London School of Hygiene and Tropical Medicine

Background

The London School of Hygiene & Tropical Medicine (LSHTM) is a public research university that specialises in the field of public health and tropical medicine, and is a constituent college of the University of London. The main site is based in the Bloomsbury area of London, with two research units located in sub-Saharan Africa – the 'MRC Unit The Gambia at LSHTM', in Bakau, The Gambia, and the 'MRC/UVRI & LSHTM Uganda Research Unit' in Entebbe, Uganda.

The Research Data Management Service

The need for good data management infrastructure was first recognised in 2002 as part of a project to develop an institutional retention schedule for paper and digital records, and re-visited in 2009 when a Research Data Working Group (RDWG) was setup to review researchers' data management practices.

The LSHTM Research Data Management Service was established in 2012 as part of the Archives & Records Management Service. It was supported using funds provided by the Wellcome Trust Institutional Strategic Support Fund during 2012-15.

It was initially proposed that operation of the post-2015 service would be funded by charging costs to externally funded projects, and a draft charging model was developed. However, the RDM Steering Group took the decision that institutional support should be centrally funded and provided to researchers for free.

The service provides RDM support and training on how to:

- Write a data management plan
- Identify and calculate data-related costs
- Allocate resources for data collection and management
- Choose appropriate storage solutions
- Appraise research outputs for long-term retention
- Prepare research outputs for sharing
- Preserve digital resources
- Cite data and software

Although LSHTM is relatively small as an organisation, there are significant data management challenges. Research data often contains personal details of human participants from around the

STAFFING LEVELS

Total research-active staff   = 840

RDM staff = 1.0 FTE

Ratio of staff to RDM support = 840 : 1

MAJOR FUNDING SOURCES

- Gates Foundation
- Wellcome Trust
- UK Research Councils

RDM SUPPORT

- Research Data Manager (1 FTE)

world. As a result, many ethical, legal, regulatory & funding requirements apply to the data, some of which can result in conflicting obligations.

Making the service sustainable

Although the RDM service at LSHTM enjoys the security of institutional funding, it is staffed by one staff member who cannot support all researchers and projects to the same level. A three-stage triage process is used to prioritise where RDM support should be directed:

In addition to this triage process, thought has gone into optimizing the delivery of specific information and services.

> 1. Staff working on funded research are given priority access to the service
> 2. Existing expertise within the school are utilised when possible
> 3. Details of every RDM support query are logged, as an evidence base to plan future development work

Data Management Plans

LSHTM initially considered introducing a requirement that all research projects write a data management plan (DMP) and obtain sign-off from the RDM service at the pre-award stage. However, this was recognised as unsustainable and resource-intensive for both researchers and the RDM service.

A review of DMP criteria and assessment of the projected number of plans that would be produced resulted in a decision made that it should be a post-award requirement to write a DMP using the LSHTM institutional template in the following circumstances (see [6] for details):

1) LSHTM is the lead institution or responsible for managing research data
2) The project is externally funded, but the funder has not ask for a DMP to be written

Making the DMP a post-award requirement does raise the risk that the project has not allocated sufficient resources to data management, but this balanced approach makes the DMP review process easier to manage for an RDM service with limited resources.

Webpages & Guides

RDM support material was initially available on the LSHTM website, but following a site redesign has been moved to the institution's SharePoint instance and TOPdesk service management tool. PDF copies of the guides and tutorials are available in LSHTM Data Compass [7].

RDM support requests are monitored by the Research Data Manager to identify common factors and emerging themes. For example, a large number of RDM support requests were generated when PLoS updated its data sharing policy [3], resulting in the creation of a guide to understand requirements [4] and procedures to be following when sharing restricted data where access decisions must be evaluated by a Data Access Committee [5]. Guidance material is also added to the TOPdesk service management tool and dynamically presented to the user based upon words they type into a support request form. For instance, if 'dmponline' was typed into the subject or message body, they would be presented with a DMPOnline tutorial. This reduces the processing time needed for common RDM queries by helping the researcher to answer their question without submitting a request or, following request submission, enables a sample response to be easily copied into the reply and tailored to the specific question.

Training events and workshops

The analysis of RDM support requests is also used to inform event planning, and has led to the organisation of training on data management plans for specific funders, data sharing to meet journal requirements and, more recently, the implications of open science for research reproducibility. The format of training events has been refined over time, in response to user needs. For instance, the organisation of training events on a specific stage of the research lifecycle (planning, project start-up, finalising a grant), which bring together professional support staff with a role in these activities to provide training, as well as the use of online delivery methods for researchers based overseas.

Repository

LSHTM commissioned Cosector (formerly the University of London Computing Centre) to develop and host an institutional repository, LSHTM Data Compass [2], built on the Eprints platform developed by the University of Southampton. This repository is used to host research outputs, such as data, code and scripts, search strategies, research instruments and other resources intended for reuse. Hosting and development costs are covered by institutional funds.

To prioritise support and minimise curation costs, researchers are encouraged to use subject, funder, or other repositories where these exist. LSHTM Data Compass is referred to as "a home for homeless data", which can be used for resources that cannot be hosted elsewhere or require additional institutional support, e.g. the author requires additional guidance when making data access decisions or help with tailoring their data transfer agreement.

Costing RDM activities

The LSHTM RDM Policy [1] encourages researchers to add data-related costs to their grant application where permitted by their funder. This is often essential when purchasing equipment or acquiring other tangible resources, but researchers are often wary of costing time for activities associated with preparing data for sharing, such as anonymisation and documentation enhancement, fearing that it will make their funding bid appear uncompetitive. In cases where RDM activities are costed into a bid, it is often these costs which are the first to be removed when a funder asks for the budget to be reduced. As a result, researchers sometimes perform these tasks outside work hours.

Acknowledgements

Links

1. LSHTM Research Data Management Policy https://doi.org/10.17037/PUBS.00612422
2. LSHTM Data Compass - institutional data repository http://datacompass.lshtm.ac.uk/
3. PLoS data sharing policy http://journals.plos.org/plosone/s/data-availability
4. LSHTM PLos Data Sharing Policy guide http://blogs.lshtm.ac.uk/rdmss/plos-data-policy-overview/
5. LSHTM Data Access Procedures https://doi.org/10.17037/PUBS.00612422
6. Gareth Knight, (2015) 'Building a research data management service for the London school of hygiene and tropical medicine', Program: electronic library and information systems. 49(4): 424-439 https://doi.org/10.1108/PROG-01-2015-0011
7. LSHTM Research Data Management Guides http://datacompass.lshtm.ac.uk/666/