Content Integration, e-Learning Standards and Middleware

May 6, 2005

Geoff Collier
Tyde Richards
Robby Robson

Eduworks Corporation
www.eduworks.com

This paper examines the use of middleware to support Content Integration, particularly in the area of e-Learning interoperability standards.

We would like to thank Recombo (www.recombo.com) for sponsoring this white paper.

The analysis and conclusions contained in this paper are those of the authors.

About the Authors

Geoff Collier (gcollier@eduworks.com) is a senior partner at Eduworks Corporation with many years of involvement in designing and developing student and learning administration systems.

Tyde Richards (trichards@eduworks.com) is a senior technologist at Eduworks Corporation and has been involved in the development of e-learning technology and interoperability standards while working with several companies including Eduworks, IBM, Macromedia and Apple.

Robby Robson (rrobson@eduworks.com) is president of Eduworks Corporation, Chair of the IEEE Learning Technology Standards Committee and director of the National Digital Science Library Reusable Learning project.

Copyright 2005 Eduworks Corporation

This work may be used under the terms of a Creative Commons Attribution-NonCommercial-ShareAlike 2.0 license (see http://creativecommons.org/licenses/by-nc-sa/2.0/)
# Table of Contents

1 Summary ........................................................................................................................................ 1  

2 Learning Technology Specifications and Standards ................................................................. 3  
   2.1 The Learning Technology Standards Community ................................................................. 3  
   2.2 Learning Technology Standards for Content Integration ....................................................... 4  
      2.2.1 The CMI / SCORM Model ................................................................................................. 4  
      2.2.2 CMI/SCORM Adoption ....................................................................................................... 5  
      2.2.3 New Developments ............................................................................................................. 6  
   2.3 Future Trends ......................................................................................................................... 7  
      2.3.1 Shift from data-centric to service-centric ........................................................................... 7  
      2.3.2 Impact of more generic specifications and standards ......................................................... 8  
   2.4 Challenges of Implementing Support for Standards ................................................................. 9  

3 Content Integration .................................................................................................................... 10  
   3.1 Approaches to Content Integration .......................................................................................... 10  
   3.2 E-learning Content Integration Middleware Vendors ............................................................. 13  
   3.3 Enterprise Content Integration ............................................................................................... 16  
   3.4 Content Integration for Learning Applications ........................................................................ 17  
      3.4.1 Learning Content Management in Large Enterprises ......................................................... 17  
      3.4.2 Enterprise Content Management vs. Learning Content Management .......................... 18  
      3.4.3 Emerging Trends in the Large Enterprise Sector ............................................................... 18  
      3.4.4 Higher Education Market Sector ....................................................................................... 20  
      3.4.5 Compulsory (K-12) Sector .................................................................................................. 20  

4 Evaluation of the Middleware Approach ....................................................................................... 21  

5 Conclusion ..................................................................................................................................... 23  

Glossary of terms and abbreviations .............................................................................................. 24  

References ......................................................................................................................................... 25
1 Summary

Interoperability among e-learning content and learning technology components is key to the successful implementation of a complex learning environment. The e-learning standards community has been working for many years on this challenge, with the goal of supporting meaningful interoperability through the development of a comprehensive and workable set of standards.

E-learning interoperability standards developed to date are intended support the following capabilities for interoperability between content and learning technology components. This is a highly summarized view of those standards:

- **Package** and unpackage content so it can be sent from system to system
- **Describe** content using metadata that can be interpreted by different systems
- **Communicate** between the content and the learning delivery environment, sharing information about the learner and learner interactions with the content
- **Navigate** between content modules based on learner characteristics, learning results and specified sequencing rules

Standards exist and are quite widely adopted by content developers and software vendors for all these capabilities, but standards alone cannot provide ‘out of the box’ interoperability for e-learning any more than they do for other types of technology. For example ‘plug-and-play’ desktop hardware and applications are based on well defined computer technology standards, but ‘helper’ software such as installation shields and specialized drivers are still needed before installation is successful and operating system / component interoperability is achieved.

E-learning interoperability standards are complex, they are continuing to evolve, and new standards are being developed to address additional interoperability capabilities. When these standards are adopted in any specific situation, they must be interpreted to fit the requirements of the technology, the industry and the subject matter being addressed. Before content developed and running in one location can successfully run in another location there is almost always a need to reformat, repackage or reinterpret the content in some way. This might include actions such as reinterpretation of run parameters, cross walking metadata taxonomies or converting content communication from one standard to another.

Content integration requirements are not unique to e-learning. Content integration is a topic that also applies to any content management technology dealing with digital content from more than one source.

There are three general solutions to these e-learning content interoperability challenges:

1. **Point to point integration**: Programming is developed by each vendor (or worse, by each implementation site) that performs the necessary content transformations / repackaging for each and every content and delivery platform variation that is encountered. This solution is expensive, error-prone, time consuming, tightly coupled and fragile.
2. **Single content management system:** The enterprise implements a common Content Management System (CMS), or Learning Content Management System (LCMS), and imports all learning content into that environment where it is repackaged and converted. The content management software manages the delivery of content to the user. This is a workable solution where it is desirable to bring all content into a central location.

3. **Content integration middleware:** This solution is the primary focus of this paper. Content integration middleware for e-learning is a relatively new class of software, offered by companies such as Recombo, Trifus and eLearning Union. The middleware provides a service that dynamically converts the content from one implementation of e-learning interoperability standards to another, at the time the content is needed. This allows content to be served to many learning delivery environments; in the form those environments need for interoperability.

The analysis in this paper leads the authors to several conclusions:

- Content integration middleware is a good addition to the options available for e-learning content interoperability. Middleware is particularly suitable to situations where the content being used originates from a number of distributed sources containing content that is frequently updated.

- Middleware is a good option for content management and learning management system vendors to consider for addressing their compliance with e-learning interoperability standards. By integrating middleware into their offerings, these vendors eliminate the need to constantly release new software to stay abreast of new developments and implementations of standards, and they can take advantage of the specialized content integration knowledge of middleware vendors.

