

RESEARCH DATA MANAGEMENT POLICY

Document Title	RESEARCH DATA MANAGEMENT POLICY
Version	1.1
Document Review Date	September 2022
Document Owner	Director of the Research Office
Revision Timetable / Process	Regular review of these aims is required to ensure they evolve to support College needs. The document will be reviewed at least biannually.

Contents

1	Introduction
2	Definitions
3	Principal Investigator Roles and Responsibilities
4	Faculty Roles and Responsibilities
5	Library Roles and Responsibilities
6	ICT Roles and Responsibilities
7	Research Office Roles and Responsibilities

1. INTRODUCTION

Imperial College is committed to promoting the highest standards of academic research, including excellence in research data management. The College is developing services and guidance for its academic and research community in order to foster best practice in data management and to facilitate, by way of a robust digital curation infrastructure, free and timely open access to data so that they are intelligible, assessable and usable by others. The College acknowledges legal, ethical and commercial constraints with regard to sharing data and the need to preserve the academic entitlement to publication as the primary communication of research results.

1.1 Description

Application: The aims in this document apply to all data generated through research undertaken by Imperial College staff and students solely or in collaboration with others. This includes, but is not limited to, research sponsored by grants or contracts awarded to Imperial College and where Imperial is a partner on awards held by other institutions.

This framework will aid Imperial's transition to best practice in Research Data Management and to compliance with funders' requirements.

1.2 Aims

The aim of this document is to provide a framework within which the College, through its academic and research community, and support services can adhere to the rules and mandates of its funding and regulatory bodies and to ensure that research data are maintained and preserved as identifiable, retrievable and reusable assets as part of the overall research project cycle. This is achieved by allocating roles and responsibilities across the institution and will be reviewed once the College has finalised its analysis of requirements.

This document supports academics retaining the choice of where to deposit their data and their entitlement to publish their findings first within the constraints specified by the funder (where one exists) though also taking account of the socioeconomic benefits of promptly sharing the outcomes of research.

Management of data generated from a research grant or contract must also comply with the policies of the relevant grant funders. Where these conflict with the College policy, the policies of the funders take precedence unless specified.

Management of data obtained under licence must also comply with relevant licence agreements. The terms and conditions of any licence agreements take precedence over this policy.

2. DEFINITIONS

- 2.1 **Research Data:**
Data that are collected, observed, generated, created or obtained from commercial, government or other sources, for subsequent analysis and synthesis to produce original research results. For the purposes of this policy, this is limited to digital data.
- 2.2 **Shareable Research Data:**
The subset of Research Data that must be preserved for long-term storage and shared.
- 2.3 **Data Management Plan (DMP):**
Plans typically state what data will be created and how, and outline the plans for protecting against loss, sharing and preservation, noting what is appropriate given the nature of the data and any restrictions that may need to be applied. More details are available on the College RDM webpages¹.
- 2.4 **Data Repository:**
A place where data may be stored and made available. This includes but is not limited to a College Repository, a discipline specific repository, a funder repository or a commercial system.
- 2.5 **Funder:**
An organisation that provides a financial contribution to support a research project.

¹ <http://www3.imperial.ac.uk/researchsupport/rdm/applying/dmp>

- 2.6 Principal Investigator (PI):
The person ultimately responsible for the research project; as such the PI takes responsibility for the intellectual leadership of the research project and the overall scientific and financial management of the research.
- 2.7 Digital Object Identifier (DOI):
A reference that enables permanent identification and location of digital objects, such as datasets and journal articles. This means a reference to a dataset elsewhere (e.g. in the journal article it supports or in the College's catalogue) is not going to break or become outdated as a web address might do.

