A Distributed National Electronic Resource? MODELS workshop 6 report

5-6 February 1998, Bath

Rosemary Russell and Lorcan Dempsey, UKOLN, University of Bath

Introduction

A Distributed National Electronic Resource? was the sixth MODELS workshop and one of the largest to date. Around 50 invited participants[i] gathered at the Stakis Hotel in Bath on 5-6 February 1998, to discuss management and access approaches to the growing mass of currently unconnected resources provided by libraries, data centres, archives, subject gateways, electronic journals, clumps and others. We are grateful to Richard Heseltine of University of Hull, for once again chairing a workshop which posed a number of challenging management questions.

The workshop did not attempt to define what the Distributed National Electronic Resource (DNER) comprises. Discussion focused on what the 'distributed' might mean. Discussions on the 'resource' itself have been taking place within other fora. In a higher education context, the Content Working Group of the Committee on Electronic Information (CEI) has described the parameters of a national collection and has set out JISC policies for developing a DNER[ii].

The workshop discussed the MODELS Information Architecture (MIA). This is a framework for talking about distributed information resources with a shared vocabulary and set of concepts. It suggests some components of such systems and arranges them within a logical architecture. MIA encapsulates the design and development issues which have animated the MODELS series to date. MODELS 7, in May 1998, will apply the framework to a range of concrete service scenarios.

MIA assumes an environment in which users begin to access resources through 'intermediate systems', 'gateways' or 'brokers'. Examples of such intermediate systems currently being developed are:

- The clumps – these are creating systems which broker access to library catalogues.
- The systems support for some of the hybrid library projects (for example, the 'glue routines' proposed by the BUILDER project which will broker access to mixed media resources).
- Document delivery systems such as EDDIS.
- The system being developed to provide access to data sets at The Data Archive at the University of Essex.
- The Arts and Humanities Service Gateway – which will broker access to AHDS service providers.
- The Aquarelle system, which brokers access to museum resources.
- The ROADS cross-searching system – which brokers access to the subject gateways.
- A variety of commercial products which allow parallel searching of Z39.50 targets.
- A variety of systems in other areas – e.g. earth observation data.

What is common about these systems is that they present the user with a 'landscape' which hides some of the underlying differences between resources, collates returned results, and supports a higher quality of service than unmediated access to the end-resources themselves would. Resources accessed may be 'metadata' (library catalogues, subject services, …) which point to a mixture of print and electronic resources, or, in due course, a range of resources themselves. Such intermediate systems will be provided by libraries, 'aggregators', publishers, data centres, and others. Some will be quite simple (a parallel searching facility); some will add value in other ways (incorporating user profiles, service descriptions, and so on).

The aim behind MIA is to establish a shared view of some of the components of such 'intermediate' systems and the 'end' systems to which they broker access. In this way, we will move to a 'plug and play' environment for services providing they conform to the architecture agreed by consensus. MIA brings together existing best practice and standards, as well as highlighting areas which still need shared solutions. Many of the potential players were present at the workshop.

It is argued that without such 'brokers' or intermediate systems, the problem of dealing with multiplying, heterogeneous resources will become tedious and time-consuming for the user, and a large overhead for the libraries or other intermediary services.

There is little point in developing a DNER which is restricted to the higher education sector. It is significant that the JISC strategy document refers to cooperation and collaboration with other organisations and sectors[iii]. There have also been recent recommendations from various sources including Dearing, on developing a cross sectoral view, together with recommendations for resource sharing (eg the Anderson Report[iv]). *The New Library: the People's Network*[v] report emphasises the need to develop services in the context of a 'Common Information Framework. The boundaries between sectors and user types are likely to become increasingly blurred in future, as lifelong learning becomes a reality.

Significant discussion themes

Several particularly important themes emerged during the course of the workshop which were recognised as fundamental to the construction of the DNER and therefore formed the basis of some recommendations. These are discussed in context later in the text.

The first of these was 'information landscapes', a term used to describe the organised view of resources presented to users. Although there was some doubt about how landscapes might be constructed and who would 'own' them, the 'landscaped' view of information resources was seen to offer a compelling service structure for libraries and information providers. Landscapes are likely to move from very simple models initially (existing web-based resource guides), to more sophisticated implementations (customised, adaptive views based on user profiles and service descriptions), as supporting services develop. The landscape is a way of bringing users and resources together in helpful ways.

