

A GUIDE TO THE KRDS BENEFITS FRAMEWORK v3 (JULY 2011)

1. INTRODUCTION

The Keeping Research Data Safe (KRDS) Benefits Framework is a tool for identifying, assessing, and communicating the benefits from investing resources in the curation/long-term preservation of research data. The Framework employs a simply structured, easily understood format, and is intended to aid internal discussions amongst research project staff, personnel from central IT or research data centre and archive services, Higher Education or funding organisation strategy and policy makers, as well as to support discussions between these groups. The Framework can assist in prioritizing alternative curation investments, and justifying data curation costs within funding applications.

The Benefits Framework organises potential benefits from the curation/preservation of research data along three broad dimensions: the outcome achieved; when the outcome is achieved; and who benefits from the outcome. Assessing a data curation activity's benefits, as well as communicating these benefits to stakeholders, requires a clear understanding of the fundamental elements of the activity's value proposition. In short, the "what", "when", and "who" of the value proposition must be identified and described.

2. THE BENEFITS FRAMEWORK

Each dimension of the Benefits Framework is divided into two categories:

Dimension 1: What are the outcomes?

- *Direct benefits:* positive impacts obtained from investing in a data curation activity.

Examples: Create new research opportunities through re-purposing/re-use of data; fulfil long-term data management mandates from funding organisations; strengthen integrity of scholarly record by facilitating verification of research results.

- *Indirect benefits:* negative impacts avoided by investing in a data curation activity.

Examples: Avoid cost of re-creating data at a later time; reduce overall cost of long-term preservation by curating data early in the digital life cycle; protect return on investment in data creation.

A GUIDE TO THE KRDS BENEFITS FRAMEWORK v3 (JULY 2011)

Dimension 2: When are the benefits received?

- *Near-term benefits*: benefits expected to be received up to five years from the present.
- *Long-term benefits*: benefits expected to be received beyond five years from the present.

Dimension 3: Who Benefits

- *Internal benefits*: benefits which impact stakeholders internal to/affiliated with the organisation undertaking the data curation activity.
- *External benefits*: benefits which impact stakeholders external to/not affiliated with the organisation undertaking the data curation activity.

The three dimensions of the Benefits Framework, along with the two-part subdivision within each dimension, are applicable to nearly all research data curation/preservation activities, and therefore can be used without modification in most contexts. However, the Benefits Framework is extensible, in the sense that each dimension can be subdivided into as many categories as appropriate for a particular curation/preservation activity. The categories listed above reflect the distinctions found to be most widely applicable and useful across a range of curation activities and contexts. For example, the near-term/long-term distinction in the second dimension reflects the planning horizons of most project and post-project timescales or organisational business plans. However, there may be circumstances where a tri-partite distinction of near-term, medium-term, and long-term is more appropriate (See Section 1.5 for more information on this topic).

Each of the dimensions in the Benefits Framework provides a different yet complementary view of a particular source of value – i.e., a benefit – generated from a data curation activity. Put another way, any benefit associated with a data curation activity can be characterised according to the three dimensions presented in the Framework: i.e., what the outcome of the benefit will be; when the outcome will be achieved, and who will benefit from the outcome. The Framework should *not* be construed as suggesting that some benefits are characterised

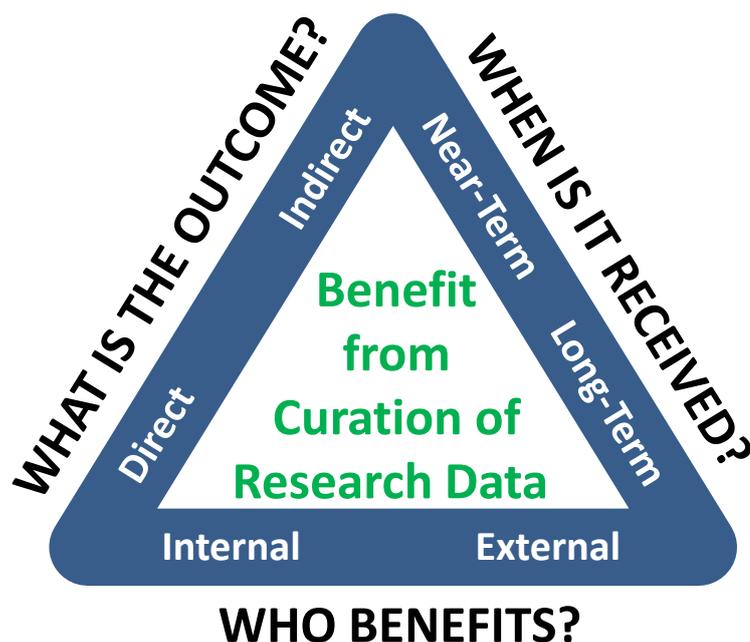
A GUIDE TO THE KRDS BENEFITS FRAMEWORK v3 (JULY 2011)

by the first dimension, others by the second dimension, and so on. Instead, *each* benefit is characterised by the three attributes of *what*, *when*, and *who*.

