

Goldsmiths', University of London  
Roadmap for ESPRC's Research Data Management  
Expectations

**Live document: Version 1.1 published 07.08.2015**

**Acknowledgements:**

We would like to acknowledge that the University of Exeter's Roadmap for EPSRC's Research Data Management Expectations was influential in the development of this document

# Contents

Context

Appendix A - RCUK Common Principles on Data Policy

Appendix B - EPSRC Data Principles

Appendix C - EPSRC Expectations for Research Data Management

Appendix D - Clarifications of EPSRC expectations on research data management

Appendix E – EPSRC Notification Letter February 2012

## Context

### EPSRC Requirements

- A. As part of their open access mandate the EPSRC (Engineering and Physical Sciences Research Council) require that research data generated on research projects funded by them is adequately and ethically managed (Appendices B and C)
- B. A letter of notification was sent to VCs **in February 2012** to announce these new EPSRC research data management requirements (Appendix E)
- C. The EPSRC, like other RCUK councils, define 'data' broadly and are not simply referring to scientific data-sets but also other outputs such as: non-digital outputs; programming code; digital surveying materials and similar resulting from research projects they have supported.
- D. The February 2012 notification letter signalled that the EPSRC would reserve the right to audit research data management practices, policies and procedures. Where this is deemed necessary, **the EPSRC reserve the right to introduce sanctions** if compliance is deliberately obstructed.
- E. Whilst the EPSRC are the only research council that have explicitly set out compliance guidance for research data management, in effect all RCUK funders and many other funders like the EU and Wellcome Trust expect HEIs to operate on an identical basis with regards to managing research data and open access outputs.
- F. The EPSRC have set out a list of expectations and consequently a clarification of said expectations to help institutions comply with expectations (Appendices B, C and D)

### Audit Requirements:

- A. The EPSRC expected institutions to self-certify the level to which they were in compliance with the EPSRC requirements for research data management strategies. No formal document had to be submitted to the EPSRC directly as part of this self-certification process.
- B. The EPSRC requested that by **May 2012 institutions produce a road map** to research data management where they were not yet fully able to comply with research data management requirements.
- C. **By 1 May 2015** the EPSRC are expecting institutions to be fully compliant with EPSRC research data management expectations.
- D. The EPSRC are working with **RCUK's Audit & Assurance Group** <http://www.rcuk.ac.uk/about/aboutrcuk/aims/units/aasg/> to formally evaluate institutional compliance as part of regular auditing practices.

- E. The EPSRC will be sending a *self-assessment questionnaire* to pro-VCs in May 2015 for them to assess current activities
- F. The EPSRC are *planning to analyse EPSRC-funded articles published after May 2015* to determine if accompanying data is available. An exact time frame for this audit is not available.

**Overview Current Status Goldsmiths', University of London:**

- A. Goldsmiths, University of London holds several EPSRC research awards and operates an EPSRC doctoral training centre. Goldsmiths must therefore ensure it complies with EPSRC research data management and open access expectations.
- B. The overall implementation of the open access and research data management agendas is led by the Open Access Working Group (OAWG), reporting to the Goldsmiths', Research and Enterprise Committee (REC). REC is chaired by the Pro-Warden for Research and Enterprise
- C. Core members to the OAWG include representatives from the Library, Research Office and a Senior Member of Academic Staff.
- D. Advisory members to the OAWG include representatives from IT services; Research Ethics and Integrity Committee; Governance and Legal Services; Research Finance
- E. OAWG are responsible for overseeing the drafting and ratification of institutional policy and governance for research data management and open access in consultation with other Goldsmiths' staff.
- F. Goldsmiths has a ratified Roadmap for Research Data Management and is working towards its implementation on an on-going basis. The Roadmap for Research Data Management is a live document.
- G. Goldsmiths has a ratified Open Access statement  
<http://research.gold.ac.uk/openaccess.html>

**The measures currently being taken or planned, to ensure compliance with EPSRC and other funders' policies include:**

- A. A review of the institutional research data management policy and associated procedures in collaboration with Governance and Legal experts
- B. Associated with the above, clarification of organisational and individual responsibilities and obligations.
- C. Provision of a robust, extensible, actively-curated research data repository (<http://data.gold.ac.uk/>) in collaboration with University of London Computing Centre (ULCC, <http://ulcc.ac.uk/>) and Arkivum (<http://arkivum.com/>) available to all Goldsmiths' researchers for the preservation and dissemination of their research data.

