A Guide to the KRDS Value-Chain and Benefit Impact Analysis Tool v1 (July 2011)

1. Introduction

This tool is the second and most advanced tool in the Digital Preservation Benefits Analysis toolset. It is intended to support a value-chain analysis of benefits and also to help you to identify quantitative metrics and qualitative indicators for their impact.

Value-chain analysis is a powerful tool for assessing where value is added to outputs in a chain of activities and for use in strategic and organisational planning. It has been widely used in industry for several decades but the industry activity models employed are less suitable for research data and curation/preservation activities. Our tool uses the KRDS Activity Model as a starting point for the value-chain analysis, so it is better suited to the specific needs of research data and its curation/preservation.

The impact component of the tool helps identify potential quantitative metrics or qualitative indicators for the value of the benefits identified. As such it goes a step further towards assessing impacts than the KRDS Benefits Framework. There are several fields in common between the two tools, so in applying it you can build on outputs from your initial use and adaptation of the KRDS Benefits Framework (tool 1).

This tool is based on an approach developed for the I2S2 project and used by the I2S2 project partners in the I2S2 Benefits Case Studies to assess benefits from I2S2 changes in research data management for the physical sciences. It has been tested and developed further in conjunction with partners from a range of backgrounds and disciplines in the Digital Preservation Benefits Analysis Tools Project.

2. Using the Tool

Two worksheets accompany this guide:

- **Benefit Impact worksheet**. This focuses solely on helping you to identify quantitative and qualitative impacts from your benefits without any reference to the research data lifecycle phases and activities, or value-chain analysis. It is relatively simple to implement if you have already used the KRDS Benefits Framework and have a number of outputs and experience from this which can feed into the worksheet. It is

likely to be most useful to those working on research projects or with short time horizons.

- **Value-chain and Benefit Impact worksheet**. This adds the research data lifecycle phases and activities and allows you to place benefits and their impact into a context of related activities and to see and evaluate a value-chain. It incorporates all of the fields from the Benefit Impact worksheet that help you to identify quantitative and qualitative impacts from your benefits. It is likely to be most useful to those working on strategic and organisational planning in research data services and archives or on research projects with longer-time horizons such as cohort studies. It is likely to require a higher level of experience and range of skills to use to maximum effect.

We recommend 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 (tool 1), other KRDS Outputs such as the KRDS Activity Model, and similar assessments of value and impact.

The tool has been designed to be generic but easily configurable by the user for their specific needs or application. You can edit or add to all pre-populated data fields once you have understanding of the tool, for example to “translate” terminology which may be unfamiliar to your target audience, or to add a benefit not already identified. You can also hide any components of the tool not relevant to your specific needs. You should add to all relevant clear cells in the tool shown with white background but you can also leave cells blank if appropriate.

Help text and definitions are available for components and terms in the tool. Click on any cell marked with an information symbol to reveal the help text or definition.

There are accompanying guides and introductory leaflets that explain the components of the toolset and uses of each tool and also worked examples of the tool for different case studies on the Digital Preservation Tools Benefits Analysis project web page. We recommend looking at these before using the tool.

Further information on KRDS (including a Factsheet and User Guide) and the wider range of KRDS tools and applications is also available on the KRDS web page.

3. GETTING STARTED

To use this tool, you should first select which worksheet most closely matches your needs. You may find it helpful to print out this guide for reference whilst completing the worksheet.

Further information on using each of the major fields in the worksheets is as follows:

KRDS Lifecycle Phase and KRDS Activity: If you are using the Value-chain and Benefit Impact worksheet, there is further information on the KRDS Activity Model on the KRDS web page.

The KRDS Activity Model is organised in a nested hierarchy of levels beginning with the main high-level phases of the research data lifecycle: Research (Pre-Archive); Archiving and Data Sharing (Archive); and Support Services. Typically Pre-Archive activities relate to all activities related to data creation and management for research projects in universities or other research institutes prior to archiving, and Archive activities to data archiving repositories run by universities or third-parties. Activities in Support Services can support either Pre-Archive or Archive activities and typically will be part of the existing infrastructure for finance, IT, and other common services. Estates are the category for buildings and other infrastructure.