- Content integration middleware is dependent on the development and adoption of e-learning interoperability standards to be successful. Without standards, middleware is just a collection of point-to-point solutions. It is standards that allow middleware to be effective and flexible and integrations to be cost effective.

- The work performed by content integration middleware is necessary, no matter how mature and complete e-learning standards become. Therefore, just as e-learning standards allow middleware to be successful, the logic contained in content integration middleware (or in content management systems) allows e-learning standards to be successful by aligning differing, localized adaptations of those standards. As IBM says in its corporate positioning, “Middleware is Everywhere.”

In this paper, the authors examine the complexities of content integration for e-learning, and of implementing the standards and specifications that have been developed to support this integration. They then consider the applicability of middleware solutions to this problem space.
2 Learning Technology Specifications and Standards

Some interoperability standards developed by the learning community have gained wide acceptance and use, but these standards allow considerable flexibility in their implementation. As a result, technology vendors tend to interpret and adopt them in slightly different ways. Also, the existing standards continue to evolve, resulting in multiple versions that need to be supported, and standards addressing new areas of integration continue to be developed and rolled out to the learning community. Technology vendors each have their own product development cycles and release schedules, which means that different vendors are at varying points of adoption for the different standards.

In successful e-learning implementations, content from many different sources needs to be accessed and presented as an integrated part of courses and performance support environments. This situation is more complex in organizations that have multiple learning management systems from different vendors and need to share learning content across these different systems. This ‘many’ to ‘many’ scenario (many sources of content running in many learning management environments) is not an uncommon situation in larger organizations.

The adoption of standards can get an e-learning environment part of the way towards integration, but there are always final hurdles that have to be cleared before each integration point works successfully. Different vendors may have implemented different standards, different versions of the same standard, or different interpretations of the standard. So, even if an organization has selected content and technology vendors who do support standards, content integration work always remains to be done. In many cases there are also legacy technologies that do not provide support for any standards.

This section of the paper provides a short history of the evolution of key e-learning content interoperability standards, and takes a look at future directions for these standards. This forms the context within which e-learning content integration solutions must operate now and in the future.

2.1 The Learning Technology Standards Community

In 1992, the Aviation Industry CBT (Computer-Based Training) Committee (AICC) was formed to respond to unique requirements for the sharing and interoperability of computer-based training content and technologies in the aviation industry. This industry invests a great deal of money in training on aerospace technologies that can have life-spans of more than 30 years. They wanted to avoid the need to redevelop digital content each time the underlying training delivery technologies changed.

The AICC was the first widely recognized international group to develop e-learning interoperability guidelines. The group also provides an open forum for the discussion of training technologies and interoperability that is not limited to the airline industry.

Over time, a more widespread interest in e-learning (triggered by the growing ubiquity of Web technology and other trends) resulted in multiple efforts to develop and standardize e-learning interoperability standards. This included initiatives such as the IMS Global Learning Consortium\(^1\), the Advanced Distributed Learning Initiative (ADL)\(^2\), the Ariadne

\(^1\) IMS Global Learning Consortium - [www.imsglobal.org](http://www.imsglobal.org)
\(^2\) Advanced Distributed Learning - [www.adlnet.org](http://www.adlnet.org)
Content Integration, e-Learning Standards and Middleware

Foundation\(^3\), the IEEE Learning Technology Standards Committee (LTSC)\(^4\), the Schools Interoperability Framework (SIF)\(^5\), and others.

This organizational complexity is at least partly caused by a variation in interoperability requirements along the dimensions of geography and market segment.

### 2.2 Learning Technology Standards for Content Integration

#### 2.2.1 The CMI / SCORM Model

The ADL’s Sharable Content Object Reference Model, or SCORM (ADL, 2004) is a direct evolution of the AICC Computer Managed Instruction (CMI) model (AICC, 2004). Collectively, SCORM and CMI comprise the dominant model for learning content integration. This model assumes that:

- There is a need to share units of reusable content that are smaller than a complete course.
- Content can communicate with learning management and delivery software through a common data model and data transport mechanism.
- External documents can be used to describe the structure of a learning experience or course of instruction in terms of its content components and external documents can be used to ascribe descriptive or behavioral metadata to that content.
- Conformance to the model can be determined through software-based conformance testing.

The original CMI model was defined by the AICC in 1992. It is intended to support computer-based training that tracks student performance. Tracking is an issue of critical importance to the airline industry where tracking an individual learner’s activities and results is needed to meet stringent regulatory requirements. Prior to the development of the CMI specification, this need was met through proprietary content communication models. This resulted in content being tightly coupled to the learning delivery software for which it was developed. Content developed to run in one environment could not be run on software developed by a different company. The goal of the CMI specifications is to enable CMI-compliant content developed by any vendor to be run on any other vendor’s CMI-compliant learning delivery software.

First released by the ADL in January 2000, SCORM provides a Content Aggregation Model (CAM) and Run-Time Environment (RTE) for learning objects. CAM supports the sharing of content objects between environments and RTE allows these content objects to communicate with the environment that is running them.

In 1999 ADL adopted the CMI approach as the base upon which SCORM was built, but the two approaches have diverged technically over time. The AICC has a significant ‘installed base’ of software and content vendors, and has therefore been conservative in making changes to the CMI model. On the other hand, the ADL has made significant changes to SCORM by replacing components with other solutions that are intended to be more modular and extensible. This includes the adoption of the IEEE Learning Object

---

\(^3\) Ariadne - [www.ariadne-eu.org](http://www.ariadne-eu.org)

\(^4\) IEEE Learning Technology Standards Committee (LTSC) – [ltsc.ieee.org](http://ltsc.ieee.org)

\(^5\) Schools Interoperability Framework (SIF) - [www.sifinfo.org](http://www.sifinfo.org)
Metadata structure for descriptive metadata (IEEE, 2004), IMS Content Packaging (IMS, 2004) as the framework for resource aggregation, and more recently the adoption of IMS Simple Sequencing (IMS, 2004) for controlling the sequence in which a learning delivery environment adaptively presents learning objects to a learner.

