

RCUK Responses to questions raised at the DCC/RDMF Special Event on funding for Research Data Management held at Aston University on 25th April 2013.

(NB only the questions in **bold** font below were discussed at the meeting)

Preamble:

RCUK accept that proper research data management has significant resource implications, and that research organisations are understandably concerned about the costs of implementing effective data management processes and systems. However, provision of support for the management of publicly funded research data is integral to the research environment that research organisations are responsible for maintaining and making available to their researchers. A variety of streams of public funding are available to research organisations to contribute to the costs of such provision, including QR from the funding councils and direct/indirect costs on research grants from the research councils (subject to the normal rules governing expenditure of grant funds). In addition, a number of subject-specific data repositories are directly supported by one or more research councils and researchers are expected to use these where appropriate.

A. What RDM services and infrastructure are in scope?

- 1) Is it appropriate for institutions to develop research data management services and e-infrastructure using research funding streams?*

It is certainly appropriate – see the preamble above. Research data management services and e-infrastructure are an essential component of the modern research environment which institutions are responsible for making available to their researchers. RCUK also recognise that it may be appropriate for some institutions to outsource aspects of research data management. Clearly, the balance between what RDM services/infrastructure to provide internally and what to source externally is a matter for each institution to decide according to its specific needs/circumstances.

- 2) Can research funding be used to develop and sustain long term institutional infrastructure or is it restricted to the support of individual funded projects?***

Research grant funding can be used subject to normal rules on grant expenditure, i.e. Direct Costs must be fully justified in the case for support and must be incurred before the actual end date of the grant; internal research data management services may only be charged as a direct cost to a project if they reflect fully auditable charges made in respect of data produced by the project; research grant funding under the heading Indirect Costs may be used by the institution to support centrally provided services that are necessary to maintain the overall research environment. However, at an institutional level, no charging model should allow grants to be charged (whether through directly incurred or directly allocated costs) for services where the cost of providing those services is recovered through the institution's indirect cost rate.

- 3) Can research grant monies be used to fund human infrastructure (i.e. institution support staff) or is it confined to expenditure on equipment and facilities?***

Research grant funding can be used to fund 'support staff'. If they are applied for as 'Direct Costs' on a research grant they must be fully justified in the case for support.

- 4) *Large capital investments for infrastructure are welcomed and are critical for maintaining leading edge research, however such investments risk not being fully realised due to lack of clarity between permitted costs for tangible assets like hardware and intangible assets like professional services to commission and maintain the hardware to end-of-life decommissioning. Are there efforts across councils to make clear the distinction between permitted costs and non-permitted costs?*

It is legitimate to use public funding streams to make provision for the sustainable management of publicly funded research data. The fEC funding model is based on the principle that Research Organisations identify the full costs to them of undertaking proposed research: the Research Councils do not maintain lists of 'permitted'/'non-permitted' expenditure that is chargeable as direct costs. Subject to the use of an auditable charging model, which may include charges to cover storage for a defined or an indefinite period, it is therefore permissible for an institutional RDM service to seek to recover operational costs by:

- i) Defining in grant applications what RDM-related services will be required, justifying their contribution to the delivery of the research and demonstrating their value for money; and
- ii) Applying for 'Directly Allocated Costs' or 'Directly Incurred Costs' to cover such services provided they are delivered before the actual end date of the grant

An institution may not directly charge a grant (whether through directly incurred or directly allocated costs) for services where the cost of providing those services is recovered through the institution's indirect cost rate. Where data arising from a research grant is destined for a subject-specific data repository directly supported by one or more research councils, the allowable direct expenditure on the research grant would be limited to cost of preparing the data for deposit and ingestion by that repository.

B. Open Access and data sharing

1. ***RCUK's recent guidance on open access indicates it has no plans to extend funding to cover Article Processing Costs for 'additional types of research output' such as data. Does RCUK or individual funders view favourably the inclusion in grant bids of any anticipated costs of depositing in non-institutional repositories such as Dryad?***

Anticipated costs of depositing research data arising from a project – whether to an internal or an external repository - may be included as direct costs in a grant application. Adequate justification should be provided within the grant application and the expenditure must take place before the actual end date of the project and be fully auditable.

2. ***Given that traditional publication costs are considered valid costs for research whether direct, indirect, green or gold OA, what are Research Council intentions for supporting data***

archiving and preservation? Particularly regarding matched or other subsidy type funding between the Institute and the particular or RCUK Councils?

