

中国科学院文献情报中心科技信息政策中心  
系列编译资料

# 决定保存哪些数据的五个步骤

## Five steps to decide what data to keep

原著者：Angus Whyte

组织策划：顾立平

编译者：江娴、王微

校对者：江娴、王微、刘晶晶、罗娇

审核者：顾立平

2015 年 3 月



## 使用须知

中国科学院文献情报中心为促进学术交流、促进文献情报服务创新发展，特组织编译《决定保存哪些数据的五个步骤》，供个人学习和研究使用。文献情报机构可以在保证《决定保存哪些数据的五个步骤》完整性和所有编译者信息完整准确的条件下在网站上整期上载和传播《决定保存哪些数据的五个步骤》的 PDF 版本，并明确说明来源。

任何机构或个人在引用《决定保存哪些数据的五个步骤》内的具体中文编译内容时，请按照学术规范注明来源，包括原始文献的著者、题名和来源网址等，也包括编译者姓名和编译内容来源。如果任何机构或个人要直接整条采用具体中文编译内容（包括仅对文字进行非实质调整后的采用）、或者要对较长中文编译内容直接采用其较大篇幅内容（例如超过五百字以上），应事先征得编译者的同意。任何机构和个人，未经中国科学院文献情报中心许可，不能直接把《决定保存哪些数据的五个步骤》的内容大规模直接编撰为新的作品或作品的一部分。

中文读者使用本份文件时，请尽量参照原作，并且引用原著及其文献来源出处。根据国家标准 GB（7714-2005），引用建议如下：

Angus Whyte.Five steps to decide what data to keep[EB/OL].

<http://www.dcc.ac.uk/resources/how-guides/five-steps-decide-what-data-keep>



本作品采用[知识共享署名-非商业性使用-禁止演绎 3.0 中国大陆许可协议](https://creativecommons.org/licenses/by-nc-nd/3.0/)进行许可。

# 目录

前言.....	4
简介——您的选择是什么? .....	5
需要存储什么数据, 以及需要存储多久? .....	6
第一步: 确定数据能够实现目标.....	7
第二步: 确定必须保存的数据.....	8
第三步: 确定应该保存的数据.....	10
第四步: 权衡成本.....	11
第五步: 完成数据评估.....	13
科研数据保存策略指南.....	14
注释.....	15
鸣谢.....	17



## 决定保存哪些数据的五个步骤

请按如下格式引用：DCC（2014 年）。“决定保存哪些数据的五个步骤：评估科研数据的清单 V.1”。爱丁堡：数字保存中心。请访问：<http://www.dcc.ac.uk/resources/how-guides>。

本文执行苏格兰知识共享署名许可 2.5，除第 4 节外，均改编自许可证 CC-BY-NC-SA。来自英国的数据存档（2013 年）数据管理成本计算工具和清单。请访问：<http://www.data-archive.ac.uk/create-manage/planning-for-sharing/costing>

## 前言

本指南旨在辅助英国高等教育机构帮助其研究人员在决定保存哪些科研数据方面作出明智的选择。该内容补充了其他的 DCC 指南：如何评估和选择需要保存的科研数据<sup>[1]</sup>，以及如何发展科研数据管理服务<sup>[2]</sup>。该指南将指导有关研究人员在逐个项目的基础上作出决定，或制订部门工作指南。指南假定：研究人员对特定数据集作出的决定，通常会考虑适当工作人员（如学术联络馆员）的建议，从而制定科研数据管理（Research Data Management，简称 RDM）的机构政策，并在他们自己的领域内提供指导。因此，该指南还可以指导高等教育机构、专业或学术团体或类似学术机构的相关政策制定人员。

该指南假设他们研究项目的负责人，或者负责数据管理的其他研究人员，需要选择保存哪些数据，同时作出共享或保存数据（例如，在数据管理计划）的承诺。评估单位是“数据集”，这可能包括：所采集的不同访问权限和或许可条件下的不同文件。