### 2.2.2 CMI/SCORM Adoption

In general, the interoperability standards developed by the AICC, ADL, IMS and associated groups address the following areas:

- **Package** and unpackaged content so it can be sent from system to system
- **Describe** content using metadata that can be interpreted by different systems
- **Communicate** between the content and the learning delivery environment, sharing information about the learner and learner interactions with the content
- **Navigate** between content modules based on learner characteristics, learning results and specified sequencing rules

During the 1990’s the AICC model for packaging, description and communication was adopted by vendors providing learning content and learning management technology to the U.S. aviation industry. In the late 1990’s there was a spike in broader adoption driven by a new generation of ‘enterprise-level’ learning management systems that were being sold to the corporate / government training markets. These systems are intended to provide centralized management of learning and learning content for large organizations, and as such they need to use and communicate with content developed by various vendors and authoring technologies. The AICC CMI model was seen as a key technology enabling a move away from the tight coupling of content authoring with proprietary delivery platforms.

In the early 2000’s the SCORM Content Aggregation Model and Run Time Environment were rapidly adopted by those vendors targeting the corporate / government training markets. Vendors who had already adopted the CMI model extended that solution to include the variations incorporated in SCORM. New software products will typically support SCORM and most also support the AICC specifications. SCORM is a key specification in the corporate / government space because conformance with SCORM is required to sell technology to the U.S. Department of Defense and other government departments (DOD, 2001). The current SCORM models cover all four of the interoperability areas summarized above (package, describe, communicate, navigate). Most vendors typically try to support all four areas, but Navigation is the most recent and untested area of standardization, and as such is less widely adopted than standards in the other three areas.

Some SCORM adoption is occurring outside of North America and in learning market segments other than corporate / government.
2.2.3 New Developments

There are a number of new initiatives underway that will continue to complicate the standards picture over time. The following is not an exhaustive list, but it provides a quick overview of some initiatives likely to impact the structure of e-learning standards in the relatively near term.

As these initiatives mature, the successful ones will allow some unaddressed interoperability issues to be resolved. This creates opportunities, and a new set of challenges to be addressed by middleware and content management software, and will open up new roles for the software.

CORDRA

The ability to search for and retrieve relevant content from repositories is a critical capability missing in SCORM. The ADL is defining a repository framework known as CORDA that recommends the use of various specifications and standards, including the CNRI handle system (CNRI, 2005). CORDRA is an acronym for Content Object Repository Discovery and Registration/Resolution Architecture.

This work is in the early research phase and is a long way from being ready for adoption, but it is being positioned by the ADL as the ‘next big thing’ after SCORM. See (Rehak, 2004) and other documents on the LSAL website for information on CORDRA.

SCORM/S1000D Integration

There is a significant activity underway within the ADL in the Department of Defense context to integrate SCORM with S1000D-compliant technical documentation for performance support applications. See (S1000D, 2005) for information on the S1000D specification.

Reusable Competency Definitions

There is work underway within the IEEE LTSC group to define an interchange format for competencies and learning objectives. This effort is based on earlier work by the IMS Global Learning Consortium. Competencies and reusable competency definitions can also be used as a key way to link competencies and learning resources to an instructional strategy.

XML Binding for the CMI Data Model

The IEEE LTSC has a draft standard that is nearing the status of an official standard and that is intended to act as a key component in a future Web service for SCORM runtime communication.

Offline Player

The AICC is developing a specification intended to work with IEEE XML Binding of the CMI data model to support disconnected use.
**PENS – Package Exchange Notification Services**

The AICC has developed a draft specification for Package Exchange Notification Services (AICC, 2005). PENS helps automate the process of transferring content from authoring systems to multiple systems that need to acknowledge the receipt of a content package, such as learning management systems and digital repositories.

**ADL Research on XML Content Formats**

The ADL is sponsoring research into XML content formats to define the internal structure of sharable content objects, and to enable the style and presentation of content to be automatically modified in different learning environments.

**SSP – Shared State Persistence**

The IMS has developed an extension to the CMI data model called Shared State Persistence. SSP makes it possible to share data across multiple Shareable Content Objects or SCOs. It will take time for all learning management technologies to implement SSP, and because it greatly increases memory requirements some developers may be reluctant to do so. This kind of functionality is easy to add through middleware, which can also be used as a place where the additional and fluctuating memory demands can be met.

### 2.3 Future Trends

Through most of its history the Learning Technology community developed standards without significant input from external, yet related, standards development initiatives. This is beginning to change in some key areas.

There are some important trends towards cooperation and the sharing of standards with other communities. As these come to fruition, some existing e-learning standards may be de-emphasized in deferral to more general standards, and in other cases more options will open up for addressing interoperability challenges. In any case, this increasing complexity and depth of standards argues for solutions such as middleware, that consolidate this complexity and evolution into one place where it can be managed and resolved by an organization and a product focused on nothing but content interoperability solutions.

*It’s not about interoperability standards; it’s about the interoperability of standards.*

**2.3.1 Shift from data-centric to service-centric**

Early work on Learning Technology interoperability focused on defining data models. Recent initiatives demonstrate an increased interest in thinking of the problem in terms of services. This view is in keeping with a general move towards a services view in the wider technology community. These services are becoming more distributed and loosely coupled (e.g. Web services).
Research and development is underway on standards for learning technology services, but this work is in the early phases of the standards development life cycle. Frameworks under development include the Open Knowledge Initiative (OKI, 2005), Sakai (Sakai, 2005), and the IMS General Web Services (IMS, 2005).