See the preamble; management of available public funding streams in such a way that allocates appropriate budgets to support archiving and preservation of publicly funded research data is the responsibility of research organisations. Where external subject-specific data repositories are directly supported by Research Councils these should be used as directed. The Research Councils do not have access to additional funding (analogous to the RCUK block grant made available for Open Access journal publications) to support data archiving and preservation.

- 3. If a project intends to use RDM infrastructure that is funded or owned by an industrial partner, will that exempt them from any policy provisions on sharing data stored in that infrastructure?*

No; the policy provisions on sharing research council funded research data will apply regardless of the physical location or legal ownership of infrastructure used to store that data.

C. Specific and allowable RDM cost elements

- 1. Which elements of research data management are allowable costs that may be included in grant proposals?***

As long as adequate justification is provided within the grant application any element of research data management may be included as a direct cost; if the grant is awarded, all expenditure on direct costs must take place before the actual end date of the project and must be fully auditable. The costs of services/infrastructure recovered through an institution's indirect cost rate may not also be applied for as directly incurred or directly allocated costs on a grant. Where data arising from a research grant is destined for a subject-specific data repository directly supported by one or more research councils, the allowable expenditure on the research grant would be limited to cost of preparing the data for deposit and ingestion by that repository.

- 2. Do any funders endorse specific categories of grant expenditure beyond the project period? For example data storage/curation for x years after the project, on-going digital preservation after the project, the costs of data publication, etc.?***

With regard to longer term storage the research councils are open to a range of options – we wish to avoid being too prescriptive as we recognise that business models that sustainably cover the costs of such storage are still evolving. For example, some models may involve regular (e.g. annual) renewal and include elements of fixed and variable costs related to data volume, access frequency etc., while others (e.g. 'pay-once-store-forever' (POSF) models) may offer much longer term storage and backup in return for a single prior payment dependant only the data volume.

Direct costs must be incurred before the actual end date of the project and must be fully auditable. It would be appropriate to charge costs associated with data sharing/preservation

(e.g. preparation and ingestion of research data to a data repository, including on the basis of POSF) provided that such activity is undertaken before the actual end date of the project.

However, when considering alternative data storage options, research organisations are reminded of the need to undertake due diligence with respect to the long term security, accessibility and integrity of the data. Note also that:

- i) costs of publication in peer reviewed academic journals and conference proceedings are not eligible due to the introduction of in April 2013 of RCUK block grants to support Open Access;
- ii) costs of long term storage should not be included in respect of any data that is expected to be deposited in a Research Council funded data repository;
- iii) where the costs of services/infrastructure are recovered through an institution's indirect cost rate, the same services/infrastructure may not be applied for as directly incurred or directly allocated costs on a grant.

3. *Is there any proportion of funding awarded (e.g. 10% of a grant) that – as a rule of thumb – is acceptable to be spent to fund RDM services?*

The value of a research grant is not generally a good proxy for the volume or complexity of the data it may generate, and therefore a flat rate charging structure for RDM services is not considered appropriate. Furthermore, research data management services are an essential component of the modern research environment which institutions are responsible for making available to their researchers. The fEC funding model is based on the principle that Research Organisations identify the full costs of undertaking proposed research projects (including provision of a high quality research environment), and research councils expect grant applications to state the full costs either direct or indirect costs

4. *If a business model to support research data management activities includes a one- off charge for deposit into an institutional data repository, will research councils accept these costs as a) incurred during a research project and b) additional to institutional provision?*

Business models to support institutional research data management activity may include one- off charges for deposit and recurrent costs. One-off charges may be based on a scale which in turn may refer to 'fixed' and 'variable' charges that depend on e.g. the volume of data being managed. Researchers are encouraged to ensure that appropriate provision is made in grant applications to cover the research data management costs that can be predicted to arise during the proposed research. Such provision should be appropriately justified, and is also subject to the guiding principle that costs of services/infrastructure which are recovered through an institution's indirect cost rate may not also be applied for as directly incurred or directly allocated costs on a grant.

Projects where the data produced is expected to be deposited in a Research Council funded data repository should not include charges for deposit in an institutional repository.

