

**Interoperability between Information and Learning Environments
– Bridging the Gaps**
*A Joint White Paper on behalf of the IMS Global Learning Consortium
and the Coalition for Networked Information*

DRAFT – Version of June 28, 2003

**Neil McLean
Clifford Lynch**

*The authors acknowledge the contributions of Lorcan Dempsey, Kerry Blinco, Jon
Mason and other members of the IMS and CNI communities*

1. Introduction

The primary purpose of this paper is to explore potential interactions between information environments and learning environments, with emphasis on work that needs to be done involving standards, architectural modelling or interfaces in order to permit these two worlds to co-exist and evolve more productively. The emphasis here is on mapping the territory at a relatively superficial level, to identify questions and issues and to make some suggestions about what groups might most usefully undertake work in the various areas.

This is a difficult task for (at least) the following reasons:

- Both information and learning environments are complex in themselves, and are now rapidly evolving, which makes any mapping between them very challenging.
- Stakeholder groups have very different views of both the problem space and potential solutions; there are political and cultural issues, not to mention issues of control, that need to be considered alongside the technical questions.
- There is a constant tension between the pressures to solve immediate problems with “quick fixes” and the need to build sustainable, long-term open systems frameworks
- Terminological confusion abounds around the term “repository” and around the terms associated with managing digital content. Repositories are a particular problem in that numerous groups are building or defining systems that they describe as repositories – all of which are different!
- The library communities are largely unfamiliar with the work done by IMS and allied groups; IMS, in turn, is often unaware of relevant work happening in library automation, digital libraries, or related areas.
- The library communities are unfamiliar with the processes and terminology used within the IMS community.
- There is really no process or organization even roughly analogous to IMS that spans the relevant library communities, so it is difficult and complex for IMS to determine how to most effectively engage the many groups concerned with the development of information environments.
- The library perspective is only one dimension of the input required to further develop the IMS DRI framework. In particular, the remit of IMS is not limited to higher education; it also involves corporate, government and K-12 settings where very different assumptions about both infrastructure and the roles of libraries apply.

- The library and digital library worlds do not have a well-articulated, well-developed consensus reference model similar to the IMS Digital Repositories architecture.
- The cultural contexts are not well understood, yet have deep implications for more technical design choices. For example, it is not clear to what extent mutual distrust among participating entities needs to be designed into the work.

This paper is intended to be used as a basis for collecting further feedback on the issues, and also as the basis for a workshop scheduled for July 2003, after which it will be updated into final form.

2. Scope

Given the exploratory nature of this White Paper in addressing interactions between information and learning environments, the scope is necessarily fairly broad. The task of refining the requirements in terms of work items for specifications (within the process context of the IMS or other venues) is considered to be beyond the scope of the White Paper.

The discussion will build on the technical challenges identified in the 2001 IMS White Paper on Digital Repositories, namely:

- Identification of common functions which can be used across services to present a common interface.
- Harmonisation and/or mapping of meta-data at appropriate points, and mechanisms for handling multiple meta-data records describing a single object.
- Common functions for components to interoperate within institutional (single) learning and information management environments.
- Common approaches to authentication, intellectual property management, and trade.
- Bi-directional search gateways and use of parallel protocols for interoperation between environments and institutions.
- The role of supporting services including Directories and Registries, People and Profile information in access management, resource discovery, and trade in assets.

There will be discussion of the problem space in terms of networked information discovery and retrieval, and interfaces between repositories in the library domain and current learning management systems, interfaces between third party commercial information systems licensed by the library and learning management systems, and common standards to source people information across enterprise systems and library and learning management systems.

There will be some discussion of institutional repository management, metadata issues, authentication and authorisation, and digital rights management, however it is acknowledged that these areas go far beyond the scope of either the instructional management world or the digital library world alone; indeed, in many cases – for example authorization and authentication and DRM -- our conclusions are that work is best left to other groups, with some modest effort from the learning management and library communities to articulate specialized requirements and track progress.

