

„Breaking With Conventions”: What Does Digital Spatial Data Require?

by SUE ANTONELLI

INTRODUCTION

The 10th LIBER conference raised questions about the archiving of digital spatial datasets, and of allowing users access to such data with increasing ease. Digital image libraries are commonplace in the worlds of photography and the media, they may be readily accessed by the public using the Internet. United States Geological Survey, Automobile Association and Bartholomew now among others all have sites from which digital spatial data may be downloaded, provided the user has the appropriate ‚plug-ins’.

The Ministry of Defence Map Library, situated in the outskirts of London, is administered by Military Survey and is the principal UK Government Map Library. The Map Library’s remit is to maintain a world-wide collection of ‚current’ topographic and thematic mapping to meet the needs of the UK Ministry of Defence. It is not a historic collection and mapping no longer required is passed on to other collections, such as the Public Record Office and the British Library.

Within Military Survey, records of digital and map products are currently held on different databases, at separate locations and utilising local expertise. I have drawn upon my experiences in this environment to write this paper. I will outline what I see as the conventional map library. A quick look at our current automated system in use will reveal some of the map library conventions broken so far. I will address two further questions. What is digital spatial data? And what information do we need to store in order to be able to access that data? In order to meet the needs of Users in the 21st century we clearly need a different approach. I will end by outlining two possible solutions, both of which require the breaking of further conventions.

„Breaking With Conventions”: What Does Digital Spatial Data Require?

CONVENTIONAL LIBRARY

The Map Library collection used to be accessed using a traditional card referencing system. Record cards were used to record the details of each individual map. However, if maps were part of a series then an additional series card was created. (Figure 1 illustrates this relationship).

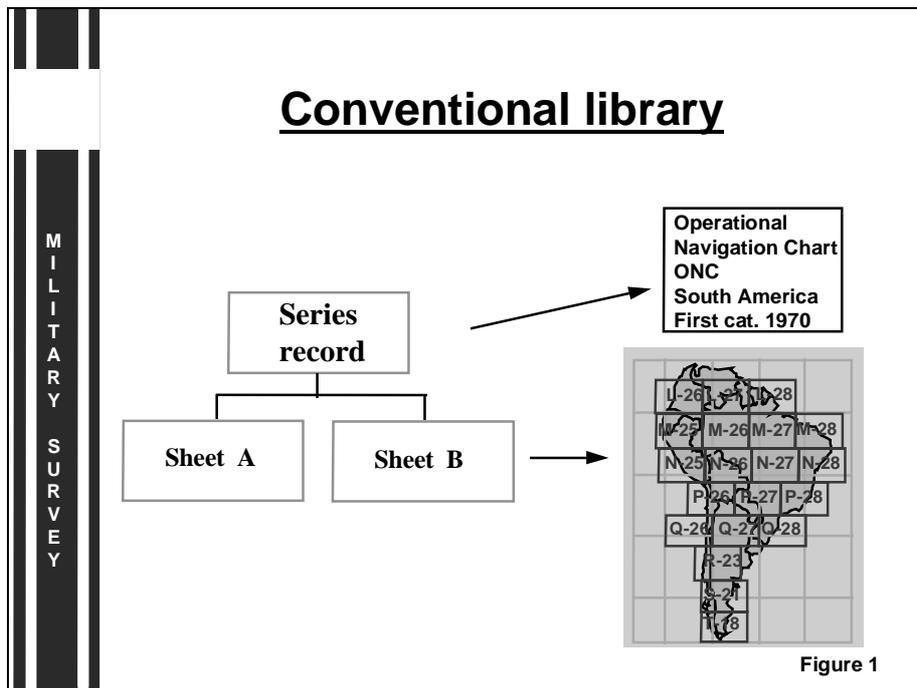


Figure 1

The Parsons Classification System (1946) has been used to allocate shelfmarks to all maps and atlases in the MOD map collection. This system divides the mapping into geographical groups. The majority of maps exist on the medium of paper but may also be found in the collection as dyelines, bromides or on plastic. Each map can be identified by a title, sheet number, name and edition or date. It will portray information about an area, which I will call its spatial reference. Any one map will usually have a single scale, a single specification and may depict one or more themes.