KRDS caters for potential dual application of the Activity Model with two “versions” presented at different levels of detail. This tool uses the KRDS “Lite” version of the Activity Model consisting of just: the main Phases, e.g. Archive; and the Activities e.g. Ingest. This provides a high-level granularity of the Model suitable for allocating generic benefits to either phases as a whole or individual activities within them.

Generic Benefit: Both worksheets have been pre-populated with a selection of common generic benefits also used in the Framework tool but you may review, delete or add more to the selection (or also move/repeat common benefits if you are using the Value-chain and Benefit Impact worksheet). This will need you to insert additional row(s) in the worksheet you are using.

Your Expression of Benefit: You should adapt the generic benefit to match the specific benefit for your organisation, and explain how that benefit is of value to your organisation and stakeholders, i.e., for "No re-creation of data" your expression of benefit might read "Curating data will avoid the costs of collecting new data as it is already held within an existing data collection".

Action(s) to realise Benefit: Complete this field to ensure that you can evaluate work and resources needed to achieve the benefit e.g. licences/protocols will need to be/are in place to allow the reuse of data.

KRDS Outcome Type: Enter either direct or indirect as an outcome type for each benefit. This is derived from Dimension 1 of the KRDS Benefits Framework. Direct benefits are positive impacts obtained from investing in a data curation activity. Indirect benefits are negative impacts avoided by investing in a data curation activity. For example in the case of “no re-creation of data”, the benefits are indirect because it is an avoidance of a negative impact.

Years to benefit: Record here the numbers of years to when you think the benefit will begin. This field is adapted from Dimension 2 of the KRDS Benefits Framework.

Stakeholders who principally benefit: List internal and/or external stakeholders who will be the main beneficiaries. For example for data re-use this may be a research funder. This field is derived from Dimension 3 of the KRDS Framework.

Quantitative Impact of benefit: list any quantifiable impact(s) of the benefits e.g. cost savings. For example in the case of data re-use, this is a potentially quantifiable high impact for the funder (if they create useful data in the first place) and for the service this is likely to be a medium to low impact, depending on the amount of repository costs avoided. Note further guidance on impact assessment is available in section 4 of the guide.

Qualitative Impact of benefit: list any unquantifiable impact(s) of the benefits that might be illustrated in other ways e.g. in a case study. Note further guidance on impact assessment is available in section 4 of the guide.

Weighting of impact: You should finally assign an impact weighting (from 1[low] to 5 [high]) based on careful consideration and balancing of data in all the columns. The assessment of

impact weightings is necessarily subjective but multiple reflections and skilled input from a team considerably strengthens its rigour and value. It is to be expected that assessment of value may vary between different individuals and stakeholders (for example between researchers, research support services and archives, or funders) and it is desirable to indicate which perspective is being used or to conduct complementary assessments from key perspectives and synthesise them (an approach used in I2S2 for researcher and service provider perspectives).

The weighting is done to aid prioritisation for investment when planning future work, and/or selection of benefits/changes with greatest impact for further analysis if resources do not permit all to be examined. This is a subjective weighting of potential impact written from the perspective of the tool user(s). Activities may not be of equal scale in terms of benefit, cost to implement, or risk, and alongside consideration of timescales for benefits and feasibility of demonstrating impact, this can be reflected in the overall impact weightings. As noted above this weighting may best be done as a team exercise, involving senior manager(s) and independent support which has knowledge of the toolset and other similar assessments as appropriate.

Comments: The Comments column allows you to record any free-text comments on the process for others or notes for future reference.

4. NEXT STEPS: DEMONSTRATING IMPACT OF BENEFITS

Demonstrating the impact of benefits for research data curation/preservation either directly via metrics (quantitative impacts) or qualitatively (qualitative impacts) via illustration in case studies is still a relatively novel area. You may find the following sources valuable in working through how to demonstrate the impact of benefits and implement capturing the relevant measures/illustrations you have identified in the worksheets.

There are still only a relatively small number of impact studies focussing specifically on data services or research data infrastructure. For a recent review of such studies see the JISC Research Data Management programme "Guide on cost/benefit analyses for research data".