Another point to keep in mind is that a particular benefit can sometimes be characterised in multiple ways. For example, a benefit can create value in the short term, as well as in the long term. Similarly, a benefit can bring value to those internal to the digital preservation activity, but also to the community at large. Users of the Benefits Framework should not view the categories within the dimensions as mutually exclusive. However, in most cases, the benefit will be primarily associated with one category or another within a particular dimension: for example, primarily short-term in regard to the timing of the value received or primarily internal in regard to the community impacted by the benefit.

The Benefits Framework is summarised graphically in the figure below.

Anatomy of a Benefit:



2.1. THE BENEFITS FRAMEWORK: EXAMPLES

Example A: Fulfil research grant data management obligations from funding agencies.

Dimension 1 (Outcome): Direct

Dimension 2 (When): Near-term

A GUIDE TO THE KRDS BENEFITS FRAMEWORK v3 (JULY 2011)

Dimension 3 (Who Benefits): Internal

Explanation: The benefit is direct, in the sense that creating the capacity to fulfil funding agency requirements to preserve the data products from research projects represents a distinct, positive source of value. The benefit is primarily near-term, in the sense that it permits current researchers to fulfil the terms of their grant awards. The benefit is internal, in that it primarily extends to those with the appropriate affiliation/authorisation to deposit data into the repository.

Example B: Strengthen integrity of scholarly record by facilitating verification of research results.

Dimension 1 (Outcome): Direct

Dimension 2 (When): Long-term

Dimension 3 (Who Benefits): External

Explanation: The benefit is direct, in that investing in the long-term curation of research data permits verification/replication of published research results; this has the positive effect of improving the quality and integrity of the scholarly record, which in turn benefits the research and learning activities that rely on this record. The benefit is primarily long-term, in the sense that the ability to verify or replicate previous research results increases confidence in the permanent scholarly record, and enhances its value to catalyze future research. The benefit is external, because the scholarly record is by and large a globally-shared resource, and improvements in its quality benefit researchers everywhere.

Example C: Avoid cost of re-creating data at a later time.

Dimension 1 (Outcome): Indirect

Dimension 2 (When): Long-term

Dimension 3 (Who Benefits): Internal

Explanation: The benefit is indirect, because it represents a “cost avoided”: by incurring the cost of preserving a research data set now, the potentially much larger cost of re-creating that data at a later time, should it be needed, is avoided. The benefit is long-term, in that once the research data is properly preserved, it can continue to be re-used (thereby avoiding

A GUIDE TO THE KRDS BENEFITS FRAMEWORK v3 (JULY 2011)

the costs of re-creation) long into the future. The benefit is internal, because it is likely that the organisation responsible for investing in the curation of the data would also be responsible for the cost of re-creating it should it become lost.

2.2. THE BENEFITS FRAMEWORK AS A TOOL

The Framework is a starting point for projects wishing to identify and evaluate the benefits from preservation/curation activities. The dimensions serve as a high-level framework within which to organise thinking about preservation benefits. The structure of the Framework can be customised and extended as needed, given the circumstances of specific projects or institutions. This is described in greater detail in section 1.6.

The Framework is especially useful as a tool for supporting and organizing early-stage brainstorming on the benefits associated with a particular activity. The Framework will assist in identifying potential benefits from a project or activity, and articulating these benefits to a broad audience of stakeholders. It can also aid in customising the expression of these benefits to address different stakeholder audiences.

