



Tanya Szrajber, British Museum, London, UK

## **Collection Online: the British Museum Collection Database goes public**

Keywords: British Museum, collection, database, Merlin, online catalogue

### **Introduction and Project outline**

The British Museum collection of art and antiquities, from ancient as well as living cultures, is one of the most extensive worldwide. The museum is committed to making the collection as accessible as possible to the widest audience. Apart from the obvious criterion of millions of people actually visiting the British Museum, there is a full programme of in-house and touring exhibitions and loans, collaborative research projects, educational programmes and coverage in the broadcasting media. In addition, its website (which was recently redesigned) provides remote access to the collection. The latest development in the context of public accessibility and engagement is the subject of this paper. 'Collection online' is the project to make all the records in the collection database (called 'Merlin') publicly available on the website. This database is one of the largest of its kind, with over 1,700,000 records for objects from the widest variety of cultures and historical periods.

This is not the first time that object records in the British Museum collection have been directly accessible to the public on its website. COMPASS, a selection of 5000 of the museum's best-known objects, was launched in June 2000. The COMPASS records, now available through the 'Highlights' option on the website, were created specifically for public use, and descriptions were commissioned for this particular purpose. This has never been the case with records on the Merlin database (nor its predecessors). The collection database was intended as an inventory of the collections, with catalogue information included where possible. Over the years, the scope of the collection database project has expanded, with the addition of academic information, as well as conservation data (with science data to follow). Nevertheless, the public will see the records as they are, rather than as they might be re-written for a general public audience. Merlin is used by British Museum staff for many purposes: as a basis for academic research, to select objects for display, exhibitions and publications, and to answer public enquiries. The decision to make the entire database available as it stands, rather than release selected areas after carefully vetting the records, is significant. It indicates the Museum's commitment to sharing information about the collections in the most open way possible, and despite the potential hazards associated with presenting raw data. Thus the public will be made fully aware of the range and breadth of the collection, with the obvious caveat that the computerisation programme is not yet complete, and therefore not all objects are recorded on the database. In addition, descriptions are not

always adequate nor reflect the most up-to-date state of knowledge. One of the important practical advantages of this project will be a reduction in the physical handling of the objects by visitors to the study rooms. In instances where the visitors are unsure of the material they need to see, the initial selection can be made through the website.

'Collection online' has a specific brief and timescale: all existing Merlin records will be made publicly available over a three year period, from 2007 to 2010, and any new ones added thereafter. It is intended that 250,000 images will be added to the Merlin records, by digitising all existing resources of transparencies, as well as scanning two-dimensional objects (such as drawings and prints, etc.) directly where this procedure is considered safe for the objects. The only restriction, obviously, concerns images of works still protected by artistic copyright (meaning works by an artist who is still alive or has died in the past seventy years). These will not be posted on the Web unless specific permission has been given by the copyright holder of that artist's work, although they are available internally on Merlin. Apart from the addition of images, the records are being improved by documentation staff, curators and assistants to the fullest extent possible, given other constraints on curatorial time. In addition, a number of catalogues are being scanned by OCR, and the text transferred to the records. The first phase of the project, released on 18th October, concerns records for two-dimensional works of art (mainly on paper). The number of records published is 262,565 records, of which 98,754 have images (Fig. 1). The data will be updated weekly from the Merlin database, so this figure will change. Subsequent releases, beginning around April 2008, are in the following sequence: three-dimensional collections from the Departments of Asia and Ancient Egypt and Sudan, the remaining three-dimensional collections (those from the Department of Coins and Medals last). By 2011, Conservation and Science records, and finally records for the collection of photographs, will be added. More detailed information is available on the website.

The screenshot shows the British Museum's website interface. At the top, there is a search bar and navigation links: Home, Visiting, What's on, Explore, Research, Learning, The Museum, Join in, and Shop online. Below the navigation is a breadcrumb trail: Home > Research > Search the collection database. The main content area is titled 'Collection database search' and displays the following information:

- Title:** Thirty-six views of the Imperial Summer Palace at Jehol (澄波叠翠 (cheng bo die cui))
- Image:** A thumbnail image of a Chinese landscape print.
- Object types:** print (all objects)
- Title (series):** Thirty-six views of the Imperial Summer Palace at Jehol
- Title (object):** 澄波叠翠 (cheng bo die cui)
- Materials:** paper (all objects)
- Techniques:** engraving (scope note | all objects)
- Production person:** Print made by Ilaliteo Ripa (biographical details | all objects)
- Production place:** Made in China (scope note | all objects)
- Place (findspot):** Found/Acquired China (scope note | all objects)
- Date:** 1714
- Department:** Asia
- Registration number:** 1955.0212.0.1.30
- Description:** Plate from a set of thirty-four prints (incomplete), mounted in an album; view of a lake or river, in the foreground, on the shore, a pavilion on a platform by the water's edge, surrounded by trees and small hills; on the opposite shore, trees and tall hills; rocky landscape in the distance. Engraving on very thin China paper.
- Inscriptions:** inscription | top right hand corner  
30 (Plate number in pencil)

Additional links include 'Larger image', 'Back to search results', and 'Back to search'.