- D. Quality control mechanism for reviewing and approving submissions to the repository.
- E. Option to submit metadata-only data records where placing data on Open Access is inappropriate (where the Data Protection Act applies, or for reasons of commercial sensitivity, for example).
- F. Upskilling those currently providing advice to researchers to ensure research data management knowledge and skills are widely dispersed (for example, Research and Knowledge Transfer staff, Subject Librarians, those responsible for overseeing Research Ethics and Integrity clearance).
- G. Ongoing programme of research data management training delivered via a number of options, aimed at increasing the knowledge and skills of researchers and PGRs.
- H. Development of a range of discipline-specific tools and resources for embedding good practice in management of research data and its associated metadata and documentation in the research lifecycle.

#### **Roadmap Document Status:**

- A. This is a living document, responsive to RCUK and EPSRC guidelines and official policy and governance communication from these bodies.
- B. Responsibility for updating this document resides with the Open Access Working Group (OAWG). The Roadmap's purpose is to provide information and guidance to Goldsmiths' research engaged staff (including administrative staff) and EPSRC stakeholders about how to manage research data.
- C. Its purpose is also to address EPSRC compliance and audit requirements and to provide evidence of this compliance by Goldsmiths', University of London. It is designed to guide and inform Goldsmiths staff and external auditors of actions undertaken to date to ensure compliance in support of research engaged staff.

#### **Timeframe:**

**Phase 0: 1<sup>st</sup> May 2012** – Ratify Research Data Management Policy (via <http://www.gold.ac.uk/research/ethics/>); Professional Services Re-Structure and Recruitment Library leads; change Senior Management Team; REF submission; implementation open access mandates.

**Phase 1: 1<sup>st</sup> May 2012 – 1<sup>st</sup> May 2015** – Continued Professional Services Re-Structure and Recruitment Library leads; change Senior Management Team; REF submission; implementation open access mandates; begin implementation research data management strategy with new Professional services staff.

**Phase 2: 1 May 2015 onwards** – Maturation of services and continued roll-out. Review services through regular self-assessment and reporting to Research and Enterprise Committee. Continue to respond to new developments in research data management and open access policies and practice.

**Document Authorship:**

This EPSRC Roadmap was authored by: Dr Muriel E. Swijghuisen Reigersberg; Ms Caroline Lloyd; Mr Andrew Gray; Prof Andrea Phillips of the Goldsmiths', Open Access Working Group.

# EPSRC EXPECTATIONS

## EXPECTATION 1.

Research organisations will promote internal awareness of these principles and expectations and ensure that their researchers and research students have a general awareness of the regulatory environment and of the available exemptions which may be used, should the need arise, to justify the withholding of research data.

### Proposed Actions:

Objectives	Actions	Responsible Teams	Target Date
1.1 Develop the data management skills and knowledge of researchers and relevant professional services staff	<p>Include research data management training in following areas:</p> <p>Academic Practice Course Graduate School</p> <p>Research Ethics and Integrity Training</p> <p>Research admin training</p> <p>Researcher Inductions/ grant writing skills</p> <p>Library</p> <p>IT/IS services</p> <p>Legal and Governance</p>	OAWG Research Office Library Data Protection Officer	Phase 1 and 2
1.2 Develop Research Data management guidance and website area for staff	Develop guidance and keep updated	Research Office Library IT/IS	Phase 1 and 2
1.3 Develop and implement communication plan	Develop and implement Communications Plan	Research Office Library OAWG Internal & External Communications	Phase 0 and 1
1.4 Provide clear guidance on the regulatory environment and exemptions which may be used to justify the	Integrate training on Data Management Plans; Research Integrity; FOI requests; IP, Copyright, sensitive data	Research Ethics and Integrity Committee Research Office IT/IS	Phase 2

withholding of research data			
---------------------------------	--	--	--

## EXPECTATION 2.

Published research papers should include a short statement describing how and on what terms any supporting research data may be accessed.

### Proposed Actions:

<b>Objectives</b>	<b>Actions</b>	<b>Responsible Teams</b>	<b>Target Date</b>
2.1 Advocate the benefits of increasing data discoverability	Training and guidance on the need to include statement on accessibility of data and development of model data access statement  Include requirement for this statement in policy	Library Research Office	Phase 1 and 2
2.2 Enable the linking of data and publications in the Goldsmiths' repository	Ensure that the data and publications can be linked in the Goldsmiths' repository	Library IT/IS	Phase 2

### EXPECTATION 3.

Each research organisation will have specific policies and associated processes to maintain effective internal awareness of their publicly-funded research data holdings and of requests by third parties to access such data; all of their researchers or research students funded by EPSRC will be required to comply with research organisation policies in this area or, in exceptional circumstances, to provide justification of why this is not possible.

