

Benefits Factsheet



Introduction to benefits

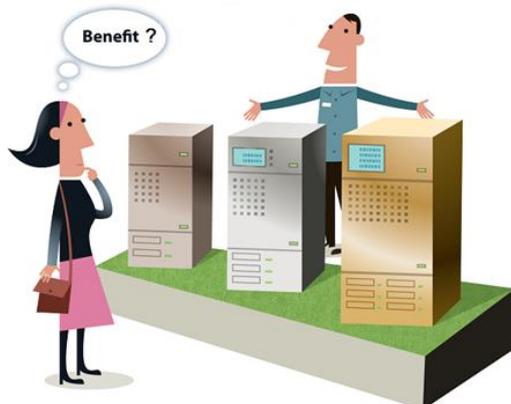


Illustration by Jørgen Stampdigitalbevaring.dk
CC BY 2.5 Denmark /modified from original

Identifying the benefits from curating social science research data is a critical task for a data archive if it is to obtain the support that it needs from funders, staff, depositors, and users.

This 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.

Key approaches

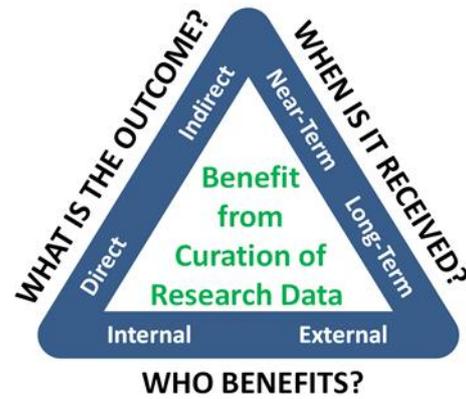
Keeping Research Data Safe (KRDS) Benefits Framework

To assist institutions in identifying and structuring benefits, the Keeping Research Data Safe (KRDS) project created the KRDS Benefits Framework. It serves as a high-level framework within which thinking about benefits can be brainstormed and organised. It aims to help institutions identify the full scope of benefits from management and preservation of research data and to present them in a succinct way to a range of different stakeholders (e.g. when developing business cases or advocacy).

The Framework organises benefits along three broad dimensions: the **outcome** achieved; **when** the outcome is achieved; and **who benefits** from the outcome. It helps identify the “what”, “when” and “who” of the value proposition for these activities.



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. Any benefit associated with a data curation/preservation activity can be characterised according to these three dimensions.



There is a KRDS Benefits Analysis Toolkit to help you apply these approaches (KRDS 2011). Each tool consists of a more detailed guide and worksheet(s). The KRDS Benefits Framework (Tool 1) is the “entry-level” and most widely adopted tool requiring less experience and effort to implement. The Value-chain and Benefits Impact Analysis (Tool 2) based on data lifecycle stages is the most advanced tool and requires more experience and effort to implement. It is more of a niche application.

The Anatomy of a Benefit
(KRDS User Guide 2011 figure 10).

Illustration by Charles Beagrie Ltd ©2011. CC-BY licensed

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

Illustration of the Benefits Summary for a Data Archive worksheet
Illustration by Charles Beagrie Ltd ©2017. CC-BY licensed.

- There is a tool in the CESSDA SaW Cost-Benefit Advocacy Toolkit that can help you start to apply KRDS Benefit Framework to your archive. The UK Data Service used the KRDS Framework to summarise and illustrate in qualitative terms the benefits arising from its activities as a social science data archive. This has been updated to provide a worksheet in the Toolkit (Benefits Summary for a Data Archive 2017).
- Using the KRDS methodology (deleting/adding/making benefits more specific to your archive, moving your key benefits to top of the lists) it can be used and modified by other social science archives to brainstorm and summarise the benefits from their archive in an accessible way.
- In KRDS, the UK Data Archive also developed a short concatenation formula to explain the impact of selected benefits to key stakeholders (UKDA 2011). It ended up looking something like this:

Given that [GENERIC BENEFIT] is a desirable benefit to [STAKEHOLDERS WHO PRINCIPALLY BENEFIT] the repository can [ACTION TO REALISE BENEFIT] in order to [EXPRESSION OF BENEFIT].

