

The Content Management Interoperability Services (CMIS) Standard

The AIIM iECM CMIS Demo

Looking Back at 2009 and Ahead to 2010

Thomas Pole
Laurence Hart

January 2010

At the 2009 AIIM Expo, AIIM's Interoperable ECM (iECM) committee presented a live demonstration of the proposed Content Management Interoperability Services (CMIS) standard. Built to illustrate the power of the proposed standard, the demonstration allowed users to search multiple content repositories with a single interface.

What is CMIS?

Originally announced in September of 2008 by seven Content Management vendors, the proposed CMIS standard is defined as:

Defines a domain model and Web Services and Restful AtomPub bindings that can be used by applications to work with one or more Content Management repositories/systems.

Originally developed by EMC, IBM, and Microsoft, CMIS was evaluated and endorsed by Alfresco, Open Text, Oracle, and SAP prior to the public announcement. All seven vendors joined together to announce and release the standard to the community and OASIS.

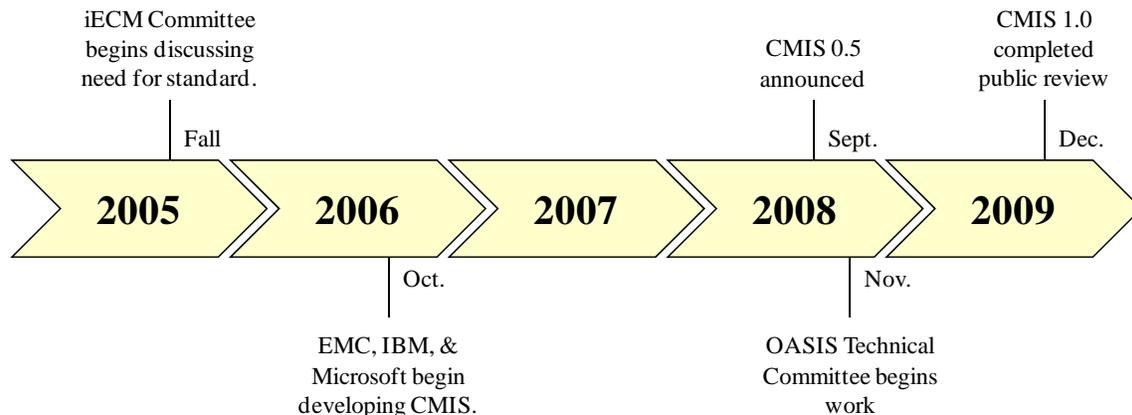


Diagram 1: The CMIS Standard Timeline

The goal of CMIS is to define a common domain model for content management systems, allowing different applications to communicate using the same terminology and language. Specific technology bindings to the domain model based upon Web Services and Restful AtomPub were developed as part of the proposed standard as well. Other bindings, such as WebDAV, may be added to future iterations of CMIS.

There are three fundamental use cases for CMIS:

1. **Repository to Repository:** This is the direct communication between repositories. This most often occurs when content needs to be managed from a central, “master”, repository.
2. **Application to Repository:** This is when an application that uses content stores that content in an external repository. Through the use of CMIS, the application can focus on providing its core business solution without managing the content or integrating with different repositories that may each have their own API.

3. **Federated Repository:** This is the presentation of a collection of repositories to a user or application as a single repository. This model occurs frequently when users need to combine the services and information available from multiple systems to satisfy searches for content across a scope larger than any individual department.

It is believed that the introduction of CMIS will have a similar impact on the content management industry that the introduction of SQL had on the database industry. By providing a common language to communicate with content repositories, application vendors will be able to focus on developing their core applications without spending a significant portion of their resources designing and developing interfaces for different repositories.

Conceiving the Demo

When the CMIS standard was announced, the iECM committee began an evaluation of the proposed standard in order to provide guidance to the AIIM community. After studying the specification, the iECM committee decided to develop a CMIS demo to evaluate the effectiveness of using CMIS to build an application that leveraged multiple Content Management (CM) systems. The effort was designed to gauge the reduction in effort using CMIS would offer in comparison to building the same integration to multiple CM systems without using the CMIS technology.

The challenges faced while implementing the demonstration were three-fold.