### 2.3.2 Impact of more generic specifications and standards

The learning technology community has tended to develop its own standards from scratch, rather than by extending or ‘profiling’ standards that have been developed by other industries to address similar requirements. Examples of this include content packaging and learning object metadata, areas that have also been addressed by the wider library management and content management communities.

More recently there has been recognition that many interoperability requirements relevant to learning applications are not unique to learning. Some specifications or standards developed by other communities can be applied to learning, with appropriate adaptations. This allows the learning community to leverage more generic standards that may already have significant support on major technology platforms. An example of this trend is the approach being taken to Rights Expression Languages (IEEE, 2004). Rather than developing its own standard for expressing and managing digital rights, the learning community is examining standards already in use by other communities, such as the ISO-approved MPEG Rights Expression Language, and determining how these can be applied to the requirements of the learning community.

Other external specifications and standards that are relevant to e-learning include:

- **JSR 170** – A platform independent Java API for content repositories. JSR 170 models content trees made up of nodes having arbitrary properties that may be either data (content) or metadata. This specification is expected to be widely supported.

- **MPEG 21 DIDL** – A specification for packaging and transmitting content packages between system components. Similar to the IMS content packaging specification used in SCORM but also offers superset functionality.

- **HR-XML Schemas** – XML schemas defined to support the exchange of Human Resources data. There are several schemas that touch on areas similar to some e-learning specifications, including: Assessments, Education History, and Competencies.

- **RSS** – RDF Site Summary (RSS) is emerging as a powerful way to syndicate dynamic content. RSS and other content distribution tools are starting to be applied in the learning world. RSS is gaining in popularity, and it is easy to use. (RSS, 2001). **Atom** is a rival format that does similar things (Atom, 2004). When syndication services are integrated in to learning environments, they will broaden the potential sources of learning content, and increase the need for dynamic content integration capabilities.
2.4 Challenges of Implementing Support for Standards

The use of interoperability standards is only one feature in a long list of development priorities for e-learning technology vendors. In general, interoperability is a feature that customers want to work smoothly, behind the scenes. Achieving interoperability between content and learning delivery environments should not require undue effort from a vendor's clients. From the vendor perspective, continually having to assist clients with interoperability problems can add greatly to the cost of ongoing customer support and software maintenance.

This challenge is compounded by the continuing release of new and revised specifications and standards from the ADL, IEEE, AICC, IMS and other organizations. Learning management software vendors tend to be relatively small companies when compared to organizations like Oracle or SAP, and as such they must focus their engineering priorities on critical differentiating features, leaving support for standards as a lower priority during a product release development cycle. As a consequence, many vendors do not support all e-learning interoperability standards, or the latest versions of those standards.

Each version of a standard that a technology vendor chooses to support adds considerable cost and time to the testing of product releases. Historically, LMS (learning management system) providers have been responsible for testing and tweaking their products for every content provider they intend to support. Even though most of these content providers may support a version of SCORM or AICC, the technology has not proven to be plug-and-play due to different implementation assumptions made by each content provider.
3 Content Integration

Content Integration is a "process that organizes content into meaningful packages and then integrates these into enterprise software systems for people to use." (Forth, 2004a) This is a capability relevant to any organization dealing with content stored in many different places, in multiple formats.

The goal of enterprise-level content integration is to make the different types of information needed by an individual in an organization accessible through a single search and retrieval interface. This includes content stored on internal servers, content accessible to the organization through external servers, and information available on the general Web. Some information is structured (i.e. databases) and much more is unstructured (i.e. documents).

This general content integration challenge is directly relevant to e-learning. Content from many different sources needs to be accessed and presented as an integrated part of e-learning courses and performance support environments. Larger organizations may have multiple learning management systems from different vendors implemented in different departments, but they still need to share learning content across these different systems.

The diagram below (Forth, 2004b) provides a visual overview of the content integration process:

3.1 Approaches to Content Integration

As discussed above, while standards can get an e-learning environment part of the way towards integration, there are the final hurdles that have to be cleared before each ‘vendor to vendor’ integration point works successfully. Different vendors may have
implemented different standards, different versions of the same standard, or different interpretations of the standard. In many cases there are also legacy technologies that do not contain support for any standards. So, even if an organization has selected content and technology vendors who do support standards, content integration work always remains to be done.

There are three general approaches that can be taken to content integration.

**Point to point integration**

The first approach to content integration is to provide unique code to integrate each variant of content with each system that uses it. This approach results in software and content vendors developing proprietary interfaces for the interaction of their technology with technology from other ‘key’ e-learning software and content vendors. E-learning implementation teams then modify these proprietary interfaces to meet local requirements, and write new code to deal with the less common interfaces that were not included in the delivered vendor software.

This solution is expensive, error-prone, time consuming, tightly coupled and fragile.

**Single provider content management system**

The second approach is to implement a single enterprise content management environment designed to catalog and store (or link to) all content of interest to an enterprise. This provides an enterprise-wide solution for finding and delivering content to other systems in the organization.

In the e-learning space this is done by using a single learning content repository. The learning content repository can come with an “e-learning suite,” or be provided by a stand-alone content management system.

A single “e-learning suite” with an integrated content management solution does work for some organizations, but in many organizations the complex and varying needs of various departments leads to the adoption of different learning management solutions that provide different features and benefits. It is not feasible to expect all organizations to standardize on a single learning management environment.

Implementing a single system to manage learning content requires all an organization’s learning content to be imported into that environment before the content can be used by the learning delivery software. But the reality of digital content is that the number of internal and external digital repositories used by enterprises is growing, not consolidating. In order to successfully scale and adopt these different content sources, learning management architecture should be able to draw content dynamically from a number of different sources. This architecture should clearly separate the functionality and services provided by a content repository from that provided by learning platforms. Furthermore, even if content is drawn from a single repository, it may come in multiple formats, each of which need to be integrated separately into one or more learning delivery systems.
Middleware

A more recent solution to the content integration challenge is to implement a middleware solution such as those supported by Recombo Router\(^6\), Trifus Zelos\(^7\) or (e)Learning Union Connect\(^8\). These middleware solutions act as ‘brokers’ or ‘routers’ between content servers (internal or external to the organization), and learning management environments.

