

History DMP

Managing data for historical research

In brief...

The Department of History at the University of Hull was the top performing department in the RAE2008 at the University. This high standing has resulted from the broad range of research activity across the department, with over 30 full-time members of staff contributing to the overall body of work. The Department's focus on teams engaged in the production of databases was specifically recognised in feedback from the RAE2008 panel.

The History DMP project will build on the existing data management best practices within the Department. It will use the experience in data management demonstrated and recognised to frame a departmental approach to data management, to enhance and build on the individual activities that have been the basis of data management thus far. This departmental data management plan will be used to support future research strategy and provide a coherent platform for data sustainability in future research.

The work to be undertaken will develop an overall plan, and will then implement this using past, present and future research activities as case studies. The role of local technical provision, in the shape of the University's digital repository, will be explored and enhanced in specific ways to deliver a platform that not only manages the data, but allows for its exploitation and access as well through repository, VRE, and linked data interfaces.

The data

The following are examples of the types of dataset generated by the Department in recent years:

- ✧ The compilation of the **Voyages and Names databases** from multiple original sources on the transatlantic slave trade, by the Wilberforce Institute for the study of Slavery and Emancipation (WISE) and international partners. Both are delivered through dedicated websites and are valuable resources in the understanding of the scale of this trade.
- ✧ The international coordination of historical data relating to the long-term impact of fisheries on fish stocks generated by the **History of Marine Animal Populations (HMAP)** project. This extensive array of data sources is used to inform current marine management.
- ✧ **Shipping in the 14th Century**, an ongoing project specifically geared towards the compilation of data from multiple sources on this subject into a coherent whole to aid future analysis.
- ✧ The **English County Administrative Database, 1290-1410**, a database of 40,000 records compiled using older database technology, raising issues about data preservation.

The technology

Three approaches will be investigated locally to investigate their merit in supporting local data management:

- The digital repository at Hull, based on Fedora¹ and presented using Hydra², will be enhanced to receive and manage datasets according to need. The repository already holds the Domesday dataset and Repertorium of Middle English Prose Sermons, and this project will adapt Hydra to offer appropriate interfaces for easier capture of metadata and the data files themselves.
- The JISC-funded CLIF project³ developed a mechanism for linking Sakai and Fedora. The potential for using Sakai as a VRE and supporting data management between this and the repository will be investigated, implementing the CLIF technical solution.
- Datasets are structured, and this structure can be exploited in parts just as much as with the whole of the dataset, particularly when properly cited and linked to ontologies describing elements of the data. The project will make use of existing, openly available tools and APIs, including Google Refine and Protégé, to demonstrate how the structure of a dataset can itself be an asset in managing the data and thus potentially an important part of any data management plan.

The data management plan for these and other related datasets will be based on the DCC checklist, but adapted for use on a departmental basis.

The plan will be open-minded on the merits of local vs. external data management provision, whilst seeking to understand the value of local capability.

Credits and Further Information

Chris Awre, Richard Green, John Nicholls, Peter Wilson, and David Starkey, University of Hull

Plus many thanks to the members of staff and postgraduate students of the department for their input

This work is taking place under the auspices of the JISC-funded History DMP project within the JISC Managing Research Data Programme 2011-3