This manual cataloguing system required the ability to *search* by area, theme, scale, date or publisher. Spatial searches were made by country in the first

instance, then within a series, and indexes were used to identify the exact sheets required. This mechanism relied heavily upon the accurate cross-referencing capabilities of the cataloguers and a clear understanding by users of how the system operated.

„BROKEN CONVENTIONS” IN THE MAP LIBRARY

The Demise of the Card Index

MODMAP, Military Survey’s database of Map Library records was designed in the early 1980s to automate the MOD requirements described above. It is simply a text database not unlike many automated book library systems. In addition to capturing designated information recorded in the marginalia of each map, *MODMAP* also stores the geographical co-ordinates of the area covered by each map. The powerful retrieval mechanism is then capable of running complex spatial queries, the results of which may then be used to produce reports and graphic indexes, as will be shown below.

By the end of the 1980s the capture of the Library’s 1,5 million maps was well underway. The system was designed to mimic the card index system, with *SERIES* and *SHEET* records in place of the cards. Considerably more information is captured, but the data is used to support a vast array of enquires to the Library from the UK Defence community and Military Survey production areas. *MODMAP* Users may define a search using one or a combination of any of the metadata captured.

Demise of the SERIES/SHEET Relationship

MODMAP initially mimicked the manual catalogue relationship between series and sheet records. However, sheet records contain all the relevant data that appears in series records. There were only 3,000 series records as opposed to 860,000 sheet records, and the lack of computing power initially made querying the series records more efficient. In addition, cataloguing a sheet record with up to 70 data fields took much longer. Subsequent improvements to the computing power of *MODMAP* freed us from those constraints and switched the emphasis to sheet based enquiries. Maintaining both the sheet and series records became a burden and so we designed a new report in which *MODMAP* summarised data from all sheets within a series. There was no longer a separate requirement to maintain the series metadata.

„Breaking With Conventions”: What Does Digital Spatial Data Require?

This break in convention may cause some concern to those libraries who maintain a series record, but no sheet records. However, it should not. The data generated by the ‚Summary report’ can be guaranteed to be the most up to date information about Library holdings, which was never the case for a series record.

Relaxing the Cataloguing Rules

The blind obeying of the existing cataloguing rules, which enforced the capture of all the information about a map, was relaxed some years ago. But the problem was exacerbated when the atlas collection was prepared for MODMAP data capture. Atlases include mapping of single or multiple scales, themes and even specifications. In an ideal world, it might be considered appropriate to catalogue each plate of an atlas as a separate record. But, how many of our collection of 4,200 atlases have topographic coverage of the World? How many of these may be found within what is designed to be a national atlas? The cataloguer is allowed to use their knowledge of the overall collection to pose the question „Do we already have similar coverage of a similar publication date?”.

The rigid application of the classification rules would place many such atlases into the world category. The Map Library has been less than purist by deliberately ignoring this general, small scale mapping and cataloguing them with reference to the plates of the principal countries depicted. This prevents the MODMAP User from wasting time identifying large numbers of atlases when running enquiries.

Extending Access to Users Based Outside the Library Environment.

The last break with convention I wish to mention concerns the user. Using MODMAP they can gain access to the library via our internal computer network. They can request information based upon the metadata stored, without having to visit the library or even having to talk to the Library staff. Requesting the loan of a map without ever seeing it could result in it being returned immediately, unused because it did not provide the information they were seeking. Will the user then request another map, give up or seek advice? I will return to this issue later.

WHAT IS DIGITAL SPATIAL DATA ?

I have described a number of cataloguing conventions we have had to break so far. Will any further changes be required to accommodate digital spatial data? It has been acquired and catalogued within Military Survey for many years now. The collection currently exists as a distinct database, partly because it is held at a separate location. I will now briefly outline the nature of digital spatial data held in Military Survey.