Fig. 1. The British Museum Collection Database search (15 November 2007).

## **History of the computerisation of the BM collections**

This began in 1976 with a scheme in the Department of Egyptian Antiquities (now the Department of Ancient Egypt and Sudan). Information for a sample of objects was transferred to a bespoke system called 'BMUSE'. This was followed in 1978 with a pilot project to computerise records for collections of the Ethnography Department (now Department of Africa, Oceania and Americas). A significant development occurred in 1979, when the BMUSE records were transferred to the Museum Documentation Association (MDA) multi-user GOS system, on Hewlett Packard computers. At this time the Collections Data Management Section (CDMS) was created, with documentation staff dedicated to the task of creating records, and later terminologies (vocabularies). GOS was a batch-processing system, laborious for the inputting staff, who had to fill in forms manually for each field of information, sending them to the Department of Information Systems for processing. In the early 1980s, the National Audit Office and the Public Accounts Committee took an interest in the project and its purpose became the creation of an inventory. The scope then increased to include other departmental collections. In 1988, 'MAGUS' was introduced, an online system with terminals, and the data was transferred to it from GOS. An important change in British Museum procedure occurred in 1993, when online registration of new acquisitions by curatorial staff was introduced, replacing the system of recording them by hand in the Registers (the old accession ledgers). However, although MAGUS was a vast improvement on its predecessors, allowing data entry and editing on the database directly, and with a sophisticated report-writing facility using, it too had drawbacks. It did not support images, non-Latin character sets, nor standard word-processing features such as 'cut-and-paste'. Furthermore, printing of the reports had to be carried out using special technology in the Department of Information Systems. So in 1998, tendering for its replacement began, and Merlin, created by System Simulation Ltd (SSL), was selected and in operation in 2000. Merlin is a bespoke system using index+ (a product based on non-relational database technology) and operates on Windows PC. It is compliant with SPECTRUM (a national and international standard developed by MDA) regarding most of its data fields, although the British Museum has a few additional fields due to the specific requirements for recording parts of the collection, e.g. coins and banknotes.

## **Data quality issues**

In terms of the accuracy and amount of information entered, the records vary enormously in content, ranging from very basic inventory to those reflecting current academic knowledge. This is due to several factors, including the size and scope of the database, the span of the project (just over twenty years) and the fact that the initial systems used were far less complex and sophisticated than the current one. Each major conversion created new problems, as more specific fields were introduced, and back-editing was required. In more complex areas, this required the assistance of curatorial

staff. In addition, record creation by the documentation staff was and still is based on record-cards and Registers (spanning 250 years of accessioning), the objects themselves, and where available, published catalogues. The information contained in these written sources is nowadays regarded as out-of-date, brief, and at times difficult for the non-expert to interpret and extrapolate into data fields. Until widespread departmental access to searching and editing was implemented with the advent of MAGUS, there was little curatorial involvement in the database. For many years now, curators and other collection department staff have been trained to edit and retrieve records, resulting in an increased awareness of its potential and enhancement of the content. However, the task is enormous, and it will take many years before all records reach the required standard.

## **Terminology**

One of the most beneficial aspects of Merlin is the range and complexity of its various forms of terminology (or vocabulary) control. They include simple drop-down lists, thesauri, and files known as 'Authorities', all of which link to the catalogue (object) records. The thesauri and Authority files were all developed in-house and two were published jointly with MDA (Object Names and Materials). The Documentation Group (formerly the Collections Data Management Section) is responsible for their creation and maintenance, and consult closely with curators and scientists in the museum. The most challenging ones are those concerned with the names of people and places. The former, the Biographical Authority, includes every name recorded in connection with the collections, whether acquisition (donors, previous owners, etc.), producers (artists, publishers, etc.) or other associations (such as the names of sitters of portraits). The aim is to identify each person or institution, rather than to provide a full biography, and the main problem is to merge all instances of variants of names referring to the same individual into one record, with provision for alternative nomenclature. Such a resource will be invaluable to scholars due to the large number of less well-known people whose names are recorded, in all cultures. Placenames are more difficult to sort out, since there is a need to integrate archaic and modern terms, and identify those recorded phonetically in the older Registers and provide up-to-date equivalents. Many of the places recorded no longer exist and are hard to track down. Another challenging area is that of historical periods and cultures since the same term can mean different things in different contexts, e.g. 'Bronze Age'. Another important thesaurus records Ethnic Names, mainly for the Department of Africa, Oceania and the Americas. This year the Documentation Group has expanded the use of the Subject field to almost all departments (the exception being Ancient Egypt and Sudan) although the resulting editing will take some time to achieve. More specialised forms of terminology are restricted to some collections, for example, 'denomination' (coins and banknotes) or 'escapement' (clocks and watches).