Note that we have focused on technical matters here – standards, architecture, interfaces and the like. However, it is essential to understand that unless these efforts are accompanied by a reassessment of broader questions of organizational culture and of

roles and responsibilities within higher education institutions, and a renewed commitment for the various units involved to work together, it is unlikely that efforts to achieve technical interoperability will be sufficient to accomplish the real goals of more coherent management, access to, and stewardship over, the broad range of institutional assets relevant to teaching, learning and research.

3. Information environments: the library perspective

Libraries have accumulated considerable experience of working in distributed information environments, though these environments are still far from fully successful. Historically, most of the library focus has been on where they actually spend most money – on ‘published’ resources, which they buy (typically tangible materials such as books, videos, etc) or license (typically digital materials such as abstracting and indexing services, electronic journals, and so on). More recently, however, libraries have been investing more heavily in bringing other materials such as digitized rare and historical materials, and institutional research and learning resources into the distributed information environments.

There is growing acceptance that simply making resources available on the network without an additional layer of services may not be very effective. There are some clear reasons for this, arising from the characteristics of the current generation of network resources. In general, many of these characteristics flow from the fact that resources are made available at interfaces with low levels of interconnectedness between them. This in turn puts the burden of interconnection back on the user, and it means that in many cases the potential value of interconnection is not realized.

The following issues have constrained optimal networked access to information resources:

- Resources have been traditionally organized by supplier interests, rather than by scholarly or pedagogical value or user preference.
- Services have different access characteristics, may require individual login, use diverse data schema and exchange formats, and so on.
- Resources and services are autonomously managed; they have been developed independently with particular service and business goals – and this is particularly difficult when the collection of information resources includes commercial resources on publisher-operated systems and library-provided local materials. This means that within any information process, it may be necessary to interact with several services which do not coordinate their activities. Until recently, these services have been conceived and designed as standalone systems, rather than as parts of a fabric of information resources on a network. So, for example, there are services which allow people to discover the documents of interest to them, there are networked services which accept document requests, and there are packages which can format requests for dispatch to such services. These may not be linked up in such a way that an end-to-end process can be automated. Data may not cross boundaries, or may have to be re-keyed or transcribed by the user or by staff.
- The information providers wish to protect the value of resources they make available so in most cases they are tightly controlled with a variety of different terms and conditions. For the libraries, there is usually a need to confirm the identity of users or the integrity of resources. At present identity, access, and rights management services are provided on a service-by-service basis, thus creating significant

impediments to use. Indeed, it is not just the value, but the brand identity for many information providers, which means that they are actively motivated to work against seamless federation of resources.

- There are different aggregations of function. For example, a journal aggregator may allow people to discover, request, and have delivered a particular selection of journal articles from a particular selection of publishers, thus promoting the concept of a ‘one stop shop’. But in most cases they are still just components since no server will meet all coverage or quality of service criteria. This is not to say that ‘one stop shops’ are not locally useful, but they inevitably provide partial solutions.
- Resources and services typically say little about themselves to a potential user. Services require the user to have significant advance knowledge of what is available, as well as persistence if they wish to use multiple resources.

Many of the functional requirements discussed in this paper are common across the base of materials that libraries are attempting to manage and form into a coherent whole, as well as to learning management systems. These include identification, resolution, metadata harvesting, and distributed query. Given the variety of resource characteristics and the variety of user requirements it remains unclear as to what it means to create sustainable digital information environments.

To the extent that libraries simply want to incorporate collections of learning objects into their familiar distributed information environments (and we will argue that libraries will want and need to do much more than this!) there isn’t too much that’s unique to the learning object case – unfamiliar metadata that needs to be mapped, possibly new searching interfaces, questions about the granularity and aggregation of learning objects and their sub components, and the like.

4. Library interactions with the e-learning space

Library systems and e-learning systems actually will, in our view, need to interact in a rich variety of ways. We have already discussed the relatively straightforward case of simply incorporating collections of learning objects into the traditional library “world-view”. Here we consider some of the other types of interactions between the two worlds.