#### Proposed Actions:

Objectives	Actions	Responsible Teams	Target Date
3.1 Develop policies on research data management	Institutional Policies on RDM developed, ratified and updated regularly	IT/IS Research Office Library	Phase 0
3.2 Track third party requests for access to data	Develop inventory of research data held at Goldsmiths.  Develop procedures for tracking 3 <sup>rd</sup> party requests for access to data	Library IT/IS Research Office	Phase 2
3.3 Raise awareness of data management and archiving issues in early stages of drawing up contracts with research partners	Review and specify RDM requirements in new contracts with research partners to ensure that appropriate legal clauses are in place and maximise accessibility and impact of data outputs where appropriate	Research Office Enterprise Office IT/IS	Phase 1 and 2

## EXPECTATION 4.

Publicly-funded research data that is not generated in digital format will be stored in a manner to facilitate it being shared in the event of a valid request for access to the data being received (this expectation could be satisfied by implementing a policy to convert and store such data in digital format in a timely manner).

### Proposed Actions:

Objectives	Actions	Responsible Teams	Target Date
4.1 Investigation of strategies for improving access and visibility of non-digital data.	<p>Create inventory of non-digital data</p> <p>Investigate strategies to enhance storage of non-digital data, to include cost estimates</p> <p>Investigate strategies for potential digitisation of non-digital research data to include cost estimates</p>	Library IT/IS	Phase 2
4.2 Develop policies on non-digital data	<p>Include non-digital data in Goldsmiths' RDM policy and clarify that it should be registered with the Goldsmiths' repository and how it can be accessed</p> <p>Develop specific policy on the storage, digitisation and disposal of non-digital research data</p>	IT/IS Library Research Office	Phase 2

## EXPECTATION 5.

Research organisations will ensure that appropriately structured metadata describing the research data they hold is published (normally within 12 months of the data being generated) and made freely accessible on the internet; in each case the metadata must be sufficient to allow others to understand what research data exists, why, when and how it was generated, and how to access it. Where the research data referred to in the metadata is a digital object it is expected that the metadata will include use of a robust digital object identifier (For example as available through the DataCite organisation - <http://datacite.org>).

### Proposed Actions:

Objectives	Actions	Responsible Teams	Target Date
5.1 Seek to increase impact of Goldsmiths' research data through increasing discoverability of data using metadata	Training and guidance on how to document research data as well as how to complete the metadata record when uploading data sets to the institutional repository to enhance discovery	Library IT/IS Research Office	Phase 2
5.2 Include requirement for metadata record in policy even when data is not available on open access	Requirement for metadata record included in policy	Library Research Office	Phase 2
5.3 Data repository to assign robust digital object identifiers to digital objects	Ensure that repository can assign digital object identifiers to digital objects	Library IT/IS	Phase 2

## EXPECTATION 6.

Where access to the data is restricted the published metadata should also give the reason and summarise the conditions which must be satisfied for access to be granted. For example 'commercially confidential' data, in which a business organisation has a legitimate interest, might be made available to others subject to a suitable legally enforceable non-disclosure agreement.

Proposed Actions:

<b>Objectives</b>	<b>Actions</b>	<b>Responsible Teams</b>	<b>Target Date</b>
6.1 Ensure that the conditions for restricted access to data are clear	Develop guidelines for the metadata records of restricted access data e.g. use of embargo periods, commercialisation and research integrity reasons	Research Ethics and Integrity Committee Research Office IT/IS Library Enterprise	Phase 2

## EXPECTATION 7.

Research organisations will ensure that EPSRC-funded research data is securely preserved for a minimum of 10-years from the date that any researcher ‘privileged access’ period expires or, if others have accessed the data, from last date on which access to the data was requested by a third party; all reasonable steps will be taken to ensure that publicly-funded data is not held in any jurisdiction where the available legal safeguards provide lower levels of protection than are available in the UK.

