A Digital Curation Centre
'working level' guide

Five things you need to know about research data management and the law

DCC Checklist on Legal Aspects of RDM

Mags McGeever, Angus Whyte and Laura Molloy, 2015

Preface

What is Digital Curation Centre (DCC)?

The Digital Curation Centre (DCC) is a world-leading centre of expertise in digital information curation. Our primary aims are

- Building research organisations' capacity, capability and skills in managing research data.
- Supporting knowledge exchange and development of good practice in digital curation.

What guidance publications does DCC produce, and for what audiences?

We aim to help a broad audience including research support, library and IT support staff, and researchers. Normally the guidance is not specific to disciplines. It includes

- **Briefings** outline current topics, explaining the importance, outlining relevant roles and responsibilities, current issues and upcoming challenges.

- **How-to Guides** provide 'working level knowledge' comprimising background concepts and practical steps towards implementing data magement capabilities in your oragnisation, or better aligning of these with best practices.

- **Checklists** also provide working-level knowledge, but cover less of the relevant background. Each aims to ensure practitioners have addressed the full scope of a challenging curation topic, and provides further sources.

- **Fold-outs** are 'quick start' guides for a researcher audience, summarising good practice in relation to a specific data management topic.

- **Case Studie**s describe how specific otganisations develop and deliver curation. Each aims to complement a How-to Guide or Checklist, and illustrate practical challenges and lessons learnt.

- **Examples** are shorter and more structured than a case study, offering 'who, what, where, how, when' summaries of approaches to RDM service delivery.

- **Catalogues** offer half page entries profiling key aspects, currently including the Tools and Services Catalogue, and Disciplinary Metadata Directory.

Introduction

Who is this Checklist intended to help?

Researchers and support staff involved in developing or delivering support for research data management (RDM). For shorthand we use the term "research data professionals". Please note the content refers mostly to UK legislation. Readers based elsewhere should consult guidance specific to their own legislative and regulatory contexts.

What does this Checklist cover and what does it exclude?

The checklist covers the main challenges for RDM support services, and the services and sources of legal information available to them. The Checklist summarises the following common elements of the main challenges for RDM support:

1. Protection of Personal Data
2. Freedom of Information (FOI) and Environmental Information (EIR)
3. Intellectual Property Rights (IPR) in Data and Databases
4. Data Sharing, Licensing and Re-use
5. Legal Considerations of Cloud Service Provision

The Checklist aims to supplement more detailed guidance, including any advice available through your institution. Likely sources would include information governance or records management specialists, the research ethics committee, and research and enterprise office colleagues specialising in contracts, licensing and legal agreements.

For many researchers, especially those in health and medical fields and in social sciences, the legal aspects of RDM overlap with ethical frameworks and procedures. The Checklist does not include any specific guidance on those, or on funding body policies that relate to them. We include some references in the section on 'Protection of Personal Data'.

Challenges for RDM support

There are many drivers for improved research data management, including opportunities presented by new technology, expectations of open access, and the need to meet research funder policies. Institutions are expected to develop a broad range of capabilities to ensure research data is properly managed and made accessible to its users[1, 2]. Legal issues have a bearing on how institutions respond to these challenges. The following RDM service areas will be particularly affected

a. Policy guidance: To ensure researchers are aware of the regulatory environment, data policy principles and expectations, and of appropriate use of exemptions that may justify withholding research data.

b. Data management planning support: To help researchers effectively manage legal risks arising from sensitive data, access requests and intellectual property rights (IPR).

c. Institutional governance processes: To ensure that sensitive data, access requests, and IPR are correctly handled at all stages of the research lifecycle.

d. IT and Computing support services: To ensure that sensitive data is held securely, and that publicly funded research data submitted to external repositories for long-term preservation is subject to legal safeguards that are equivalent to UK jurisdiction.

e. Staff development and training: To ensure content is relevant to providing secure and quality assured data management and curation.

Disclaimer – and further guidance

The DCC does not provide legal services of any kind. We provide this checklist 'as is', and aim to ensure the content is accurate and relevant, but we cannot accept any liability arising from its use. For each topic we include recommended sources of further guidance, many of which are available on the archived Jisc Legal site. Please note that Jisc Customer Services now provides support with legal aspects of data management. Also consult your institution's specialist staff for further guidance, and in case of doubt take professional legal advice.

---

[1]For an introduction to the services many institutions have established see How to Develop RDM Services - a guide for HEIs, available at: www.dcc.ac.uk/resources/how-guides/how-develope-rdm-services

[2]The DCC Policy and Legal pages offer an overview of relevant funding body policies on research data management, available at: www.dcc.ac.uk/resources policyand-legal

1. Protection of personal data

Compliance with the UK's Data Protection Act (the DPA) has been required since 1998. The Act is designed to strike a balance between the interests of the individual in maintaining privacy over their personal details and the potentially competing interests of those with legitimate reasons for using other people's personal data (see box 1). The DPA places obligations on you and your organisation if you process personal data, and in addition gives individuals certain rights in relation to data pertaining to them. The definition of 'processing' is broad and will include transfer, storage, alteration and deletion – i.e. it covers all interaction with the data concerned.

