The Data Management Journey

The Digital Humanities Hub aims to conduct exemplary research and foster the use of digital research methods in the humanities. We work closely with communities, researchers and cultural institutions across disciplines such as anthropology, visual arts, social and environmental history and literary analysis. Drawing on our experience, one of our goals is to develop a set of workflows and tools to achieve better data management processes in the research environment.

Key components of our digital workflow:

Acquisition
We help communities and researchers record cultural and environmental information. In a project in Arnhem Land, GPS-equipped PDAs are used to record text, images and geospatial data. We also receive and process information from a great variety of sources, ranging from current researchers’ digital files to archival glass plates and wax cylinders.

Import
This component involves the import of data from multiple sources into databases. Harvesting technology was used in the IDIG project to import data from cultural institutions and to develop a search engine. XSLT was used to import data from researchers’ Filemaker and Endnote files into databases in the AUSTLANG, BIDWERN and ASEDAL Projects.

Storage
For the purpose of sustainable preservation and the usage for analysis and discovery, data needs to be stored in the right format, media and with sufficient metadata attached. It should also have proper access control and history management.

Analysis
This process helps researchers to categorise data, add annotations, comments, relationships and other research information. The stored research information can then be used in data discovery or further analysis.

Discovery
This process aims to publish data in accessible ways and to provide diverse and flexible interfaces for discovery. Advanced query, Fuzzy search function and Geo mapping interfaces were developed in the AUSTLANG, BIDWERN and AUSTKIN projects.

Online Cultural Collections Analysis and Management System (OCCAMS):
Occam’s (or Ockham’s) razor — entia non sunt multiplicanda praeter necessitatem — is the principle that the simplest explanation or strategy tends to be the best one.

OCCAMS is a collaborative research tool for people working with and creating cultural collections. It is an online database that allows people to organise, annotate and link data in standard formats, either for depositing in an archive or publishing on the web. Its functions address the needs of our data management processes.

Acquisition
- Provide field tool for working offline in field. • Provide file editing tool to edit metadata for digital files. • Support standard formats for documents, images, video and audio.

Import
- Batch import. • Automatically retrieve metadata from files. • Map metadata. • Support various metadata standards. • Automatically create or link digital files to database records. • Import bibliographic records.

Storage
- Support standard formats for digital files. • Support various record types, eg event, people, rockart site, object and location. • Cutomisable metadata fields. • Version control and history management. • Access and user management.

Analysis
- Add annotations to digital files. • Add comments to records. • Enable relationships between records. • Support metadata standards and default metadata mapping. • Create controlled vocabulary lists. • Flexible category hierarchies.

Discovery
- Publish collections. • Create online exhibits. • Create geospatial map view and timeline view. • Simple search and advanced query. • Flexible API to retrieve data for various applications and environments.

Email: dhh@anu.edu.au
URL: dhh.anu.edu.au/occams