1. A readily understandable scenario, with available content, needed to be developed.
2. CM vendors needed to be recruited to support the effort with CMIS compliant repositories.
3. The entire effort had to be completed in two months.

The concept was to create a federated search interface that would allow users to search multiple repositories from a single search screen. The original idea was to leverage the presentations and supporting content from the AIIM conference. Divided into logical subsets, the content was to be distributed among the various CM systems. The access to the entire collection would be via a web page interface integrated with a CMIS based federator.

While the availability of the content was being confirmed, the iECM committee reached out to vendors participating in the OASIS CMIS Technical Committee. Several vendors showed interest in assisting the iECM committee and participated on the initial planning calls. Each vendor was asked to host instances of their ECM/CMS repositories and expose their CMIS interfaces (even those that were not yet available to the general community) to the repositories through the Internet.

It was then learned that the conference presentations would not be available. The decision was made to use old articles from AIIM's E-DOC, now Infonomics, magazine for the demonstration system. Vendors were encouraged to supplement the Infonomics

articles with their own content. The tagging of all content was left to the vendors, though the iECM committee assisted with tagging the content when possible. Each vendor was provided articles from two issues so as to randomly distribute the ECM topics upon which users could search.

The User Scenario

The iECM CMIS Demo had one basic use case, the search and retrieval of content. The user experience was designed to be straight-forward and readily understood by attendees of the AIIM conference.

1. A person would initiate a search for content using a combination of free text and metadata attributes that had previously been assigned to the content. The actual query was then constructed automatically by the system.
2. The system would then search multiple repositories by issuing the query to each repository connected to the demonstration system and return a list of documents. The list was to be a constructed by merging the result list of the individual systems.
3. The user would then select the document that they wished to retrieve.
4. The system would then request the document from the source repository and return the document to the user.

This scenario was designed to represent a typical “Federated Repository” use case in order to clearly demonstrate the communication to multiple repositories. The limiting of the demo to search and retrieval functionality was done to lessen the CMIS support level from each of the vendors given the early state of both the standard and CMIS implementations.

The Design

The first step was to design a metadata model with which to tag the content. A simple model was developed that specified a custom Object Type named "AIIM Content".

Name	Type	MaxLen	Cardinality	OpenChoice	Req.
Title	String	100	single	N/A	TRUE
Sub-Title	String	100	single	N/A	FALSE
Authors	String	25	multi	N/A	FALSE
Publication Date	DateTime	N/A	single	N/A	TRUE
Keywords	String	50	multi	N/A	FALSE
Source Repository	String	25	single	N/A	TRUE
Information Management Topics	String	60	multi	FALSE	FALSE
IT Topics	String	40	multi	FALSE	FALSE
Industries	String	30	multi	FALSE	FALSE
Lifecycle Stages	String	25	multi	FALSE	FALSE

Table 1: The CMIS Demo Metadata Model

For the value for "Source Repository", all pieces of content in a repository were to be assigned the same value. The value was to clearly identify the source CM repository for each piece of content to the end-user.

The next decision was to determine which technology platform upon which the search federator was to be built. It was decided to use .NET based upon simple economics. The iECM committee had access to a free, reliable, IIS Server to host the federation application.

From there, it was decided that the demo would leverage the Web Service binding for CMIS. While the Restful AtomPub binding would have been more appropriate for the application the iECM committee was developing, it was felt that the Web Services binding would be easier for the iECM development team to implement given the timeframe. The team had access to sample code utilizing the Web Services binding in .NET and wanted to take advantage of the deeper support for Web Services in the Visual Studio development platform.

Due to the decision to use the Web Services binding, some vendors were unable to provide a host repository as they did not anticipate having the Web Services binding implemented in time for the conference. Others were unable to set up a CMIS enabled repository that was available through the Internet in time for the show. Some vendors faced both challenges.

Given the tight timeframe, getting more than one vendor to participate was a challenge, but commitments were collected from Alfresco, EMC, and Nuxeo.