Your institution will have a framework in place to ensure the security of all personal data held for administrative reasons. The DPA also applies to personal data used for research, albeit with some exemptions, but it is a popular misconception that there is a *blanket* exemption for research. Institutions should ensure that researchers, staff and students are aware that the majority of the principles still apply to research.

Box 1. How the Data Protection Act defines 'personal data'

"Personal data means data which relate to a living individual who can be identified –
(a) from those data, or (b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual.

Sensitive personal data means personal data consisting of information as to -

(a) the racial or ethnic origin of the data subject,

(b) his political opinions,

(c) his religious beliefs or other beliefs of a similar nature,

(d) whether he is a member of a trade union (within the meaning of the Trade Union and Labour Relations (Consolidation) Act 1992),

(e) his physical or mental health or condition,

(f) his sexual life,

(g) the commission or alleged commission by him of any offence, or

(h) any proceedings for any offence committed or alleged to have been committed by him, the disposal of such proceedings or the sentence of any court in such proceedings.

Source and further details: Information Commissioner's Office www.ico.org.uk/for-organisations/guide-to-data-protection/key-definitions/

It is crucial for researchers and data professionals to be aware of the legal constraints. Advances in technology and in public expectations of data sharing have changed how researchers generate data and make it accessible. They also increase the potential for collaborative research. Increasingly global collaboration raises the issues of privacy and data protection more acutely, and poses the risk of unintentional legal infringement.

Under the DPA, eight data protection principles apply to handling personal data, shown in Box 2.

Box 2. Data Protection Principles

1. Personal data shall be processed fairly and lawfully. In practice this means that you must

- Have legitimate grounds for collecting and using the personal data.
- Not use the data in ways that have unjustified adverse effects on the individuals concerned.
- Be transparent about how you intend to use the data, and give individuals appropriate privacy notices when collecting their personal data.
- Handle people's personal data only in ways they would reasonably expect, and
- Make sure you do not do anything unlawful with the data.

2. Personal data shall be obtained only for one or more specified and lawful purposes, and shall not be further processed in any manner incompatible with that purpose or those purposes.
3. Personal data shall be adequate, relevant and not excessive in relation to the purpose or purposes for which they are processed.
4. Personal data shall be accurate and, where necessary, kept up to date.
5. Personal data processed for any purpose or purposes shall not be kept for longer than is necessary for that purpose or those purposes.
6. Personal data shall be processed in accordance with the rights of data subjects under this Act.
7. Appropriate technical and organisational measures shall be taken against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data.
8. Personal data shall not be transferred to a country or territory outside the European Economic Area unless that country or territory ensures an adequate level of protection for the rights and freedoms of data subjects in relation to the processing of personal data.

Source and further details: Information Commissioner's Office ico.org.uk/for-organisations/guide-to-data-protection/data-protection-principles/

Principles 7 and 8 need special attention if you use cloud services (see Section 5).

Research exemptions

Data used for research purposes is exempt from data protection principles 2 and 5. Also there is an exemption from the data subject's right of access where

- Personal data is not processed to support measures or decisions with respect to particular individuals.
- Personal data is not processed in a way that substantial damage or distress is likely tobe caused to any individual.
- The research results, or any resulting statistics, are effectively anonymised.

See Protection of Personal Data: Further Resources for guidance, especially point (f).

European General Data Protection Regulation

The forthcoming European General Data Protection Regulation (GDPR) is a major change and is expected to come into force in 2018. The main provisions relevant to research include

- Data controllers (including HEIs) need explicit consent to process personal data.
- Mandatory to appoint a Data Protection (DP) officer for a public body or organisation with over 250 employees.
- Mandatory obligations regarding transfer of personal data to countries outside the European Economic Area (EEA)
- Data subjects have a new 'right to be forgotten' and to have personal data erased.
- Data subject's right to obtain their data in a structured, commonly used electronic format.
- Introduces principles of privacy by default and privacy by design.
- Obligation to notify personal data breaches within 24 hours, if feasible, or without undue delay.
- Obligation to carry out a DP impact assessment prior to some processing.
- Specific protection for children.
- Clarification on processing of personal data for the purposes of historical, statistical or scientific research.

See Protection of Personal Data: Further Resources for guidance, especially points (j) and (k).

Protection of personal data: Checklist

1. Data protection is covered in any RDM guidance or training materials that you are responsible for, and:
   - Includes signposting to further guidance on research ethics ☐
   - Will be updated to reflect changes e.g. the European Data Protection Regulation ☐
2. When planning research, the following points are covered, e.g. in a Data Management Plan:
   - All data is stored and labelled with the appropriate level of security/ confidentiality, and this is documented before depositing the data in a repository ☐
   - Protection of any personal data that is collected or created ☐
   - Compliance with the institution's data protection guidance, and any IT or information security policy ☐
   - Informed consent request should include consent for data preservation and sharing, unless an ethics committee deems this inappropriate. ☐
   - Sensitive data (if any) is securely stored and transferred ☐
3. There is a clearly worded deposit agreement and adequate checks on any data deposited in any repository, so that it is only shared openly if it has been fully anonymised, and it is adequately protected otherwise ☐
4. Any data repository that you are responsible for maintaining is operating consistently with the institution's information security policy and relevant IT service guidance ☐
5. Any personal data already retained is regularly appraised to determine whether it is still needed for the purposes it was retained for, or it could be anonymised (so it doesn't come under the DPA), or securely disposed of ☐
6. Data subjects (or participants) can meaningfully exercise their right to object to the processing of data, on the grounds that it would cause them significant damage or distress ☐
7. The institution makes plans for additional technical and organisational changes that may be needed to enable it to comply with the European General Data Protection Directive ☐