### Proposed Actions:

Objectives	Actions	Responsible Teams	Target Date
7.1 Ensure that research data is securely preserved according to EPSRC expectations	<p>Update Goldsmiths’ Code of Good Research Conduct according to RDM policy</p> <p>Develop standards and training on data retention and disposal</p> <p>Track third party access to and preservation of EPSRC research data</p> <p>Ensure that data in the Goldsmiths’ repository (ULCC/ Arkivum) will be stored and backed-up in the UK</p>	<p>IT/IS Library Research Ethics and Integrity Committee</p>	Phase 2

## EXPECTATION 8.

Research organisations will ensure that effective data curation is provided throughout the full data lifecycle, with ‘data curation’ and ‘data lifecycle’ being as defined by the Digital Curation Centre. The full range of responsibilities associated with data curation over the data lifecycle will be clearly allocated within the research organisation, and where research data is subject to restricted access the research organisation will implement and manage appropriate security controls; research organisations will particularly ensure that the quality assurance of their data curation processes is a specifically assigned responsibility.

### Proposed Actions:

Objectives	Actions	Responsible Teams	Target Date
8.1 Clarify research data management responsibilities	Define cross-institution responsibilities for research data management, including responsibility for quality assurance of the data curation process	Library Research Office OAWG	Phase 0, 1 and 2
8.2 Sensitive data managed according to standards of good practice	Training and guidance on the management of sensitive data for researchers and professional services staff	Research Ethics and Integrity Committee Research Office Graduate School Library IT/IS	Phase 0, 1 and 2

## EXPECTATION 9.

Research organisations will ensure adequate resources are provided to support the curation of publicly-funded research data; these resources will be allocated from within their existing public funding streams, whether received from Research Councils as direct or indirect support for specific projects or from higher education Funding Councils as block grants.

### Proposed Actions:

Objectives	Actions	Responsible Teams	Target Date
9.1 Include research data management infrastructure requirements in the institutional planning lifecycle	Ensure that research data management is included in high-level institutional budget planning	OAWG IT/IS Research Office Library	Phase 1 and 2
9.2 Development of business plan for sustainable research data management service	IT/IS, Library and Research and Enterprise strategies to include mechanisms for creating a sustainable research data management service	OAWG	Phase 2
9.3 Ensure sustainable funding resources for research data management actions	Identify and apply for further funding opportunities	Library Research Office	Phase 0,1 and 2
9.4 Encourage researchers to include costs of research data management in data management plans when possible	Raise awareness about including costs in data management plans in data management plan training	Research Office Research Finance Enterprise Research Admin	Phase 0, 1 and 2

# APPENDICES

## Appendix A - RCUK Common Principles on Data Policy

<http://www.rcuk.ac.uk/research/datapolicy/> (accessed 22.04.2015)

Making research data available to users is a core part of the Research Councils' remit and is undertaken in a variety of ways. RCUK are committed to transparency and to a coherent approach across the research base. These RCUK common principles on data policy provide an overarching framework for individual Research Council policies on data policy.

### Principles

- Publicly funded research data are a public good, produced in the public interest, which should be made openly available with as few restrictions as possible in a timely and responsible manner that does not harm intellectual property.
- Institutional and project specific data management policies and plans should be in accordance with relevant standards and community best practice. Data with acknowledged long-term value should be preserved and remain accessible and usable for future research.
- To enable research data to be discoverable and effectively re-used by others, sufficient metadata should be recorded and made openly available to enable other researchers to understand the research and re-use potential of the data. Published results should always include information on how to access the supporting data.
- RCUK recognises that there are legal, ethical and commercial constraints on release of research data. To ensure that the research process is not damaged by inappropriate release of data, research organisation policies and practices should ensure that these are considered at all stages in the research process.
- To ensure that research teams get appropriate recognition for the effort involved in collecting and analysing data, those who undertake Research Council funded work may be entitled to a limited period of privileged use of the data they have collected to enable them to publish the results of their research. The length of this period varies by research discipline and, where appropriate, is discussed further in the published policies of individual Research Councils.
- In order to recognise the intellectual contributions of researchers who generate, preserve and share key research datasets, all users of research data should acknowledge the sources of their data and abide by the terms and conditions under which they are accessed.
- It is appropriate to use public funds to support the management and sharing of publicly-funded research data. To maximise the research benefit which can be gained from limited budgets, the mechanisms for these activities should be both efficient and cost-effective in the use of public funds.