该文还假定机构将提供以下功能：

公众资助数据的长期价值的机构目录或注册表，确保潜在的用户能找出保存了什么数据，为什么保存、什么时间保存、数据是如何生成的，以及如何访问数据

如果外部存储库不可用，提供设施以保持所选择数据的潜在长期价值，在存在有效外部请求时，辅助任何非数字资料的数字化

有关上述任一功能将如何提供不存在任何假设；例如，他们可能是存储库或托管存储服务，不同于或集成了出版物资料库或 CRIS（当前研究信息系统）。在任一情况下，可以以内部或外包形式提供这些功能，例如通过 Janet 云服务<sup>[3]</sup>。该指南可以调整，以反映本地服务，并指导数据存储时外部存储库的选择<sup>[4]</sup>。DCC 可以提供帮助，以满足定制机构的需求，并提供视觉设计<sup>[5]</sup>。（数据保存中心，Angus Whyte）



# 简介——您的选择是什么？

作为一名研究人员，您可能会从研究周期的不同时期中选择可用的数据。您将从数据源中选择可用的数据，用于您的初始研究，从收集的数据中选择以进行分析，然后利用分析的结果数据对已有结论加以阐述，一些结论可能已经发表。

随着越来越多的数字化技术应用于研究，人们有越来越强的需求，需要进一步选择：应长期保存什么、什么数据可用、什么数据应该处理。如果数据是纵向的或引用的数据集，进行数据评估的最佳时间是在项目结束之前，或周期性地进行。

本指南旨在帮助您对选择什么样的数据来保存作出抉择，以实现自己的目的，并符合您的机构和外部资助者的要求。关于用什么来保存您的数据这一问题，您可能有许多种选择：

使用已经创建的有关您的研究领域的外部存档或资料库

使用数据共享平台，如 [Figshare.com](https://figshare.com) 或 [Zenodo.org](https://zenodo.org)

作为研究文章的补充材料，提供给出版商

根据您的学科公认的标准，使用您的研究小组已经建立的数据管理设施保存数据

如果基于科研数据资源库的机构（见下文）可用

如果您所在领域，甚至您的研究小组或部门，有既定的数据管理工具<sup>[6]</sup>，则很容易进行选择。您的研究资助者可能会建议选择数据中心或自存档案。例如，英国的数据服务为社会科学家们提供共享工作簿存档（[reshare.ukdataservice.ac.uk](https://reshare.ukdataservice.ac.uk)）。选择时需要考虑多种因素，如存储库是否：

给您所提交的数据集持久唯一的标识符

为每个数据集提供一个查询页面，利用元数据帮助别人找到数据、了解数据以及引用数据

帮助您跟踪数据如何被使用

响应社会需要和或认证为“值得信赖的数据存储库”

提供明确的条款和条件，符合法律要求，例如，拥有数据保护、允许重复使用，同时避免不必要的许可条件

即将颁布的 DCC 指南有助于进一步选择外部资源。您所在的机构可能提供科研数据管理支持，以帮助您解决这些问题，并使您的研究获得最大的投资。这可能涉及：

在该机构的数据目录中注册数据集，以扩大研究可见性



用机构知识库保存数据集，以保持其存储期间的长期记录，如果该数据是公开的，保存其访问和下载统计。

确保选定数据在您机构用于长期保存的专用存储库中的安全性

建议采用合适的外部存储库

该指南将指导您完成以下五个步骤：

考虑潜在的重用目的——该数据可以满足什么样的目标？

考虑法律或政策合规风险

确定保存哪些可能具有长期价值的

数据  
权衡成本——已经产生了哪些数据管理成本，且这些成本是否有助于它的价值实现，以及后期计划且负担得起的成本是多少？支付这些费用的资金从何而来？这些问题是您的数据评估的成本要素，都需要考虑到，并有助于识别任何外部咨询需求，例如，如何应对任何预算短缺。

完成对您数据的评估——这将列出数据必须、应该还是可以实现哪些潜在的重用目的。评估还应该总结需要为存储数据作出的任何准备行动，以及不保存它的理由。

本指南主要借鉴了现有的 DCC 指南“如何评价和选择需要保存的科研数据”<sup>[7]</sup>、英国自然环境研究理事会(Natural Environment Research Council, 简称 NERC)数据价值清单<sup>[8]</sup>和布里斯托尔大学(University of Bristol)的科研数据评估指南<sup>[9]</sup>。第 4 节是改编自英国数据服务的数据管理成本计算工具和清单。

## 需要存储什么数据，以及需要存储多久？

本指南采用科研数据的广义定义，“出于研究或学术用途，作为证据代表观察结果、对象或其他实体”<sup>[10]</sup>。在选择需要保存的数据时，您需要考虑以下哪两大类<sup>[11]</sup>是否适合重用：