Protection of personal data: further resources

a. Data Protection Act (1988), available at: www.legislation.gov.uk/ukpga/1998/29/contents
b. Data Protection Act (1988), Section 33 (the 'research exemption'), available at: www.legislation.gov.uk/ukpga/1998/29/section/33
c. Data Protection Report (blog), available at: www.dataprotectionreport.com
d. University of Edinburgh (n.d.) 'Researcher's Guide to the Data Protection Principles', available at: www.ed.ac.uk/records-management-section/data-protection/guidance-policies/research-and-the-data-protection-act/guide-principles
e. Jisc Legal (n.d.) 'Data Protection', available at: www.jisclegal.ac.uk/LegalAreas/DataProtection.aspx
f. Jisc Legal (n.d.) 'Code of Practice for the HE & FE Sectors on the DPA', available at: www.jisclegal.ac.uk/Portals/12/Repository/Data Protection Code of Practice for FE and HE.pdf
g. UK Data Archive (n.d.) 'Overview of Anonymisation' available at: www.data-archive.ac.uk/create-manage/consent-ethics/anonymisation
h. Information Commissioner's Office (n.d.) 'Topic Guide on Anonymisation', available at: ico.org.uk/for-organisations/guide-to-data-protection/anonymisation/
i. Jisc legal (n.d.) 'Data Protection Regulation' available at: www.jisclegal.ac.uk/ManageContent/ViewDetail/ID/2308/Data-Protection-Regulation.aspx
j. European Data in Health Research Alliance (2015) 'Frequently Asked Questions', available at: www.datasaveslives.eu/faqs-and-glossary/
k. Jisc Legal (2014) 'Data Protection and Research Data: Questions and Answers', available at: http://goo.gl/HfCqZM

2. Freedom of Information (FOI) and Environmental Information (EIR)

Higher Education Institutions (HEIs) in the UK have a legal duty to comply with freedom of information (FOI) legislation. This legislation promotes greater openness and accountability in public bodies, which include HEIs, despite their increasing income from private and charitable sources. It gives a general right of public access to all forms of 'recorded' information held by public authorities, sets out exemptions from that general right, and places a number of obligations on public authorities. Note that Scotland and the rest of the UK have different FOI legislation. There is also separate, and stronger, legislation covering information of relevance to the environment - the Environmental Information Regulations 2004 and Environmental Information (Scotland) Regulations 2004. See Further Resources below for more guidance.

FOI is a wide ranging and multi-faceted requirement which is relevant to staff at all levels of your institution, including researchers and support staff. Under the legislation, HEIs have two main responsibilities: providing a publication scheme and handling requests for information.

Publication Schemes

HEIs should adopt and maintain a Publication Scheme (typically on the web). This links to documents proactively made available to the public by the HEI. Your institution is likely to provide this centrally.

Requests for information

The second main responsibility of HEIs is to respond to requests for information. This is extremely wide-ranging although so are the possible exemptions. The applicant requesting the information can be an individual or organisation from anywhere and does not have to be the subject of the information or be affected by its holding or use. Applications must be made in writing but are not required to mention the legislation. Requests must be dealt with promptly within a maximum time frame of 20 working days. See Further Resources (g) for more detail.

Box 4. Definitions of 'information' and 'datasets' in FOI and EIR

Information can include "any recorded information …a public authority may hold. This includes information held on computers, in emails and in printed or handwritten documents as well as images and video and audio recordings. A request may be written in the form of a question, rather than a request for specific documents; it may be addressed to any person in a public authority. For a request to be valid for the purposes of FOI, the request must be made in writing, state the name of the applicant and an address for correspondence and describe the information requested. Under the EIR, verbal requests are also valid." Note that the research data or information 'held' may include that which is funded by a HEI or produced under contract with it. See further guidance (b).

Dataset is defined in section 6 of the Protection of Freedoms act as follows:
" 'dataset' means information comprising a collection of information held in electronic form where all or most of the information in the collection (a)has been obtained or recorded for the purpose of providing a public authority with information in connection with the provision of a service by the authority or the carrying out of any other function of the authority, (b) is factual information which—(i) is not the product of analysis or interpretation other than calculation, and (ii) is not an official statistic …and (c) remains presented in a way that (except for the purpose of forming part of the collection) has not been organised, adapted or otherwise materially altered since it was obtained or recorded." See further guidance (c).

Environmental information is very broadly defined in the legislation, and could include research data, for example where it consists of "… written, visual, aural, electronic or any other material on… the state of the elements of the environment, such as air and atmosphere, water, soil, land, landscape and natural sites including wetlands, coastal and marine areas, biological diversity and its components…". See further guidance (j).

Exemptions – absolute and qualified

The legislation tends to presume in favour of disclosure, although there are exemptions. The most relevant exemptions for research data professionals are likely to be data or information that has one or more of these characteristics:

- produced in research intended for future publication;
- people's health and safety would be affected by its disclosure;
- commercial interests of any person or organisation would be prejudiced;
- it contains personal data.

