS Activity Model |
| | DANS Cost Model and Balanced Score Card |

Core components and key linked external resources in the toolkit can support a number of different functions in cost-benefit analysis and advocacy. As a finding aid for these, the cost-benefit advocacy toolkit components are also mapped on to an outline of the Archive Development Canvas as illustrated below. The Archive Development Canvas itself and how to use it, are described in greater detail later in this user guide.

## The Cost-Benefit Advocacy Toolkit mapped onto the Archive Development Canvas

| Key Partners | Key Activities | Benefits | Beneficiaries | Beneficiary Relationships |
|---|---|---|---|---|
| | | Archive Development Canvas (Detailed Version)<br>User Guide<br>ROI Factsheet<br>Benefits Factsheet<br>Benefits Summary for a Data Archive<br>UK (UKDS) Case Study<br>Lithuania (LiDA) Case Study<br>Finland (FSD) Case Study<br>ESDS Impact Study | User Guide<br>Slovenia (ADP) Case Study<br>ADP User Satisfaction Survey<br>ESDS Impact Study (User and Depositor Impact Surveys) | |
| **Key Resistances** | **Key Resources** | | | **Channels** |
| | Benefits Factsheet | | | |
| **Cost Structure** | | **Key Metrics** | **Funding Streams** | |
| User Guide<br>KRDS Cost Model<br>4C Cost Comparison Tool | Costs Factsheet<br>Finland (FSD) Case Study | User Guide | | |

**Key** CESSDA SaW | External Resource



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

***Illustration of cost-benefit advocacy toolkit components mapped on to an outline of the Archive Development Canvas, showing the different cost-benefit functions they can support***

## The Factsheets

**Use the factsheets as an overview to gain an understanding of benefits, costs, and return on investment. You can also use them as a source of key evidence and infographics to support your advocacy and funding cases**




***Illustrations of the cover pages of the Benefits, Costs, and Return on Investment Factsheets***

The 3 factsheets on benefits, costs, and return on investment, are intended to be individually free-standing but an inter-linked and complementary series that can be read and used together. They are 7-8 pages in length and aim to explain the fundamentals of cost-benefit advocacy for social sciences data archives in an easily assimilated and usable way. They summarise, analyse, and visualise, the evidence base from which the appropriate potential case can be made locally by data archives. Each factsheet selects and presents the key evidence and describes the key tools and approaches available. Links are provided to other relevant components of the toolkit and to relevant external tools, studies, and reports.

## Benefits Factsheet

The Benefits Factsheet sets out key approaches you can use to think about and identify benefits for different stakeholders; some of the main arguments for benefits and the evidence for them; and finally potential metrics and case studies for benefits. The approaches outlined should be seen as incremental, building up in steps and increasing in complexity as you move from qualitative to quantitative evidence of benefits. The main focus is on cost-benefit and economic approaches which should be seen as complementary to other measures of benefits such as citations to data and services in the academic literature.

Communicating benefits is most effective if you can consider it together with the investment (costs) required, and if you can quantify and explain the value (benefits in relation to the investment). This factsheet is therefore intended to be used with other components of the CESSDA Saw Cost-Benefit Advocacy Toolkit particularly the Return on Investment (ROI) Factsheet, the Costs Factsheet, and the Archive Development Canvas, to help you make the case for your archive.

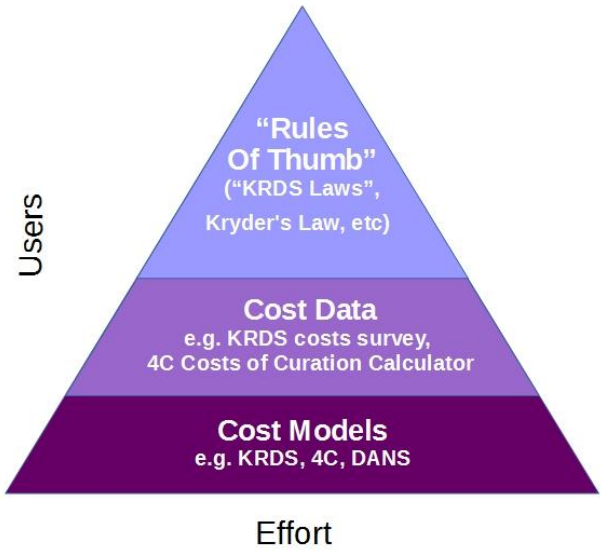
The Finnish Social Science Data Archive (FSD) has produced a case study detailing its experience of costs and benefits tools. This is also highly recommended reading in conjunction with this factsheet.

## Costs Factsheet

In the Costs Factsheet, it is important to understand that the costs of digital preservation and curation are not a simple topic and in practice can be very complex. This complexity means that the effort threshold for some costing activities is very high and therefore direct use by individual data archives may be limited.

