Plan Overview

A Data Management Plan created using DMPonline

Title: Pay now or wait a year: embargoes as a selective barrier to access

Creator: Liam Bullingham

Principal Investigator: Liam Bullingham

Data Manager: Liam Bullingham

Affiliation: Edge Hill University

Template: DCC Template

ORCID iD: 0000-0003-1777-5825

Project abstract:

In the absence of funds for 'gold' open access, publishers may allow universities to share the author's manuscript on a 'green' basis, with or without an embargo. The former means the work can be openly accessed typically after a period of 12-24 months. Such time gives publishers opportunity to build revenue, but creates inequalities. Firstly, there is a division between those who can afford a fee or can access a library subscription to benefit from the research and those that cannot; but there is also a divide between universities that have funds to pay publishing charges and less research intensive institutions that have no such fund. This case study considers journal articles published by Edge Hill University in 2019.

ID: 64609

Last modified: 10-01-2022

Grant number / URL: n/a

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Pay now or wait a year: embargoes as a selective barrier to access

Data Collection

What data will you collect or create?

A secondary data source will be the project's primary source of data. This will be collected from Pure, the Edge Hill University research information repository. This will be supplemented with additional information taken from the Sherpa Romeo website and also publisher website in some cases. All information is in the public domain.

This will be bibliographic data (article title, publisher, journal, embargo period, etc.) and will be stored in Microsoft Excel, in the .xlsx file format. Once the data is shared publicly in a data repository, a version will also be published as a .csv file to support long-term storage. I have chosen this format as it is widely used, can be exported from Excel and is non-proprietary.

Data visualisations (pie charts, bar charts) will be created based on the dataset. These will be stored as image files, most likely .png or .jpg.

In total, the data will not exceed 50MB.

How will the data be collected or created?

The data will be collected from the Pure, the Edge Hill University research information repository. It will be exported as an .xls file. An Excel workbook will form the sole file for the dataset, and this will be arranged into multiple worksheets. As the file will be stored in OneDrive, versioning will be maintained by the functionality in Office 365, so for example if the file becomes corrupted, a previous version can be accessed.

Documentation and Metadata

What documentation and metadata will accompany the data?

As the dataset will be shared and stored in Figshare, The standard Figshare metadata will accompany the data:

- author
- title
- category (ie academic discipline)
- item type
- group (place in the university hierarchy)
- description
- keywords
- direct links to related papers
- Creative Commons licence (CC BY)
- DOI

A short guide to the data will also be produced and shared alongside the dataset in Figshare. This will explain how the data was collected and refined, its value, and suggest possible applications for re-use

Ethics and Legal Compliance

How will you manage any ethical issues?

The project will be submitted to Edge Hill University for ethical approval. Unless approval has been granted, the data will not be shared.

There will be no research participants, no is the data sensitive in nature.

No ethical issues are anticipated as all data is already publicly available.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

The University's Research Information Policy notes that metadata on all content in the Pure portal can be accessed free of charge and records can be:

"Re-used under a Creative Commons licence in any medium in accordance with the terms of the specific licence, provided...Edge Hill University’s research information repository is mentioned".

As such, when the dataset for this project is shared, Edge Hill Pure will be acknowledged as the original source.

Storage and Backup

How will the data be stored and backed up during the research?

The dataset will be stored on my personal Edge Hill University OneDrive account and also privately in Edge Hill Figshare during processing. This means it will be kept in two separate cloud storage locations. There is more than sufficient storage capacity.

For OneDrive, Edge Hill University's IT Services will be responsible for backup and recovery, and for Figshare, Digital Science (the holding company) will be ultimately responsible. I act as the site administrator for Edge Hill Figshare. In the event of an incident, the data will be recovered from the primary or backup source, or as a last resort the report can be retrieved from Pure.

One processed, the dataset will be stored and publicly shared in Figshare. Both locations are GDPR compliant.

How will you manage access and security?

Access to the pre-processed data will be password protected, in line with its being kept on a personal Edge Hill University staff account. The data will not be encrypted however.

As the pre-processed data will be stored on Edge Hill Figshare, its safe transfer when moving it to being shared publicly will be ensured by its being housed within the same platform. If the dataset needs to be transferred from OneDrive to Figshare, this will be done via direct transfer between cloud systems, and a 'go between' version will not be stored locally on my personal computer.

The risk is substantially limited here as the data is already publicly available.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

The processed dataset will be of potential value for re-use so will be shared on Figshare, with the intention that it is kept for at least 10 years. In addition, the visualisations will also shared here.

The data could be potentially used to provide a comparison with further datasets, e.g. a comparison between publications in 2019 and 2020.

What is the long-term preservation plan for the dataset?

The dataset and visualisations will be shared on Edge Hill Figshare, with the intention that it is kept for at least 10 years.

Data Sharing

How will you share the data?

After the data has been processed (January 2021) it and the subsequent visualisations will be shared on the Figshare repository using a Creative Commons CC BY 4.0 license. The files will be organised under a single Figshare repository record, and a DOI will be reserved, meaning each version of the dataset will be saved too.

Potential users may find out about the data if I or other individuals share it via social media, email, etc. or if it is cited in any publications emerging from the project.

Are any restrictions on data sharing required?

No restrictions on data sharing are necessary, nor is a data sharing agreement.

Responsibilities and Resources

Who will be responsible for data management?

I (Liam Bullingham) am responsible for implementing the DMP and ensuring it is reviewed and revised. There are no consortium agreements or partner contracts to consider.

What resources will you require to deliver your plan?

I do not require access to additional expertise, hardware, or software. Similarly, no charges beyond institutional provision will be incurred.