The exemptions fall into two categories: absolute and qualified.

Absolute exemptions are cases where the information seeker's request can be disregarded. Examples of these include data or information that has one or more of these characteristics:

- otherwise accessible e.g. in your institution's publication scheme, or institutional repository;
- confidential material whose disclosure would lead to an action for breach of confidence (marking it as 'confidential' is not enough to fit this exemption);
- personal information covered by data protection regulations, which still need to be upheld (if in doubt consult your FOI officer).

Qualified exemptions include unpublished research and are invoked on a case-by-case basis through a two-stage procedure

1. Institution decides whether the exemption could be used.
2. Institution applies the 'public interest' test to decide whether disclosing the information is more in line with the public interest than applying the exemption.

Exemption for unpublished academic research

The exemption for academic research prevents the premature disclosure of research data genuinely intended (rather than just vaguely planned) for publication. This is exempt from disclosure if it relates to information obtained in the course of, or derived from, a programme of continuing research that is intended for future publication. A further subsection, however, provides that the information will be exempt only if disclosure would, or would be likely to, prejudice a matter listed in that subsection. Check with your institution's FOI practitioner for help with this.

Datasets

The Protection of Freedoms Act 2012 (POFA) changed FOI legislation in England, Wales and Northern Ireland (but not Scotland). Under the new act, public authorities must proactively release datasets in a reusable format. The creation of this new 'right to data' means that datasets requested from public authorities must be provided in a useable format and thereafter published on a regular basis through the publication scheme, ensuring that all data published by authorities is made available in an open and standardised format so that it can be used easily and with minimal cost by third parties.

The new provisions are about *how* information is released, rather than *what* information is released. They only relate to information that the institution holds as a dataset (as defined earlier). There is no new duty to provide any information in response to an FOIA request that was not previously accessible, and there are no new exemptions from that duty in POFA. Any of the exemptions available offer a route to deciding that releasing a dataset would be 'inappropriate'.

Institutions can also take into consideration other factors to decide whether it is 'reasonably practical' to convert a requested dataset into a reusable form. The legislation does not define 'reasonably practical' but the Information Commissioner's Office (ICO) guidance says relevant factors may include the time and cost of conversion, technical issues and the resources of the public authority.[3] POFA also allows data-holding institutions to charge for the fulfilment of requests for datasets.

[3] *Para 47, ICO Guidance on FOI and Datasets-http://www.ico.org.uk/news/blog/2015/~/media/documents/library/Freedom_of_information/Detailed_specialist_guides/datasets-foi-guidance.pdf*

Freedom of Information: Checklist

1. The institution has data governance processes in place to deal with FOI requests for research data, and relevant support staff are aware of applicable exemptions ☐
2. The institution is responsible for categories of research information or that fit the FOI definition of a 'dataset' ☐
3. Researchers are given appropriate guidance on the application of FOI to data, including the applicable exemptions ☐

Freedom of Information: further resources

a. ICO 'Freedom of information and environmental information regulations', available at: ww.ico.org.uk/for-organisations/guidance-index/freedom-of-information-and-environmental-information-regulations/

b. ICO (n.d.) 'FOI Legislation and Research Information: Guidance for the HE Sector', available at: https://ico.org.uk/media/for-organisations/documents/1133/foi-legislation-and-research-guidance-for-the-higher-education-sector.pdf

c. ICO (n.d.) 'Guidance on FOI and Datasets', available at: www.ico.org.uk/news/blog/2013/~/media/documents/library/Freedom_of_Information/Detailed_specialist_guides/datasets-foi-guidance.pdf

d. ICO (n.d.) 'FOI Definition Document for Universities and other HEIs', available at: www.ico.org.uk/news/current_topics/~/media/documents/library/Freedom_of_Information/Detailed_specialist_guides/definition_document_for_universities_and_higher_education_institutions.pdf

e. University of Edinburgh (n.d.) ''Freedom of Information in Scotland and the rest of the UK', available at: www.ed.ac.uk/schools-departments/records-management-section/freedom-of-information/about/scotland-uk

f. Jisc (2010) 'Freedom of Information and Research data: Questions and answers' - Andrew Charlesworth and Chris Rusbridge, available at: www.jisc.ac.uk/publications/programmerelated/2010/foiresearchdata.aspx

g. ICO (n.d.) 'Guidance on the exemption for information intended for future publication', available at: ico.org.uk/media/for-organisations/documents/1172/section_22_information_intended_for_future_publication.pdf

h. ICO (n.d.) ' Guidance on commercial Interests exemption', available at: - ico.org.uk/media/for-organisations/documents/1178/awareness_guidance_5_v3_07_03_08.pdf

i. ICO (n.d.) 'Guidance on how to apply the FOIA exemption relating to personal data', available at: ico.org.uk/media/for-organisations/documents/1213/personal-information-section-40-and-regulation-13-foia-and-eir-guidance.pdf .

3. Intellectual Property Rights (IPR) in data and databases

Intellectual property rights (IPR) are rights granted to creators and owners of works that are the result of human intellectual creativity. Many of these rights, e.g. copyright, apply automatically without their creator having to apply for them. Knowing what rights are attached to data, and their implications, is crucial to being able to legally share, use or reuse data.