3. PRINCIPAL INVESTIGATOR ROLES AND RESPONSIBILITIES

- 3.1 Principal Investigators have overall responsibility for the effective management of research data generated within or obtained for their research, including by their research groups. The Library and ICT will provide training, guidance and services to support PIs in the following roles and responsibilities.
- 3.2 Principal Investigators will provide a data management plan (DMP) for all research grant and contract proposals from the date mandated by the funder where applicable, and in any case by 1st January 2016. This should be in the DMP format provided by the funder if one has been defined, and otherwise the College standard format² should be used. See 5.1 and 5.2 for details of the support provided by the Library for this task.
- 3.3 During the course of a research project, the PI must ensure their own and their group's data (and use reasonable endeavours to ensure their collaborators' research data) are stored and described in ways that will minimise the risk of loss and maximise the likelihood of future discoverability and re-use.
- 3.3.1 Minimum standards of data safety to be followed are published by ICT³.
- 3.3.2 ICT will provide shared network storage that complies with these standards (see 6.1).
- 3.3.3 Research data are to be accompanied by adequate metadata and/or documentation to facilitate its identification and re-use. See 5.1 and 5.6 for details of the support provided by the Library for this task.
- 3.4 Principal Investigators are responsible for identifying and preserving all their research data (whether shareable or not) that would be required to validate published research findings in the future.
- 3.5 Principal Investigators are responsible for determining what their shareable research data comprises in view of the considerations in 3.6. The minimum requirement is to share all relevant data to support and underpin published findings including e-theses. Any funder requirements must also be met.
- 3.6 Principal Investigators must deposit their shareable research data in a publicly-available repository of their choosing no later than the time of publication of the findings. Where the funder requires deposit earlier, the timeframes specified by the funder take precedence. Where the funder requires the data to be deposited in a different repository, the funder's requirements must also be followed although deposit in multiple repositories is acceptable.
- 3.6.1 Publication – the PI's entitlement to be the first to publish based on data they have generated must be preserved. If the PI will not be able to publish their findings by the funder's deadline for data sharing, the PI should request an extension of the embargo period from the funder.
- 3.6.2 Copyright – the PI will not share any data that would infringe any right of title or copyright. Where third party research data are used, access conditions should be sought and followed and the source of that data will be acknowledged where required.

² <http://www3.imperial.ac.uk/researchsupport/rdm/applying/dmp>

³ <http://www3.imperial.ac.uk/secretariat/collegedgovernance/provisions/policies/infosystems/guidelines/copguide2>

RESEARCH DATA MANAGEMENT POLICY

- 3.6.3 Licences – the PI is entitled to select the licence under which they publish their data provided the chosen licence permits sharing and re-use in accordance with the appropriate funder’s terms and conditions (if any exist) and this policy.
- 3.6.3 Intellectual Property Rights (IPR) – the PI will not share any data that would infringe the IPR of others.
- 3.6.4 Commercial Interests – the PI is entitled to delay sharing of their data where early release may harm commercialisation opportunities arising from their research until appropriate protection is in place.
- 3.6.5 Privacy – the PI will take all appropriate steps to ensure the anonymity of any personal data prior to sharing. For patient data, further requirements are detailed in “Data Protection Policy – Code of Practice 2: Handling of Patient Data”⁴ and must be met.
- 3.6.6 Ethical Concerns – the PI should bear in mind potential ethical considerations in making the decision to share their data.
- 3.6.7 Cost Effectiveness – where computer-generated data may be reliably recreated at a cost less than that of storing raw output data, then the inputs and human-readable outputs of the relevant programme may be stored instead along with a reference to or copy of the software version used.
- 3.6.8 Legal Safeguards – where data are shared with restrictions, e.g. if they identify human subjects, they would normally be stored in repositories located in the EU to prevent their full disclosure via Freedom of Information and Data Protection safeguards. See 5.6 for details of the support provided by the Library for this task.
- 3.7 If software is developed as part of a research project, Principal Investigators must archive the particular version of the software used to generate or analyse the data in a repository and inform the Library of its location, taking account of the points raised in 3.5 above. Principal Investigators are encouraged to follow the Sustainability and Preservation Framework of the Software Sustainability Institute.⁵
- 3.8 Principal Investigators will inform the Library of the location of their shared research data and specified metadata including the Digital Object Identifier (DOI) for that data package at the time of deposit.
- 3.9 Principal Investigators will ensure that all journal articles include a means to obtain the supporting data by referencing the appropriate DOI.
- 3.10 Appropriate funding for RDM should be requested at application stage where this is an eligible cost item allowed by the funder [TBC based on most effective means of covering these costs under FEC methodology].
- 3.11 Principal Investigators will ensure that shareable research data are made publicly available for 10 years after publication or for the minimum period required by the funder if longer than 10 years.
- 3.12 If the PI leaves Imperial College or circumstances (such as death, incapacitation or leaving academia) mean the PI is no longer able to fulfil their responsibilities then:
 - 3.12.1 If there is an Imperial College Co-Investigator, they inherit the responsibilities of the PI;
 - 3.12.2 Otherwise, if the PI and the award move to another institution, the responsibilities follow the PI;
 - 3.12.3 Otherwise, whom so ever is appointed by the department to complete any of the PI’s requirements for the research project will become the PI from an Imperial College perspective and as such inherit the responsibilities outlined in this section.