1、源数据：收集、创建或其他地方研究已使用的数据

2、集成的数据集：从上述（1）中提取或衍生的数据

3、引用的数据：为了进一步分析或得出结论，加工上述（2）中数据子集所得的数据集。

与您所在领域研究成果一致的“补充材料”。

本指南还使用了数据收集这个术语，适用于任何收集对象，这些对象时访问或者解释上述内容是所必需的。例如笔记本电脑、协议、软件或一系列校准仪器，无论其是全部或部分



数字化。在某些情况下，可能有理由保持软件使用以操作和分析产生的数据。这些理由可能强如保持数据本身，例如，如果要求软件可以使结果再现。在其他情况下，它可能更加强大，例如，如果结果只能由该算法转分析得出。

一个单一的研究项目可以很容易地生产出许多“数据集”，每个匹配不同的潜在用途。正如研究结果可能有不同的读者和出版商，数据集可以制作并存放于不同的存储库。每个数据集本身可包括各种数字文件，它们需要不同的访问权限和/或许可条件。您需要计划如何将它们打包成集合，同时考虑您预期存储库存放数据的术语和条件。如果您需要存入不同种类的产生以提供不同的服务，您可能需要计划将它们组织在多个数据集合中。

本指南使用的“长期”指“超出您的研究项目期间”。或者，如果研究数据有助于参考集合或纵向研究，其长期价值将定期评估，例如，每3年评估一次。结合您所在机构或资助者的科研数据管理策略，该指南对数据应保存多长时间，提供具体的指导意见。DCC还对资助机构提供政策指导<sup>[2]</sup>。如果您的项目中存在，其他可以帮您评估保存什么数据的消息来源，包括研究机构构想的生产数据管理计划，这些可以识别数据的长期使用可能性<sup>[3]</sup>。虽然可能有影响评估结果的其它途径，但对研究产出长远目标存在一些想法。

## 第一步：确定数据能够实现目标

考虑目标或“重用”创建或收集的数据可以超出其研究范围以外。满足以下7条之一，即可以说明保存的数据可以进行长期访问。许多（但不是全部）涉及超出您的研究小组范围访问，您至少有一次机会初次使用。

验证：使别人能遵循产生所公布结果的过程，并可能复制或验证这些结果

进一步分析：增加对数据进一步分析的机会，例如，使用新的方法，与其他来源的元数据分析整合，可以通过新的合作或第三方分析

建立学术声誉：可发现的数据具有较大的知名度，可以提高已出版结果的引用率

社区资源开发：向已知用户组公布数据资源价值，例如，一个引用数据集、方法试验台，或领域数据库

进一步出版物：数据条款出版物<sup>[4]</sup>将促进学术交流，讨论有关数据管理问题或数据在您所在领域重用

学习与教学：在学习/教学或公众参与资源嵌入数据，以提高其交互性，吸引用户参在学习或研究



私人使用：更容易找到几年内的数据集，以挖掘其他潜在用途

原因 1 和 2 与资助机构的政策目标一致，因为这些机构通常侧重于确保发表的研究成果的完整性，并最大限度地提高数据的投资回报。资助者主要关心的是图表背后的数据<sup>[15]</sup>，但对数据承担责任的是研究人员和那些直接支持它们转化为有意义的政策指南的人。合同和其他法律义务也可能会发挥作用。我们重新回到“哪些数据必须保存”这个问题上。

为了帮助您作决策，您可以针对需要类型的数据（如下表 1<sup>[6]</sup>所示），匹配与您研究最相关的重用目的。

重用情况	保护方针
进一步出版物	引用其它文档的引用数据
学习 & 教学	样本源 & 集成数据的分析脚本
验证	引用数据以及分析脚本
进一步分析	所有源数据以及用来采集的软件

表 1、示例保存指南

## 第二步：确定必须保存的数据

关于“必须”保存什么数据的决定，通常依赖于数据创建者的优先级，即该数据如何有价值，能满足上述确定的目的，考虑到长期保存的成本。但这一决定还需要考虑到法律、监管或政策合规问题。在决定保存什么数据时，关注最多的是，数据是否应该公开或限制访问，在什么条件下可以访问，并确保不遵守规定的风险得到化解。