Secondly, it was recognised that the current network information environment is one of 'standalone' services, where the typical user is a human visitor. This means that services compete on 'user interface' issues as well as on content. Each service provider is a publisher of data and of a user interface. We are moving to a situation where there is more emphasis on access from intermediate systems. This means that services will compete on 'application or server interfaces', which support access through Z39.50, WHOIS++ (the subject gateways), or other protocols. For example, a 'clump' wants to hide individual OPACs; it wants programs to visit the OPAC and pull back records which may be merged with others before being presented to the user. The clump is interested in getting data out of the remote OPAC; it is not interested in its user interface. Currently competition between services is based at the user interface level – users tend to distinguish between databases by the interface and become accustomed to using one particular service provider. Increasingly, services also need to compete on the *server* interface level, so that they can be effectively accessed by intermediate systems. Another way of putting this is that the current environment is one of 'technical' access (services can be visited over the network); we are moving to an environment which will support 'information' access (services can make their information content available to other systems, which can use it to support other operations or re-present it to users).

A third important area of discussion was based around the need for market 'pull': libraries should start exercising more influence and control on the direction of future development. Now that the networked library and information environment is more mature, the issues that MODELS is raising are becoming more real and consequently the need for solutions is more pressing. However solutions will not come about unless there is market pull. The more informed the market becomes, the more it becomes a force for change. This reinforces the need for wide dissemination of MODELS results.

The MODELS Information Architecture: MIA

MIA has been developed progressively over the life of the project, with the input of a wide range of MODELS participants. It represents a high-level blueprint for how distributed information services are being designed. The library broker (or trading place) model lies at the core of the design. This clusters the components that have emerged as most desirable during the MODELS series and in surrounding discussions.

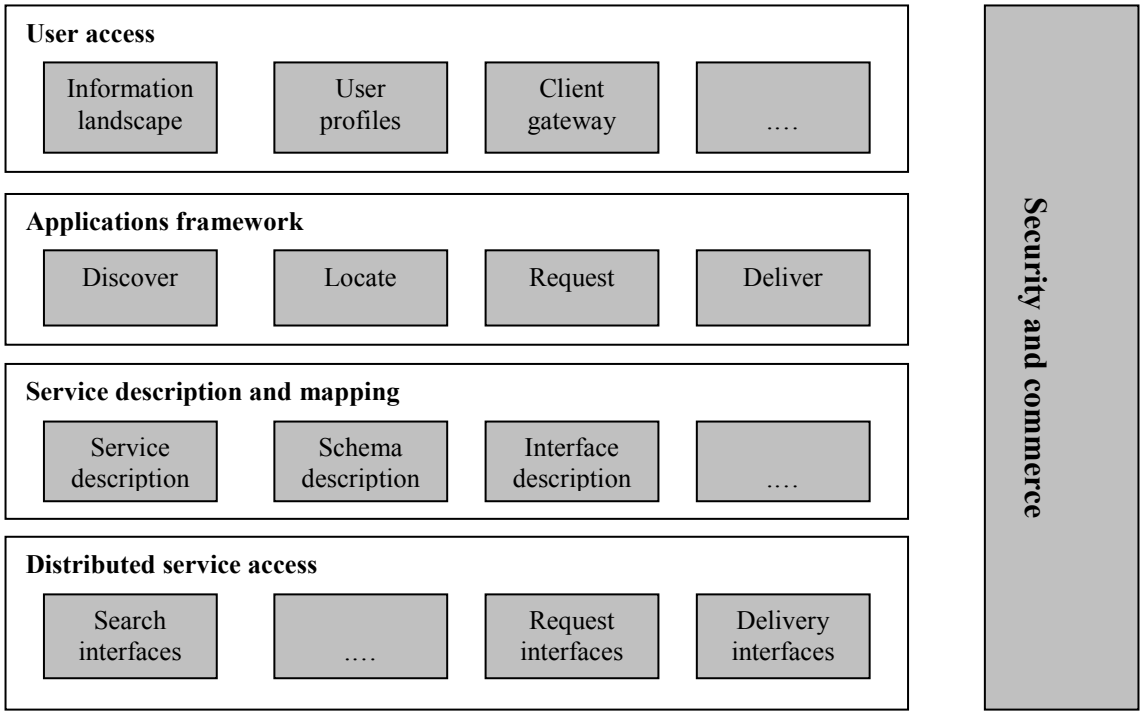

Figure 1: Library broker component examples

Again, this is not meant to be any particular system; it is meant to show the types of function that will improve the quality of service offered by brokers, and where agreement is useful. A clump might include the functionality represented by the boxes which align vertically along the left hand side (landscape, service descriptions, search interface). A richer broker (the EDDIS system for example) will include other functions. It might be implemented in different ways: as a monolithic system, in a more distributed environment, and so on.