4.1 Linkage: Cross-system searching, import, export

Dynamic interactions, made possible by increased technical interoperability, means that teachers and learners begin to find new ways of developing learning activities, which in turn influence the way they use, or wish to use, learning and information content. Any analysis of library service interactions with learning systems environments must take into account the potential transformation of learning activity. This is no easy feat because there is so little documented analysis of e-learning activities.

Some analysis of the current state-of-play is necessary in order to provide a context for exploring new levels of functional interaction. Until recently, most learning and information content was tightly bound in learning management systems, although URL references to internal web sites were frequently used. Through this phase of development content has been incorporated in learning management systems in a variety of ways, mostly “hand-created” by the lecturer.

Transparent links between library systems and learning management systems have been rudimentary, and students have had to either logout of the learning management system to pursue a reference held in the library in an ereserve-type repository, or the lecturer has had to obtain an html or pdf document for mounting within the LMS. Note that particularly in the context of the TEACH act in the U.S., for example, there are going to be requirements to handle streamed audio and video, not just static documents.

Both approaches mean that the learning activity is relatively static as an exercise and there is no opportunity to move between information resources and the learning activity in runtime mode. Given this situation, there has been a lot of interest over the past year or so in developing links between library systems and learning management systems to facilitate seamless interaction during the uploading into the LMS by the lecturer of required resources.

IMS is currently reviewing limited areas of functionality relating to this technical challenge. Much of the current thinking is based on a fairly library-centric view of being able to “push” information resources into the LMS. There has been little thought given to the learner activity perspective where the learner may wish to draw on any number of information resources either prescribed, or of his or her choosing, at any given moment in the learning activity.

There is a need therefore, to develop more innovative use scenarios in order to map the dynamic functionality required in a “pull” runtime environment.

There are other linkage needs; for example, one should be able to initiate and manage searches of library resources from within the LMS easily.

Perhaps more interesting, and more challenging, than simple linkages, are the import and export of objects that to some extent inhabit both the library and learning management worlds. These might include information structures such as reading lists or personal profiles, for example. A much less clear case is the emergence of electronic portfolios by students; there seems to be little consensus about where these are housed, or indeed about the information policy considerations that surround them. But at the very least they need to reference and incorporate materials from both learning management systems and library environments.

4.2 Repositories and Stewardship

There is a major problem developing around the entire area of repositories. From the library perspective there is an emerging concept of institutional repositories; these are archival, stewardship and dissemination systems for content that have a fairly heavy policy component in terms of who can deposit, what metadata is required for deposit, acceptable formats and the implications of format choices for institutional preservation guarantees. Currently, there is no standardization of the interfaces to such repositories at a technical level (though MIT’s DSPACE seems to be positioning to be a widely-deployed system and as such might offer an example of a deployed de facto set of interfaces).

Again, speaking at a high level, there will be needs to place both learning objects and larger composite constructs (such as entire courses – in various sense, including both the course “frameworks” and actual populated “instances” of courses within such frameworks -- exported from learning management systems) for archival and records management purposes.

IMS has a concept of repositories for learning objects which is a much lighter-weight concept from a policy perspective, but specified more tightly from a technical point of view. It is clear that institutions will need learning object repositories that are more “transient” than institutional repositories; they will need access to commercial learning object repositories. We need to understand what needs to be done to make an institutional repository capable of functioning as an IMS learning object repository, but to recognize that we will be in an environment where from the IMS perspective there will be multiple repositories (which different policy frameworks); we will need capabilities to allow learning management systems to determine which repositories are available and which are appropriate for different purposes (i.e. transient vs. archival storage of objects). The current OKI repository interfaces do not understand this distinction well, and offer another case study of the problem that urgently requires attention.

It is much worse than a simple dichotomy between institutional repositories and learning object repositories. We also have the emergence of enterprise file systems, content management systems (at the enterprise or more granular level), and the appropriation of the term “repository” by a number of other peripherally related standards efforts as additional sources of confusion.