5. *We will need to develop charging models to cover the cost of RDM throughout the lifecycle of the data. Which RDM-related activities (e.g. cost of storage and backup, hardware replacement, software replacement, maintenance, data ingest into a repository, staff time etc.) are eligible for inclusion in such a cost model?*

All elements of research data management specific to a research project are allowable. It is for institutions to decide which elements to meet through provision of services supported indirectly (e.g. by using QR or Indirect Costs), and which to recover as direct costs on research grants. The guiding principle is that costs of services/infrastructure recovered through an institution's indirect cost rate may not also be applied for as directly incurred or directly allocated costs on a grant.

D. Long term storage, preservation and archiving

1. ***Do funders have any thoughts on how the costs of long term storage requirements should be met? Would they envisage funding these as part of the on-going overhead for each HEI?***

See responses to C2 above and D4 below

2. ***Where needed, can the cost of long-term storage (25 years +) be retrieved from the research grant?***

See responses to C2 above and D4 below

3. ***Where institutions do not treat RDM as a Research Facility, are funders happy to pay for storage costs as Directly Incurred charges on a grant? i) where there are exceptional storage needs on a project ii) for general storage needs on a project***

The distinction between 'general' and 'exceptional' is subjective. The guiding principle is that costs of services/infrastructure recovered through an institution's indirect cost rate may not also be applied for as directly incurred or directly allocated costs on a grant. Therefore, if an institution chooses to recover the costs of meeting normal data storage needs through the indirect cost rate then these may not also be applied for as direct costs on a grant. However, if a particular project will have exceptional data storage needs that significantly exceed the norm then the costs should be applied for and justified. Directly incurred costs on a grant need to be incurred *before the actual end date of the project*.

4. ***What is the funder view on POSF (pay once, store forever) price models?***

With regard to longer term storage the research councils are open to a range of options – we wish to avoid being too prescriptive as we recognise that business models that sustainably cover the costs of long term data storage are still evolving. 'Pay once, store forever' (POSF) price models have obvious attractions, however research organisations are reminded of the need to undertake due diligence with respect to the long term security, accessibility and integrity of the data. It is permissible to apply for POSF costs and to charge them to a grant provided they are incurred before the actual end date of the project.

5. *Many funders (e.g. AHRC section 4 of the 2013 Research Funding Guide, section 2.b of the Technical Plan) recognise that they will cover costs for resources *additional to the*

institutional provision. If a University currently provides a quota of research storage free to funded research projects, but finds this model to be unsustainable and therefore starts charging all researchers for data storage and archive, which funding councils will accept these costs as “additional to institutional provision”?*

See response to question C1 above. The research councils expect institutions to manage provision of research data management in a sustainable way and accept that this may include business models that charge projects for specifically incurred costs. However, whatever charging model is adopted should be applied consistently within an institution. The guiding principle is that costs of services/infrastructure recovered through an institution's indirect cost rate may not also be applied for as directly incurred or directly allocated costs on a grant. Therefore, if an institution chooses to recover the costs of meeting normal data storage needs through the indirect cost rate then these may not also be applied for as direct costs on a grant. However, if a particular project will have exceptional data storage needs that significantly exceed the norm then the costs should be applied for and justified. Services/infrastructure included in an institution's indirect cost rate should be equally available to RCUK-funded researchers; it is not acceptable to charge RCUK grants for access to facilities or services that are freely available to researchers not funded by RCUK.

- 6. Use of institutional storage provision using centrally managed backed up systems is more expensive than say, storage of data on external hard drives. How do research councils consider this fact in relation to the value for money argument that each application for funding needs to make?*

RCUK recognizes that research data management throughout the data life cycle amounts to more than simple storage. Value for money is not simply a matter of lowest cost; data must be appropriately stored and curated in a cost effective manner.. thus there are likely to be considerable benefits to the use of centrally managed and backed up systems with regards to delivery of appropriate assurance versus the use of 'local' backup, e.g. individual hard drives.

- 7. Several institutions are investigating data archive-as-a-service options for the long term storage of research data. Service offerings are available which provide near-line access to the data in JANET-connected data centres and insurance backed guarantees protecting against data loss, as well as ISO27001 data security certification. Do funding bodies view long term data archive as an allowable cost, on the basis that institutions fund the repository service, local disc storage and costs associated with data preservation and curation activities? (Indicative costs are £3.75 per Gb for 20 years storage).*

Long term data storage/archive is allowable subject to the guiding principle that costs of services/infrastructure recovered through an institution's indirect cost rate may not also be applied for as directly incurred or directly allocated costs on a grant.