- Digital spatial data can be raster, vector, matrix or imagery. Such *types* of data may exist singly or combined with others to produce a complete suite of data, of a single area, on a single medium. This allows a User to combine one or more aspects of the data to meet a their own requirements. Cataloguing such a dataset poses many problems.
- The *scale* of the data may vary within a dataset. The „Amtliche Topographische Karten (ATK)“ raster data of each German Land displays an area at the scales of 1:50,000, 1:200,000 and 1:1 million. But the scale may not be specified at all.
- The area of coverage of digital data is *seamless* and not sheet based as it is with maps. It may be broken down into tiles but this is not a sheet boundary.
- Data usually resides on one of the following *media*, CD, magnetic tape, on the hard drive of any mainframe or Personal Computer (PC). Alternatively it can be accessed via the Internet/Intranet.
- *Access* to data often requires the purchase of associated software, although this may differ according to the supplier. In many instances the PC hardware needs to have a high specification.
- The use and display of such data can be *flexible*, it allows for its display at various scales and on different datums and grids. The whole or part of the dataset may be selected for output, allowing for annotation if required.

CAN A MAP DATABASE PROVIDE FOR DIGITAL SPATIAL DATA?

Users

Who are the users of digital data, do they have different requirements to users of the map collection? Each Military Survey Library has the same MOD customer base. It consists of our own:

- a. Librarians managing the collections,

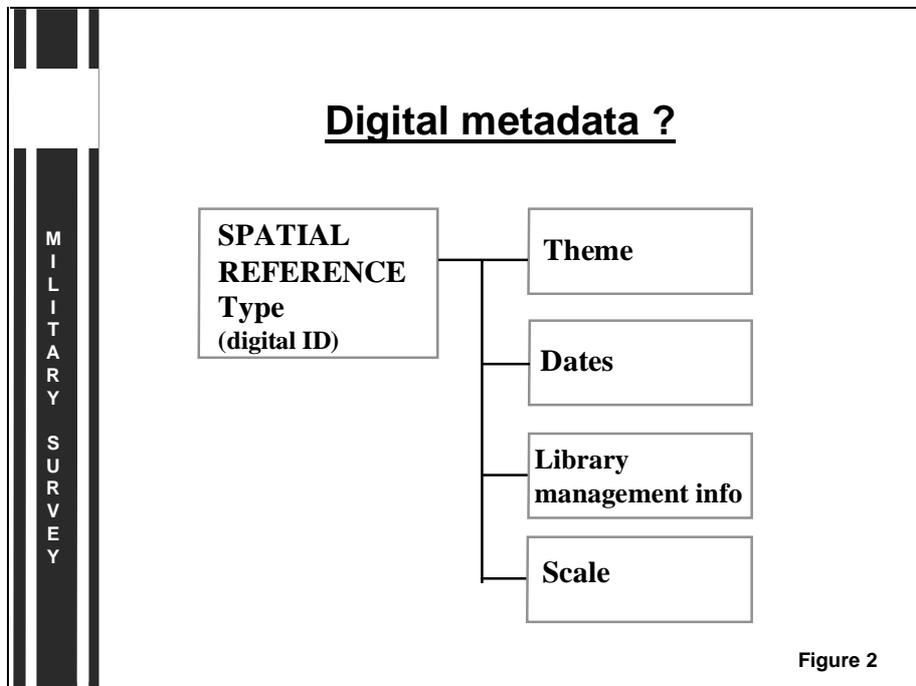
„Breaking With Conventions”: What Does Digital Spatial Data Require?

- b. Military Survey staff accessing the collections via databases or by direct access to the datasets,
- c. the defence community who can receive a service from the same librarians who are specialists in their regional areas.

This customer base will NOT change, but it increasingly requires rapid and direct access to geographic information.

Metadata

The description of digital data I have given illustrates that its metadata does not fall neatly into the existing map cataloguing conventions. Like the atlas collection, digital data may vary according to type, theme and scale within each medium. All such variations would require a separate SHEET record to be created, which is time consuming. Figure 2 illustrates that for digital data the metadata requirement might be much more succinct. It is anticipated that only thirty fields will be required, half of which would be equivalent to the map metadata set. In part, this will be because the ,digital' cataloguer will not actually see the data when the metadata is recorded, unlike the ,map' cataloguer. They will be more dependent upon the information supplied by the producer.