To help archives understand the range of costs tools available and how they fit together, we have used a pyramid structure to explain the levels of effort, use, and knowledge, required for costs tools. The factsheet uses a tripartite pyramid (Costs Models, Cost Data, “Rules of Thumb”) as a means of understanding existing work, each building on (and requiring the existence of) the other in terms of a knowledge-base, and each requiring different levels of effort.

In terms of effort, costs models are the most demanding, cost data and then rules of thumb progressively less so. It is not possible or necessary for every data archive to make the considerable investment required to develop a cost model. This pyramid illustrates how other data archives can then use the cost model with their own cost data. It still takes time and effort to collect the cost data, but much less than it would to develop the cost model. If enough data services can collect and share their cost data, patterns will emerge to allow the formulation of “Rules of Thumb”, which can then be applied by everyone.




***Effort and Use Knowledge Pyramid for Costs Tools***



Cost analysis should be accompanied by an analysis of the anticipated benefits. The Costs Factsheet should therefore be read and used in conjunction with other components in the Cost-Benefit Advocacy Toolkit, particularly the Benefits Factsheet and the Return on Investment Factsheet. The Finnish Social Science Data Archive (FSD) has produced a case study detailing its experience of costs and benefits tools. This is also highly recommended reading in conjunction with this factsheet.

### Return on Investment Factsheet

Return on Investment (ROI) is a metric used to evaluate the merit of a single investment or to compare the relative merits of a number of different investments. It measures the amount of quantifiable benefit (return) relative to the investment's cost. As a result, you may find reading the Benefits and Costs Factsheets helpful prior to reading the ROI Factsheet.

The Return on Investment Factsheet explains return on investment and how it is used; the key ROI findings from the impact study of the Economic and Social Data Service (ESDS) in the UK (currently the only example of a fully developed quantified economic impact study and ROI metrics for social science research infrastructure); the evidence we have that data collections are usually appreciating assets: returns can increase over time as collections reach a critical mass and user awareness of them grows; and how to consider "the cost of inaction" and the counter-factual position if no archive exists.

ROI is not an easy metric to calculate for data archives as economic returns are difficult to capture and can grow over time. If calculating a ROI for a specific data archive is not feasible, then citation of results from other studies such as ESDS can be helpful if they are relevant to your specific context. For new and emerging archives, it may be more useful to focus on the "cost of inaction" and counter-factual metrics.

The UKDS case study looks at how the ESDS impact study has been used since in funding advocacy to Government. This is also highly recommended reading in conjunction with this factsheet.

## The Worksheets

**Use the worksheets as practical tools to brainstorm the value proposition and benefits for your data archive**

### The Archive Development Canvas

The Archive Development Canvas (Detailed Version) is the CESSDA SaW implementation of the Business Model Canvas. The Business Model Canvas is used by companies of all sizes worldwide. It is openly licensed (Creative Commons By Attribution, Share Alike: CC-BY-SA) so it has also been adapted for use in not-for profit and social enterprises.

The Archive Development Canvas has adopted some changes and additions from not-for-profit social

versions of the Business Model Canvas, particularly in thinking about the language used (e.g. changing its title, using words such as funding streams instead of revenue streams, beneficiaries instead of customers, etc.), and adding new categories such as Key Metrics and Key Resistances. It also includes prompts under each category tailored to social science archives or taken from other components and approaches recommended in the CESSDA SaW cost-benefit advocacy toolkit.

You can use the Archive Development Canvas as a brainstorming tool for the value proposition and development of new data archives or new services. It helps identify the building blocks and steps required, and presents them clearly on one page. Other component tools in the cost-benefit advocacy toolkit can help complete it. Prompts are in grey text. As you complete each section you should begin to see connections to the others. The value proposition (benefits) is central and it is the pivotal part of the Canvas.

The Canvas gives an overview of your activity from a business perspective, setting out the key elements and showing how they interrelate to support the analysis and evaluation of your own strategy. You can also use the Canvas to model how the elements would change if you adopted a different approach.

To make best use of the Canvas you should:

- have a small, focused team working with it;
- remember the "value proposition" in the benefits section is pivotal to its successful use;
- If you can, involve someone with previous experience of using the Canvas.

Each use of the Canvas will have its own specific requirements and you should feel free to select and use those elements and prompts within the Canvas that are relevant to your needs. For example, a simpler version and illustration of the Archive Development Canvas is used in the CESSDA SaW Task 3.3 Guide for the elaboration of national data service development plans. The Archive Development Canvas (Detailed Version) is included as a worksheet in the Toolkit for you to use. To help you complete it, there is also a Cost-Benefit Advocacy Toolkit Mapped to the Archive Development Canvas finding aid with active links to all the relevant toolkit components.