## Appendix B – EPSRC Data Principles

<https://www.epsrc.ac.uk/about/standards/researchdata/principles/> (accessed 22.04.2015)

EPSRC's Charter is to support high-quality basic, strategic and applied research and related post-graduate training in engineering and the physical sciences, and to communicate and disseminate the resulting outcomes and knowledge. As such, EPSRC believes that the following guiding principles, which are aligned with the agreed RCUK principles on sharing of research data, should inform all decisions relating to the management of all research data that has arisen as a result of EPSRC funding:

- i. EPSRC-funded research data is a public good produced in the public interest and should be made freely and openly available with as few restrictions as possible in a timely and responsible manner.
- ii. EPSRC recognises that there are legal, ethical and commercial constraints on release of research data. To ensure that the research process (including the collaborative research process) is not damaged by inappropriate release of data, research organisation policies and practices should ensure that these constraints are considered at all stages in the research process.
- iii. Sharing research data is an important contributor to the impact of publicly funded research. To recognise the intellectual contributions of researchers who generate, preserve and share key research datasets, all users of research data should acknowledge the sources of their data and abide by the terms and conditions under which they are accessed.
- iv. EPSRC-funded researchers should be entitled to a limited period of privileged access to the data they collect to allow them to work on and publish their results. The length of this period will depend on the scientific discipline and the nature of the research.
- v. Institutional and project specific data management policies and plans should be in accordance with relevant standards and community best practice and should exist for all data. Data with acknowledged long term value should be preserved and remain accessible and useable for future research.
- vi. Sufficient metadata should be recorded and made openly available to enable other researchers to understand the potential for further research and re-use of the data. Published results should always include information on how to access the supporting data.
- vii. It is appropriate to use public funds to support the preservation and management of publicly-funded research data. To maximise the scientific benefit which can be gained from limited budgets, the mechanisms for managing and providing access to research data should be both efficient and cost-effective in the use of such funds.

## Appendix C – EPSRC Expectations for Research Data Management

<https://www.epsrc.ac.uk/about/standards/researchdata/expectations/> (accessed 22.04.2015)

EPSRC has the following clear expectations of organisations in receipt of EPSRC research funding:

- i. Research organisations will promote internal awareness of these principles and expectations and ensure that their researchers and research students have a general awareness of the regulatory environment and of the available exemptions which may be used, should the need arise, to justify the withholding of research data.
- ii. Published research papers should include a short statement describing how and on what terms any supporting research data may be accessed.
- iii. Each research organisation will have specific policies and associated processes to maintain effective internal awareness of their publicly-funded research data holdings and of requests by third parties to access such data; all of their researchers or research students funded by EPSRC will be required to comply with research organisation policies in this area or, in exceptional circumstances, to provide justification of why this is not possible.
- iv. Publicly-funded research data that is not generated in digital format will be stored in a manner to facilitate it being shared in the event of a valid request for access to the data being received (this expectation could be satisfied by implementing a policy to convert and store such data in digital format in a timely manner).
- v. Research organisations will ensure that appropriately structured metadata describing the research data they hold is published (normally within 12 months of the data being generated) and made freely accessible on the internet; in each case the metadata must be sufficient to allow others to understand what research data exists, why, when and how it was generated, and how to access it. Where the research data referred to in the metadata is a digital object it is expected that the metadata will include use of a robust digital object identifier (For example as available through the DataCite organisation).
- vi. Where access to the data is restricted the published metadata should also give the reason and summarise the conditions which must be satisfied for access to be granted. For example 'commercially confidential' data, in which a business organisation has a legitimate interest, might be made available to others subject to a suitable legally enforceable non-disclosure agreement.
- vii. Research organisations will ensure that EPSRC-funded research data is securely preserved for a minimum of 10 years from the date that any researcher 'privileged access' period expires or, if others have accessed the data, from last date on which access to the data was requested by a third party; all reasonable steps will be taken to ensure that publicly-funded data is not held in any jurisdiction where the available legal safeguards provide lower levels of protection than are available in the UK.
- viii. Research organisations will ensure that effective data curation is provided throughout the full data lifecycle, with 'data curation' and 'data lifecycle' being as defined by the Digital Curation Centre. The full range of responsibilities associated with data curation over the data lifecycle will be clearly allocated within the research organisation, and where research data is subject to restricted access the research organisation will implement and manage appropriate security controls; research organisations will particularly ensure that the quality assurance of their data curation processes is a specifically assigned responsibility.
- ix. Research organisations will ensure adequate resources are provided to support the curation of publicly-funded research data; these resources will be allocated from within their existing public funding streams, whether received from Research Councils as direct or indirect support for specific projects or from higher education funding councils as block grants.

## Appendix D

<https://www.epsrc.ac.uk/files/aboutus/standards/clarificationsofexpectationsresearchdatamanagement/> (accessed 22.04.2015)

OFFICIAL 9 October 2014

Page 1 of 7

### Clarifications of EPSRC expectations on research data management.

#### Expectation I

Research organisations will promote internal awareness of these principles and expectations and ensure that their researchers and research students have a general awareness of the regulatory environment and of the available exemptions which may be used, should the need arise, to justify the withholding of research data;

#### Clarification

This is expected to be an on-going part of the normal business of organisation-based training and awareness for staff and students in all research organisations funded by EPSRC. For example, research data management may be addressed as part of the normal induction process for new staff/research students, and awareness of the issues may be subsequently refreshed through regular online and/or facilitated training.

