

# Manage it locally to share it globally: RDM and Wikimedia Commons

This proposal is to link RDM with the open science movement via the Wikimedia suite of tools. We will invite experts from the Wikimedia Foundation to co-facilitate an editathon across the White Rose Consortium focussed on a specific subject<sup>1</sup> and featuring high profile academics from each of the three universities (Leeds, Sheffield and York).

## Research data lifecycle: leading by example

In order to encourage researchers to manage their research data effectively we have found it useful to 'start at the end'. That is, to begin with an example of a dataset that is properly curated in a repository before systematically describing the research data lifecycle in order to get to that point.

## What's in a repository?

There are a wide range of repositories for researchers to deposit their research data. These include institutional repositories (Research Data Leeds, Edinburgh DataShare), disciplinary repositories (The Cambridge Structural Database) and generic repositories (Zenodo, DataDryad) as well as data centres supported by research funders (e.g. NERC).

Research data can comprise a wide range of file types from numerical and textual, software and multi-media material and is increasingly made available under an open licence. Often esoteric, there is perhaps a presumption that such material is of interest only to specialists. However, the open science movement demonstrates that there is potentially a broader audience amongst non-specialists and the general public.

While data deposited in repositories will be discoverable on the public internet, functionality is typically focussed on formal curation and preservation. File presentation can often be limited, especially for media files which may be packaged in zip files for example.

Most repositories are unlikely to have the resource to develop a more sophisticated presentation layer. However, the Wikimedia suite of tools potentially provides a well-established infrastructure that can be leveraged to disseminate appropriately licensed media files. Any incoming links will help to improve search engine ranking of the hosting repository.

Example: [https://commons.wikimedia.org/wiki/File:Baxter\\_full.webm](https://commons.wikimedia.org/wiki/File:Baxter_full.webm)

## Wikimedia Commons

Wikimedia Commons is a repository of openly licensed media files – images, video and audio – for use in education and, like Wikipedia itself, anyone can upload or edit material. It also makes is very easy to embed media files across Wikimedia projects.

---

<sup>1</sup> TBC. A discipline or related range of disciplines with established expertise across the White Rose Consortium e.g. Climate Science, Environment, Conservation

Intrinsic to Wikimedia are the concepts of reusing and repurposing content and there is an established community of enthusiasts.

### The Proposal

The editathon will be hosted in the Research Hub at University of Leeds and will last a full day:

- An RDM session in the morning focussed on good practice and linking to the open science agenda. It will include examples of research data files that have been uploaded to Wikimedia Commons and embedded in Wikipedia.
- The afternoon session will focus on the mechanics of uploading media files to Wikimedia Commons and identifying appropriate Wikipedia pages to edit and improve with a properly formatted citation that will be picked up by altmetric.com.

Participants will be asked to upload at least one of their own data files to Wikimedia Commons with a view to improving at least one Wikipedia page.

### Beneficiaries

- Participating researchers – participants will have the opportunity to learn from experts in research data, Wikimedia and associated technologies to manage and share their data appropriately
- White Rose Consortium – the expertise in the room will be valuable to all three Universities, to contribute to REF impact case studies for example
- Wikimedia Foundation – the outcomes of the editathon will contribute directly to the objectives of the charitable Wikimedia Foundation, furthering their mission
- Global public – the activity undertaken in the workshop will make high quality research data easily available and reusable and improve the quality of Wikipedia

Budget	
Venue	Free
Promotional	£200
Catering	£500
Travel Expenses (x2 Wikimedia Foundation)	£500
Travel Expenses (x3 WR Consortium)	£150
<b>TOTAL</b>	<b>£1350</b>

Feasibility (25%)	An editathon is relatively straightforward to run with freely available resources from the Wikimedia Foundation
Impact (25%)	Potential for considerable impact to embed good practice across the WR consortium and beyond
Transferability (25%)	A similar editathon can be run at any university or similar institution with minimum resource implications
Value for Money (25%)	The only costs are associated with promoting and running the event, representing optimum value for money