5. Institutional perspective: Libraries and Learning Management in the Enterprise Context

Most academic institutions are grappling with the implications of developing the digital campus which includes the two important concepts of digital information management and e-learning management. Central to both of these key management challenges, is the need to organize and manage the creation flow and use of content. In most institutions content is managed in silos which have little institutional interoperability, and as yet, the dream of aggregated reusable content remains unfulfilled. Given that this situation pertains across information management and e-learning management within the institution, a much more holistic view is required of the functional and technical service layers required to sustain required levels of service.

Part of the challenge here will be to appropriately define *institution-wide* infrastructure components and insist that all sectors within the institution share them. This includes authorization, basic identity management, authentication, perhaps persistent identifier systems at a minimum. Library systems and learning management systems should both be prepared to employ institution-wide services here rather than to create their own. Note that this raises some issues for IMS; the infrastructure services that one might reasonably expect to evolve within a higher education institution, or across a community of such institutions, is quite different that what one might reasonably expect from some of the other client constituencies for IMS (i.e. K-12 or corporate settings).

6. Overview of IMS Digital Repositories Interoperability Problem Space

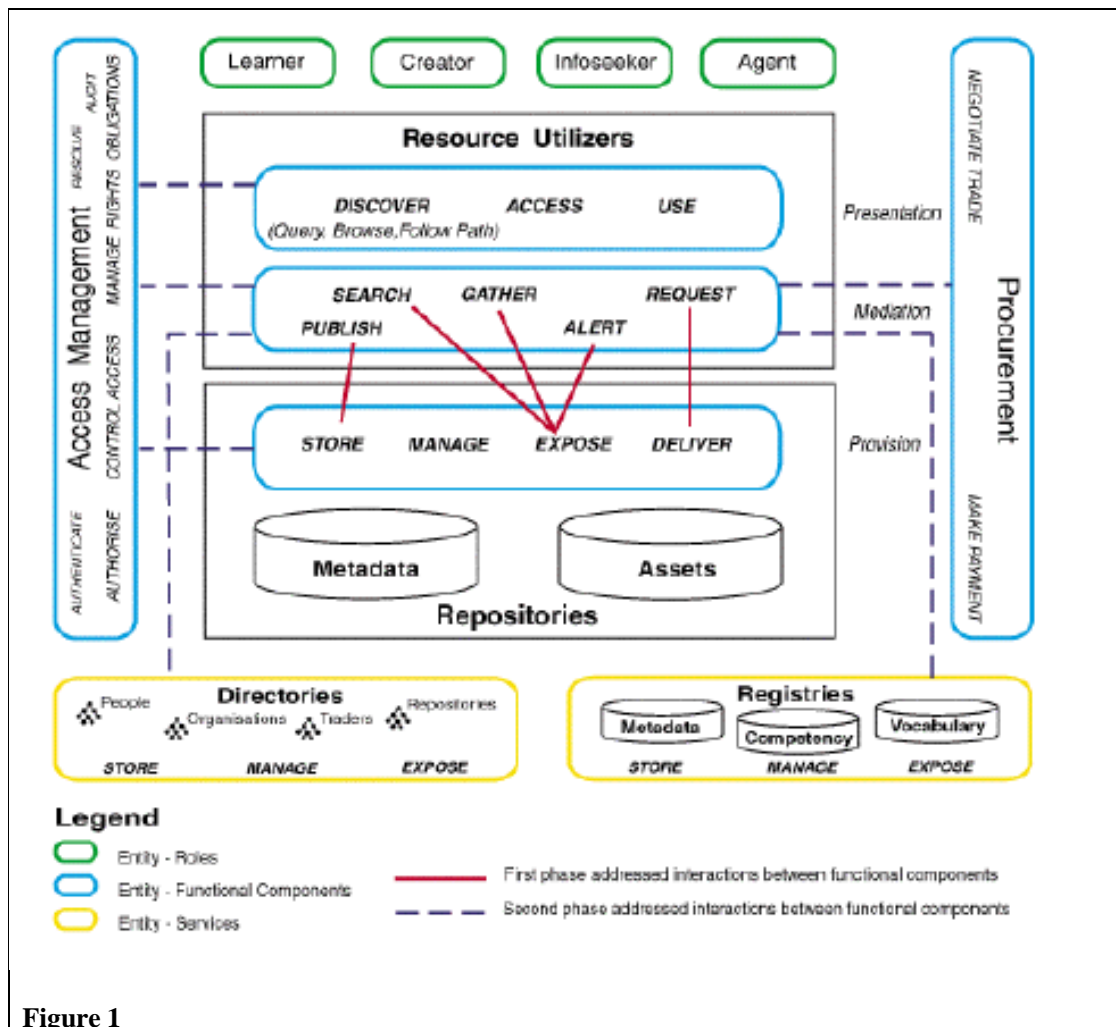
The principal focus of the IMS Digital Repositories Interoperability work falls into two categories:

- integration of e-learning with existing and emerging online digital services, and;

- newer repository technology to support the configuration, presentation and delivery of learning objects required for learner-centric learning to become a reality.

The Digital Repositories Working Group concentrated on identifying common functions, which can be used across services to enable them to present a common interface.

These common functions cover both desirable and necessary features such as authentication and authorisation, enrolment search, location and retrieval, IPR management, user preferences and profiling and transactions and search gateways across services. It was noted that “Learning object repositories share all the above (either directly or via the LMS they serve) but have the added dimension of supporting contextualised sequencing and navigation and potentially, dynamic branding of objects to a service in runtime.”



A generic functional architecture was developed as depicted in Figure 1 and this remains a principal point of reference for future work.

Of fundamental importance in developing this functional architecture, was the use of service models as the basis for identifying functional components. It is acknowledged however, that these service models were mostly based on “what-is” rather than “what might-be”, and it is important that the library communities engage in some quite lateral thinking with regard to potential service provision in the emerging interactive learning environments.

It is worth noting that the DRI Working Group decided to tackle only a small chunk of the functional architecture as the basis for the first specification and a great deal of work remains to be done to develop a complete set of DRI specifications.

7. IMS Digital Repositories Framework

From the IMS perspective, there is still much work to be done to “flesh-out” the Digital Repositories Interoperability Framework.

Much of the work to be done is interdependent with activity being pursued in other IMS Special Interest Groups and in other standards forums. A great deal of coordination is required to avoid the creation of yet more standards “silos”. A number of these areas are touched upon here; many bring together common interests of the library and IMS communities.

7.1 Search Services

The IMS DRI specification specifies only a very basic level of “search” and much more work is required in this area, including:

7.1.1. Search definitions for LOM metadata.

There are two approaches that could be conducted in parallel, namely:

- Look at the LOM as it exists now and define recommended practices for searching some, or all of the specified fields. This could be done in two stages:
 - The “abstract definition” of the searches, indexes required to support these searches and the “logic” one applies to the LOM to construct the indexes.
 - Implementation bindings e.g. Z39.50, SRW/SRU and XMLQuery, the Google API
- Conduct further research on what searches are really required (The Google vs. specialized search services issue). There is a need to start with use cases so as to find out how people really want to discover learning objects and other assets that may be developed or incorporated into learning objects. There is likely to be quite different requirements for people vs. software agents doing searches (see search services (7.1.2), and broad searches vs. need for precise searches to identify

small components. More granular models for Learning objects and “Lego” construction of metadata will influence different search services.

7.1.2. Definition of Search as a “Service”

The IMS Learning Design Group has expressed use cases for learning management systems accessing search services e.g. rather than having a static reading list, clicking on the “reading list” would trigger a request to a search service that would return a current list of materials that meet a predefined search criteria. Again this would be a two-part activity - defining the abstract search service, and developing implementation bindings. There is a need to look seriously at Web Services bindings where appropriate. With regard to a Z39.50 binding, it is probably a profile of Z39.50.