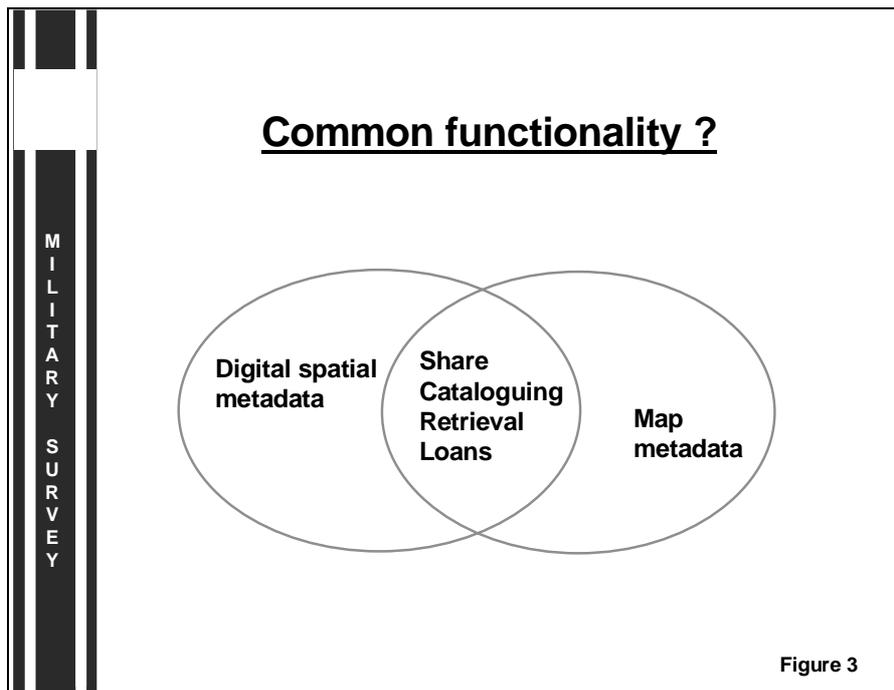
Unlike maps, most digital data will be identified by its spatial reference and type, for example Arc Standard Raster Product, rather than any sheet numbering system. A digital dataset could equate to one or more sheets of a map series, for example the Tactical Pilotage Chart series. But as the number of sheets on one disk is variable, there is little or no value in storing such metadata. The area covered on one disk is more dependent upon the need to optimise the use of disk space, which itself will continue to evolve.

Functionality

Similar search requirements to those used in MODMAP are required. Queries to a digital dataset will usually require a spatial reference, such as latitude and longitude or grid co-ordinates. They should also allow for additional text criteria to make the query more specific, for example by theme, scale, date or publisher. You may recall that an identical statement was made earlier about the search requirements of the manual system. Common library management functions would be needed to support the capture of metadata, facilitate loans

„Breaking With Conventions”: What Does Digital Spatial Data Require?

and provide a data enquiry/request mechanism. Integration of the datasets would be a logical progression to share common functionalities between the different systems, which could remain physically disparate (see figure 3).



HOW MIGHT WE ACHIEVE THIS ?

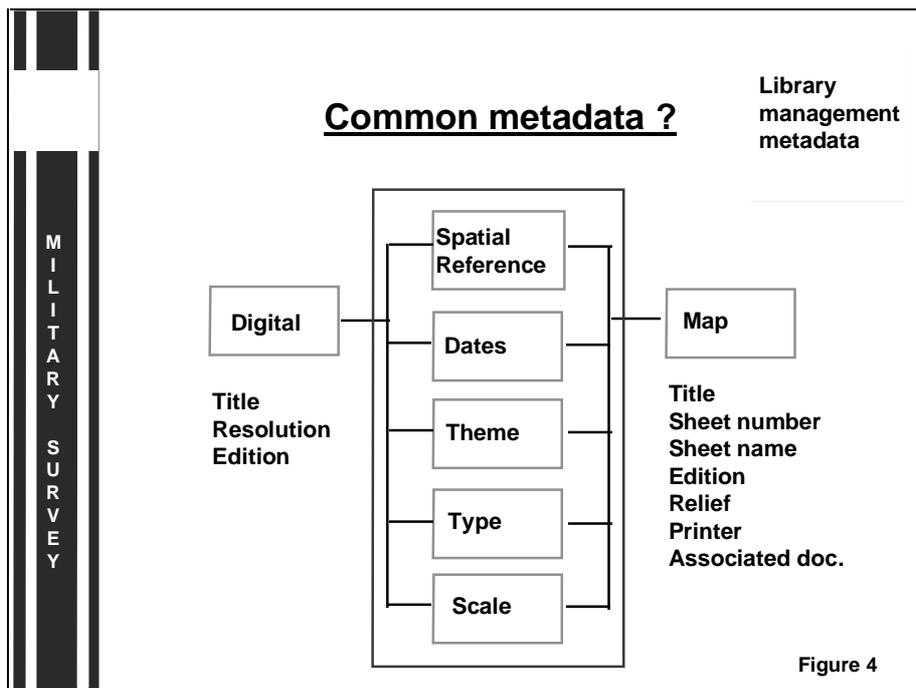
There are a number of possible solutions to ensure that we meet the increasing and changing requirements placed on our Libraries. I will outline just two. The first will be to *create a database specifically for digital data*.