⁴ <http://www3.imperial.ac.uk/legalservicesoffice/dataprotection/codesofpractice/cop2>

⁵ <http://www.software.ac.uk/resources/guides/sustainability-and-preservation-framework>

4. FACULTY ROLES AND RESPONSIBILITIES

- 4.1 As part of the routine approval of grant applications, Faculty Research Services and Contracts / Joint Research Office will ensure that all research grant and contract applications include a DMP and that this is attached to the relevant record in the appropriate College management system. They are not, however, responsible for ensuring the content is appropriate for the research.
- 4.2 Faculty Research Services and Contracts / Joint Research Office must ensure that data management responsibilities reflected in this policy are also applied to research partner organisations and the data they generate on research undertaken in collaboration with the College in research partner contracts.
- 4.3 Faculties are responsible for ensuring their PIs adhere to their obligations as detailed in this policy.

5. LIBRARY ROLES AND RESPONSIBILITIES

- 5.1 With the support of ICT, the Library will provide training and advice on all aspects of research data management including minimum standards of metadata description, the writing of DMPs, copyright and licences.
- 5.2 The Library will provide guidance and support towards the generation of generic DMP templates to aid PIs, and will maintain a library of the latest recommended version(s).
- 5.3 The Library will maintain a College data catalogue using the information provided to them.
- 5.4 The Library must be able to reassure the College that it has confidence in the longevity of any external repository used for data sharing purposes. Where the Library is not confident in the longevity of a repository, Imperial data stored within it may be copied into an Imperial-approved data storage facility for safe-keeping.
- 5.5 The Library will publish sufficient appropriate metadata describing every shared dataset on a publicly-available catalogue for that dataset to be discovered, accessed, understood and re-used. Where research is externally funded, the funder of the grant that generated the data and the funder's grant reference number must be included within the metadata.
- 5.6 The Library will provide assistance in determining the physical location of discipline repositories where a dataset should not be shared without restrictions.

6. INFORMATION AND COMMUNICATIONS TECHNOLOGY (ICT) ROLES AND RESPONSIBILITIES

- 6.1 ICT will make available appropriately resilient storage for research data during their collection and throughout the data lifecycle.
- 6.2 ICT will provide technical advice and support data management including data storage, backup and archiving.
- 6.3 ICT will ensure a system is available for storing and sharing of research data for which there is no appropriate external publicly-available repository.
- 6.4 ICT will support the Library in the technical aspects of development and provision of training and advice.
- 6.5 At Library's request, ICT will provide advice on if a repository is suitably robust and may need to facilitate copying of the data from this external storage to an approved storage facility should this not be the case.
- 6.6 ICT will ensure software to support effective research data management is available.

7. RESEARCH OFFICE (RO) ROLES AND RESPONSIBILITIES

- 7.1 RO is responsible for quality assurance of the College's research data curation processes.
- 7.2 RO is responsible for ensuring stakeholders understand their obligations.
- 7.3 RO is responsible for ensuring funder and College requirements are communicated to stakeholders.