None of this work is about creating yet another search engine; it should take advantage of existing work and be implemented using existing technologies. The goal is to be able to obtain consistent results in distributed searches.

7.2 Modeling learning objects/learning activities and information assets.

There is a need to go back to basics for the development of a conceptual model for learning objects, which includes the concept of a distributed service.

In the context of this paper there is also a need to explore a model that shows how an object moves from being an information asset to being a learning object This would help to clarify what is unique about learning objects; in addition, the process model might be helpful in workflow, metadata and even IPR issues. It is interesting to note that we are beginning to see some signs of the emergence of a discussion within the digital library community about the need to develop common conceptual and architectural models for the “digital objects” that populate various library collections, and this discussion might be usefully connected to an examination of the foundational model for learning objects.

7.3 Identifiers, Locators, Persistent Locators, Resolving Services

The above terms tend to be used interchangeably within the IMS community, which is probably clouding the issues; e.g. the difference between an LCCN, URL, PURL and open URL.

Within IMS, the latest discussions have focused on the DOI which has features across the set and, whilst this has been a useful exchange, it overlooks the difference between these different basic purposes. Libraries and learning management environments should share common technology and infrastructure in this area.

7.3.1. Identifiers

In IMS, the term Identifiers has been used mostly to refer to locators. There is probably a need for many different identifiers. There is a need to do (4.2) in order to really understand what the extent of identifier requirements and the services and metadata needed to support them. The library community

has long experience in developing and maintaining identifier services and should have much to offer in this area.

7.3.2. Persistent Locators, Resolving Services

The critical issue in interoperability between repositories and other applications for both libraries and learning applications is having a persistent mechanism for retrieving the referenced object. Like authentication, this is not something that one community will resolve, or where there will be one solution, but one where the communities need to be well informed and need to apply influence so that a small set of well implemented solutions are used. There is a need also to maintain the metadata that is required to support persistence.

7.4 Metadata

Much of the debate in the learning communities standards arena revolves around the use of metadata to manage learning objects.

Having achieved a common standard, the IEEE Learning Object Information Model Standard (LOM), there is now a great deal of attention being paid to the means of implementing the standard.

As a result of the recent exchanges between DCMI/IMS/LOM experts, the following agenda has emerged;

- 7.4.1. To develop workable implementations of the LOM standard, based on generic types of learning objects.
- 7.4.2. To consider mechanisms for harmonizing the production and publication of good practice use studies.
- 7.4.3. To explore the conceptual and practical challenges of decoupling instructional context, learning objectives and content in learning object metadata, with particular reference to secondary or contextual metadata.

This is an ambitious agenda and a great deal of goodwill and persistence will be required in developing an action program over the next few months. It is clear that the library community will need to be extensively involved in these discussions, as well as in some of the mappings between this work and more traditional bibliographic structures.

7.5 Access management (authorization and authentication)

Access management in the library communities is a label for a range of activities involving identity management and authorization in the context of accessing systems and services. The two most common terms associated with access management are authentication and authorization and there are multiple approaches at the institutional level in terms of providing solutions. One very important development is the Shibboleth system for distributed authorization; this builds upon institutional authentication and identity management systems. The library community, at least in

higher education, has now largely developed a consensus that authentication and authorization should be treated as enterprise rather than library infrastructure, although the systems deployed must be capable of supporting library requirements for user privacy and highly controlled release of personal information, for example. We suggest that the learning management system community, at least in higher education, follow the same approach.

7.6 Digital Rights Management (DRM)

DRM remains a confusing, much-discussed, and controversial topic. There is a growing recognition that some specific types of rights enforcement may be needed in environments where learning objects and other library materials are shared across distributed networks in an e-learning context. In addition, there is a widely-held need in both the library and learning management communities to document rights for materials, and to provide automated tools to document rights for new, composite or derivative works by drawing upon rights documentation associated with existing works. In addition, there is a desire to use rights documentation metadata as part of the searching criteria for both library materials and learning objects.