The following diagram provides a high level overview of Recombo’s Router technology, as an example of a middleware architecture. (Provided by Steve Forth of Recombo.)

![Middleware Diagram](image-url)

These products represent a class of learning content integration solutions that provide the following general capabilities:

- Single point of integration between multiple learning management systems and multiple content servers
- Common interface for content search and retrieval, including adaptors to content sources to handle differences in query languages and metadata used to request content
- Normalization of metadata between different sources and mapping between different metadata vocabularies
- Adapters that ‘wrap’ content in a consistent way for widely used learning management systems, allowing the system to retrieve, run and communicate with the content

---

6 Recombo Web site – [www.recombo.com](http://www.recombo.com)
7 Trifus Web site – [www.trifus.com](http://www.trifus.com)
8 (e)Learning Union Web site – [www.elearningunion.com](http://www.elearningunion.com)
o Support for varying versions and interpretations of e-learning and content management standards
o Support for a heterogeneous set of content sources and learning management environments
o Support for proprietary technologies that have not yet adopted the standards

This paper focuses on one feature of content integration middleware, support for content interoperability. It should be noted that the e-learning / content management middleware solutions discussed in this paper are not designed solely to solve interoperability issues. Other functions provided by the middleware include things such as the ability to:

- Collect data for Business Information and Data Warehouse systems
- Listen for events that can be used in Complex Event Processing and Business Activity Monitoring
- Integrate multiple taxonomies and ontologies used to organize large content sets and to maintain these multiple ontologies and their links to content sets (competencies are one early example of this)
- Support the integration of content into other applications such as workflow management (thereby enabling workflow learning)

### 3.2 E-learning Content Integration Middleware Vendors

Recombo, Trifus, and the eLearning Union all have products designed to mediate between differences in content and learning management system support for AICC, SCORM and similar content integration technologies.

**Recombo “Router”**

Recombo positions the value proposition of their “Router” product as follows:

- Learning Managers can leverage existing content without needing precious IT time and resources
- Product Managers can get faster and deeper product adoption using customers own learning delivery systems
- IT Managers can quickly and easily manage migration from one delivery system to another

The Recombo router technology is middleware that routes learning content and data in enterprise environments. It is designed to provide interoperability by

---

9 The information in this section is provided by the vendors. There has been no independent verification of capabilities by the authors. We thank Steven Forth of Recombo, Ciaran O’Connell of (e)Learning Union, and Tom Reilly of Trifus for their input to this section.
creating a standards-based abstraction layer that can be used by cooperating systems.

The Recombo Router is designed to provide learning management systems and learners with access to multiple content sources, combining content with data (using the SCORM RTE), and then packaging the assembled content and data as SCORM for delivery through a learning management system or other SCORM enabled application.

The Router is built around four core processes: Label, Assemble, Package and Deliver.

The **Label** process associates metadata with identified units of content. It also supports both content search and content syndication (notification of regular content updates). In advanced implementations, the metadata can also provide hooks for content assembly and packaging, as well as guide the delivery.

The **Assemble** process takes content from multiple sources and assembles it into a meaningful organization. The process includes chunking of content into units that are relevant to learning, reorganizing these according to the SCORM Content Aggregation Model, applying sequencing rules, and linking of data into the content.

The **Package** process formats the assembled content so that it can be used within a specific enterprise application. Recombo generally uses SCORM for content packaging. In certain cases, the older AICC HACP approach is employed instead (HACP is a specific implementation of the AICC communication protocol over the Internet). Other packaging standards being developed include MPEG 21, which is intended to be more generally applicable.

The **Deliver** process delivers the content to the learner. The process captures information about the ongoing communication between SCORM content and the learning management system.

**Trifus “Zelos”** (Trifus, 2004)

As a leading provider of flexible managed learning solutions, Trifus created the Zelos™ platform to enable content integration in a distributed e-Learning infrastructure. Trifus Zelos offers several key advantages:

1. **Efficiency**: communication to the LMS is limited to when the course is launched, and when the course is exited and tracking information is returned, greatly reducing utilization of the LMS system for certain courses.

2. **Modularity**: since the interface with Zelos is standard and consistent, you no longer need worry about varying course types. SCORM, AICC, and custom content all communicate with your LMS in the same way.

3. **Security**: if your LMS is integrated using the proprietary ZAP format, the launch and tracking requests are all secured. When combined with securely designed courseware this means you can be sure the results sent to your LMS are actually coming from the course.
4. Distribution: multiple LMS’ may use the same Zelos server, allowing you to maintain your courseware at a single location. In addition, the Zelos solution can launch and track content on a remote Zelos courseware server with your LMS securely behind a firewall on an intranet.

5. Flexibility: the Zelos solution is designed to be easily customized. Trifus can create customized and new integrations to content with ease.

The Zelos platform consists of several modules and libraries that are combined or extended into content integration solutions:

The **Zelos Access Protocol** (ZAP) is the common, secure, extensible protocol used by Zelos components to communicate with one another.

The **LMS Integration** translates ZAP messages into a consistent format that is understood by the LMS (often AICC or SCORM). The LMS Integration is also capable of relaying local messages through firewalls without breaching corporate security policies.

The **Courseware Integration** is responsible for launching and tracking course content and persisting user sessions. Courseware integrations are optimized for the specific type of courseware whether AICC, SCORM, or proprietary. The **Remote Courseware Integration** allows the Courseware Integration to operate in a distributed mode where portions of the courseware may be located behind a secure firewall and would not be otherwise available.

The **Trifus Learning System** (TLS) is an example of an LMS with full ZAP integration. Full ZAP integration allows richer data about course progress to be stored in the LMS and more secure data interchange than would possible through standards alone.

The **Zelos Hub** is distributed across all the Zelos components and allows each component to send, receive and broker ZAP requests.