A key point is that a copyright work can be legally used and reused if there is a licence attached to it that permits these acts. Licensing, which is discussed in the section 4, 'Data Sharing, Licensing and Re-use', can remove many of the headaches around copyright. The complex law in this area makes it difficult even for the experienced lawyer to establish conclusively whether a data resource attracts IPR or not and/or whether a certain use falls within an exception/permission. Sadly, researchers or institutions may unintentionally curtail data use and reuse, and this can be a genuine barrier to scientific progress.

In this section we focus on copyright and the database right, a form of IPR similar to copyright. It is possible for any one database to enjoy copyright, the database right, both or none.

Copyright

This protects the expression of ideas or information. Some definitions of research data applicable in social science and humanities[4] are broader than the legal definition of datasets (see above under FOI) but would be included under legal definitions of 'information'. For many scientists and some funders (e.g. EPSRC) however, research data are 'facts' rather than interpretation or creative expression.

Considered as 'facts', data may be seen as falling outside the realm of copyright protection. Nevertheless many research outputs that would in practice be treated either as data, metadata or contextual information fall within the copyright law definition of a 'literary, artistic, dramatic or musical work, sound recording, film or broadcast or a typographical arrangement of published edition'. Typical examples likely to attract copyright include dataset annotations, lab notebooks, experimental protocols, questionnaires, interview transcripts, coding guides, and analytical reports.

While copyright law does not distinguish between factual data and creative output in any absolute way, it is still a useful distinction. It is especially important for databases because the contents of a database may have different copyright to the collection as a whole. Either the content or the collection as whole may or may not attract copyright

- Firstly, copyright might exist independently in the *contents* of the database, for example in a database of images each image could attract its own copyright as an 'artistic work'. This is less likely to be so if individual database items can be identified as factual.
- Secondly, copyright may protect the *structure* of a database, for example, if the selection or arrangement of the contents can be said to constitute the database author's own intellectual creation.

If copyright does exist, the copyright holder has the exclusive right to perform certain actions in relation to the work, and can prevent others from doing these. The most significant actions for research data professionals are: copying the work, making an adaptation of it or communicating it to the public. The latter would include depositing in a repository or putting the work on a website.

There are certain things prospective users can do without permission. They can use an insubstantial amount of the work (measured as much by quality or importance as it is by quantity). Academic researchers and data users may be able to use the 'fair dealing' provisions (see further guidance (g)) to allow an individual to make a copy of all or a substantial part of a work without having to ask permission or pay fees, as long as certain conditions apply. One of these conditions is 'non-commercial research'. A second condition is that the copying is fair, meaning that it must not damage the legitimate commercial or other interests of the rights holder.

The database right

This is a form of IPR existing only in EU countries and applies to a database if there has been substantial investment in obtaining, verifying or presenting the contents. It is important to note this only applies to collecting materials that already exist, not to the creation of the database contents. Of course it may be practically difficult to distinguish between creation and collection, especially where the same body or person is responsible for them both.[5]

The database right applies to databases, whether or not copyright applies to their arrangement or to individual items in its contents. It applies where the contents have been wholly or substantially taken out and re-arranged (generally by a computer) so as to provide a quite different organisation but to essentially the same material — a re-organisation which would not necessarily amount to infringement of copyright in the original arrangement.

The right lasts for 15 years from making the database, or from publication (if that is later). However, further substantial investment in additions, deletions or alterations starts time running afresh. The potential duration of the right in the case of dynamic databases being continually re-published has been questioned and this is not yet resolved.

The database right prevents extraction or reutilisation of a substantial part of the database. Again, there is a lack of certainty about the meaning of "substantial". A community could draft its own guidelines and these would likely exercise a strong influence on most users. However, exact interpretation would remain with the courts.

There is a relevant but limited exemption, in that the extraction (but not reuse) of data from the database is allowed for illustration in teaching or research, but not for any commercial purpose.

---

[4]For example research data has been defined as "representations of observations, objects, or other entities used as evidence of phenomena for the purposes of research or scholarship". CL Borgman (2015). "Big Data, Little Data, No Data. Scholarship in the Networked World, MIT Press (p.28)

[5]Dr C.Waelde and M.McGinley (2005). 'Public Domain; Public Interest; Public Funding; focussing on the 'three Ps' in scientific research', available at: www2.law.ed.ac.uk/ahrc/script-ed/vol2-1/3ps.asp

Ownership of IPR in UK law

- Copyright in the contents or structure of database: First owner is normally the creator of the work. For single-institution projects ownership will usually lie with the University. However there are exceptions (see below)
- Database Right: First owner is the 'maker' of the database (the 'investing initiator'), i.e. whoever takes the initiative in obtaining, verifying or presenting the contents

Institutional policy can affect IPR ownership in data, metadata or contextual information. If someone creates the data or database as part of their employee duties, it is their employer who owns the copyright, unless there is an agreement to the contrary. If an institution's employment contracts (or other IP agreements?) are not explicit about IPR, the law would recognise 'custom and practice' as an implicit agreement. That means the researcher may own any copyright that exists in the data, metadata or contextual information if it is common practice that they take the initiative in obtaining, verifying or presenting it. This may also apply if it is not clear if the person is an employee or not.