Some of the common challenges identified in these impact assessments are listed below. A list of possible metrics for measuring benefits arising more generally from research data management, produced by infrastructure projects in the JISC Managing Research Data Programme, is also provided together with a brief introduction to impact in research council projects. These sources may be helpful in considering how to select and further develop possible metrics and qualitative indicators for your specific use case when you have completed a worksheet included in the tool.

4.1. THE CHALLENGES IN DEMONSTRATING IMPACT FOR RESEARCH DATA CURATION/PRESERVATION

There are many challenges in measuring the scale of identified benefits and the impact measurements or indicators for them. These include:

Absence of Data Citation

There are still real barriers of existing conventions and practice around citation that prioritise secondary works (articles etc) and tend not to include data in the list of references, or limit the length of such lists so multiple data sources are excluded. These current limitations clearly mitigate the application of data citation metrics for demonstrating impact. Furthermore any changes may take considerable time to reach a critical mass of adoption and citations.

Timescales and Attribution

Digital curation/preservation is a long-term activity. Hence measurements of benefits are challenging. Many benefits will only emerge over a longer timescale. Longer-term effects tend to arise from a complex combination of developments and circumstances which can be difficult or impossible to disentangle and attribute to use of a single data repository, dataset, or digital curation/preservation activity.

Maturity and Critical Mass

Usage of repositories and acceptance of data sharing often grows very slowly at first and may only become significant when a critical mass of data has been assembled. The length of time before or the point in time when a repository or dataset's impact is assessed can therefore be critical.

4.2. ASSESSING BENEFITS AND METRICS IN THE RDMI PROJECTS

Identification of benefits using the KRDS Benefits Framework was successfully applied by a range of the Research Data Management Infrastructure (RDMI) projects. Despite the short timescale of the RDMI projects and other challenges, many were able to identify potential metrics that have or would provide measurement of specific benefits for different stakeholders summarised in the table below:

Summary of Metrics Identified by the RDMI Projects
Benefits Metrics for Institutions • New research grant income • Number of research dataset publications generated • Number of research papers • Improvements over time in benchmark results (e.g. repeats of AIDA benchmarking or surveys of awareness of relevant support services or funder requirements) • Cost savings/efficiencies for central services and/or departments • Re-use of infrastructure in new projects
Benefits Metrics for Researchers/Research Teams • Increase in grant income/success rates • Increased visibility of research through data citation • Percentage improvement in routine back-up of data • Average reduction in waiting time (time latency) for data requests • Average time saved in research data management and grant proposal activities • Percentage improvement in range/effectiveness of research tool/software
Benefits Metrics for Research Support Services • Percentage of potential user community that takes up services • Number of data deposits with a repository • Number of downloads of a dataset(s) within a repository • Activity based costing methods (e.g. using KRDS activity model to benchmark activity based costs over time) • Results of user feedback forms • Number of times different researchers collectively create/maintain a dataset via the repository

Benefits Metrics for Scholarly Communication and Access

- Number of citations to datasets in research articles
- Number of citations to specific methods for research data management
- Number of datasets deposited with enhanced metadata
- Percentage increase in user communities
- Number of service level agreements for nationally important datasets

4.3. ASSESSING IMPACT IN RESEARCH COUNCIL FUNDED PROJECTS

Assessing impact is an important component of the strategies of the UK research councils and in their expectations of projects and services that they fund. However it is defined in a very specific way to focus on stakeholders outside of academic research institutions. You will need to take this into account if using the tools for research council funded work. Impact is defined by Research Councils UK (RCUK) as the demonstrable contribution that excellent research makes to society and the economy; it embraces all the extremely diverse ways in which research-related knowledge and skills benefit individual, organisations and nations by:

Fostering global economic performance, and specifically the economic competitiveness of the United Kingdom;Increasing the effectiveness of public services and policy; and enhancing quality of life, health and creative output.

RCUK's impact strategy was launched in March 2010 and its webpage provides links to the impact and knowledge exchange activities of all the research councils. These are extremely informative sources on impact work in, and guidance for, their different domains.