For example:

Given that **Demonstrating research data integrity** is a desirable benefit to **Data Users and Funders**, the repository can **ensure ease of citability** in order for researchers to **check the outcomes of others' research**.

This doesn't work perfectly (or grammatically) every time, but it is a useful starting point to provide more detailed explanation of the impact of selected brief KRDS benefit statements for funders and other key stakeholders.

- The KRDS Framework tool has also been used by FSD, the Finnish social science archive, and is discussed in their case study of benefits and cost tools (FSD case study 2017).

Physical and intangible assets

A traditional view of research data infrastructure and assessing its impact has tended to focus on its physical assets: its buildings and equipment, IT networks, and staff.

In recent years, the importance of research data as an intangible asset has been increasingly recognised. However, arguably this still provides only a very partial view of the work and benefits of data archives. A data archive is not solely about data: there is a broad spectrum of value-added activities.

Another helpful way of thinking about and advocating for data archives focuses on all the intangible assets and value-added activities undertaken and the benefits these bring to stakeholders.

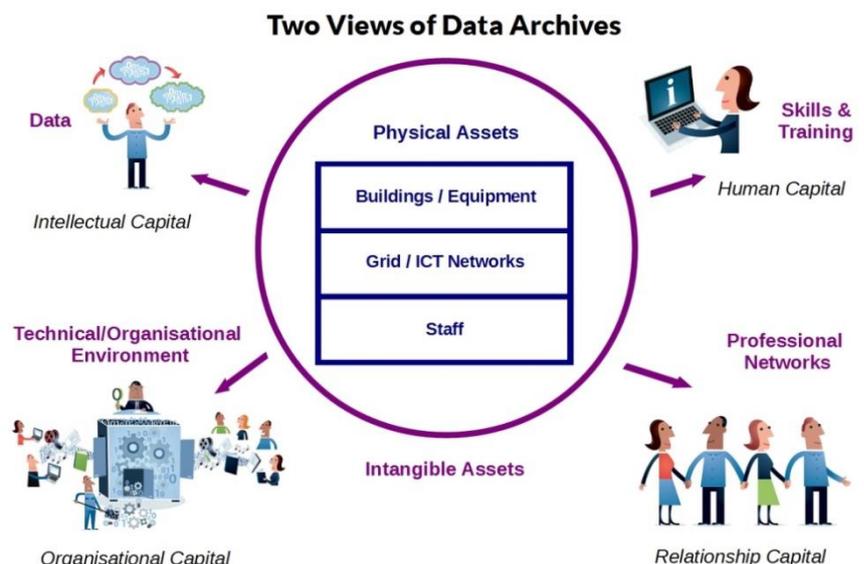
A broad four-part division of intellectual capital/human capital/organisational capital/relationship capital proposed for valuing in economic terms the intangible assets covered by digital preservation or digital curation (Hunter 2006) can provide, with some adaptation, a suitable definition for research data service infrastructure (Beagrie et al 2012, p41).

The intangible assets of a research data archive would be:

- **Research Data** [Intellectual Capital]
- **Skills and Training** [Human Capital]
- **Technical and Organisational Environment** e.g. software tools or ontologies [Organisational Capital]
- **Professional Networks** [Relationship Capital]

The roles of data archives as “competency centres” in fostering skills and training; developing tools, standards and ontologies; and disseminating innovation in research data management practice, within their countries and disciplines, can be as important as the data they hold.

Professional networks and relationship capital are critical in fostering innovation and best practice internationally. The role of CESSDA as a professional network for social science data archives in Europe is particularly important in this context.



Physical Assets and Intangible Assets.

Illustration by Charles Beagrie Ltd ©2017 incorporating images by Jorgen Stamp digitalbewaring.dk. CC-BY licensed.