Institutional guidance should identify how an institution's IPR policy applies if a research funder's contractual terms conflict with it, i.e. which policy should take precedence.[6] Clarity on this is especially important as it is good practice for researchers to state in a data management plan (DMP) who will own the copyright and other IPR in any data that will be collected or created, and how the data will be shared or licensed for reuse.

Student ownership of IPR

Students are not normally employees of their institution, so their IPR needs particular attention, especially where research funders (e.g. the EPSRC) expect institutions to make publicly accessible the data from studentships they fund. Students can be asked to assign or license materials they create to the institution. Care should be taken to do this fairly, otherwise any written agreement may be deemed illegal if challenged in court. See points (g, h) in Intellectual Property Rights: Further Resources .

International considerations

For multi-partner projects, especially international ones, a consortium agreement should cover IPR ownership of data and research outputs.[7] Jurisdictions vary, especially in how copyright and the database right apply. For example the database right exists in the EU but not the US. In the US databases can be protected by copyright if the collector applied creativity in producing a "compilation". In other jurisdictions e.g. Australia, copyright is legally defined in terms of originality. This can be judged on factors including skill and labour. See further resources (i).

Intellectual Property Rights: Checklist

1. The institution's policy guidance on RDM has explicit statements about ownership of IPR in data and research records
2. Collaborative research projects with external partners, especially international ones, have effective support to produce a clear written agreement about ownership of IPR in data and research records
3. The institutions' IPR and RDM policies are consistent with each other, and identify which terms take precedence in any conflict within a research contract
4. DMP guidance is consistent with relevant institutional guidance on IPR
5. Any internal repository that is used to hold research data has a workflow (e.g. a checklist) to ensure that depositors are not depositing any third-party data without the appropriate permission and copyright clearance
6. Guidance to researchers and students on depositing data in any internal or external repository, deals effectively with the licensing of any copyright they own

[6]See for example 'University of Oxford Policy on the management of reserach data and records', available at: www.researchdata.ox.ac.uk/university-of-oxford-policy-on-the-management-of-research-data-and-records/

[7]See 'Example of a Basic Consortium Agreement' available at: www.web2rights.org.uknavigator/content/documents/3.6_Template%20Consortium%20Agreement_1.1.pdf

Intellectual Property Rights: further resources

a. ICO Topic Guide on Data Sharing, available at: www.ico.org.uk/for-organisations/guide-to-data-protection/data-sharing/

b. Jisc IPR Toolkit, available at: www.jisc.ac.uk/media/documents/publications/scaiprtoolkitoverview.pdf

c. Intellectual Property Office Copyright pages, available at: www.ipo.gov.uk/types/copy.htm

d. Crash Course on EU Database Rights, available at: www.iusmentis.com/databases/crashcourse/

e. Web2Rights IPR and Legal Issues Toolkit, available at: www.web2rights.org.uk/documents.html

f. Jisc Legal (2014) 'Copyright Law Overview', available at: www.jisclegal.ac.uk/ManageContent/ViewDetail/ID/3588/Copyright-Law-Overview-12-June-2014.aspx

g. Jisc Legal (2004) 'Intellectual Property and Electronic Theses, available at: www.jisclegal.ac.uk/Portals/12/Documents/PDFs/ethesepaper.pdf

h. Jisc Legal (2007) Investigation into Student Work and IPR, available at: www.jisclegal.ac.uk/portals/12/documents/pdfs/jisclegalstudentiprreport.pdf

4. Data sharing, licensing and re-use

Anyone holding IPR in a dataset needs to provide its users with clear guidance on what they are permitting its users to do with it, and effective licensing does this. Research and innovation are hampered by problems interpreting and acting in accordance with copyright and license restrictions, especially when datasets are combined for larger-scale analysis.[8]

In this section we identify the need for institutions to offer guidance to researchers on the use of open licenses for data. More detailed help on this is available in a DCC guide *How to License Data* (see further resources (a) below). Deposit agreements for repositories are also discussed below. As researchers are most likely to face data licensing issues when accessing a repository or depositing in one, the institution will need to have such an agreement if it has its own repository.

Licensing

Writing a bespoke licence for data that is to be shared publicly is in most cases unnecessary, not least because of the potential pitfalls of creating one yourself, but also because of the cost of getting a professional involved, and the range of standard licences available as an alternative. It is good practice to use licenses that are open and standard. These include Creative Commons licenses (preferably version 4), Open Data Commons, or Open Government Licences. These offer the benefit of legal interoperability and therefore increase the usability of research data. There may be exceptions – if there is a plan to exploit commercial value in the data, a need to protect any database rights (see section 3), or unusual rights to be defined. An institution's commercialisation specialists may give further guidance in these circumstances.

Commercial considerations aside, the main 'good' reason to use a bespoke licence is to comply with legislation that safeguards data subjects, e.g. Data Protection legislation. For example, the UK Data Archive requires users to sign up to an End User Licence, which imposes 'safeguarded' access to data about human research subjects or participants. Where an institution has a repository that accepts research data, an 'end-user license' may be an appropriate way to define the terms under which it is accepted (e.g. the UK Data Archive). However an 'end-user agreement' would be the appropriate term only if the institution's researchers are allowed to deposit in the repository, and the institution already owns the IPR to any data they create.