In order to assist researchers and the less specialised users in their searches, the thesauri and authorities include a large number of search terms ('Non-Preferred' or 'Use for' in the thesauri and AKAs in the Biographical Authority) and explanatory Notes (Scope Notes). Although this process is far from complete, progress has been made. The main problem with Scope

Notes is to establish how a term is interpreted and used in various contexts (e.g. European and non-European) rather than set a definition based on one source, such as the OED, which might not apply to all cultures and time periods. Work continues on the thesauri and Authority files, and this will take some years to complete.

### **Collection online and future projects**

Collection online is accessed on the British Museum website from the 'Research' option on the Home page, then the option 'Search the collection database'. The public then has four search choices as well as Search help text and background information: a Basic search (the default), an Advanced search, a Museum number and provenance search, and a Publication reference search. The search screens will be modified and additional ones added in due course to take account of requirements applicable to records for three-dimensional objects. In addition, some field names will change, such as 'Support materials' to 'Materials'.

The Basic search operates across all fields of the database - both the free text and the index fields - and will retrieve every record that matches the word (or words) entered in the Search box. This Basic search can be refined by selecting from the options: 'Title and subject' (title, description or subject), 'Physical attributes' (type of object, material or support, or technique), 'Production information' (e.g. artist, publisher, etc.) and 'Images only'. In addition, a Date option (years only) is available. For those seeking a more sophisticated search mechanism, and more accustomed to databases, the Advanced Search is preferable. Users can select a Category from a drop-down list (the categories listed are essentially the thesauri and Authority files) and proceed from there. The search can be built up using several Category terms, returning to the Advanced Search screen after each term selection. It is possible to browse the terminology records from within the terminology screens (ie. those accessed from Category), moving from one term to another. The search can be refined on the Advanced Search screen using free-text and Date information, and selecting records with images.

The immediate search results are presented as brief records with thumbnail images and limited text. By selecting a record, the user is then shown the full set of fields and an image which can be enlarged (Fig.1). All fields in the database are shown, apart from price, valuation, storage location within the museum and NGR references for three-dimensional works. In addition, addresses of private individuals are withheld.

Apart from building up a search methodically, returning to the Advanced Search screen each time, in a linear fashion, users can embark on new, more random, routes through the database using hypertext links associated with the terms displayed. More detailed information is included in the online Help text and will not be repeated here.

A free digital image service (for images of 2500 pixels at the longest edge), stipulating certain terms and conditions, will be available soon for non-commercial, usually academic, purposes. The terms will be clearly stated on the website and the effect will be that academics with print runs of less than 4000 should not have to pay any reproduction fees to the British Museum in

their books and articles. Users will be required to register details and they will be sent the image as an e-mail attachment. For higher resolution images and commercial purposes, the users are referred to [bmimages.com](http://bmimages.com).

The public is actively invited to participate in 'Collection online', since there is a prompt, on each record page, to do so by e-mail: 'Noticed a mistake? Have some extra information about this object? Please contact us'.

Future website plans include publishing electronic catalogues (referred to as 'Collection online catalogues'), which will use Merlin records as entries, and possibly invitations to submit for publication, after peer review, scholarly articles about the British Museum collection. It is even envisaged that a system for remote addition to entries by outside experts might one day be introduced!

## **Conclusion**

Addressing the Conference issue, the so-called 'top-down' institutions clearly include the British Museum, along with other museums, libraries and galleries, etc. It can indeed be said that, traditionally, museums have mediated access to their collections. This to some extent is still the case. After all, the objects are curated, researched and cared for by specialist staff who select the objects for display, special exhibitions and publications. This is to be expected and encouraged, since people with such expertise have to be the link between the collections and the public. But rather than 'controlling' such access, one might equally say that they 'facilitate' it. After all, the collective expertise of museum staff regarding the collections under their custodianship results in their correct identification, function, and contextualisation. In addition, outside expertise, both in Britain and abroad, is regularly sought to enhance this knowledge and share it. Such improved knowledge is of obvious benefit to the public, who gain a better understanding of their cultural heritage, and an invitation to share and enjoy it through visits, studies, special events, and browsing the website. It is hoped that this paper has shown how the 'bottom-up' access to such information and collections is clearly welcomed by the British Museum, not only through traditional means but now through its website. Rather than seeing it as a threat, it is viewed as an opportunity to attract and include a wider audience, and this increased awareness and interest in museum collections safeguards their future