Digital metadata doesn't easily fit into the existing MODMAP database structure. To accommodate digital metadata the software would need amendment. A database designed specifically for digital metadata would ensure that the minimum time is spent capturing metadata, possibly onto one or two pages of a form rather than MODMAPs current eight pages!

A Geographical Information System (GIS) might be considered the ideal tool for accessing and working with a selected digital dataset on screen; metadata may be stored as attributes ,attached' or linked to the spatial reference. In the case where digital data is stored on one hard disk, this approach would work. GISs, however, are essentially designed for production needs rather than the retrieval needs of a metadata system.

Single Database Option

The digital dataset could be placed on the same database as the map and atlas collection. A modified data structure would be required. It could present the opportunity to merge the metadata sets as illustrated in figure 4, or to keep them as separate entities with a common spatial reference if required.



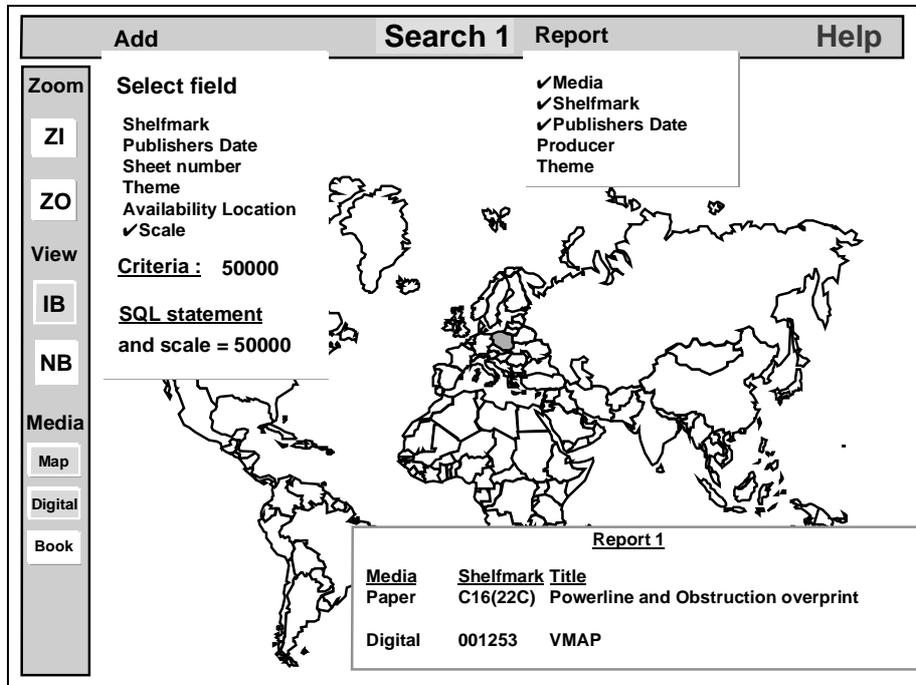
A large Relational DataBase Management System (RDBMS), would allow common functionality such as loans, cataloguing and retrieval to be applied to

„Breaking With Conventions”: What Does Digital Spatial Data Require?

all metadata sets. The adoption of the same software for all databases would also minimise the requirement for multi-skilled staff to support it. Many efficiencies could also be achieved by unifying the collections. Merging the libraries would create a nucleus of skills and expertise, and allow for the adoption of common operating procedures throughout. This is one aspect where Military Survey Libraries could be considered as being ‘behind the times’. Combining the storage of media poses other difficulties, for example providing map presses and shelving together could not be considered as an efficient use of space.