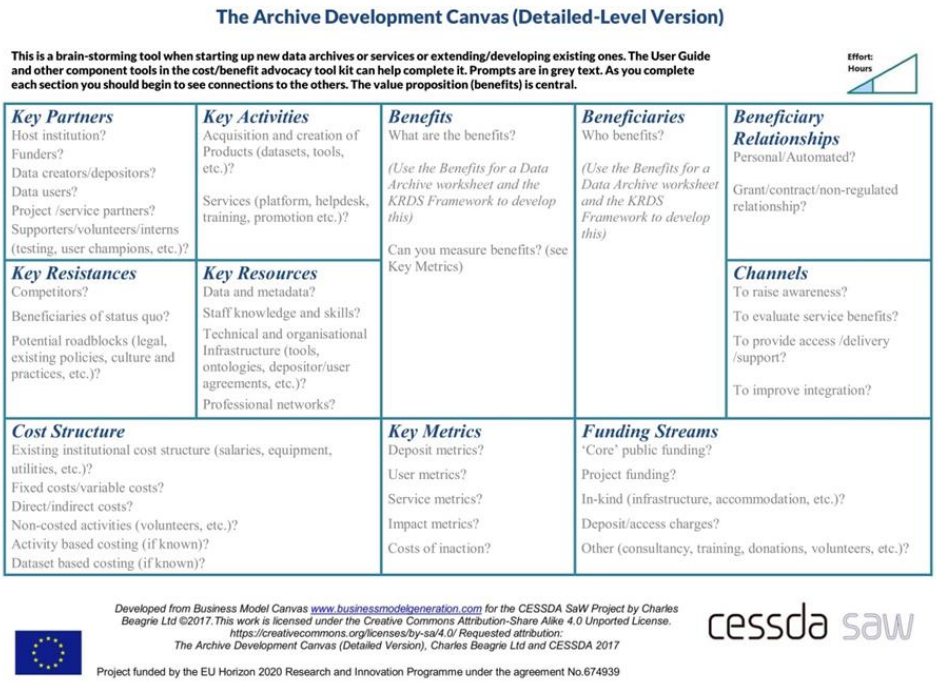
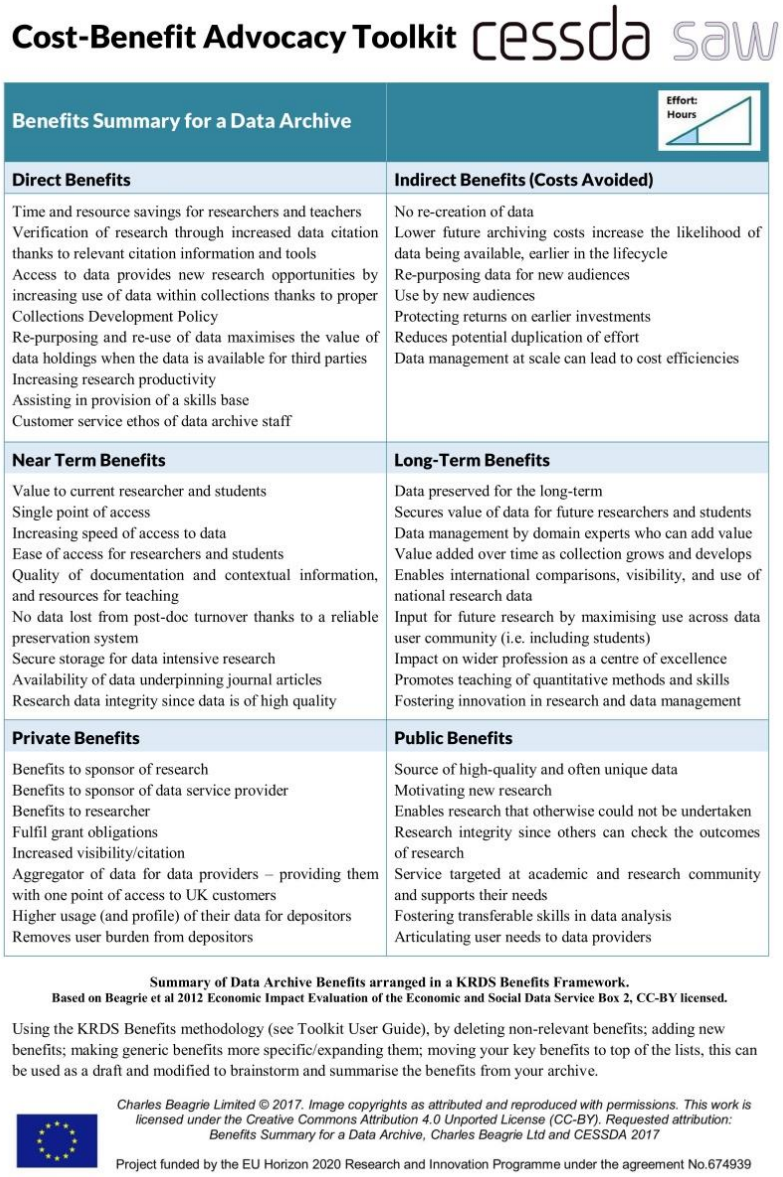