In the area of rights documentation, it seems fairly clear that libraries and learning management communities share common cause and might well take some leadership, perhaps in collaboration with efforts like MPEG21 and the Creative Commons, among others.

The area of rights enforcement is much more complex and controversial. Here the issues involve so-called Rights Expression Languages (sometimes called rights enforcement languages, or restriction expression languages, or even restriction enforcement languages) and systems that interpret and apply expressions written in these languages. Here there is a great deal of activity taking place in other standards forums, such as OASIS, IEEE DREL Study Group, MPEG 21, etc. It is unclear what actions should be taken here by either the library community or IMS, together or jointly.

7.7 Portal Standards

Many developers working in the IMS space are very familiar with the capabilities of portal technologies, but there is limited knowledge within either the IMS or the library communities of standards in the portal space or the potential for integration with other architectures including web services, and information feeds (i.e. RSS). It is not clear how important it is to invest effort in this area, at present, however, particularly as a joint library/IMS undertaking.

7.8 People information for directories

The IMS communities and the library communities need common standards to source people information across enterprise systems and LMS's. Work is being done in the IMS Enterprise Working Group but the libraries need to have a lot of input on functional requirements from their perspective. They very often have the most comprehensive people database in the institution. This also needs to be connected closely to enterprise-level work on directories and related areas, such as

the EDUPERSON initiative. Note that this interacts as well with authorization and authentication.

8. **Shifting to New Conceptual Models**

The challenge inherent in this analysis is to find common ground that transcends the seemingly endless problems that pervade the current interactions between information and e-learning environments. A principal objective is to define common services and abstractions required to sit over multiple repository types so that they can be used effectively from within either learning environment applications or information environment applications. Of equal importance to interoperability within and between the information and learning environments is the definition and development of common systems and services, such as directories and authentication services.

This approach requires a conceptual shift away from a traditional systems architecture viewpoint to one where applications become defined by the services provided and the services that can be accessed. Groups of functions or services can then be put together more flexibly and new architectures developed which are not restricted by traditional views of systems applications and services.

For example, library systems have long had an integrated library management system connotation when it is now likely that there will be a series of (interoperable) applications covering OPAC's, ereserve, portals and document delivery services which may, or may not, be part of an integrated systems platform.

From another perspective, an OPAC search is a search service over a traditional library metadata repository, and is an example of the same service requirement in both an information environment application and a learning management application.

It will be vital that both the information and e-learning communities look to the possibility of applying service-oriented architectures such as Web Services, which are now the focus of attention in many other industries. All the familiar problems are being addressed in this service-oriented architecture environment, namely: security, identity management, authentication, authorisation, administration, privacy, confidentiality, transactions, registry solutions, web services orchestration and workflow solutions.

The IMS DRI functional architecture is capable of embracing these latest developments but there remains a need to identify common services that apply to both information and learning environments.

9. **Conclusion**

Where do we go from here? In this paper, we hope that we have supplied a reasonably comprehensive summary of the current key areas of necessary interaction between the e-learning and library worlds, and also some perspective on common services that both of these worlds should draw upon rather than re-develop parochially. The summary has of necessity been quite superficial in terms of technical details, but we hope that it will help to facilitate collaborations between people from the different communities that will need to work together to address the issues.

As we evaluate this draft white paper with stakeholders from the various communities, we hope that we will be able to sharpen the recommendations for action in the various standards and architecture areas, and perhaps also suggest some sense of priorities, given that much of the work discussed here will compete for scarce resources.

Finally, it is important to recognize that while we have stressed architectural and standards efforts in this paper, those are only a means to an end. As they move forward, they must be complemented by experimental implementations, test beds, and other deployment efforts to validate and refine the standards and architecture work. It is hoped that this white paper will also help to lead to the formulation of collaborations to advance such prototype deployment efforts.