***(e)Learning Union “Connect”*** (eLearning Union, 2005)

The **(e)Learning Union Connect** value proposition is as follows:

- Learning managers have more time to focus on the success of their training programmes.
- New content providers can be added in a significantly shorter time than manual processes meaning organizations can dynamically respond to training needs.
- Valuable IT resources can be employed elsewhere in the business because content set up, support, integration and delivery management can be outsourced and managed by (e)Learning Union.
- Learners have more opportunity for taking their training because (e)Learning Union Connect provides access to content across multiple modes and formats.
(e)Learning Union Connect consists of modules that can be implemented individually or combined according to the organization’s needs.

(e)Learning Union Connect Integrator is the base module that provides:
- Access to content from multiple content vendors.
- Interoperability bridging for varied implementation of standards (AICC\SCORM) between content and the LMS.
- Consistent user experience across multiple content sources.
- Real-time content tracking to the LMS independent of network domains.
- Tracking support for legacy (non-compliant) content.

(e)Learning Union Connect Mobile Learner provides:
- Offline access to training content from any standards-based LMS.
- SkillSoft & Netg content libraries are supported out-of-box.
- Replication of progress records to the LMS is included.

(e)Learning Union Connect Global provides:
- Efficient delivery of content for organizations with challenges in infrastructure or for deployment across large geographies.

In summary, the (e)Learning Union Connect suite removes the need to manually adapt content for crucial application areas such as online integration, mobile use, collaboration and security. In addition the content it supports doesn’t have to be standards compliant as Connect bridges the standards gap that can exist with the LMS.

### 3.3 Enterprise Content Integration

Learning content integration must be seen as one part of a larger enterprise level content integration challenge. There is a continual expansion of the digitization of information, including documents, audio, video and other content. Vast digital repository projects have been undertaken in recent years, and these projects continue to grow and evolve. It is only logical to conclude that corporate, government and educational organizations will need to tap into these multiple content repositories. So, the trend is moving away from single, monolithic enterprise content repositories, towards a world where enterprises access a large, heterogeneous set of content servers, inside and outside the corporate boundaries.

When seen in this light, content integration middleware has much more value than if only considered as relevant to e-learning. Although content specifically developed for e-learning is likely to conform to some version of e-learning interoperability standards, content that has been developed for more general purposes will not. Therefore, the existence of middleware services that can ‘broker’ and ‘wrap’ arbitrary content in a way that allows it to interoperate with learning management systems could be vital to the ability of learning environments to take advantage of this much wider base of digital content.

An example of this demand is the work being done by the ADL to integrate SCORM with the S1000D technical documentation standard. The idea is that the content developed for technical manuals should provide a significant amount of the content needed for training, and that it should be possible to link to the content management structure for
S1000D documents, to reduce the effort of developing training content and to be linked to the technical documentation repository, so that updates to technical documents are propagated to training materials.

It can be argued that it makes sound architectural sense for the technology that implements any approach to supporting this type of cross-walking should be developed as a separate category of technology, i.e. as middleware, rather than baking the underlying translations into point-to-point solutions.

In the future, Content Integration middleware may be integrated with other enterprise capabilities within enterprise-wide service-based architecture, accessible via Web services. At present a comprehensive architecture such as this is a vision and there are no agreed upon, openly developed service interfaces to support content management and integration at this level. Services for content integration is an emerging technology area that will continue to see vendor-specific ‘bridges’ for the foreseeable future, much like those provided by Recombo, Trifus and the (e)Learning Union.

3.4 Content Integration for Learning Applications

Content integration requirements differ for each learning market sector.

3.4.1 Learning Content Management in Large Enterprises

It is in the large enterprise corporate / government sector that there is the most obvious need for considering learning content integration as a subset of the larger content integration architecture. However, learning content integration has some characteristics and requirements that differ from general content integration.

The corporate / government sector frequently uses content that is designed to be delivered as a formal, self-paced learning experience. Ideally, this training content will communicate learner performance to a learning management system. Also, learning content may be described by metadata that links the content to enterprise skills and competency taxonomies. Aggregations of content units may include logic for sequencing their presentation to a learner. There is frequently a need to support the movement of learner-related performance and skills data among content units, learning management systems, and Human Resource (HR) systems that manage skills and competencies.

The learning content integration challenge is to allow a learning management system to deliver a consistent learning experience using content from a number of sources that vary in their support for e-learning interoperability specifications and standards.

In a large enterprise, content is stored in many formats, each of which is appropriate for the department that created or is managing the content. The same is true of metadata, rights and other properties associated to the content. But content is often moved and repurposed, for example from a technical manual to a training course. The applications that use the content each require different formats and properties. A learning application, for example, might want information about competencies associated with content and might want the content to be imported in the format of a SCORM aggregation. On the other hand, a Web Content Management system might be more interested in the technical parameters of content and in the right to display it.
Solutions that employ a single format or that cater to single applications will not satisfy these needs, and point-to-point translations and integrations are difficult and expensive to create and maintain. It is therefore logical to seek solutions that handle the problem of accessing, formatting and interpreting content from a variety of sources and for a variety of applications in a flexible way. If the formats and system requirements are based on open standards, then such solutions can be constructed and have a fighting chance of keeping up with future changes. If the formats are non-standard, then "middleware" becomes a euphemism for collecting point-to-point solutions in one place.

### 3.4.2 Enterprise Content Management vs. Learning Content Management

The primary challenge for “enterprise” content management systems is to present a common search and retrieval capability into heterogeneous information sources, each potentially having their own query methods and metadata. Often the content is unstructured and an important enterprise content management function is to infer and collect useful metadata for purposes of efficiently finding and retrieving content at a later date. This diverse collection of content is then ‘pushed out’ through relatively standard channels, e.g. as Web content, as publications, and through syndication. Enterprise content management is deeply concerned with version control, compatibilities across metadata (i.e. ontologies) and with understanding the structure of incoming content, which may be quite idiosyncratic.