***Illustration of the detailed version of The Archive Development Canvas***

## Benefits Summary for a Data Archive



**Cost-Benefit Advocacy Toolkit** cessda saw

| Benefits Summary for a Data Archive | Effort: Hours |
|---|---|
| **Direct Benefits** | **Indirect Benefits (Costs Avoided)** |
| Time and resource savings for researchers and teachers<br>Verification of research through increased data citation thanks to relevant citation information and tools<br>Access to data provides new research opportunities by increasing use of data within collections thanks to proper Collections Development Policy<br>Re-purposing and re-use of data maximises the value of data holdings when the data is available for third parties<br>Increasing research productivity<br>Assisting in provision of a skills base<br>Customer service ethos of data archive staff | No re-creation of data<br>Lower future archiving costs increase the likelihood of data being available, earlier in the lifecycle<br>Re-purposing data for new audiences<br>Use by new audiences<br>Protecting returns on earlier investments<br>Reduces potential duplication of effort<br>Data management at scale can lead to cost efficiencies |
| **Near Term Benefits** | **Long-Term Benefits** |
| Value to current researcher and students<br>Single point of access<br>Increasing speed of access to data<br>Ease of access for researchers and students<br>Quality of documentation and contextual information, and resources for teaching<br>No data lost from post-doc turnover thanks to a reliable preservation system<br>Secure storage for data intensive research<br>Availability of data underpinning journal articles<br>Research data integrity since data is of high quality | Data preserved for the long-term<br>Secures value of data for future researchers and students<br>Data management by domain experts who can add value<br>Value added over time as collection grows and develops<br>Enables international comparisons, visibility, and use of national research data<br>Input for future research by maximising use across data user community (i.e. including students)<br>Impact on wider profession as a centre of excellence<br>Promotes teaching of quantitative methods and skills<br>Fostering innovation in research and data management |
| **Private Benefits** | **Public Benefits** |
| Benefits to sponsor of research<br>Benefits to sponsor of data service provider<br>Benefits to researcher<br>Fulfil grant obligations<br>Increased visibility/citation<br>Aggregator of data for data providers – providing them with one point of access to UK customers<br>Higher usage (and profile) of their data for depositors<br>Removes user burden from depositors | Source of high-quality and often unique data<br>Motivating new research<br>Enables research that otherwise could not be undertaken<br>Research integrity since others can check the outcomes of research<br>Service targeted at academic and research community and supports their needs<br>Fostering transferable skills in data analysis<br>Articulating user needs to data providers |

**Summary of Data Archive Benefits arranged in a KRDS Benefits Framework.**
**Based on Beagrie et al 2012 Economic Impact Evaluation of the Economic and Social Data Service Box 2, CC-BY licensed.**

Using the KRDS Benefits methodology (see Toolkit User Guide), by deleting non-relevant benefits; adding new benefits; making generic benefits more specific/expanding them; moving your key benefits to top of the lists, this can be used as a draft and modified to brainstorm and summarise the benefits from your archive.



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

***Illustration of the Benefits Summary worksheet***

The CESSDA SaW Benefits Summary for a Data Archive worksheet is a social science specific implementation of the Keeping Resource Data Safe (KRDS) Benefits Framework in the KRDS Benefits Analysis Toolkit. It utilizes a summary produced during the ESDS Impact Study, together with updates and additions to this suggested by the project partners, to provide a template that can be used in social science data archives.

This worksheet can be used as a draft for identifying and assessing benefits using the KRDS Benefits methodology. You can modify it by deleting non-relevant benefits; adding new benefits; making generic benefits more specific or expanding them; moving your key benefits to top of the lists, to summarise and present the benefits from your data archive.

The Finnish Social Science Data Archive (FSD) has produced a case study detailing its experience of costs and benefits tools including the KRDS worksheet. This is also highly recommended reading in conjunction with this worksheet. FSD recommended the KRDS Benefits Framework as the optimal starting point for archives analysing benefits for the first time. The effort needed is relatively small (several hours) and you can use the Benefits Summary for a Data Archive. You can then proceed to any of the other tools presented here, depending on your needs and on your resources.