Contract law – applying terms and conditions to data access or deposit

Repositories and databases can achieve similar results to a license by specifying contractual terms and conditions ('T&C's') that a user must agree to in order to access or deposit data. An institutional RDM service should be able to help researchers judge whether it is appropriate to agree to the T&Cs a repository attaches to data access and use, as this may affect later data sharing. The service also needs to guide researchers later on, when they come to share their data, on whether to agree to T&Cs an external repository attaches to depositing data. The service may also need to amend the T&Cs for any in-house repository or database it expects researchers to put their data in.

[8] *e.g. Jisc 'Value and Benefits of Text Mining', (2012, updated 2015) available at: - www.jisc.ac.uk/reports/value-and-benefits-of-text-mining; and Reimer, T. (2012) 'Text mining, copyright and the benefits and barriers to innovation', insights 25(2)*

Repositories may have 'good' reasons to attach access conditions (i.e. other than self-interest). For example a repository has a legitimate need to require users to agree to any access conditions imposed by legislation or by the funder of the data. Other reasons, which may be 'less good' in terms of meeting Open Access principles, are conditions on the sharing and reuse of data that is behind a subscription paywall, e.g. commercial data. This may be entirely legal, and there may be good research reasons for using commercially sourced data. However if there is an alternative source of the data that may be used without incurring the same limitations on reuse, doing so may reduce any risk of non-compliance with a research funder's contractual terms.

Support for depositing data to an institutional repository should ideally include a deposit checklist, and guidance on the T&Cs (see University of Edinburgh DataShare, for example). As well as guiding depositors on an appropriate choice of open license, the T&Cs should require some 'due diligence' on the depositor's part. They may for example require depositors to warrant that data for public sharing is free of personal data, and does not infringe the copyright of any other person. It is also important to check any human subjects have given informed consent to the deposit, as a consent form can itself be considered a contract.

Personal data and other legislative considerations

Data protection principles must be upheld when sharing any data considered personal (see section 1, Protection of Personal Data). The EU Human Rights Act is also relevant in sharing personal data, in particular Article 8, which states: "Everyone has the right to respect for his private and family life, his home and his correspondence."[9] If data sharing complies with the Data Protection Act, it will most likely comply with this, as the safeguards in it derive from the Human Rights Act.

One additional concern around data sharing is potential breach of confidence. A 'duty of confidence' arises when a party either knows or ought to know that the other party could reasonably expect their privacy to be protected. In an institutional environment this may apply where research has been carried out with a commercial partner.

There is further legislation applicable to sharing, e.g. Official Statistics (regulated by the Statistics and Registration Services Act, 2007), and 'administrative data'. This includes personal data made available for research by Government departments and agencies. The Administrative Data Research Network (ADRN) provides legal advice for researchers, particularly those who have to be accredited and trained to use data safely and lawfully before they can access any personal data in a secure environment. See section 'Data Sharing, Licensing and Re-use: Further Resources', points (f and j).

Future reform

Reform of the law on data sharing between public bodies has been recommended by the UK Law Commission. Their 2014 report was based on public consultation on whether the current law is sufficiently clear and certain, whether there are inappropriate obstacles, and whether the right balance exists between the ability to share data and the protection of privacy[9]. Hopefully reform will result in better sharing and linkage of administrative data, which would in turn have a positive effect on research activities in this area.

Practitioners with an interest in administrative data should keep track of changes in this area, which may also come about through European law.

Data Sharing, Licensing and Re-use: Checklist

1. Institutional policy gives clear guidance on the use of open licenses for data, and how to deal with any conflicts with research funders' licensing requirements ☐
2. Researchers are advised about the fact that using non-open, non-standard licenses will make it difficult or even impossible for others to reuse their data ☐
3. Institutional repository deposit workflows include checks that are appropriate for data, e.g. that the depositor is not in breach of any pre-existing licence or contract terms and conditions, and that the repository offers a choice of licenses that are appropriate to the data ☐
4. Any repository holding data covered by the Data Protection Act ensures that users agree to a license, or a statement of terms and conditions, committing them to ensure any subsequent data access and reuse complies with the Data Protection principles. Data in the repository is protected by institutional policy on information security, and operating procedures that comply with information security standards ☐
5. RDM support professionals monitor changes in legislation affecting data sharing, including changes in EU Data Protection regulations ☐

[9] *www.lawcommission.justice.gov.uk/docs/cp214_data-sharing.pdf*

Data Sharing, Licensing and Re-use: Further Resources

a. Ball, A. (2014). 'How to License Research Data'. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available at: www.dcc.ac.uk/resources/how-guides

b. Open Data Commons (n.d.) Licences FAQ, available at: www.opendatacommons.org/faq/licenses/

c. Open Data Commons (n.d.) 'Making Your Data Open – A Guide', available at: www.opendatacommons.org/guide/

d. University of Edinburgh (n.d.) 'DataShare repository: Checklist for deposit', available at: www.ed.ac.uk/schools-departments/information-services/services/research-support/data-library/data-repository/checklist

e. UK Data Archive (n.d.) 'End User Licence', available at: www.data-archive.ac.uk/conditions/data-access

f. UK Data Archive (n.d.) 'Create & Manage Data: Legislation', available at: www.data-archive.ac.uk/create-manage/consent-ethics/legal?index=1

g. Digital Curation Centre 'Where to Keep Research Data' (forthcoming 2015), will be available at: www.dcc.ac.uk/resources/how-guides

h. Law Commission (2014) 'Data Sharing between Public Bodies report', available at: www.lawcom.gov.uk/document/data-sharing-between-public-bodies/

i. Open Knowledge (n.d.) 'Open Definition - Conformant Licenses', available at: www.opendefinition.org/licenses

j. Administrative Data Network (n.d.) 'Protecting Privacy – Legal framework', available at: www.adrn.ac.uk/protecting-privacy/legal