#### Expectation II

Published research papers should include a short statement describing how and on what terms any supporting research data may be accessed.

#### Clarification

This expectation is consistent with the RCUK policy on Open Access, and applies to all papers which acknowledge EPSRC funding with a publication date after 1<sup>st</sup> May 2015. It is also consistent with the Principle that EPSRC-funded research data should be made freely and openly available with as few restrictions as possible in a timely and responsible manner. We acknowledge that not all research papers are supported by research data, and will therefore rely on researchers making informed judgements about when it is appropriate to include such a statement.

The expectation is that data relied on in published research findings will, by default, be available for scrutiny by others. The expectation could be satisfied by citing such data in the published research and including in such citations direct links to the data or to supporting documentation that describes the data in detail, how it may be accessed and any constraints that may apply. It is important that any such links are *persistent* URLs such as DOIs.

If compelling legal or ethical reasons exist to protect access to the data these should be noted in the statement included in the published research paper. A simple direction to interested parties to 'contact the author' would not normally be considered sufficient.

It is not recommended to publish findings that cannot be validated by others. If published research findings rely on data from third parties every effort should be made to ensure that this is also available for scrutiny. It is accepted that ethical or legal constraints may apply – see Expectation VI.

### **Expectation III**

Each research organisation will have specific policies and associated processes to maintain effective internal awareness of their publicly-funded research data holdings and of requests by third parties to access such data; all of their researchers or research students funded by EPSRC will be required to comply with research organisation policies in this area or, in exceptional circumstances, to provide justification of why this is not possible.

#### **Clarification**

The expectation that requests to access research data will be logged is pragmatic and applies primarily to data held by the research organisation; it will help build evidence of impact and inform future decisions that may need to be taken about whether to keep specific datasets (data that has not attracted any interest for more than 10 years is not expected to be retained – see discussion of expectation VII below).

EPSRC does not expect research organisations to log access requests to research data originated by their researchers but deposited in an external repository (e.g. subject specific or journal/publisher specific); however, we believe research organisations will find such information a valuable indicator of impact.

### **Expectation IV**

Publicly-funded research data that is not generated in digital format will be stored in a manner to facilitate it being shared in the event of a valid request for access to the data being received (this expectation could be satisfied by implementing a policy to convert and store such data in digital format in a timely manner);

#### **Clarification**

We do not expect research organisations to digitise all research data, but we also do not expect requests for access to research data to be rejected on the basis that it is physically too awkward to share. See also Clarification under Expectation IX

## Expectation V

Research organisations will ensure that appropriately structured metadata describing the research data they hold is published (normally within 12 months of the data being generated) and made freely accessible on the internet; in each case the metadata must be sufficient to allow others to understand what research data exists, why, when and how it was generated, and how to access it.

Where the research data referred to in the metadata is a digital object it is expected that the metadata will include use of a robust digital object identifier (For example as available through the DataCite organisation - <http://datacite.org>).

## Clarification

The emphasis in this expectation is on publishing 'metadata for discovery' (rather than publishing the actual data itself) with a further expectation that this metadata will be published 'normally within 12 months of the data being generated'. However, it can be difficult to be precise about the exact 'date of generation', particularly for data recorded over a period of time or which requires significant processing before it can be interpreted. Given the uncertainties that can arise, the following guidance may be helpful and is intended to help research organisations to prioritise their activity and to establish effective data management as part of the normal research process:

1. If the data directly supports research findings published after 1<sup>st</sup> May 2015 the expectation will be met if the published findings contain appropriate data citations (see clarification under Expectation II). In such cases the cited data/supporting documentation is expected to be accessible online no later than the date of first online publication of the article. This should not be onerous: RCUK Policy on Open Access, which took effect from 1<sup>st</sup> April 2013, already states that published research papers must include 'if applicable, a statement on how the underlying research materials... can be accessed'.
2. In the case of other retained research data with a clear 'date of generation' being on/after 1<sup>st</sup> May 2015, and where there is no intention to publish findings which rely on that data, metadata for discovery should be exposed online within one year of the 'date of generation'.
3. If retained research data has no clear 'date of generation' and the relevant EPSRC grant ends on/after 1<sup>st</sup> May 2015, and if there is no intention to publish findings which rely on that data, the end date of the EPSRC grant may be used instead, i.e. the metadata for discovery should be exposed online no later than one year after the end date of the EPSRC grant.