## Case studies

**Use the case studies for practical examples and lessons in benefits, costs, return on investment, and advocacy**

Four case studies were developed with partner archives in task 4.6. Topics were selected on the basis of local experience of different topics and value and interest to a wider CESSDA community. They are 3-7 pages in length. The focus of the four case studies are as follows:

## ADP (Slovenia) – Case study on user satisfaction surveys

The ADP (Slovenia) – Case study on user satisfaction surveys examines user satisfaction measurement via online surveys. It focuses on the overall picture for surveys in social science data archives and related organizations in general, and the specific experience of the Slovenian Social Science Data Archives (ADP). User satisfaction surveys can have an important part to play in helping services maximise their usefulness and impact. Related materials include an English translation of the ADP user satisfaction survey, and a section with generic guidance on survey questionnaires later in this user guide.

## FSD (Finland) – Case study on using cost and benefit tools

The FSD (Finland) – Case study on using cost and benefit tools is based on their experience of using some pre-existing tools. It examines how some existing benefit and cost tools could be used to determine the benefits of data archiving and the costs of this kind of research data infrastructure. It aims to add insight to what is already known through previous research or from other components of the CESSDA Cost-Benefit Toolkit such as the factsheets.

## LiDA (Lithuania) - Case study on toolkit advocacy focus groups

The LiDA (Lithuania) - Case study on toolkit advocacy focus groups reports on testing in focus groups of the emerging cost-benefit advocacy toolkit during 2016. The focus groups were of two types: for staff from the social science data archive; and for their key stakeholders (typically senior staff from the host university, government ministries, national statistics offices, representative researchers and depositors). It presents an analysis of the discussion and feedback from the Lithuanian focus groups, and lessons learnt.

## UKDS (UK) - Case study on use of the Economic and Social Data Service (ESDS) economic impact study

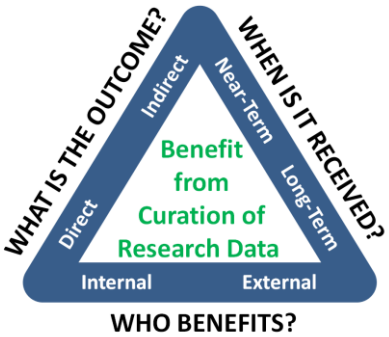
The UKDS (UK) - Case study on use of the Economic and Social Data Service (ESDS) economic impact study examines how the economic impact study of the ESDS published in 2012 has been used in funding advocacy to Government by the UK Data Service and its principal funder the Economic and Social Research Council. Key lessons learnt and its relevance for other countries and archives are considered.


***Illustration of the cover pages from the case studies***

## Selected external tools

**Use the selected external tools as/if required to support more detailed learning and analysis**

### KRDS Benefits Analysis Tools




***The Anatomy of a Benefit***

**(KRDS User Guide 2011 figure 10). Charles Beagrie Ltd ©2011. CC-BY licensed**

The Keeping Research Data Safe project (KRDS) created two benefit analysis tools: the KRDS Benefits Framework; and the Value-chain and Benefits Impact tool.

The KRDS Benefits Framework organises benefits along three broad dimensions: the outcome achieved; when the outcome is achieved; and who benefits from the outcome. Each of these dimensions can be subdivided into two categories: direct and indirect benefits, near-term and long-term benefits and internal and external benefits respectively.

In the CESSDA SaW toolkit, we have created a new version of KRDS Benefits Framework, the Benefits Summary for a Data Archive worksheet, that is customised specifically for a social science archive. It requires little experience and effort to implement and is an ideal starting point when identifying your benefits. It can be used as a stand-alone tool in many tasks. It is also a recommended tool when working with the Archive Development Canvas in the toolkit.

The KRDS Value Chain and Benefits Impact Tool is a more advanced tool (a day or more of effort) and is available with an accompanying guide from the project website. It requires more experience and effort to implement. The Tool consists of a detailed user guide and two worksheets; the Benefits Impact worksheet and the Value-chain and Benefits Impact worksheet. It is recommended that both worksheets in the Tool are used by a team with a senior member of staff or independent support (e.g. consultancy). For maximum effectiveness in applying the Tool, ideally at least one person in the team should be very familiar with the KRDS Benefits Framework, other KRDS Outputs such as the KRDS Activity Model, and similar assessments of value and impact.

The Finnish Social Science Data Archive (FSD) has produced a case study detailing its experience of costs and benefits tools including the KRDS benefits analysis worksheets.

### Curation Costs Exchange

The Curation Costs Exchange (CCEx) is an online tool created by the 4C project and hosted by the Digital Preservation Coalition as a community owned resource. CCEx provides comparison information of what you have spent (e.g. per gigabyte) on various activities, with similar institutions. You first submit profile information about your organisation, and cost data for specific activities (e.g. pre-ingest, ingest, storage, access) and projects. You add your data by year for each project/dataset/collection and say what it consists of (e.g. 10% databases, 25% video) and what you spent on different activities such as digitisation/preservation specialists etc. The tool then combines all your datasets for analysis. If you input your own information the tool will help find similar organisations so you can compare your costs either to

the average costs of a group of organisations, or one-to-one with a single peer organisation.