The user access group of components looks out to the user, providing the 'landscape', capturing their interests, and providing the entry point for web browsers and other user systems. Among various functions it will convert user input (probably entered via a web browser) into an appropriate format to be used by the applications (discover, locate etc) in the next component group.

An applications framework is necessary to manage operations, to select interfaces and services, to pass data from function to function, and so on.

The service description and mapping group of components will provide a range of metadata for human users (eg collection descriptions), as well as machine-readable data for client systems. It will also support semantic interoperability, when a user wishes to search across several services, together with conversion of results between different formats.

The service access components include interfaces to support interaction with services (eg Z39.50 for searching, ILL for requesting an item). The broker might support several interfaces for services: for example, it might support the ILL protocol, the British Library's ART interface, and a bespoke email requesting system. Running across all these layers are security and commerce requirements, such as authentication and billing mechanisms.

The AGORA (eLib-funded) and Aquarelle (European Commission-funded) projects were presented at the workshop as illustrations of how MIA might map onto service structures. Whereas the AGORA project clearly originated alongside MODELS developments, Aquarelle was designed around two years ago in isolation from MODELS. It was therefore

interesting that the Aquarelle architecture was shown to fit reasonably neatly onto MIA. The Aquarelle *access server* is functionally equivalent to the MIA broker. The project reported that a range of problems had been experienced because the Aquarelle broker did not separate out the service description functionality. If the MODELS architecture had been available earlier, the broker components would have been designed differently. The new system being developed by the Data Archive was also noted to map onto the architecture.

The next MODELS workshop, to be held in May 1998, is specifically addressing the deployment of MIA. The overall aim of the workshop is to 'concretise' MIA – to take it one stage further into something practical, which can be *used* by the various parties involved: system vendors, information providers, and libraries purchasing services. An 'embodiment' of the architecture is needed, so the workshop plans to develop demonstrator scenarios.

Some of the issues which arose in discussion of components were:

Information landscapes

Information landscapes were introduced above. Whereas the concept had previously arisen at MODELS 5[vi], this workshop discussed the opportunities offered by landscapes in more detail. The concept bears different interpretations, but from a user point of view it essentially means that instead of being offered a mass of undifferentiated resources and services, a personal view is presented, based on the individual user's information needs. This view, or landscape may be constructed in different ways, and initially it is likely to be fairly simple – it could be just a set of web links. A more sophisticated service might dynamically match user profiles against service descriptions, to present an up-to-date view of available resources, taking into account subject interests and access privileges.

User profiles

These were seen as a place to record user preferences and access rights. Current systems have very limited interaction with the user: typically it is limited to a single search. The system does not remember anything about the user. There may be some merit in having consistent ways of creating user profiles, as users are increasingly mobile and variously affiliated.

Collection description

Each MODELS workshop builds on the work of the preceding workshops, and a number of issues explored previously re-emerged in a DNER context. Collection description in particular was recognised as one of the major stumbling blocks in the construction of the DNER. In a distributed environment, a mechanism for discovering relevant databases to search is required; it is impractical to search across all available services, so a way of providing information describing individual collections is needed. The metadata will be used in some cases by systems, and in others by end-users themselves.

Despite the standard differentiation between 'collection' and 'item' the workshop acknowledged that the real situation is often more complicated. Depending on the context, collections and items may exist more on a continuum, with no clear distinction between the two. There may be several layers of collections within collections. Collections will also appear differently to different users.

There are several possible approaches to creating and providing collection descriptions, although none may meet all the various requirements for access. A range of current options is outlined in the eLib study of collection description which is being coordinated by UKOLN[vii]. The use of Z39.50 Explain is included in the study. Explain does contain a collection description record. However there are no cataloguing rules to govern how it is used – the type of guidance provided by AACR is needed. Alternatively, the use of centroids (database representations based on indexes) would allow services to be self-describing. The ROADS project has implemented centroids, although it is not yet known if this option would be scaleable. UKOLN expects to carry out further technical work in this area: it is currently finalising discussions with a group of vendors about a sponsored post to carry work forward here.

Protocols and profiles

Although it is one protocol among several, there was considerable emphasis on Z39.50. It was noted that it was not widely implemented, and where implemented it was not widely used. Some reasons for this were discussed. It was recognised that there needed to be further agreement in the form of profiles, and that it would not support all required functionality. However, at the current stage of development there did not seem to be an alternative for search and retrieve which was in widespread use. Other protocols and formats were less discussed. However, there was considerable support for the proposed Interoperability Agency and the urgency of setting it up was emphasised.

MIA: overcoming the barriers?