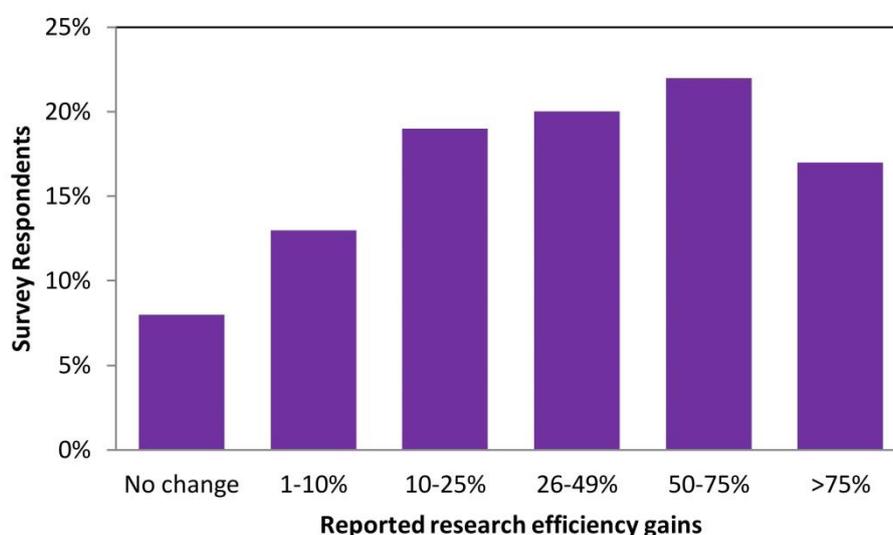
Key arguments

Research, teaching and study efficiencies

One key argument for data archives is that they help users to work more efficiently, saving time, and making the best use of public investment in research, teaching and learning. They save time and money, by making it easier for users to find what they need and by eliminating unnecessary re-creation of data. This is reflected in the responses of users in studies of the value and impact of data archives. Although these studies present emphatic responses they are usually qualitative results (e.g. RIN 2011).

The Value and Impact Study of the Economic and Social Data Service (ESDS) however sought to quantify the efficiencies reported (Beagrie et al 2012). It estimated major efficiency impacts for ESDS's active registered user community (excluding school and under-graduate students) of £68 million to £112 million per annum, which might translate to at least £100 million per annum or more for the wider user community.

The ESDS impact user survey had asked if respondents had any sense of the extent to which their use of data and services from ESDS had changed their research efficiency. The use-weighted mean responses suggested a 46% increase in research efficiency and a 23% increase in teaching efficiency.



Impact of using ESDS data and services on research efficiency (after Beagrie et al 2012, p77, Figure 15)

Economic and Social Research Council © 2012. CC-BY licensed

In a value and impact study of the Archaeology Data Service (which also examined the effect on students' learning) similar major efficiency gains of 44% for research and 32% for teaching were noted. In addition when asked to what extent their use of ADS data and services had changed their study/learning efficiency students reported a mean 44% efficiency gain (Beagrie and Houghton 2013a).

Further studies have looked at the British Atmospheric Data Centre (Beagrie and Houghton 2013b), and the European Bioinformatics Institute (Beagrie and Houghton 2016). Note although all these studies considered efficiency gains from research, consideration of teaching and study efficiencies was not as extensive.