You will need an understanding of cost models and OAIS functional categories, to capture your cost data or to use/compare cost data from other organisations appropriately.

The Finnish Social Science Data Archive (FSD) has produced a case study detailing its experience of costs and benefits tools including its experience of using the CCEx. They used the CCEx tool to analyse the costs of pre-ingest and ingest, and data access. The tool itself was easy to use and if you have the required cost data at hand you can get results very quickly. However, often this kind of detailed cost information is not readily available and many archives should anticipate use of the tool will require days of effort to assemble costs information in the required format.

## KRDS Activity Model and Costs Framework

The Keeping Research Data Safe (KRDS) Costs Framework is an example of a life-cycle costing method applied to research data. It is a generic model designed for adaptation to local needs and is well documented. However, potential users should be aware developing any local cost model may involve a month or more of effort.

It models a life-cycle for a specific process(es) and then identifies measurable component activities, cost drivers (variables that affect the costs of the activity e.g. volumes, formats etc.), and resources (staff time, equipment etc.) to provide an understanding of costs for that process.

KRDS sets out the broader cost framework and guidance within which the KRDS Activity Model can be applied. That cost framework consists of three parts:

- **KRDS Activity Model.** A generic activity model for research data identifying activities with cost implications for preservation and ordering them in a nested hierarchy of Phases, Activities, and Sub-activities.
- **Cost Drivers**. Key variables (e.g. salary levels or rates of inflation), which affect the cost of preservation activities. The cost drivers are divided into two major groups: economic adjustments and service adjustments.
- **Resources Template**. This presents categories ("resource pools") of cost (e.g. staff or equipment) and duration (year 1, year 2, etc.) in a simplified, generic form closer to that used in the cost methodologies of UK universities based on the Transparent Approach to Costing method (TRAC).

You can find a full description of how to use the KRDS model in the KRDS User Guide. For the Activity Models see the KRDS website.

## DANS Cost Model and Balanced Score Card

The CMDA model is a cost model for implementation in a specific archive – it was developed and customised specifically for use by DANS, the Dutch data archive. It is an activity-based costing model utilising the OAIS Reference Model for its functional categories. It identifies activities and a set of costing components of each activity. It also takes the varying data complexity of datasets into account. Based on these factors the model estimates costs per dataset in "euros per dataset". Implementation of the CDMA, like many other cost models, required months of work.

A Balanced Score Card (BSC) was used as a tool to clarify a vision and strategy and translate them into action. The BSC also acted as a measurement system and communication tool. The BSC tool defines Success Factors which describe the strategic objectives of the organisation. They are described further by a set of Performance Indicators which are an indication of how it shall be known that the outcome has come

to pass. Success Factors and Performance Indicators can be used to connect measured benefits to activities and costs.

Details of the CDMA cost model and its use of the BSC are available in a published article. It is primarily a study of implementation and experience of a costing model and use of the BSC within DANS and lessons learnt.

## ESDS Impact Study

At the time of writing this guide (March 2017), the Economic Impact Evaluation of the Economic and Social Data Service (ESDS) is the only example of a fully developed quantified economic impact study of a social science archive. As such it is cited extensively in the toolkit.

The report provides a detailed description of the methods used and findings from the study and may take a day or more to fully read and digest. Although the methods used are well established, obtaining reliable impact data to use in them is more challenging. The study itself took several person months to complete.

The study combined quantitative and qualitative analytical approaches in order to quantify value and impacts in economic terms and explore other, non-economic benefits. It covers both users and depositors of data, and the survey of depositors that was undertaken was the first of its kind. The key economic findings from the ESDS impact study were that the quantifiable benefits significantly exceeded the value of the funding invested in the Service.

The Economic and Social Data Service (ESDS) received additional responsibilities in 2012 and was renamed the UK Data Service.

Selected findings from the ESDS impact study are discussed in the Return on Investment Factsheet. The UKDS case study looks at how the ESDS impact study has been used since in funding advocacy to Government; and finally, the FSD case study describes their scenarios developed from the ESDS impact study to assess improved access to FSD's data holdings.

The ESDS impact study was the first of a series conducted on data archives in different disciplines. The others are described in the Value and Impact Studies section of the user guide.

## ADP User Satisfaction Survey

An English translation of the 2016 Slovenian Social Science Data Archives (ADP) User Satisfaction Survey questionnaire is included in related external materials for the ADP case study. The ADP case study describes this survey and user satisfaction surveys from other archives and related organisations.