在这一步中考虑的问题基本如下，以帮助确定保存的数据。从机构的科研数据管理服务中寻求进一步的建议，或类似的技术支持人员，例如，记录管理器。保存或者处理数据之后，如果您不确定风险是否得到最好的解决，在任一种情况下，确保安全的完成这一步。

### 有科研数据政策因素推动保存数据吗？

英国研究理事会的科研数据政策原则强调的是“……具有确切的长期价值”的数据应该保存。<sup>[17]</sup>期刊、学术团体和专业协会积极参与确定这在各个学科的含义。关于“必须”存什么数据的决策，将需要考虑任何相关的出资人或机构的政策。<sup>[18]</sup>但究竟什么才算“具有确切长期价值”的数据，将基于创作者对该数据的深度认识。所以，您必须保存该数据最基本的指标是，如果以下问题，您的回答是“是”：

将文章的支撑数据提交到杂志，有政策要求它可供使用吗？





通过 RCUK 资金产生的数据能支撑公布的研究成果吗？

这里的“是”表明您要保存“图表背后的数据”（如上所述）。下面的步骤 3 给出了更多的帮助，指导找出可能具有“长期价值”的数据。

#### **法规要求数据可用吗？**

这里的主要问题是：

需要保存数据是否遵守信息自由或环境信息的规定？

有规定要求数据需要保存的数据作为研究记录的一部分（如健康或安全原因）而保存吗？

如果研究已经完成，且与之有关的数据仍然可用，涵盖了信息自由和环境信息的法律法规要求所提供的科研数据在请求情况下可访问。这意味着，应根据信息安全分类法案（例如，公共访问/内部/机密/秘密<sup>[19]</sup>）清楚地识别被保存的任何数据。对于可能适用于数据的任何豁免的一般指导可咨询信息专员办公室（ico.gov.uk）和苏格兰信息专员

（www.itspublicknowledge.info）网站。

#### **还有其他法律或合约因素吗？**

如果该研究对公共政策有影响，它涉及到商业合作伙伴，或具有附带应用潜力，这些都有可能。

数据是否提供商业价值的信息，抑或是在专利申请中使用到？

合同条款和条件状态是否暗示数据必须保存？

是否有理由相信该数据可能在公众查询或警方调查，或任何可能受到法律挑战的报告中使用？

#### **它是否包含与重用宗旨无关的个人数据？**

数据保护法定义了个人数据，并设置了标准，以确定它应予以保存的时间，以及存储和处理的要求。如果以下问题，您的回答是“是”，下一步就是遵守现有信息专员办公室的指导，包括在需要的情况下匿名化数据<sup>[20]</sup>。英国数据存档还提供指导，并设立一个安全的实验室，允许在严格管制下，个人资料用于学术研究<sup>[21]</sup>。

无论数据是独立的或可链接到另一个数据组，该数据是否包含直接识别个人的细节，或可以被用来推断他们的身份？

您所在机构的伦理委员会是否允许保存数据以用于进一步研究？

同意协议是否允许数据按您所设想的目的被重复使用？



资料当事人对他们数据的归档是否知情同意？

如果是这样，坚持自己同意的任何条件是否可行，例如匿名化数据的任何承诺？

这些数据是否可以安全地存储，并积极管理，以符合公认的信息安全标准，如 ISO27001？

以上述任何问题，如果您的回答是“不是”，您可以获得机构的科研数据管理服务，或得到记录管理者的帮助，以解决任何问题。

## 第三步：确定应该保存的数据

铭记您之前确定的潜在重用目标，考虑标准和以下问题，以决定哪些数据应予以保存，并明确保存的原因。作为数据被保存的一般规则，如果您已经确定了合规性原因，或者依据如下任意两个标题（标准），至少有一个问题，您可以回答“是”。