Reported Efficiency Gains from Value and Impact Studies

	Research	Teaching	Study	Source
Economic & Social Data Service	46%	23%	N/A	Beagrie et al 2012
Archaeology Data Service	44%	32%	44%	Beagrie and Houghton 2013a
British Atmospheric Data Centre	28%	15%	34%	Beagrie and Houghton 2013b
European Bioinformatics Institute	46%	N/A	N/A	Beagrie and Houghton 2016

Efficiencies (time savings) reported by users, in research, teaching, and learning as a result of using the data center/service

Illustration by Charles Beagrie Ltd ©2017. CC-BY licensed

Open data/access/science

Social science archives support open data, open access and open science, whilst ensuring the protection of personal information. In doing so they support and help implement the aims and objectives of governments, research funders and other key academic bodies. These are some examples and quotes from European and international organisations that illustrate these points:

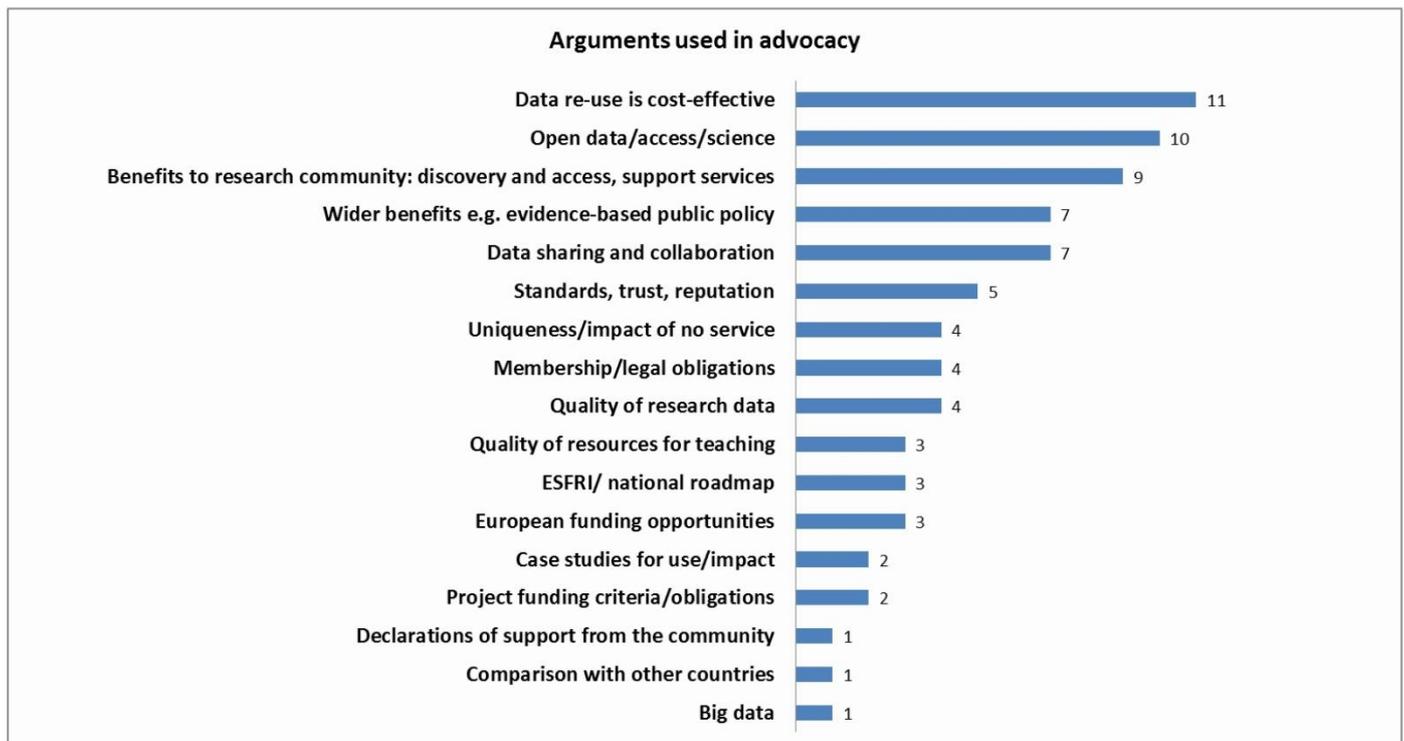
“Our vision is a scientific e-infrastructure that supports seamless access, use, re-use, and trust of data.” (High Level Expert Group on Scientific Data, 2010)