Where an intention to publish is relied on to justify withholding research data attention is drawn to the law on Freedom of Information and the importance of ensuring a publication plan is in place. The expectation that published metadata 'must be sufficient to allow others to understand what research data exists, why, when and how it was generated, and how to access it' may be satisfied by a combination of information made directly accessible online and more detailed information contained in documentation that can be identified and freely downloaded. Over time, research organisations are expected to use their best endeavours to make older research data discoverable and accessible in line with our policy framework.

The expectation that research organisations will ensure ‘appropriately structured metadata’ is published does not refer to any specific standard. This is intentional: making research data normally ‘open’ is a relatively new undertaking for all concerned, and robust standards for metadata that facilitates discovery of research data have yet to emerge. EPSRC will continue to monitor developments but does not anticipate being able to mandate use of specific standards in this area for several years. In implementing the policy at this time, we will focus on the desired outcome of ensuring metadata is available and freely accessible rather than assessing how an institution has chosen to achieve that.

### **Expectation VI**

Where access to the data is restricted the published metadata should also give the reason and summarise the conditions which must be satisfied for access to be granted. For example ‘commercially confidential’ data, in which a business organisation has a legitimate interest, might be made available to others subject to a suitable legally enforceable non-disclosure agreement.

### **Clarification**

- Research data in which third parties have a legitimate interest should be managed in a way that does not damage the research process. Personal information should not be put into the public domain without the explicit consent of the person to which it relates. Sensitive information should not be placed in the public domain. Sensitive information is information the release of which would compromise unprotected intellectual property or which, in the judgement of the security services, would result in unacceptable risk to the citizens of the UK or its allies.
- Research organisations and researchers have a responsibility to ensure that publicly funded research involving third parties is planned and executed in such a way that published findings can be scrutinised and if necessary validated by others.
- A delay in the publication of research findings and access to supporting research data is acceptable if necessary to protect intellectual property that would otherwise be compromised.
- Third parties who collaborate in publicly funded research should be made aware of the importance of ensuring that published findings can be validated by others.

## **Expectation VII**

Research organisations will ensure that EPSRC-funded research data is securely preserved for a minimum of 10-years from the date that any researcher 'privileged access' period expires or, if others have accessed the data, from last date on which access to the data was requested by a third party; all reasonable steps will be taken to ensure that publicly-funded data is not held in any jurisdiction where the available legal safeguards provide lower levels of protection than are available in the UK.

### **Clarification**

Data that has not attracted any interest for more than 10 years is not expected to be retained. It is accepted that there may be cases in which it may not be possible or cost effective to preserve research data. This will depend on the type and scale of the data, their role in validating published results, and their predicted long term usefulness for further research. For example, in the case of simulated data or outputs of models, it may be more effective to preserve the means to recreate the data by preserving the generating code and environment, rather than preserving the data themselves. Provided that the ability to validate published research findings is not fundamentally compromised, a deliberate decision to dispose of research data at an appropriate time is acceptable in these cases.

It is reasonable to use data citations, or any other metric based on reliable sources of evidence and widely accepted at the time, to evaluate if interest has been shown in a dataset. It may be reasonable for research organisations to require persons requesting access to specific research datasets to do any of the following before granting access: register; create an account; prove identity; accept terms and conditions of use of the data.

In the event of any transfer of assets (arising for example through merger, acquisition or closure) Research Organisations should ensure that appropriate arrangements are made for the preservation and continued accessibility of retained research data in accordance with these expectations.

Research Organisations are not expected to assume responsibility for the preservation and management of third party research data not generated within their own organisation; however, in the case of important research results relying on such data it may be prudent to take reasonable steps to assure the continued availability and accessibility of the data to allow others to validate the research results.

The wording 'legal safeguards' in the expectation refers to legislation governing access to, or otherwise affecting, the security of information held in digital or electronic form.

### **Expectation VIII**

Research organisations will ensure that effective data curation is provided throughout the full data lifecycle, with 'data curation' and 'data lifecycle' being as defined by the Digital Curation Centre. The full range of responsibilities associated with data curation over the data lifecycle will be clearly allocated within the research organisation, and where research data is subject to restricted access the research organisation will implement and manage appropriate security controls; research organisations will particularly ensure that the quality assurance of their data curation processes is a specifically assigned responsibility.

