



Collections Navigator: Collections Descriptions at the Natural History Museum

Rachel Perkins, Collections Description Manager, Natural History Museum

Background

The Natural History Museum has an extensive cross-domain range of resources. These include an estimated 70 million specimens, one million library volumes, the third largest collection of art on paper in the UK, 10,000 listed archive items, an extensive web site plus educational, electronic and exhibition resources.

Collections descriptions document the holdings of the Museum as collections or groups of items. This was recognized as an efficient and effective complement to item-level cataloguing given the vast numbers involved.

Objectives

When the Collections Description Project began in September 2000 the primary objective was to create a user-friendly, resource discovery tool providing one point of access to descriptive, cross-domain electronic records.

Since then the collections descriptions have been used to create links between users and collections as well as to establish links between collection resources. The latter, "virtual collections", are applicable not only within the Museum but also on a global scale.

Additionally, the role of collections descriptions as a tool to aid collections management

purposes is recognized by those concerned with collections documentation. These processes include registration of new acquisitions, retrospective cataloguing, audit and security, collections development, prioritization of items for item-level cataloguing and assessment of collections for preservation and conservation treatment.

What we did

A "collection" is defined as "a group of things with something in common". The "something in common" is identified on the basis of answers to the queries "who?", "what?", "where?", "when?", "why?" and "how?". This purposely vague definition is intended to be flexible enough to accommodate the wide range of resources at the Museum as described earlier.

Creating collection description records is a four-step process.

Step 1

Firstly, museum staff must identify specific collections. Specimens are organized into curatorial units. For example, specimens in the Zoology Department are curated as Amphibians, Birds, Fish, Mammals, and so on. Library resources are traditionally documented according to format - monographs, manuscripts, maps, photographs, serials, and so on. In the Archive,

collections are sorted by provenance, which provides the cultural context in which the records become intelligible through associations with people or places or events. Educational, electronic and exhibition resources are identified on a case-by-case scenario.

Step 2

The second step, researching and documenting the collections can be a most demanding although ultimately rewarding task. In creating the first of the collection description records every effort was made to repurpose existing data. In searching for electronically held material, files saved in old file formats were converted into current formats, saved from technological obsolescence just in the nick of time. Additionally, data produced originally in print format only is made available electronically.

Again the cross-domain range of collections has required different data capture approaches. Science department data, education, electronic and exhibition data are recorded using MS Excel templates. This system however is to be superseded. The implementation of the collections management system KE EMu will enable collections descriptions for the specimens to be captured. For miscellaneous collections a web-based template is being developed. A customized data entry template for library collections was developed in the Unicorn Library Management System. Calm 2000 software is used to record archive collections data.

Step 3

The third step is to collate all the collection description data in order to present it on the web. In the interest of interoperability the decision was made to use EAD (Encoded Archival Description) [1] data structure standard to encode all the collections data in XML (eXtensible Markup Language). The text data of the specimen, educational, electronic and exhibition collections are encoded directly into EAD XML. The library collections are exported from Unicorn and converted to EAD XML. Data

from the archive database are sent to The National Archives where it is encoded into EAD XML for the A2A project. The encoded data is then returned for re-use at the Museum.

Step 4

The most satisfying stage in the process is the on-line display of the collections description data. The Collections Navigator web site has been developed. Behind the scenes an SQL database is used to manage files which are indexed and searched using Swish-e. The collections description records are displayed using XSLT with functionality provided by Java applets.

Specific features

Working with cross-domain data presented particular challenges to be addressed. Collections description data is captured using about 50 data fields. These data fields can be grouped under the following headings:

- details about the collection
- details about the collector
- acquisition data
- material either related to or separated from the collection
- access to and use of collection details
- standardized subject key words
- links to a range of electronic resources
- administrative data

There is considerable similarity in the type of data captured across each domain. However differences in the terminology used to label each data field did present its challenges in ensuring that data was understandably associated with the correct data field.

The large number of fields used provides the opportunity for collection description records to capture a wide range of information that might otherwise be lost. Recording more rather than less data increases the possibilities of serving a more diverse user community. New and innovative ways in which to use the collections data can also be developed.

