

**University of
Cambridge
HSPS IT**



Data Management Policy (for Research Data)
[UNRATIFIED]
K Bradley, HSPS Faculty

24/09/2018

Author: kjb53@cam.ac.uk, kjb53@cam.ac.uk,

Data-Management-Policy-(forResearchData)-20180924.docx

1 Data Management Policy (for Research Data) [UNRATIFIED]

1. Introduction

This document is issued by HSPS IT, and is a guide for research staff looking to choose the most suitable data storage service. This document is intended to be of particular use to individuals writing research funding proposals who wish to include costings for data storage for the entire project lifecycle. This document is not comprehensive, but does include pointers for the preservation of your research data, technical and legal requirements, solutions and costs. This policy covers all electronic data formats, including visual data (photos, videos), documentary data (written documents, whether scanned or original) and audio data (voice recordings).

This document has not yet been ratified by the HSPS IT Committee but ratification will be sought during 2018.

You will need to consider the suitability of any storage service for the storage of the data you control (or 'own'). You will need to consider the suitability of the storage service through the full lifecycle of your project and beyond. See the Data Storage Guidelines on the HSPS Intranet for more details of your responsibilities (<https://faculty.hsp.cam.ac.uk/hsp-faculty/faculty-staff/intranet>). If you are in any doubt of where it is practical, recommended, legal and possible to store your data, please contact your research grants administrator in the first instance.

In addition, the Office of Scholarly Communication offers a number of services including a reviewing service for Data Management Plans (with 7 days' notice). The OSC can be found here: <https://osc.cam.ac.uk/>.

The Data Management Team (which includes personnel from OSC and other parts of the University) also provides guidance on storing and managing research data and about proposed Data Management Plans. The Data Management Team's website can be found here: <https://www.data.cam.ac.uk/>

The Research Operations Office (ROO) offers a number of resources, including the University's Policy on the Ethics of Research Involving Human Participants and Personal Data. ROO's policy can be found here: <https://www.research-operations.admin.cam.ac.uk/managing-research-projects/research-ethics>.

See also the HSPS Intranet (<https://faculty.hsp.cam.ac.uk/hsp-faculty/faculty-staff/intranet>) for guidance on Data Protection and Data Retention.

Additional guidance can be obtained from your funding body, your departmental administrator, the University's Information Compliance team, the University's Data Protection team or the HSPS IT team.

2. GDPR and Storage Services

The General Data Protection Regulations (GDPR) came into force in May 2018 and form part of the framework within which particular types of data, including aspects of research data, are safeguarded. Some data storage services (especially some public services) are not suitable for certain types of data (especially confidential or personal data). As mentioned already, you will need to consider the suitability of these services for the data you control or own. The following information is collated and interpreted from various sources (including a UIS presentation by Maddy Taylor, 14/2/18) *:

1. Some types of personal data (eg 'strictly confidential information') should be considered as unsuitable for storage on a cloud service of any sort (whether that 'cloud' is within Cambridge or not).
2. 'General personal information' (name, email address [or other 'online identifier'] and IP address) is currently understood as being generally suitable for storage on a 'Cambridge cloud' service. Such storage services include the following:
 - a. Cambridge versions of cloud services listed on the relevant UIS page (<https://help.uis.cam.ac.uk/service/supporting-research/servers-data-storage-and-backup/data-storage/individual-storage>),
 - b. Other approved platforms including Cambridge's Qualtrics 'EU silo' (<http://cambridge.eu.qualtrics.com/>).
 - c. Other University Information Services (UIS)/University of Cambridge (UoC) services which reside wholly in the UoC, including Hermes, CUFS etc.
3. Other public cloud services (including services not listed above or otherwise not approved by the University, including the public versions of Google, Microsoft and DropBox) are generally not encouraged as being suitable for the storage of Level 2 data (and are even less suitable for Level 3 data).
4. More guidance on storing personal data can be found on the HSPS Intranet (<https://faculty.hsps.cam.ac.uk/hsp-s-faculty/faculty-staff/intranet>) and the UIS website (<https://help.uis.cam.ac.uk/service/user-accounts-security/security/personal-data>).

Another consideration is that there are other reasons that data might be deemed unsuitable for storage in some UIS data storage services. The UIS states (on a page relating to Research Storage services : <https://help.uis.cam.ac.uk/service/supporting-research/servers-data-storage-and-backup/data-storage/research-data-storage/terms-and-conditions/security-guidance>), that the following data is also unsuitable for storage in some storage services (including RFS):

1. Data classified as Level 3 (see the 'security guidance' link above)
2. Patient Identifiable Data (including other identifiable data which is subject to the Clinical School's mandatory data security policy which can be found at <http://www.medschl.cam.ac.uk/research/information-governance/>)
3. Data that is subject to a specific contractual agreement that specifies a particular storage method (that is not research storage services).