The questionnaire had four sections. The first one covered questions about the frequency of use of materials and services from ADP, and the purpose of their use. The second section of the questionnaire covered overall satisfaction with services, research data and materials provided by ADP. The third section referred to the respondent's own research practices, and the fourth contained demographic questions about the respondent.

The ADP case study and user satisfaction survey will help those preparing a survey for their own archives. Some general guidance on preparing surveys is provided in the User Surveys section of the user guide.

# User surveys

**Use this section for preliminary guidance if you are conducting a user survey for the first time. User surveys can have an important part to play in helping data archives maximise their usefulness and impact**

This section provides general guidance on preparing user surveys that should be seen as complementary to details of user satisfaction surveys provided in the ADP case study in the toolkit.

## Why a user survey?

**Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark**

A user survey provides a snapshot of the attitudes and behaviors of your target survey population. It can help you identify and address issues and opportunities in your service provision, and sets a baseline to measure and demonstrate improvement over time, especially if repeated at regular intervals. You can show funders real examples of how the service has contributed to research and its outcomes.

## Planning issues to consider

- What is the objective of the survey?
- What information is needed?
- How will the information be used?
- What is the best way to reach the target population?

## Potential areas to cover

- Purpose of access.
- Frequency of access.
- Overall satisfaction rating for the service and its various components.
- Importance of the service to the respondent.
- How can the service be improved?
- Include key institutional and socio-demographic variables if relevant. Examples might include gender, age, income and education, employment status, research area, experience in using data services, experience/expertise in data analysis, etc.
- If previous surveys have been conducted, consider reusing questions for consistency and trend analysis.

## Survey design considerations

There are a range of online survey tools such as SurveyMonkey or LimeSurvey available, and you should start with guidance provided by your chosen survey tool. SurveyMonkey for example,offers extensive guidance on survey and question design. Supplementary guidance, providing a summary of the most common lessons learnt and advice for those designing and conducting surveys for the first time, is provided below.

| | |
|---|---|
| **Use formatting** | Group similar questions to keep your survey logical and focused. Page breaks, page titles , headings, sub-headings, comments and instructions help people understand |

| | |
|---|---|
| | what you are asking and why, and help them progress smoothly through the survey questionnaire. |
| **Ask often** | Using the same question in a series of surveys, or even using the same survey over time, is a good way to build a baseline and measure changes in respondents' attitudes. |
| **Be brief** | Keep questions and surveys as short as possible. |
| **Be specific** | Create survey questions that explore one idea. Vague, general, or multi-part questions can be confusing; it is better to ask multiple specific questions (but avoid making the survey too long). |
| **Clarify** | Spell out everything that could be interpreted in more than one way. |
| **Keep it relevant** | For online surveys a good way to do this is by using skip logic to bypass questions which are not relevant for the current respondent. |
| **Avoid yes/no questions** | They are appropriate for capturing facts, but for questions about attitudes and opinions they don't capture nuances. |
| **Avoid long matrix questions** | A matrix allows respondents to familiarise themselves with the structure of the question and thus answer more quickly and efficiently. If it is too long however, they may focus on filling in the grid rather than paying careful attention to each question. |
| **Use words rather than numbers** | When designing answer choices, phrases such as "slightly likely" or "extremely likely" to indicate degrees of preference are generally easier for people to understand. However, longer numeric scales (e.g. 0-10) where the minimum and maximum are explained can also be appropriate, and will provide data that are more ready for numerical analysis. |
| **Use a balanced rating scale** | For opinion scales – there should be an odd number of points on the scale, the middle point being neutral, with an equal number of positive and negative options on either side of it, and a Not Applicable option. |

## Balanced rating scale

| ✗ Don't | ✓ Do |
|---|---|
| Excellent (***positive***)<br>Good (***positive***)<br>Fair (***neutral***)<br>Poor (***negative***) | Highly Satisfied (***positive***)<br>Satisfied (***positive***)<br>Neither Satisfied or Dissatisfied (***neutral***)<br>Dissatisfied (***negative***)<br>Highly Dissatisfied (***negative***)<br>Not Applicable |
| *This creates an unbalanced scale, making respondents more likely to feedback a positive score.* | *This system is recommended for the both the accuracy of the results and the ease of understanding for the respondent* |

***Illustration of bad and good practice when constructing a balanced rating scale***

### Conducting the survey

**Introduce the objective** with a brief sentence that provides context. Tell survey takers why you're asking these questions and how their feedback will help. Give them a contact email address to ask for additional explanation if needed.

**Test drive** the survey with a handful of people who are a representative sample of your intended audience.

### Reporting of Survey Results

Keep in mind the audience you are preparing the report for and their background and interests. The aim should be to convey the meaning of the numbers rather than simply the numbers themselves.