Another cross-domain challenge to be addressed was the application of terminology standards. Within the Museum there is no single terminology standard in use for natural history specimens, geographic names, people and institution names or dates. For the purposes of the Collections Description Project the following standards were used from the start: Library of Congress Subject Headings; Getty's Thesaurus of Geographic Names; Library of Congress Names Authority File and ISO 8601 (for dates). Taxonomic terminology standards under consideration for use are the more recently available Integrated Taxonomic Information System and the National Biodiversity Network's UK Species Dictionary.

Progress so far

Collections Navigator has been on-line at www.nhm.ac.uk/navigator since March 2003. There are some 750-collection description records available. Of these more than half are manuscript collections found in the Entomology Library and the Ornithology Library at Tring.

The display of these records was a significant step forward in the development of the public user interface. Making available functioning XML records posed particular technical challenges requiring developments in software packages. In-house IT staff addressed these issues.

Issues of quality assurance, maintenance, file management and sustainability are now being addressed seriously. The accuracy and quality of the collections description data are reviewed by the subject specialist or Collections Description Manager. This recognizes the expertise of the staff and helps encourage ownership of the data.

The way in which the data is managed is dependant upon how the data was initially captured and converted to EAD XML. As data is modified on the original data system the records are re-exported, converted and loaded into Collections Navigator.

There are two types of files to be managed in creating collections descriptions. The first are the EAD XML files which are stored in an SQL database. Exemplar digital images used to visually describe the collections are currently stored on a network server. A bespoke Digital Image Documentation database records the metadata of the images based on VRA Core 3.

The creation of collections description records for the holdings of the Natural History Museum was funded as a project back in September 2000 with one contract member of staff. Now, four years on, a permanent post has been created, that of Collections Description Manager. With the support of a recently appointed Registrar plans are underway to create collections description records as part of the normal "new acquisition" procedures. Funding for additional cataloguing of collections is sought through retrospective cataloguing sources.

Outcomes

The value of collections descriptions is increasingly recognized within the Museum as well as nationally and internationally. Recognition of the need to include collections descriptions into the day-to-day activities of the Museum's archives, library and specimen collections speak volumes.

During the past four years we actively contributed to the UKOLN's Collection Description Focus Group. This was the most valuable consistent opportunity to share knowledge, discuss issues and exchange experiences regarding the implementation of collections descriptions. As a cross-domain project presentations have been made to archive, library and museum audiences. Additionally, European institutions have sought our advice and work has begun on a trans-Atlantic Joint Collections Description Project organized with the help of the RLG (Resource Library Group).

BioCASE is a European funded project designed to "establish a web-based information service providing researchers with

unified access to biological collections in Europe while leaving control of the information with the collection holders.” [2] The Natural History Museum is one of 35 institutions representing 31 countries involved in the project and now serves as the UK national node. Data collated for the Museum’s Collections Navigator is repurposed for use in BioCASE.

That there is great potential for the use of collections descriptions is undisputed. The range of applications is limited only by our imaginations. It is with some regret that I announce that I will no longer be directly responsible for collections descriptions at the NHM as I have accepted a new post at the Natural History Museum. However, my belief in the value of collections descriptions is such that in my new job I find myself time and again actively advocating the use of collections descriptions for a variety of purposes.

Under the direction of a new staff member **Collections Navigator** will continue to be developed and augmented. Eventually all the Museum’s holdings will be documented by collection-level descriptions. This will fulfil our primary objective to create a user-friendly, resource discovery tool for the management of the cross-domain holdings of the Natural History Museum.

References

1. EAD - <http://www.loc.gov/ead/>
2. BioCASE - <http://www.biocase.org/>

Contact details:

Rachel Perkins
Library and Information Services
Natural History Museum
Cromwell Road
London
England
SW7 5BD

020 7942 5027
R.Perkins@nhm.ac.uk

Neil Thomson
Library and Information Services
Natural History Museum
Cromwell Road
London
England
SW7 5BD

020 7942 5294
N.Thomson@nhm.ac.uk



The University of Bath
Bath, BA2 7AY
ukoln@ukoln.ac.uk
tel 01225 386580
<http://www.ukoln.ac.uk>

UKOLN is funded by The Museums, Libraries and Archives Council (MLA), the Joint Information Systems Committee (JISC) of the Higher and Further Education Funding Councils, as well as by project funding from the JISC and the European Union. UKOLN also receives support from the University of Bath where it is based.