Generally, the role of the enterprise content management system is to gather, organize and distribute heterogeneous content.

In learning applications, however, the fundamental challenge may be characterized as gathering content from similar sources and pushing it out in relatively non-standard channels – as learning content that is part of a learning experience. Underlying most e-learning is traditional and well-understood content, which means that the metadata problem is less complex than it is for enterprise content management, which typically deals with a more diverse set of content from the entire enterprise. In e-learning, the tricky interface is on the delivery side, not the gathering side, and it is on the delivery side that the e-learning standards community has worked to create sophisticated representations of content, functionality and behaviors.

The challenge for learning content and software is content delivery, more than it is gathering, organization and distribution.

So, those responsible for managing learning content delivery in a large enterprise need to find solutions that leverage enterprise content and infrastructure, while addressing the content delivery challenges that are unique to learning content management.

### 3.4.3 Emerging Trends in the Large Enterprise Sector

This section describes some trends in the large corporate and government enterprise sector that are likely to impact content integration middleware requirements.

**Workflow-based learning** is becoming more important. Many companies and government organizations are trying to augment and replace traditional formal learning with on-demand learning delivered in a work context. This trend is motivated by the
observation that as much as 80% of learning is informal. However, informal learning is not supported by most existing learning management system products. Large platform providers such as IBM want to leverage the built-in support for workflow management that exists in their software products.

While workflow-based learning may reduce the need to deliver content in aggregated structures such as courses, it will not reduce the need for content to communicate student performance information. It will also increase the importance of associating skill/competency metadata with content and using learner performance data to track learner mastery of those skills/competencies.

**General content integration** is becoming more important to learning content integration. The U.S. Department of Defense (DoD) has already identified as a priority the need to search a large number of federated repositories for learning-related content. This requirement is a key motivation for the DoD CORDRA project which is being positioned as the next evolutionary stage for SCORM. In a corporate context, as the importance of informal or workflow-based learning grows there will be an increased demand to apply general content integration technologies to learning applications. A technology driver will be the desire for large platform providers to leverage their content management offerings for learning applications by providing learning-specific extensions, such as SCORM support, in addition to core capabilities.

**Semantic markup** is gaining in importance. RDF (Resource Description Framework) is a simple data modeling language developed by the W3C that represents entities and the relations between entities. It is intended to enable the formal semantic description of Web resources so these can be processed using techniques such as logic programming. This is an important part of a larger objective referred to as the Semantic Web. W3C Web Ontology Language (OWL) is a set of extensions to RDF intended to model constraints used to represents ontologies such as “owns is the inverse of owned-by”. Research is underway in the Learning Technology field to determine how RDF can be applied to evolving interoperability standards.

RDF adoption has been slow and it is not clear when or if the vision of the semantic Web will be realized. However, the ability to formally express semantic relationships does have value for content integration middleware. IBM Research has worked on using RDF to express relationships between entities in a domain of interest, and then used this relationship graph as part of a query process to dynamically assemble content objects for delivery in a work context.

The number of semantic representations relevant to learning-related content will increase, and are likely to include not just domain representations but also representations of pedagogies and ways of organizing content.

Content management middleware currently needs to provide crosswalks between different ways of describing content using metadata, and different ways of organizing content such as SCORM aggregations, IMS Packaging or S1000D packages. The implication for the future is that middleware will need to provide crosswalks between different ways of understanding content in the context of different semantic organizations.
3.4.4 Higher Education Market Sector

Traditional higher education e-learning does not have the same emphasis on supporting formal self-paced online training that is found in the corporate / government e-learning sector. As a result, content-to-LMS integration issues are often not seen as a priority in this sector.

However, adult learning and continuing education programs do tend to approach learning in a way that is very similar to the corporate / government sector. Also, Higher Education has other content integration issues, e.g. integrating course content with the body of research being produced and stored in libraries, with data sets, with virtual laboratories, and with other forms of content. Furthermore, the ability to integrate and use content across multiple institutions is important for both cost reduction and for improving instructional quality through collaborative efforts.

Higher education also has a focus on search and retrieval from multiple digital repositories. These repositories are not primarily under control of one organization.

3.4.5 Compulsory (K-12) Sector

K/12 is similar to corporate / government learning in its desire to track learner performance in formal learning activities, and there is also a trend towards organizing content and tracking learner performance by curricular competencies.

There is a comparatively small amount of interactive content available to support online, self-paced learning in K-12 schools. The desire of country and state governments to create broadly applicable solutions may encourage single provider solutions.

In the area of interoperability standards, the Schools Interoperability Framework (SIF) consortium is well-established. SIF has defined what is essentially a middleware interface standard. Its purpose is to enable middleware to route information between heterogeneous administrative systems within a school district. The standard is being extended to support interactive learning. There is also an effort underway to localize the specification for geographies other than the U.S.
4 Evaluation of the Middleware Approach

The authors’ analysis found that there are two feasible, and potentially competing, options for meeting e-learning content integration requirements in a standards-based environment.

Which solution is best?

**Single content management system:** The enterprise implements a common Content Management System (CMS), or Learning Content Management System (LCMS), and imports all learning content into that environment where it is repackaged and converted. This solution should be used where there is a positive return on the investment needed to move all learning content into a centralized system and normalizing it so the interaction with learning delivery environments is consistent.

This is a centralization strategy. If a training department is implementing its own content management solution, and the content sources are known and relatively stable, then an LCMS may make sense.

If the training department is part of a larger organization that has a well-supported, highly centralized Content Management System, then the department would probably want to put all of its content onto the central system, and to draw content from other departments through the CMS. However, unless the central CMS provides e-learning standards support, content integration middleware may still be useful to put the appropriate ‘shell’ around the content for interaction with the learning delivery environment.

