



Cornucopia

Chris Turner, ICT Adviser, MLA

Background

Cornucopia is a searchable database of collections held by cultural heritage institutions throughout the UK. It is developed and managed by the Museums, Libraries and Archives Council and was initially established in response to the Government's *Treasures in Trust* report which called for a way to be found of recognising the richness and diversity of our collections.

The original Cornucopia site was set up in 1998 and featured data from the 62 museums in England holding collections which are 'Designated' as being of outstanding importance.

Objectives

MLA's long-term plan for Cornucopia is focussed on its ability to act as a comprehensive information resource on UK collections. This meant that the Phase 3 development has concentrated on the following key areas:

- Coverage – populating the database with information from all regions, and from other cultural domains (Libraries and Archives)
- Interoperability – enabling the widest possible access and use of the data
- Sustainability – ensuring that content is easily updateable

In addressing these areas an early decision was made to adopt the Collections Level Description schema [1] from the Research Libraries Support Programme (RSLP). It was and is felt that the use of an internationally recognised metadata standard was an essential prerequisite to achieving the objective of widening access and use of the data held within Cornucopia.

What we did

The original Cornucopia site was searchable by a variety of criteria including subject area and geographical location and was based around static-HTML pages with little or no database functionality. The pages relating to each collection could be 'grouped' according to search criteria (subject type, for example), but none of the results were assembled on the fly.

This version went online in October 1998 as part of an evaluation process, and was generally well-received. At the end of 1999 a full evaluation [2] was carried out of this pilot on behalf of MGC by Soloman Business Partners.

In the light of the evaluation, over the next two years a second version of the database was designed, with the aim of extending the model to include information from all Registered museums in the UK. Stuart Holm was brought in as a consultant to create the data structure for the new Cornucopia, using the model of the original data with amendments derived from the evaluation report. This data structure referred to a very early report from the UK Office of Library and Information Networking [3] (UKOLN) on schema, metadata and interoperable information – a precursor to the RSLP Collection Description schema.

The structure was designed to accommodate information from several sources. To save time and effort, it was decided to populate Cornucopia with data from the MGC's Digest of Museum Statistics (DOMUS) database. This provided top-level institutional information for all Registered museums in the UK. The structure also had to be sufficiently flexible to accommodate information from a variety of other sources, in particular, data from the West Midlands and South West Area

Museum Councils mapping projects. Between them, these projects cover almost 500 museums, providing detailed information about collections, access arrangements, documentation and a range of other areas.

Key to the design of the original schema was the representation of the relationship between collections and the institutions that hold them. The database reflected this structure by preserving the concept of three 'levels' of information:

- **Institutional data** (e.g. address, access, website & institutional (or 'overall') collection)
- **Collections data** (e.g. title of collection, subject area, object type, geographical and temporal coverage)
- **Collection strengths** (e.g. objects of particular importance which wouldn't otherwise be retrieved by a collections-level description)

This structure was then handed over to database designers System Simulation (SSL), a company with extensive experience of working on museum projects.

Two consultants were contracted to complete records for each of the museums in the South West and West Midlands regions on the basis of this preloaded information, drawing together information from a wide range of sources including:

- MGC Registration files
- Printed documentation of collections
- Information derived from personal visits and/or telephone interviews
- Information held at Area Museum Councils

Having established a core dataset of approximately 1250 basic records and approximately 500 detailed entries, plans were drawn up for extending the scope of Cornucopia to provide detailed information across the remaining regions.

The decision to adopt the RSLP Schema necessitated a mapping exercise to map the then existing Cornucopia data structure on to the RSLP Schema, and this in turn led to a review of the database structure.

By a happy coincidence the Crossroads [4] project in the West Midlands was using the RSLP Schema and thus offered the chance to make use of a ready-made database structure. Furthermore the Crossroads system was built using open source tools and was making use of direct data entry for remote input of data.

A decision was therefore made to migrate from the Version 2 database and to contract the developers of the Crossroads system, Orangeleaf Systems, [5] to develop Cornucopia Phase 3.

Specific features

The new Cornucopia system [6] allows the recording and maintenance of collection descriptions and details of their associated repositories, and makes these available for search and retrieval over the internet. The database can be remotely updated by more than one person, using only a common web browser such as Netscape or Internet Explorer and no extra software to install.

In addition there is an Open Archives Initiative – Protocol for Metadata Harvesting [7] (OAI-PMH) interface to the database, so that data may be harvested from Cornucopia for use in other systems. The latest addition to the system is Web Service access. A Web Services Description Language [8] (WSDL) file is available which will enable third party application developers to incorporate searches of the Cornucopia database within their own applications.