The breakout group sessions on the first day identified a range of organisational and technical obstacles to building a DNER and providing appropriate user access. The following figure summarises the principal barriers and shows which of these are addressed by the various services provided within the MIA layers. It is assumed that technology is not a barrier to implementation.

Users

Barriers	
	Addressed by MIA?
Low awareness of resources	√
Low integration of services	√
Lack of 'contextualisation'	√
Deficiencies in metadata	
Lack of user involvement	
Behavioural conservatism	√
Variable skills	
Copyright	
Authentication	√
Competitive market	√
Inequality of access	
Infrastructural capacity	
'Lumpiness' of content provision	

Information Resources

Figure 2: Barriers to access

The figure does not mean to suggest that MIA's role is to overcome all of these barriers – for example it is clearly the responsibility of individual communities to address deficiencies in their own metadata. Some barriers may be generic and others may be sector specific. However the majority are shared concerns which require interoperable solutions. 'Lumpiness' of content provision (ie good coverage in some areas but not in others) is not just a JISC problem – it has to be addressed by the whole community together. Although it may seem significant, the higher education sector in fact controls a very limited portion of the general information resource.

It is therefore essential to take steps towards getting communities to:

- talk the same 'language'
- work together to complete the unfilled boxes in the table above.

However, the principal challenge is to achieve this:

- across the sectors
- across the purchaser/provider divide

- across the end-user/internal provider divide
- and also within individual sectors.

Possible strategies

The following possible strategies for developing the DNER were again based on discussions reported from the breakout group sessions on the first day.

The workshop acknowledged the low probability of national policy emerging:

- at governmental level
- within the sectors.

Other more probable strategies may be as follows:

- working through existing and new collaborative arrangements
  - institutional; within a sector; cross sectoral; purchaser/provider; internationally
- identification of champions
- the higher education sector should show leadership by working with the following organisations:
  - the forthcoming Interoperability Agency
  - SCONUL, UCISA, SCL
  - The Library Association
  - Library and Information Commission
- the library and information community as a whole should take more responsibility – it is too often assumed that JISC will take responsibility for moving things forward
- compelling examples are needed.

Areas for action

1. Promoting MIA

1.1. **It was recommended** that MIA should be presented at appropriate conferences and meetings during the autumn, since the MODELS workshop addressing the deployment of MIA will take place in May.
Action: MODELS 6 participants to raise with meeting organisers and make suggestions to UKOLN.

1.1.1. In the interim, MIA should be presented at the SCONUL conference, 30 March-1 April
Action: Jean Sykes to raise with SCONUL Executive.

1.2. **It was recommended** that MODELS 6 participants should themselves promote MIA within relevant fora.

1.2.1. ***Action:*** Michael Breaks to write an article for the *Library Association Record*.

1.2.2. ***Action:*** Mary Auckland to raise with the Library Association IT Group Committee.

1.3. **It was recommended** that an edited version of the text based on MODELS 5 (without the current public library focus) should be printed and sent to all SCONUL librarians together with a covering letter explaining that a demonstrator will be available that translates some of the concepts into practice.
Action: UKOLN to edit the text and send by 30 April

1.4. **It was recommended** that a consultative process should be associated with the DISI demonstrator initiative (see (5) below).

1.5. **It was recommended** that MIA should be 'concretised'. The aim should be to make it more fixed and ensure that it is accessible. It is likely that this will be realised as a result of MODELS 7 in May 1998.

1.6. **It was recommended** that a MIA 'glossy' should be produced
Action: UKOLN to produce a publication after MODELS 7.

1.7. **It was recommended** that another Z39.50 'bakeoff' should be organised linked to an event such as Libtech. It would be particularly useful to demonstrate cross-sectoral searching.

2. **Information landscapes and collection description**

2.1. **It was recommended** that a workshop should be held to explore more fully how information landscapes might be constructed, together with collection level descriptions.

2.1.1. ***Action:*** David Kay to draft programme

2.1.2. ***Action:*** Mary Auckland (CEI Collection Working Group Chair) and UKOLN to explore JISC funding to support the workshop.

3. **Working with other organisations**

3.1. **It was recommended** that a short promotional document should be produced outlining why MIA and related activities are important, together with the business and civic benefits. It should make the case for interoperability

across sectors. The document should be sent to LIC in particular, but also to other organisations with influence. It is important to lobby LIC now, so that when the government response to *New Library: The People's Network* is received, part of the agenda will already be in place.

3.1.1. ***Action:*** UKOLN to produce document.

3.1.2. ***Action:*** Michael Breaks to introduce the workshop report and above document to the LIC Research Committee.