勾选任何您所期望的数据，只要这可以估计。您可以根据长期目标，选择数据需要满足的不同标准，以及您对这些数据价值如何满足这些目标的确定性。

### 这就够了吗？

说明：有没有足够的信息，例如，最新的数据管理计划，有关数据是什么，如何采集，为何采集，以及如何处理以评估其质量和实用性，是否满足您所确定的目标？

质量：就完整性、样本大小、准确性、有效性、可靠性、代表性，或领域中任何其他标准而言，数据质量足够好吗？

### 很有可能是需求吗？

已知用户：有用户正在等待此数据吗，或者有过去需求的证据吗，例如，是否向既定资源或系列中添加值？

建议：出资者、学会/专业协会或相应研究领域的机构是否建议共享此类型的或此研究主题的数据？

整合潜力：数据是否描述了其他领域中适合标准化的术语或词汇，如地理位置和时间段？

声誉：数据是由一个研究小组或项目产生的吗，其原创性、重要性和严谨性得到的评价很高吗？数据可用能显著增加该团体或项目的声誉吗？

上诉：数据具有广泛的吸引力吗，例如，涉及到具有里程碑意义的发现，新研究过程中的显著发现，或者国际政策和社会问题？



### 复制它有多难？

不可复制：数据复制困难/昂贵（或者在不可重复观测情况下不可能进行数据复制）吗？

### 进一步使用是否存在任何障碍？

清楚：数据是否根据其灵敏度分类，并不受隐私/伦理、契约、许可证或著作权条款和条件限制，公众可获取和重用？正常的研究领域是否有任何限制？

开放格式：重用统一格式的数据是否需要许可费或专有软件或硬件？

独立：如果任何专门的软件/硬件需要使用数据，这是广泛用于研究并且容易获得的吗？

### 这是唯一的副本吗？

唯一：这是数据唯一且最完整的副本吗？

危险：数据被存储在不能保证长期储存的地方吗？

## 第四步：权衡成本

此步骤有助于考虑不同经济情况下的数据保存活动。考虑数据管理成本对您的研究承诺和您组织预算的限制影响，这是很重要的。如果您最近做了这些，并且以下几个问题，都可以毫不含糊的回答“是”，您可以跳过此步骤。

资金是否可用来支付研究过程中所产生的数据管理成本（包括那些准备归档的数据）？

资金是否可用来支付超出研究期限的任何存储和监管费用？

您可以使用本部分估算任何时间短缺或预算数据管理成本<sup>[22]</sup>。已经产生的任何费用将计算在不同的经济情况下，保存数据的“价值”方面，同时，任何短缺将不利于它。关于如何履行“必须保存”类别下数据的承诺，您所在机构的科研数据管理服务研究办公室，图书馆或 IT 服务也许能够提供建议。

使用下面的标题，或数据管理计划中的任何费用类别，估计已经花费的工作人员时间、设备/硬件，或软件和服务费，以及仍需花费多少时间、金钱在这些类别上。

该表只用于自己的目的（数字并不需要向任何人透露，任何人都不会以其他方式访问它们）。它有两个目的：首先，帮助识别您数据积累的价值，其次，确定任何您可能需要寻求外部帮助以规避风险的地方，尽管这些价值不可能实现。



	到目前为止的花费	完成任务所需预算的花费	可能不足的地方?
<b>创作、采集&amp;清洗</b>			
创建合适的同意书，并同意数据共享			
数据传输或从站点、媒体或工具转录			
描述和文档			
验证、检查或清洗			
格式设置和文件组织			
纸或物理对象的数字化			
<b>短期存储&amp;备份</b>			
项目持续时间内所有工作数据的存储空间			
项目期间所有数据的备份			
<b>短期访问&amp;安全</b>			
为外部合作者或参与者提供访问和身份验证			
对数据进行在线和物理保护，防止未经授权的访问或披露			
<b>团队沟通&amp;发展</b>			
数据管理会议			
联机协作、虚拟的研究环境			
数据管理培训			
<b>保存&amp;长期存取</b>			
版权所有、许可			
分类数据的敏感性或匿名的个人数据（如果需要）			
准备存档，转换为开放文件格式			
元数据引用、发现和重用			
数据存储费用			
长期存储成本			
<b>工作时间（个人时间）</b>			
<b>设备（£）</b>			
<b>软件或服务收费（£）</b>			



## 第五步：完成数据评估

最后一步是将权衡价值和仍可能产生的任何费用，考虑到长期目标、您识别的特质、已经投入的时间和金钱以及暂时无法保存任何“必须保存”类数据的风险。填写一个表格（如下图所示），有助于您做到这一点。该表可以用于与其他参与您研究的人员交流，或者与您的机构的支持服务人员或所选数据存储库交流。