Common Search Tool

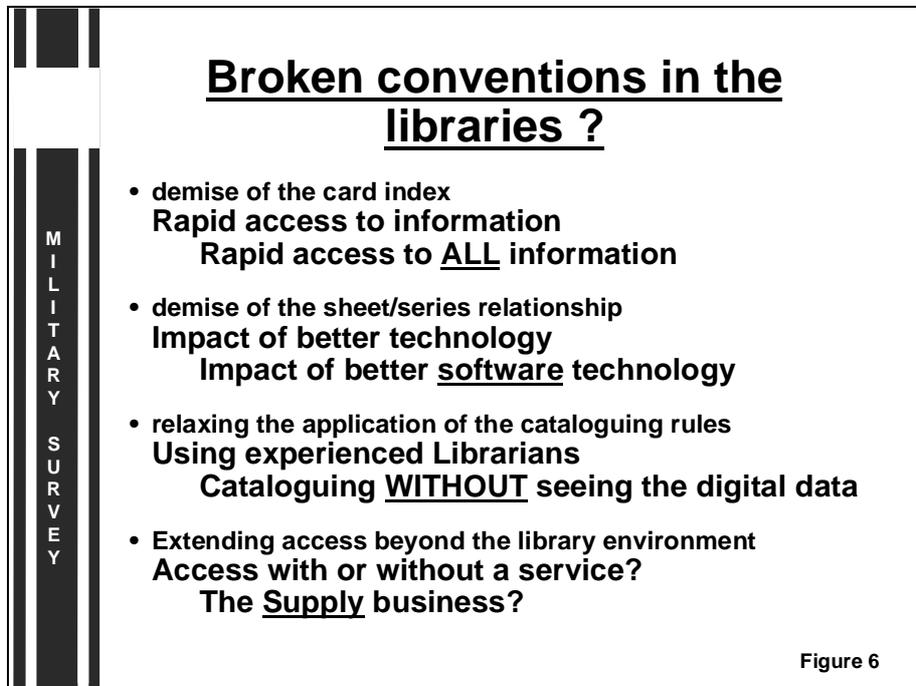
Whether storing the metadata on one or more databases, a common search tool could be utilised. The Commercial Off The Shelf (COTS) software which is available today is capable of accessing one or more databases, regardless of who produced the database. ‘User-friendly’ browser technology could be used to access the disparate collections, a world map would be an ideal starting point for each query. Zoom functions would enable a user to accurately select the required search area using a rectangle or by simply ‘clicking on’ a country. A user would then be able to add additional ‘text’ criteria using metadata associated with all the datasets, as illustrated in figure 5. A single enquiry would access *all* the metadata sets.



CONCLUSION

Geographic information is available on a variety of media. Identifying the appropriate information currently requires separate enquiries to separate metadata systems. Our goal will be to access all metadata sets with a single enquiry.

The four conventions I spoke of earlier could easily be re-written as shown in figure 6.



The use of modern computing technology is necessary to achieve this, but at what cost to

- a. the knowledge of the librarian and
- b. the service offered by the Library.

Digital metadata may force us to consider redesign of our database or databases. To the researcher, it will not matter as new search tools will provide the opportunity for them to save time by making a single enquiry.

In the same way, digital data presents no additional problem to that posed by the multi-scale and themed atlas. The further breaking with conventions is likely to be of a similar nature, it's not what we do but the way that we do it. However, the issue of providing access to metadata without sight of the product may remain. Today, our Librarians draw upon their experience and knowledge of the map collection to help a customer, our metadata systems offer no such similar advice. In a world where the statement „we accept no

SUE ANTONELLI

responsibility” is commonplace, perhaps we are moving into the world of ,supply’ rather than ,service’.

I will conclude with the question that if our metadata systems are populated without seeing the data, then what sort of service can the Librarian offer in the future?

©British Crown Copyright 1998/MOD

Reproduced with the permission of the controller of Her Britannic Majesty’s Stationery Office.

Sue Antonelli
Library Automation Branch
Military Survey
Government Buildings
Block A, Hook Rise South
Tolworth
Surrey, KT6 7AN, United Kingdom
Antonell@knox1,milsvy.gov.uk