### **Clarification**

The Digital Curation Centre ([www.dcc.ac.uk](http://www.dcc.ac.uk)) provides 'how to' guidance on the scope of services that institutions should be able to provide their researchers, whether by establishing in-house capabilities, collaborative arrangements with other institutions, or through outsourcing to external service providers.

### **Expectation IX**

Research organisations will ensure adequate resources are provided to support the curation of publicly-funded research data; these resources will be allocated from within their existing public funding streams, whether received from Research Councils as direct or indirect support for specific projects or from higher education Funding Councils as block grants.

### **Clarification**

Provided two simple rules are adhered to, all costs associated with research data management are eligible expenditure of research grant funds. The rules are:

- 1) no expenditure can be 'double funded' (a service that is centrally supported by the indirect costs paid on all research grants cannot then also be included as a direct cost on a grant)
- 2) all directly incurred expenditure of a grant must be incurred before the end date of the grant. RCUK has already clarified that all elements of research data management are eligible for research grant funding

It is suggested that research offices ensure appropriate provision for research data management is included in a research proposal before it is submitted to EPSRC. In particular:

- a) does a data management plan (DMP) exist? (EPSRC does not require DMPs with research grant applications, but our research data principles include that '...project specific data management policies and plans... ..should exist for all data');
- b) does the project plan reflect the time and resource needed for data management activity, such as preparation of datasets for deposition in an appropriate repository?
- c) if the project is likely to need more resource for research data management than may be provided centrally by the institution, has this been budgeted for in the proposal?

While research organisations have clear responsibilities to ensure, and allocate resource for, effective stewardship of the research data they hold, there is a balance to be struck in terms of meeting the costs of individual sharing requests. Provided the research data is suitably *discoverable* and (subject to any legitimate constraints) *accessible*, Research Organisations are free to choose the extent to which they support the *sharing* of specific research datasets; they may for example consider that enhanced reputational benefits justify greater investment in support of datasets in wide demand. In general, however, while researchers are clearly responsible for the quality of data and supporting documentation made available for sharing, EPSRC considers it reasonable that those requesting access to data be held are responsible for:

- a) determining the suitability for intended use of accessed data; and
- b) ensuring they have the necessary resources to be able to make use of it, if necessary including:
  1. re-formatting it to suit their own research needs
  2. access to proprietary third party software that may be necessary to process the data
  3. travelling to the place where the data is held if it is in physical form and it is impractical to digitise it.

## Appendix E - EPSRC Sample Letter sent to VCs (emphasis Goldsmiths')

EPSRC expectations

Our ref: 9558

10<sup>th</sup> February 2012

Dear Vice Chancellor,

We wrote in April last year to let you know about our Policy Framework on Research Data, which sets out our principles and expectations concerning how the institutions we fund will ensure that the research data generated as a result of our support is appropriately managed and shared. The importance of ensuring data is accessible was recently re-emphasised in the government's Innovation and Research Strategy for Growth. The Strategy explicitly required Research Councils to ensure there is compliance with current policies on access to research outputs.

I am therefore writing to remind you that the EPSRC policy timetable, whilst acknowledging that institutions would need time to review their current practice, gave two deadlines: by 1<sup>st</sup> May 2012 each institution should have a clear roadmap in place to align policies and processes with our expectations, and **by 1<sup>st</sup> May 2015 each institution should be fully compliant with our expectations.** You may also recall that the EPSRC policy indicates that if, after 1<sup>st</sup> May 2015, **an institution is found to be deliberately obstructing the proper sharing of research data, or otherwise seriously failing to comply with EPSRC's expectations, this will initiate a process that could ultimately lead to it being declared ineligible for EPSRC support.**

However, a key aspect of our policy is that it is not prescriptive about how our expectations should be met; it recognises that what may be appropriate for one institution may not be so for another. Since we are not in a position to evaluate the appropriateness of institution-specific plans I wanted to let you know that we will not require roadmaps to be formally submitted to us.

**We will however ask that institutions self-certify that they have taken and will continue to take the necessary steps to ensure they are able to comply with our expectations from 1<sup>st</sup> May 2015, and we may ask to see individual roadmaps on a case by case basis.** The full policy framework is available on our website and I would be grateful if you could ensure that the colleagues responsible for this area in your institution are reminded of our requirements. If you or these colleagues have any questions relating to the policy please address them directly to Ben Ryan, Senior Evaluation Manager

[\[ben.ryan@epsrc.ac.uk\]](mailto:ben.ryan@epsrc.ac.uk).

Yours sincerely

Attila Emezc (Director Communications, Information and Strategy)