5. Legal Considerations of Cloud Service Provision

The term 'cloud services' refers to services that a third party provides via the internet or a private network, to enable organisations and individuals to use pooled resources such as storage, software applications or computing power. Cloud services can usually be deployed 'on demand' and typically offer some level of scalability and customisation. These characteristics can enable organisations to avoid costs of building or operating services on their own premises.

Typical RDM use cases for cloud services include working storage for 'active' research data, wikis or similar tools for maintaining research records, or the underlying storage used by a data repository or digital preservation facility. They offer efficiencies and ease of use, as long as the potential legal risks are attended to.

The main risks stem from loss of control of the service provision, as would be expected with the outsourcing of any organisational function. The specific risks cover all four of the legal aspects of RDM already summarised in this guide. These risks are helpfully summarised in a guide produced by The National Archives (TNA) *'Guidance on Cloud Storage and Digital Preservation'* (see Cloud Services: Further Resources, point (a)). This identifies three categories of legal issues for archives using cloud services, as follows:

1. Any legal requirements in terms of management, preservation, and access placed upon archives and their parent organisations, by their donors and funders via contracts and agreements or via Government legislation (e.g. accessibility, availability, information security, retention, audit and compliance, Public Records Act, etc.);
2. Those legal obligations relating to third party rights in, or over, the data to be stored (e.g. copyright, data protection); and
3. The legal elements of the relationship between an archive and a cloud service provider or providers (e.g. terms of service contracts and service level agreements).

A critical point to note is that an institution is responsible for data protection compliance for personal data for which it is the data controller whether or not it uses a cloud based data processor (such as Google) to manage or store that personal data on its behalf.

For a full consideration of these issues please refer to the further resources in this section. The TNA guide is highly recommended, particularly section 7, which is summarised in section 1.5, 'Legal Issues'. Many of these issues apply to RDM services using cloud services, especially in large digital data-holding institutions such as universities and other cultural heritage institutions. Institutions that are connected to Janet[10], and buy cloud services through Jisc framework agreements[11] can be confident that these legal issues have been addressed, making it easier to check that they meet the institution's needs.

Cloud Services: Checklist

1. The notification that your institution has sent to the Information Commissioner's Office correctly identifies the countries where data might be sent to, and the uses made of it, as a result of cloud services being used for RDM ☐
2. Data protection implications of using cloud services have been considered, particularly when personal data may be stored outside the European Economic Area, to ensure the service complies with information security standards ☐
3. Freedom of Information implications of using cloud services have been considered, particularly the availability of data archived for long-term preservation ☐
4. Copyright and licensing implications of using cloud services have been considered, particularly when depositing and storing data containing third-party IPR in overseas legal jurisdictions ☐
5. Risks of serious loss, destruction or corruption of data have been assessed, to ensure there is adequate provision for data recovery and continuity of the service in the event of the provider ceasing operation ☐

Cloud Services: Further Resources

a. The National Archives (2014) 'Guidance on Cloud Storage and Digital Preservation', available at: www.nationalarchives.gov.uk/documents/archives/cloud-storage-guidance.pdf,
b. Jisc (n.d.) 'Cloud Computing', available at: www.jisc.ac.uk/guides/cloud-computing
c. Jisc Legal 'What are the FOI and DP implications of using services like Dropbox or evernote?' (2013), available at': www.jisclegal.ac.uk/ManageContent/ViewDetail/ID/2883/What-are-the-FOI-and-DP-implications-of-using-services-like-Dropbox-or-Evernote-28-January-2013.aspx
d. Information Commissioner's Office 'Information security (Principle 7)', available at: www.ico.org.uk/for_organisations/data_protection/the_guide/principle_7
e. Information Commissioner's Office 'Sending personal data outside the European Economic Area (Principle 8)', available at: www.ico.org.uk/for_organisations/data_protection/the_guide/principle_8
f. Oppenheim, C. (2012) 'The No-Nonsense Guide to Legal Issues in Web 2.0 and Cloud Computing ' London: Facet Publishing

Acknowledgments

Thank you to Rachel Bruce and John Kelly at Jisc for their comments.

[10] *See Jisc 'Janet Network', available at: www.jisc.ac.uk/janet*
[11] *Jisc 'Data Archiving Framework', available at: www.jisc.ac.uk/data-archiving-framework*

Please cite as: DCC (2015)**.** 'Five Things You Need to Know About Research Data Management and the Law: DCC Checklist on Legal Aspects of RDM'. *DCC Publications*. Edinburgh: Digital Curation Centre.
Available online: www.dcc.ac.uk/resources

Follow the DCC on Twitter @digitalcuration, #ukdcc