“Fuller and wider access to scientific publications and data can help to accelerate innovation (faster to market = faster growth); foster collaboration and avoid duplication of effort (greater efficiency); build on previous research results (improved quality of results); involve citizens and society (improved transparency of the scientific process). What is at stake is the speed of scientific progress and the return on R&D investment, and in particular publicly-funded investment, which has enormous potential for boosting productivity, competitiveness and growth.” (European Commission Horizon 2020, 2015)

“We are committed to openness in scientific research data to speed up the progress of scientific discovery, create innovation, ensure that the results of scientific research are as widely available as practical, enable transparency in science and engage the public in the scientific process... Open scientific research data should be easily discoverable, accessible, assessable, intelligible, useable, and wherever possible interoperable to specific quality standards.” (G8 Science Ministers, 2013)

“Data are the bedrock on which the scientific edifice is built. More efficient data-sharing and more open access to information and resources will make it easier for observations to be confirmed...and ultimately for answers to societal challenges to be given.” (ALLEA, The European Federation of National Academies and Humanities, 2012)

Other benefits and arguments used by social science archives



Arguments used in advocacy by European social science archives to make their funding case (N=20)
(source CESSDA-SaW Cost-Benefit Advocacy User Requirements Survey 2016). Charles Beagrie Ltd ©2017. CC-BY licensed

Metrics and case studies

Metrics or indicators requires careful evaluation of what to measure and understanding of their relative strengths and weaknesses. While some metrics employ readily countable things such as downloads, others may use values that are not truly numeric, such as rating scales. Often the easiest and most available indicators (e.g. unique web visitors, downloads or hits) are approximations for actual things we would like to measure (e.g. users and use).

These are some of the most common ones you may wish to consider using to quantify or illustrate benefits:

- Deposit Metrics
- Web Statistics
- User Registration Data
- User Satisfaction Data

Return on Investment

Return on Investment or ROI is a key economic metric quantifying benefits against costs (investment). The ESDS impact study published in 2012 is currently the only example of a fully developed quantified economic impact study and ROI metrics for social science research infrastructure. It found there was a 5.4 to 1 benefit/cost ratio for the ESDS. ROI is discussed more fully in a separate factsheet in this series.

Data citations

Funders are interested in the number of publications that cite uses of the data that they fund. Funders often consider citation counts as one potential measure of impact as they can indicate the amount of secondary analysis that is taking place as researchers continue to exploit the analytic potential of existing data collections to replicate results, extend findings, and generate new avenues of research. We know that formal archiving of data across a range of disciplines can lead to a greater number of papers based on those data (Pienta et al 2010, Piwowar et al 2007, Ember and Hanisch et al 2013).

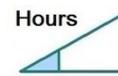
Case studies

Another and often complementary approach involves case studies, which typically follow the impacts on users and beneficiaries through surveys and/or through tracing the use of information derived from the data archive. However, the use of case studies alone to demonstrate impact is problematic as they tend to be small-scale and highly variable; they cannot be scaled-up to give a broader picture. Consequently, case studies are often used to provide illustrative examples, highlight the mechanisms through which impacts have been realised, or combined with quantitative approaches (e.g. Academy of Social Sciences 2013, Beagrie et al 2012, Woollard 2015).

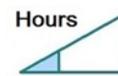
Linked toolkit resources

Effort

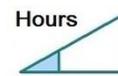
Archive Development Canvas (Detailed Version), <http://dx.doi.org/10.18448/16.0009>



Costs Factsheet, <http://dx.doi.org/10.18448/16.0003>



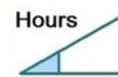
Case study on using benefit and cost tools, <http://dx.doi.org/10.18448/16.0006>



Return on Investment Factsheet, <http://dx.doi.org/10.18448/16.0002>



Benefits Summary for a Data Archive, <http://dx.doi.org/10.18448/16.0010>



Other references

Academy of Social Sciences, 2013, *Professional Briefing – impact* <http://campaignforsocialscience.org.uk/wp-content/uploads/2013/06/Professional-Briefings-2-June-2013-Impact-RGB.pdf>