The system offers faster, more efficient searching and the ability to 'collect' CLDs, print them, save them or email them to a friend or colleague. Users both professional and public will be encouraged to use the descriptions for research, to plan a visit, or simply to find out more about our cultural heritage.

Progress so far

Cornucopia now includes over 6000 collection descriptions from almost 2000 institutions. The Phase 3 development has left us with 2 aspects to Cornucopia:

- **the database** of collection descriptions covering the registered museums in the UK, and
- an open source software system for the recording, maintenance and searching of collections descriptions conforming to the RSLP Schema.

These will have distinct although inevitably interconnected futures.

The Cornucopia **database** of museum collection descriptions clearly has a close relationship with the 24 Hour Museum [9] database of museum collections and the details of how that will operate are now being worked out. Web service capability will be added to the system, as it was always intended that this database would act as a source of collections descriptions to be incorporated in other services rather than a stand alone website.

The Cornucopia **system** is already used by another project – Cecilia [10] – a database of over 1,800 collection descriptions of music materials held in some 600 museums, libraries and archives in the UK and Ireland. Cecilia offers an overview of the national music resource enabling all kinds of users to identify, locate and assess materials. The work to create the database has been funded through the British Library Co-operation and Partnership Programme, and, for preservation purposes, a copy of the data has been placed with the AHDS performing Arts.

Because the system interface can be tailored to reflect the unique identity of any project while storing the data and making it available in formats which conform to international standards – RSLP Schema, OAI-PMH, Web Services, - a number of other collection description projects, covering Libraries and Archives as well as Museums are also proposing to use Cornucopia.

Providing the support infrastructure in terms of system support, consultancy, training and system development is the next challenge for MLA in the evolution of Cornucopia.

Lessons learned

In order to manage the process of information-gathering during Phase 2, a system was established whereby the consultants were issued with 'satellite' versions of the data-entry client. These satellite copies held all of the information, edits and updates to the data, which were then exported on a regular basis to a 'master' version of the database held at MLA. This approach was adopted to allow the consultants sufficient freedom of movement to gather the information, but also to provide a stable version at MLA in case of machine failure.

In the event, progress with data collection was hampered slightly by technical difficulties with the 'satellites' of the data entry client which it was not possible to replicate on the master copy, and so were hard to resolve.

In Phase 3 a web based data entry client was provided, but this also brought its 'challenges'. Support for consultants entering collection description data is vital; especially as data entry is through the various available browsers. The final addition to the original Crossroads design has been an interactive support forum built into the Cornucopia application. Messages added to it are sent to Orangeleaf for tracking and action but are also available to read and reply to by all the other consultants with a view to fostering a 'community'. Such messages have helped us improve both the data entry and public facing areas of the web site.

Perhaps the greatest challenge of constructing the data for the two regions in Phase 2 was in ensuring consistency in the application of terminology, standards and protocols. The structure of satellite copies of the data feeding into a master copy enabled Nick Poole at MLA to edit the incoming information to ensure that information from the two regions would be

broadly comparable and consistent. This process was aided greatly by the inclusion of an 'Editor's notes' field in the database, in which the consultants would add any comments or highlight discrepancies in the available information for a given collection. Much of the editing of the data involved resolving these issues raised by the consultants.

References

1. For more information on the RSLP Schema see: <http://www.ukoln.ac.uk/cd-focus/>
2. The full text of the evaluation can be downloaded at: http://www.cornucopia.org.uk/pilot_eval.html
3. Collection Level Description - a review of existing practice, an eLib supporting study - Andy Powell, UKOLN, August 1999. <http://www.ukoln.ac.uk/metadata/cld/study/>
4. Crossroads: Discovering West Midlands Collections. <http://www.crossroads-wm.org.uk/>
5. Orangeleaf Systems Ltd. <http://www.orangeleaf.com/>
6. Cornucopia: Discovering UK Collections. <http://www.cornucopia.org.uk/>
7. The Open Archives Initiative <http://www.openarchives.org/>
8. Information on WSDL can be found at <http://www.w3.org/TR/wsdl> and the Cornucopia WSDL file can be accessed from the Cornucopia web site <http://www.cornucopia.org.uk>
9. 24 Hour Museum: The National Virtual Museum <http://www.24hourmuseum.org.uk/>
10. Cecilia: Music Collections on the UK and Ireland. <http://cecilia.orangeleaf.com/>



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.