If you are running a larger organization where content from many sources is used for many purposes, then a single CMS is not likely, and implementing an LCMS does not provide value because the source documents, which may be changing constantly, is distributed across multiple repositories.

**Content integration middleware:** As discussed throughout this paper, content integration middleware provides a service that dynamically converts the content from one implementation of e-learning interoperability standards to another, at the time the content is needed. It also allows this content to be served to many learning delivery environments; in the form those environments need for interoperability.

It makes sense to use middleware when it is preferable leave the content where it normally lives, in its native format. By accessing the content in place, the middleware approach avoids creating redundant copies of content, and allows each content repository to be managed without interference, while gaining access to all content for e-learning Purposes.

**Middleware Companies Focus on Interoperability Issues and Standards**

One of the real challenges of e-learning standards adoption, particularly for small vendors and system implementers, is allocating sufficient resources to understanding the always complex and sometimes Byzantine world of e-learning standards development.
E-learning middleware companies see integration issues as their primary competitive focus, unlike software product and content vendors to whom this is only one function point in their offerings.

**Middleware Can Provide Value to Vendors**

Middleware solutions can solve detailed integration challenges for the organizations that acquire and implement learning technology and content. For those organizations, content integration middleware increases the value and usability of that technology and content.

Responding to integration issues between product releases is a costly proposition for e-learning technology vendors, so a reduction in the number of these issues that arise will save these vendors money by reducing the need for ongoing client support.

A Content Integration middleware layer is of particular value to smaller vendors when competing with larger vendors who have more resources. Common middleware eliminates the need to develop and test support for multiple standards, and multiple vendor implementation of those standards. The vendor can focus on implementing one version of a standard for each key area of interoperability, and the middleware can be used to ‘translate’ that implementation to other versions and interpretations of the standard.

**Middleware Services are Part of the Enterprise Content Integration Solution**

Content integration for learning must be seen as an extension of overall content integration across the enterprise. There is a growing trend in the corporate training sector towards informal or workflow-based learning which will need to access content from sources that were not developed specifically to support learning applications.

Therefore, content integration middleware for learning must fit in with the content integration solutions for the entire organization. Ideally, the middleware used to support enterprise content integration would also support the special needs of the enterprise learning environments.

**The Need for Content Integration Will Continue to Grow**

The volume of digital content available and the number of content sources that need to be integrated are growing rapidly, and will continue to grow for the foreseeable future. Additionally, the importance of middleware for “learning content” integration will increase as there is increased adoption of AICC CMI and SCORM, especially now that SCORM 2004 adoption is mandated within the U.S. Defense purchasing requirements.
5 Conclusion

It is natural to ask, if standards are so good, why do we need middleware, and if we have middleware, why do we need standards? In fact, middleware and standards often go hand-in-hand, not just in e-learning but in all areas of information and communication technology.

Standards and specifications tend to set out the syntax for how systems will interoperate but they only sketch out the semantics. The syntax is the foundation of interoperability, but it is the semantics (the meaning or intent) that allows people to build business solutions that actually work together. For example, a course publisher and the user of a learning management system might both be using the IEEE LTSC Learning Object Metadata, but with different interpretations. Middleware can help to mediate these differences in interpretation.

This situation is similar in other areas where technical interoperability standards have been developed. Standards provide ‘building block’ interoperability in specified situations, but they do not address all aspects of interoperability. For example, standards may ensure that a female pipe fitting will work with a male pipe fitting of the same diameter. This allows a sink to be installed, but it does not guarantee that the sink will fit into the space, look good in the space, or that it will produce hot water if the hot water heater is not adequate, and the standard does not guarantee anything if different diameter pipes are used. In the desktop computer hardware world “plug and play” standards do not result in immediate interoperability. If they, all applications would do what we wanted and we would never have to look for an SSID, give someone a MAC address to get connected, or make use of helper software such as ‘install shield’, which plays the role of middleware.

For organizations implementing components from multiple vendors the middleware maintains interoperability among those components, even as new releases of technology are adopted. The use of middleware allows integration to occur in a much more flexible ‘loosely coupled’ way, reducing the technical dependency between components, and reducing the need for implementers to be technical ‘experts’ on all of the e-learning interoperability standards.

Content integration middleware should support and augment the value of e-learning interoperability standards, by resolving differences between versions and interpretations of those standards. The use of middleware should also speed up the adoption of new versions of standards, because those new versions can be accommodated through the middleware, rather than waiting for multiple product vendors to adopt the new versions, and for all these new versions to be implemented in the enterprise. The adoption of common standards gets e-learning components much ‘closer’ to interoperability, but there are always final steps that have to be made to resolve differences in implementation.

Conversely, standards make middleware solutions feasible, because without general agreement on standards the components of an e-learning environment would interact in purely proprietary ways. Standards get the components to the point where they are trying to interoperate in functionally and structurally common ways, at which point it is possible to develop and support middleware that takes the final step to ‘out of the box’ interoperability.
Glossary of terms and abbreviations

**ADL** – Advanced Distributed Learning Initiative

**AICC** – Airline Industry CBT (Computer Based Training) Committee

**CMI** - Computer Managed Instruction

**Content Repository**: A secure storage facility that is capable of handling a wide variety of content types, and enabling access to authorized users.

**CORDRA**: Content Object Repository Discovery and Registration/Resolution Architecture. See (Rehak, 2004).

**Digital Rights Management (DRM)**: The process of defining, tracking and enforcing permissions and conditions for digital content through digital means.

**IEEE LTSC**: IEEE Learning Technology Standards Committee

**LCMS**: Learning Content Management System

**LMS**: Learning Management System

**MPEG**: Motion Pictures Experts Group

**RDF**: Resource Description Framework – A specification developed by the W3C intended to “provide a lightweight ontology system to support the exchange of knowledge on the Web” (W3C, 2005).

**RSS**: RDF Site Summary is a lightweight multipurpose extensible metadata description and syndication format. (RSS, 2001)

**SCORM**: Sharable Content Object Reference Model
References