Application of the Benefits Framework to a range of projects over the course of its development has revealed a number of common benefits that frequently arise from preservation of research data. Often, these can be simply expressed in a generic form independent of project specifics. A list of examples of these “generic” benefits is provided below as a starting point for applying the Framework. *Please note this is not a comprehensive list of potential benefits from preserving research data.* Specific activities should feel free to add new benefits associated with their particular circumstances, modify their wording for their specific context, or delete benefits that do not apply. **A separate worksheet is provided to assist in this process.**

Examples of Common Benefits

New research opportunities	No re-creation of data
Input for future research	No loss of future research opportunities
Motivating new research	Secures value to future researchers & students

A GUIDE TO THE KRDS BENEFITS FRAMEWORK v3 (JULY 2011)

New research funding	Protecting returns on earlier investments
Increasing research productivity	Lower future preservation costs
Stimulating new networks/collaborations	Planned management from an early stage in the research life-cycle is ultimately more cost-effective than late intervention (providing proper selection of what to keep is done)
Knowledge transfer to other sectors	
Knowledge transfer to industry	
Commercialising research	Re-purposing data for new audiences
Increasing skills base of researchers/students/staff	Use by new audiences
Increasing economic growth	Re-purposing methodologies
Catalysing new companies and high skills employment	Enhancement of research tools and software by testing on a range of well-curated datasets
Verification of research/research integrity	Scholarly communication/access to data
Fulfilling organisational mandate(s)	Long-term re-use of well curated data
Fulfil research grant obligations	Short-term re-use of well curated data
Value to current researcher & students	Adds value over time as collection grows and develops critical mass
No data lost from Post Doc turnover	Increased visibility/citation
Secure storage for data intensive research	
Availability of data underpinning published findings	

2.3. THE BENEFITS FRAMEWORK IN PRACTICE

The Benefits Framework was tested and critiqued in the context of several research data preservation activities, with the results reported in Chapter 8 of *Keeping Research Data Safe 2* (April 2010). It is a flexible tool which can be used in a variety of ways to analyse and present benefits. Two potential methods of presentation are shown below.

One method of presentation can be derived directly from the worksheet accompanying this Guide, and the application of the list of examples of common benefits above to the three dimensions of the Framework. Although a benefit will always have attributes that appear in all three dimensions of the Framework, in practice a specific benefit often has a much

A GUIDE TO THE KRDS BENEFITS FRAMEWORK v3 (JULY 2011)

stronger impact in one dimension: for example in a particular timeframe or on a particular group of beneficiaries. Hence a decision can be made for presentation purposes to characterise it solely in a dimension(s) where it is particularly significant. A worked example of this for the Worksheet accompanying this Guide is shown below (using just the common benefits from the list of examples without further modification). As noted above, users are encouraged to add, delete, or modify the common benefits to fit the circumstances of a specific activity.

Dimension 1: What are the outcomes?

Direct Benefits	Indirect Benefits (e.g. costs avoided)
New research opportunities	No re-creation of data
Scholarly communication/access to data	No loss of future research opportunities
Re-use of well-curated data	Lower future preservation costs
Increasing research productivity	Re-purposing data for new audiences
Stimulating new networks/collaborations	Re-purposing methodologies
Knowledge transfer to other sectors	Protecting returns on earlier investments
Increasing skills base of researchers/students/staff	
Increasing economic growth	
Verification of research/research integrity	
Fulfilling organisational mandate(s)	

A GUIDE TO THE KRDS BENEFITS FRAMEWORK v3 (JULY 2011)

Dimension 2: When are the benefits received?

Near-Term Benefits (up to 5 years)	Long-Term Benefits (5 years+)
Short-term re-use of well curated data	Long-term re-use of well curated data
Value to current researcher & students	Secures value to future researchers & students
No data lost from Post Doc turnover	Adds value over time as collection grows and develops critical mass
Secure storage for data intensive research	
Motivating new research	Input for future research
Availability of data underpinning published findings	Planned management from an early stage in the research life-cycle is ultimately more cost-effective than late intervention (providing proper selection of what to keep is done)

Dimension 3: Who benefits?

Internal Benefits	External Benefits
Motivating new research	Knowledge transfer to other sectors
New research funding	Catalysing new companies and high skills employment
Stimulating new networks/collaborations	Increasing economic growth
Fulfil research grant obligations	Scholarly communication/access to data
Increased visibility/citation	Verification of research/research integrity
Commercialising research	

A second method of presentation is illustrated in the next figure, which displays some of the benefits selected and described in detail from real-world situations with the aid of the Framework. It demonstrates how, as a next step in characterizing benefits, the simply expressed common benefits in the list of examples above can be selected and extended into more detailed descriptions specific to particular projects if required.