Generally, the survey report should contain:

- A table of contents;
- An executive summary for easy assimilation of key points and to provide an overview of the report;
- Key findings;
- A full description of the methodology;
- Sampling description;
- Response rate;
- A copy of the survey questionnaire;
- A list of tables and figures if there are many of them in the survey report, so the reader can easily find the material of primary interest.

Present categorical and short-scale ordinal variables as a distribution, but use descriptive statistics (e.g. mean, standard deviation) for numeric and long-scale ordinal variables. Refer to the number of respondents to each question as this can vary.

A large set of possible values can be arranged in more general groups – classify the data and present the classified distribution.

If sufficiently large groups of respondents can be distinguished in terms of content, then compare them with each other.

For each thematic module, consolidate assessments evaluated according to a common scale into a single graph, arranging the results in ascending or descending order.

Decide whether to present results as a graph or figure instead of a table. A graph is more suitable for general-interest audiences who need an overall impression rather than data.

## Value and Impact studies

**Use this section to learn more about associated methods and metrics if you need to pursue these topics in depth or are thinking of commissioning your own value and impact study**

This section provides some supplementary detail on value and impact studies for users of the toolkit, who are interested in the broader context.

There is a growing body of literature on the value and impact of science facilities. The 2013 EvaRIO evaluation of the impact of Research Infrastructures final report summary and the 2013 Technopolis Big Science and Innovation Report reviewed general approaches to measuring the effects and impacts of

research infrastructures. However, the emphasis tends to be on 'Big Science' facilities rather than on data repositories and related infrastructure and services. Methodologically, these studies fall into three main groups: those using various forms of Input-Output (IO) analysis; those featuring case studies and examples; and various forms of cost-benefit analysis, typically using activity costing and/or contingent valuation to underpin the analysis. These methods can be combined, with complementary use of qualitative and quantitative approaches highlighting the various dimensions and mechanisms through which value and impact can be determined.

## Input-Output analysis

'Big Science' facilities are typically focused on the generation of research data, but they may also host and curate data. The majority of economic impact assessments of such facilities follow a broadly similar approach, wherein evaluators take expenditure and employment data and feed them into an Input-Output (IO) analysis to estimate the direct and indirect benefits of public expenditure. Such evaluations arrive at economic multipliers that typically range between 2 and 3, which is to say that every 1 million in public expenditure is generating an additional 2 million to 3 million in wider economic activity through onward purchases within supply chains and the personal consumption of employees (see Technopolis 2013, p6).

## Case studies

Another and often complementary approach involves case studies, which typically follow the innovation impacts on suppliers and users through surveys and/or through tracing the development of spin-off firms and the use of information derived from the science facilities. Such case studies are widely used in the evaluation of research facilities and activities, and can focus on the scientific, economic, and/or wider social impacts. Among studies of larger facilities, those of CERN have reported the value of supplier contracts and the ways in which these have facilitated the development of new products or processes, and NASA's Spin-off Database reports on the number and revenue of spin-off firms emerging from the space agency's research work (see Technopolis 2013, p47).

While case studies provide concrete examples and often highlight the mechanisms through which impacts can be realised, they are limited because it is not possible to scale up a case study to estimate overall impacts. Consequently, case studies are often combined with broader economic estimates and/or formal frameworks for analysis.

## Mixed method approaches to cost-benefit analysis

Among previous studies adopting a more formal framework are a series of projects named Keeping Research Data Safe (KRDS) referenced in the toolkit.

A range of mixed methods for exploring the value, benefits, and impacts of research data and services have been developed and applied by Beagrie and Houghton in studies of the economic impact of research data centres (see links below). The studies combined qualitative and quantitative methodologies to measure the value and impact of research data and associated services and tools. They also consider the indicators and metrics that provide the data for the analysis.

Qualitative approaches included the KRDS Benefits Framework, interviews, and case studies.

Economic approaches used included estimates of access and use value, contingent valuation using stated preference techniques, an activity-costing approach to estimating the efficiency impacts of data and services, and a macro-economic approach that seeks to explore the impacts of the data archive's use on returns to investment in research. The economic approaches develop a picture, beginning with estimates of minimum direct values for a data archive's user community and moving progressively toward approaches

that measure wider social and economic value.

These studies have assessed the economic value and impact of the Economic and Social Research Data Service (2012), the Archaeology Data Service (2013), the British Atmospheric Data Centre (2013), and the EMBL European Bioinformatics Institute (2016). In addition to the individual studies, a short synthesis of the first three studies of UK research data centres, The Value and Impact of Data Sharing and Curation (2014), has been produced and is a useful summary of the key findings and issues.