3. Service Offerings

A number of data storage services are available. Some of the following data storage services are available to the public, and some services are currently deemed unsuitable (by the University) for storing University data. Some services (especially some of the publicly-available services) are not suitable for certain types of data (especially confidential or personal data). All details below are believed correct at the time of writing and based on current explanations of services (especially those marked '*').

In addition, if you are choosing (or using) a data storage service, you also need to ensure you have a suitable backup method in place. Some of the data storage services below include an automatic backup, but not all do. The backup method suitable for your data might not be provided by every data storage service. For instance, if you

accidentally delete a file (or folder), can you get it back? If you accidentally damage a file (or folder) can you restore it from a backup? If the answer to either of these questions is 'no' or 'I don't know', then you need to consider whether or not that data storage service is suitable for your requirements.

It is worth stating that Research Cold Storage (RCS) might be suitable as a backup solution.

Data storage services which are not included in this document are not recommended - if you think a service should be listed, please contact helpdesk@hsps.cam.ac.uk.

Service	Description	Costs (actual)	Status	More details
Institutional storage (small data)	For small-scale data (<500GB). Backups included.	available on request	Available but possibly being phased out	helpdesk@hsps.cam.ac.uk
Institutional storage (larger data)	For larger amounts of data, but generally being phased out. Backups included (with prior arrangement).	available on request	Not preferred due to hidden costs to Dept	helpdesk@hsps.cam.ac.uk
RFS (Research File Store, UIS)	Dedicated UIS service for Research File Storage (some backups included)	https://selfservice.uis.cam.ac.uk/	Preferred solution for immediate data access and storage of data for duration of project	https://www.hpc.cam.ac.uk/research-data-storage-services
RDS (Research Data Store, UIS)	Dedicated UIS service for Research Data especially for Linux (no backups included)	https://selfservice.uis.cam.ac.uk/	Preferred solution for large-scale data stored on High Performance Computing facility (HPC)	https://www.hpc.cam.ac.uk/research-data-storage-services
RCS (Research Cold Store, UIS)	Dedicated UIS service for Cold Storage of Research data (useful as a backup service)	https://selfservice.uis.cam.ac.uk/	Preferred solution for storage of rarely-used data for finite period (normally for the duration of project)	https://www.hpc.cam.ac.uk/research-data-storage-services

Google Drive (G-Suite@Cambridge) for individuals	Google's cloud-based file hosting service gives each member of the University free, unlimited file storage that has an individual file size limit of 5TB (some backups provided by Google).	see online details	Acceptable, but Team Drives are preferred (also, a separate backup method required as this service does not include automated backups)	https://help.uis.cam.ac.uk/service/supporting-research/servers-data-storage-and-backup/data-storage/individual-storage
Google Drive (G-Suite@Cambridge) for groups (Google File Stream)	as above, but for groups	see online details	Preferred solution, but a separate backup method required as this service does not include automated backups	https://help.uis.cam.ac.uk/service/supporting-research/servers-data-storage-and-backup/data-storage/individual-storage
Google Drive (public version)	unsuitable for University data	n/a	Not suitable (see link)	unsuitable for University data*
OneDrive for Business	Microsoft's cloud-based file hosting service, which is available to people who are eligible for a University Microsoft account. It offers 1TB of storage space (some backups provided by Microsoft).	see online details	Acceptable, but limited capacity and a separate backup method is recommended for use with this service	https://help.uis.cam.ac.uk/service/supporting-research/servers-data-storage-and-backup/data-storage/individual-storage
OneDrive (public version)	unsuitable for University data	n/a	Not suitable (see link)	unsuitable for University data*

DropBox for Business	Dropbox's paid-for cloud-based file hosting service, which is available to Cambridge students and staff at a heavily-discounted rate through the University using their @cam domain email address (some backups provided by DropBox).	see online details	Not preferred - free options exist and a separate backup method is recommended for use with this service	https://help.uis.cam.ac.uk/service/supporting-research/servers-data-storage-and-backup/data-storage/individual-storage
DropBox (Free)	unsuitable for University data	n/a	Not suitable (see link)	unsuitable for University data*
Apollo/Dspace	University of Cambridge repository (no backups included).	see online price list	Preferred solution for long-term storage of project data by completion of project	https://www.repository.cam.ac.uk/
ESDS	see website for details	see website for details	Preferred solution if acceptable to funding body	https://www.ukdataservice.ac.uk/
Archaeology Data Service	see website for details	see website for details	Preferred solution if acceptable to funding body	http://archaeologydataservice.ac.uk/
IFS (Institution File Service)	Proposed UIS service for Institutions	n/a	Not yet launched (as at Sept 2018) and not intended for research data.	Not yet launched (as at Sept 2018)