完成数据评估表，总结步骤 1 到 5 的结果，如下所示：

在第 1 步识别每一个您认为具有潜在重复使用目的（例如，前面的表 1）的数据集及其创建者

从第 1 步中确定相关的再利用目的

在第 2 步中，在“价值”列中插入标题相关的关键字，即，任何合规性原因，无论是保存数据，或者处理它；在第 3 步中，标明数据可能具有重用价值的原因

依据资金和支持是否满足成本，在“预算短缺风险”一栏，标注高，中，低

将您的决定记录在“保留？”一栏。如果有帮助，使用“MOSCOW”的缩写，即，必须、应该、可能或不保存。如果决定是“不保存”，那么您该单元格中注明不保存的理由

### 项目数据评估

数据采集	重用目标	价值	预算不足	保存吗？
标题、创建者和说明			的风险	MOSCOW
1) DCC 2014 RDM 调查结果				
&分析				
Angus Whyte, Diana Sisu	验证	说明、质量清洗	低	M
针对参与科研数据管理决策的	再分析	开放格式		
英国高等教育机构经理的调查	可见性	已知的用户		
问题、文档、匿名响应与分析				
2) DCC 2014 RDM 调查未匿名响应				
Angus Whyte				
参与科研数据管理决策的英国	N/A	一致性	低	W



高等教育机构经理的未匿名调

查结果

3) 等

所需采取的行动（引用表中的行）

记录分析中使用的源和变量（1）

电子表格中的标签图表（1）

不保存数据的理由（“不保存”）（引用表格中行）

当向受访者（2）介绍概况时，将删除未匿名的调查结果

## 科研数据保存策略指南

- 数字保存中心（2014 年）出资者的数据政策概述。可访问：  
[www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies](http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies)
- 联合监委会讯网（北达科他州）高等教育行业分类计划和记录保存时间安排。可访问：  
[bcs.jiscinfonet.ac.uk/he/](http://bcs.jiscinfonet.ac.uk/he/)
- 约翰·霍普金斯大学（2008）“获取和保存科研数据材料的政策”：可访问：  
[jhuresearch.jhu.edu/Data\\_Management\\_Policy.pdf](http://jhuresearch.jhu.edu/Data_Management_Policy.pdf)
- 伦敦大学国王学院（2012 年）“应该将科研数据保存多久”。可访问：  
[www.kcl.ac.uk/library/using/info-management/rdm/res-guide.aspx](http://www.kcl.ac.uk/library/using/info-management/rdm/res-guide.aspx)
- 自然与环境研究委员会数据价值清单（2012 年）。可访问：  
[www.nerc.ac.uk/research/sites/data/dmp.asp](http://www.nerc.ac.uk/research/sites/data/dmp.asp)
- 西北大学（2012）的科研数据：所有权、保存和访问。可访问：  
[www.research.northwestern.edu/policies/documents/research\\_data.pdf](http://www.research.northwestern.edu/policies/documents/research_data.pdf)
- 布里斯托大学（2014 年）关于记录保存和归档研究需要研究赞助者研究的治理指导意见。可访问：  
[www.bristol.ac.uk/red/research-governance/practice-training/archiving.pdf](http://www.bristol.ac.uk/red/research-governance/practice-training/archiving.pdf)
- 布里斯托尔大学的科研数据服务（2013 年）。科研数据评估指南 V3。可访问：  
[data.bris.ac.uk/research/sharing-and-publishing/](http://data.bris.ac.uk/research/sharing-and-publishing/)



## 注释

[1] Whyte A. and Wilson A. (2010) How to Appraise & Select Research Data for Curation. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: [www.dcc.ac.uk/resources/how-guides](http://www.dcc.ac.uk/resources/how-guides)

[2] Jones, S., Pryor, G. & Whyte, A. (2013). How to Develop Research Data Management Services - a guide for HEIs. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: [www.dcc.ac.uk/resources/how-guides](http://www.dcc.ac.uk/resources/how-guides)

[3] Details of Janet Cloud Services available at: [www.ja.net/products-services/janet-cloud-services](http://www.ja.net/products-services/janet-cloud-services)

[4] A DCC Checklist for Evaluating Data Repository Services will be available 2014

[5] Further information at [www.dcc.ac.uk/tailored-support](http://www.dcc.ac.uk/tailored-support), or contact [info@dcc.ac.uk](mailto:info@dcc.ac.uk)

[6] Lists of data repositories are maintained by Re3data ([www.re3data.org/](http://www.re3data.org/)) and Databib ([databib.org/](http://databib.org/)) The University Library may be able to offer more discipline-specific guidance on data repositories.