E. What is the application process to recover RDM costs?

- 1. If an institution provides research data management support that is tailored to the needs of an individual grant holder, can the actual FEC for providing it, be charged to a grant?*

This question is dealt with by responses to several other questions, in particular C1 and D5.

- 2. How should allowable research data management service costs be included in grant proposals? Directly incurred? Indirects? Blanket % addition? Other?**

[not responded to here as this question is covered by so many of the other questions]

- 3. Institutions may consider a research data management service as a Small Research Facility and seek to charge use of the facility to a grant. Do funding bodies view this as an acceptable way for institutions to recover the capital costs of developing research data management services and thus avoid a proportionate increase in their indirect costs?**

Provided such facilities are transparently charged to all projects that use them there is no reason that they may not be charged to research council grants subject to the normal rule that directly incurred grant expenditure must be incurred during the project.

- 4. How many/what percentage of ESRC grant application peer reviewers are ticking 'Unable to assess' when asked to review a data management plan? Have other funders got similar data?**

This information is not available at this time.

F. Guidance and Support

- 1. Are funders developing a collective vision for research data management, to help institutions develop consistent policies, procedures and benchmarks?**

The Research Councils have already agreed the [RCUK Common Principles on Data Policy](#). We are working to consider the scope for aligning the respective policies of individual councils and issuing common guidance on research data management.

- 2. Is a business model being developed to help us plan future data storage requirements, e.g. are there any plans to develop a subscription-based cloud storage service, or regional or discipline-specific data repositories?**

The research councils already provide direct support to a number of subject-specific data repositories. The research councils are not engaged in the development of regional or cloud-based storage services.

- 3. Are funders developing guidance on what proportion of research grant funding may be used to support institutional RDM infrastructure costs?**

See response to C3 above

- 4. Are there examples of grant submissions that have been rejected because of the inclusion of research data management costs?**

Such examples are not currently available. However, we stress that while scientific/research excellence is the primary criterion the requested resources are also assessed as part of the peer review process: if they are considered either inadequate or unreasonable this may influence the outcome of the application. Researchers/research organisations should

therefore justify all requested resources included in a grant application by explaining their contribution to the delivery of the research and demonstrating their value for money.

5. *Can each funder provide examples to illustrate the successful inclusion of research data management costs in approved proposals?*

The research councils will consider this as part of the work referred to in response to F1 above. It may be possible to provide indicative examples without breaching the confidentiality of actual proposals.

6. *Many funding councils have recently updated their data policies. When this is done, please can funding councils archive all previous versions of their data policies on their websites, so that researchers whose projects started several years ago, can refer back to the policy that was in place when their funding was granted.*

Archiving previous versions of policy online risks causing confusion (e.g. people accessing and inadvertently relying on out of date material). The management of research data should at all times reflect best practice: the start date of a grant should not affect the level of care and attention given to data that is currently being generated or actively managed by research institutions.

G. Compliance

1. ***The EPSRC stated that they "may request evidence of activity to achieve compliance at any time" and if lacking, may "impose appropriate sanctions including, in an extreme case, the removal of eligibility for EPSRC funding". This has been used by many institutions to justify allocation of central resources to data management activities. How will the EPSRC demonstrate to the community that this strong motivator is real?***

The statements quoted above relate to EPSRC's policy framework on research data (see <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/policyframework.aspx>), which was published in May 2011.

EPSRC has already received confirmation from the great majority of organisations it funds that activity is indeed underway with the expressed intention of meeting EPSRC's expectations on research data management by May 2015.

The quote about sanctions, including possible loss of eligibility for EPSRC funding, refers to the a worst case scenario in which an research organisation is demonstrably and unjustifiably obstructing the sharing of publicly funded research data, despite

- i) the Freedom of Information Act(s),
- ii) the Environmental Information Regulations, and
- iii) having been given four year's notice of EPSRC's expectations that they will by May 2015 have robust policies and processes in place to ensure that they are aware of their research data holdings, at least make the existence of such data discoverable, and make it available for sharing unless there are valid reasons to withhold it. Valid reasons may include legal, ethical and commercial constraints, as noted in the RCUK

Common Principles on Data Policy (see
<http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx>)

In the circumstances it will be understood that EPSRC does not intend to review compliance at each institution it funds, but the council will certainly react if it is brought to our attention that an organisation in receipt of EPSRC funding is not meeting the published EPSRC expectations.