3.2. **It was recommended** that the formal endorsement of other stakeholders such as the BBC and SCRAN should be sought. This will demonstrate that MIA addresses the concerns of a range of communities. Leadership from within the communities is needed.

3.3. **It was recommended** that communication with TLTP and CTI should be sought. This will again demonstrate the breadth of MIA's interest.
Action: TLTP and CTI representatives to be invited to MODELS workshop 8

4. **Interoperability Agency**
It was generally felt amongst workshop participants that given the fundamental importance of the Agency's role, it is important that it should be properly resourced. It is desirable to carry out a full range of activities in order to support the communities. The annual funding allocated for running the Agency is considerably less than was recommended by the scoping study report. However one of the reasons for basing the Agency at UKOLN was that greater value could be offered due to synergies with other UKOLN work. It is recognised that some parts of the Agency's work require a long-term view; longevity is a key part of achieving its goals.

5. **Demonstrators**
Some of the existing eLib document delivery projects (eg EDDIS) and the new hybrid library projects (eg AGORA) are already working on services which demonstrate MIA-based architectures. However further demonstrator work is needed to fire the imagination of vendors and other systems developers. It is much easier to rationalise the commitment of resources to development, if prototype services have been seen to perform.

With this in mind, a range of interested organisations including UKOLN, Fretwell-Downing, LASER, ILRT, UEA and BLDSC are collaborating on DISI: Demonstrating Information Services Interworking. The aim is to demonstrate an integrated service in business and social sciences. It will include discovery, locate, request and delivery of both print documents and network resources. It is planned to have a DISI demonstrator running in Summer 1998, based on AGORA.

6. **Galvanising JISC data centres**
Given the various types and formats of data held by the data centres, a number of difficult issues need to be addressed in order for the centres' collections to fit within MIA. The current model describes how a user might discover a data set. However it is likely that a separate mechanism would be available for *using* the data. It is notable that AHDS are dealing with equally complicated data and working out ways of coping with issues such as provision of Z39.50 access.

Currently data centres (and consequently also their users) see themselves as stand-alone entities; this should not be the case – they need to be woven into a service structure based on their data and not on their organisational name. Differentiation between products now happens at the user access level; however elaborate proprietary interfaces can be a barrier to access. In the hybrid library environment the database underneath is often of most concern and users will need direct access at search engine level via a broker (although other users will continue to come via the proprietary access route). In the future it is likely that differentiation between data centre products will occur when one is conformant with protocol-based services and another is not.

It is essential for MIA to encompass all data service types, including non-bibliographic.

7. **Galvanising vendors**
In order to galvanise vendors it is firstly necessary to raise awareness of MIA amongst libraries; vendors will not start to develop services until libraries include a requirement to support MIA in their system specifications.

The influence of ARL on vendor development in the US is notable; ARL libraries are starting to specify support for ISO ILL as a requirement. Critical mass is a clear benefit. (The intention of UKOLN to collaborate with the National Library of Canada and the Committee on Institutional Cooperation in the US on looking at how well Z39.50 supports emerging operational requirements was discussed.)

Acknowledgements
We are grateful to the participants of the 6th MODELS workshop for their contribution to these discussions and recommendations. We are pleased to acknowledge the support of JISC's Electronic Libraries Programme in funding this workshop, together with technical consultancy.

UKOLN is funded by the British Library Research and Innovation Centre, the Joint Information Systems Committee of the Higher Education Funding Councils, as well as by project funding from the JISC's Electronic Libraries Programme and the European Union. UKOLN also receives support from the University of Bath where it is based.

---

[i] We are grateful to CEI members and others for suggesting names of suitable participants working in this area. A list of MODELS 6 participants is available at <URL:http://www.ukoln.ac.uk/dlis/models/models6/participants>

[ii] Committee on Electronic Information - Content Working Group. *An integrated information environment for higher education: developing the distributed, national electronic resource (DNER).* <URL:http://www.jisc.ac.uk/cei/dner_colpol.html>

[iii] Committee on Electronic Information - Content Working Group. op. cit. paragraphs 24-27.

[iv] *Joint Funding Councils' Library Review: Report of the Group on a National/Regional Strategy for Library Provision for Researchers (The Anderson Report).* 1995. <URL: http://www.ukoln.ac.uk/services/elib/papers/other/anderson/>

[v] *The New Library: the People's Network:* <URL:http://www.ukoln.ac.uk/services/lic/newlibrary/>

[vi] A paper based on workshop 5 is linked from: <URL:http://www.ukoln.ac.uk/dlis/models/models5/>

[vii] The eLib supporting study on collection description is due to be completed in April 1998 and will be linked from: <URL:http://www.ukoln.ac.uk/dlis/models/studies/>