A GUIDE TO THE KRDS BENEFITS FRAMEWORK v3 (JULY 2011)

DIRECT BENEFITS

[New research opportunities]. A direct benefit from continued access to data at the UKDA is the ability of researchers to use data which they did not create themselves. Survey data collected by government agencies in the UK may never have been accessible to the research (and/or wider) communities had it not been for their preservation at the UKDA. The re-use of government data, especially of the major surveys (e.g., British Social Attitudes Survey), has propelled research across a wide range of disciplines. (KRDS2, pp. 70)

NEAR-TERM BENEFITS

[No data lost from Post-doc turnover]. The constant turnover of post-doctoral researchers often results in lost data. Currently, there are no established mechanisms to routinely collect and organise the data that post-doctoral researchers generate. In some cases, researchers that generated data several years ago could not make sense of them now as they had not kept enough information on how the data was created. In these circumstances, well-curated data has clear near- and medium-term benefits. (KRDS2, p.60)

INTERNAL BENEFITS

[Increased visibility/citation]. A curated and preserved research data set may generate internal benefits if the research data set is made publicly available and is frequently used and re-used by external researchers, this may increase the visibility and impact of the original research, and by extension, enhance the reputation and standing of the researcher and the institution in which it was created. (KRDS2, p. 62)

INDIRECT BENEFITS (COSTS AVOIDED)

[Lower future preservation costs]. The Digitale Bewaring Project in the Netherlands, which focused on government electronic records, estimated that the creation of a batch of 1,000 appropriately-documented records during the Pre-Archive phase would cost approximately 333 euros. Conversely, once 10 years have elapsed since creation it may cost 10,000 euros to 'repair' a batch of 1,000 records with badly created metadata. (KRDS1, p.25)

LONG-TERM BENEFITS

[Adds value over time as collection grows and develops critical mass]. One advantage of archiving data over many years is that long time series of consistent data are built up. Richard Berthoud has analysed the General Household Survey between 1974 and 2005, to describe changing patterns of advantage and disadvantage in employment. The analysis was described by the civil servant responsible for commissioning the research as having made more difference to policy thinking than any other project for which he had been responsible. (KRDS2, p.72)

EXTERNAL BENEFITS

[Catalysing new companies and high skills employment]. External benefits may manifest themselves on a variety of scales: across a group of collaborating universities, across the scientific community as a whole, and even on an economy-wide scale, to the extent that long-term preservation of research data enhances the prospects for commercialising scientific discoveries, catalysing new companies, and expanding opportunities for high-skill employment. (KRDS2, p.62)

A GUIDE TO THE KRDS BENEFITS FRAMEWORK v3 (JULY 2011)

2.4. BUILDING ON THE FRAMEWORK

The Framework is a starting point for projects wishing to identify and evaluate the benefits from preservation/curation activities. The dimensions serve as a high-level framework within which to organise thinking about preservation benefits. The structure of the Framework can be customised and extended as needed, given the circumstances of specific projects or institutions.

For example, Dimension 3 (Internal/External) in the Framework could be further sub-divided by more specific groups of stakeholders if desired. An illustration of this, populated with some examples of common benefits, is provided in the figure below.

Dimension 3 (Who Benefits) Sub-divided by a University's Stakeholders					
Internal Benefits			External Benefits		
Researcher	Research Group	Institution	Research Funder	Discipline	Others (e.g. NHS, etc)
Increased visibility/citation	No data lost from Post Doc turnover	Fulfilling organisational mandate(s)	Increasing research productivity	Scholarly communication /access to data	Knowledge transfer to other sectors

3. NEXT STEPS: MEASURING BENEFITS

Once benefits are identified and organised within the Benefits Framework, further work can proceed aimed at identifying potential measures or illustrations of the value of those benefits. This next stage is supported by a Value-Chain and Benefit Impact Analysis Tool in the Toolkit. This tool helps assign identified benefits to specific activities or phases of activity (based on the KRDS Activity Model) to create a value-chain analysis. It also supports further identification of metrics and qualitative illustrations of benefits in the Framework. Outcomes from using the Benefits Framework can be re-used in applying the Value-Chain and Benefit Impact Analysis tool.

A GUIDE TO THE KRDS BENEFITS FRAMEWORK V3 (JULY 2011)



Copyright HEFCE 2010, Charles Beagrie Ltd 2010, and KRDS Partners 2011.

This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 2.0 UK: England & Wales License](#). Conditions: Attribution — You must give the originators credit (Keeping Research Data Safe (KRDS) and Charles Beagrie Ltd). Non-Commercial — You may not use this work for commercial purposes. Share Alike — If you alter, transform, or build upon this work, you may distribute the resulting work only under a licence identical to this one.