[7] Whyte A. and Wilson A. (2010) How to Appraise & Select Research Data for Curation. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: [www.dcc.ac.uk/resources/how-guides](http://www.dcc.ac.uk/resources/how-guides)

[8] Natural and Environmental Research Council Data Value Checklist (2012). Available online: [www.nerc.ac.uk/research/sites/data/dmp.asp](http://www.nerc.ac.uk/research/sites/data/dmp.asp)

[9] University of Bristol Research Data Service (2013). Guide to Research Data Evaluation v3. Available online: [data.bris.ac.uk/research/sharing-and-publishing/](http://data.bris.ac.uk/research/sharing-and-publishing/)

[10] C.L. Borgman (2015). Big Data, Little Data, No Data: Scholarship in the Networked World. MIT Press

[11] Adapted from: Peter Burnhill, Muriel Mewissen & Adam Rusbridge (2014) ‘Where data and journal content collide: what does it mean to ‘publish your data’?’ Presented at ‘Dealing with Data Conference’, 26 August 2014, University of Edinburgh Library. Available at: <https://www.era.lib.ed.ac.uk/handle/1842/9394>

[12] DCC Overview of Funders’ Data Policies. [www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies](http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies)



[13] Some funders do not ask specific questions about this in their pre-award guidance on preparing a Data Management Plan. If a DMP exists but does not include any envisaged long-term uses of data you could consider updating the DMP based on your appraisal of what should happen.

The [DMPonline](#) online tool helps you to keep an up-to-date DMP throughout your research.

[14] A data article is a peer-reviewed description of ‘what, why and how’ aspects of your data collection and may be based on a Data Management Plan. University of Edinburgh Data Library maintains a list of data journals available

at: <https://www.wiki.ed.ac.uk/display/datashare/Sources+of+dataset+peer+review>

[15] See, for example, ‘Data policy for the Shelf Sea Biogeochemistry Programme’. Available at: [www.uk-ssb.org/data/policy/](http://www.uk-ssb.org/data/policy/)

[16] Adapted from Akopov, Z. et al (2012) Status Report of the DPHEP Study Group: Towards a Global Effort for Sustainable Data Preservation in High Energy Physics. Available

at: [arxiv.org/abs/1205.4667](http://arxiv.org/abs/1205.4667)

[17] RCUK (2011) Common Principles on Data Policy. Available

at: [www.rcuk.ac.uk/research/datapolicy](http://www.rcuk.ac.uk/research/datapolicy)

[18] E.g. PLoS Data sharing policy [www.plosone.org/static/policies#sharing](http://www.plosone.org/static/policies#sharing) and the JoRD project at: [jordproject.wordpress.com/project-data/social-science-journals-that-have-a-research-data-policy/](http://jordproject.wordpress.com/project-data/social-science-journals-that-have-a-research-data-policy/)

[19] See e.g. Wikipedia

at: [en.wikipedia.org/wiki/Information\\_security#Security\\_classification\\_for\\_information](http://en.wikipedia.org/wiki/Information_security#Security_classification_for_information)

[20] Information Commissioner’s Office (n.d.) Anonymisation. Available

at: [ico.org.uk/for\\_organisations/data\\_protection/topic\\_guides/anonymisation](http://ico.org.uk/for_organisations/data_protection/topic_guides/anonymisation)

[21] UK Data Service (n.d.) Secure Lab. Available at: [ukdataservice.ac.uk/use-data/secure-lab.aspx](http://ukdataservice.ac.uk/use-data/secure-lab.aspx)

[22] RCUK has given general guidance on what costs may be included in grants, at [blogs.rcuk.ac.uk/2013/07/09/supporting-research-data-management-costs-through-grant-funding/](http://blogs.rcuk.ac.uk/2013/07/09/supporting-research-data-management-costs-through-grant-funding/)





# 鸣谢

非常感谢以下人对早期草案提出的意见：英国伦敦政治经济学院的 Laurence Horton、伦敦大学卫生和热带医学院的 Gareth Knight、布里斯托大学的 Debra Hiom、拉夫堡大学的 Helen Young 和 Susan Manuel。