ALLEA, The European Federation of National Academies and Humanities, 2012, *Open Science for the 21st Century*, <http://www.gsrt.gr/EOX/files/Open%20Access%20Declaration.pdf>

Beagrie, N., Houghton, J., Palaiologk, A., and Williams, P., 2012, *Economic Impact Evaluation of the Economic and Social Data Service*, <http://www.esrc.ac.uk/files/research/evaluation-and-impact/economic-impact-evaluation-of-the-economic-and-social-data-service/>

Beagrie, N. and Houghton, J., 2013a, *The Value and Impact of the Archaeology Data Services: A Study and Methods for Enhancing Sustainability*, Joint Information Systems Committee, Bristol and London. <http://www.jisc.ac.uk/whatwedo/programmes/preservation/ADSImpact.aspx>

- Beagrie, N. and Houghton, J., 2013b, *The Value and Impact of the British Atmospheric Data Centre*, Joint Information Systems Committee and the Natural Environment Research Council UK, Bristol and London. http://repository.jisc.ac.uk/5382/1/BADCReport_Final.pdf
- Beagrie, N. and Houghton, J., 2016, *The Value and Impact of the European Bioinformatics Institute Full Report*, <http://www.beagrie.com/EBI-impact-report.pdf>
- Ember, C., Hanisch, R. et al, 2013, *Sustaining Domain Repositories for Digital Data: A White Paper*, <http://dx.doi.org/10.3886/SustainingDomainRepositoriesDigitalData>
- European Commission, 2015, *Fact Sheet Open Access to scientific publications and research data in Horizon 2020: Frequently Asked Questions (FAQs)*, <https://www.iprhelphdesk.eu/sites/default/files/newsdocuments/Fact-Sheet-Open-Access-to-Publications-and-Data-in-H2020-FAQ.pdf>
- G8 Science Ministers, 2013, *G8 Science Ministers Statement London UK, 12 June 2013*, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/206801/G8_Science_Meeting_Statement_12_June_2013.pdf
- High Level Expert Group on Scientific Data, 2010, *Riding the wave - How Europe can gain from the rising tide of scientific data - Final report of the High Level Expert Group on Scientific Data - October 2010*, http://ec.europa.eu/information_society/newsroom/cf/document.cfm?action=display&doc_id=707
- Hunter, L., 2006, *Investment in an Intangible Asset*, Chapter in DCC Digital Curation Manual (Digital Curation Centre 2006), <http://www.dcc.ac.uk/resources/curation-reference-manual/completed-chapters/investment-intangible-asset>
- Keeping Research Data Safe (KRDS), 2011, KRDS Benefits Analysis Toolkit, <http://beagrie.com/krds-i2s2.php>
- Pienta, A.M., Alter, G.C. and Lyle, J.A., 2010, *The Enduring Value of Social Science Research: The Use and Reuse of Primary Research Data*, <http://hdl.handle.net/2027.42/78307>
- Piwowar, H.A., Day, R.S. and Fridsma, D.B., 2007, *Sharing Detailed Research Data Is Associated with Increased Citation Rate*, <http://dx.doi.org/10.1371/journal.pone.0000308>
- Research Information Network (RIN), 2011, *Data centres: their use, value and impact*, September 2011. http://www.rin.ac.uk/system/files/attachments/Data_Centres_Report.pdf
- UKDA 2011, *KRDS Benefits Framework, Value-Chain and Benefit Analysis Tools: UK Data Archive Case Study*, http://www.beagrie.com/static/resource/UKDA-KRDS_ImpactWorksheet.pdf
- Woollard, M., 2015, *Data enriched research, data enhanced impact: the importance of UK data infrastructure*, The Impact Blog LSE, <http://blogs.lse.ac.uk/impactofsocialsciences/2015/11/17/the-importance-of-uk-data-infrastructure/>