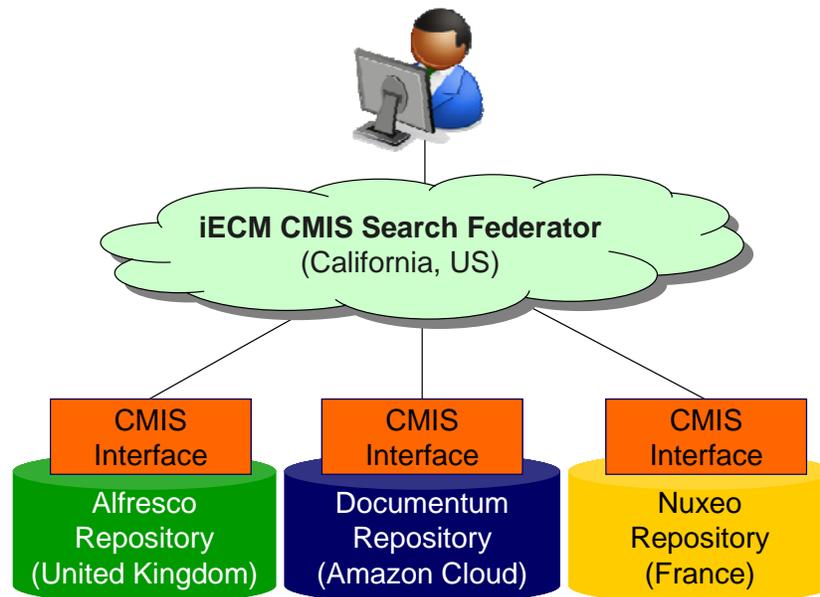


Diagram 2: The CMIS Demo Architecture

The resultant architecture is outlined in Diagram 2. It should be emphasized that the three repositories were located in England, France, and the Amazon cloud while the

interface was hosted in the United States. This provided a sound test for the Federated Repository model.

A Draft Standard Implemented by Beta Software

Many challenges were faced during the implementation of the Search Federator. The CMIS specification that vendors were referencing was the draft specification originally submitted to OASIS. The implementations themselves were beta quality software as vendors were not planning to fully develop and test the implementations until the CMIS specification itself reached a final state.

The resultant issues can be simply illustrated with the following diagram:

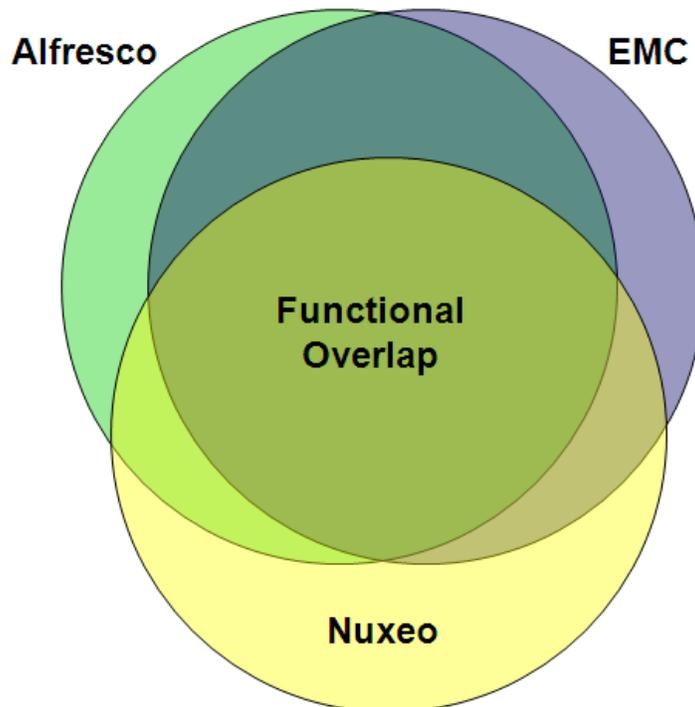


Diagram 3: CMIS Functionality by Vendor

Each vendor had ~80% of the Web Services binding implemented, but not all of it overlapped. There were variations between the implementations and some bugs, which each vendor worked quickly to fix. When the implementation didn't support the specification correctly, the iECM committee let the vendors know for future reference and used the implementation as it was.

Even working against slight variations of the CMIS standard was a significant improvement to working directly with distinct vendor interfaces. Microsoft's Visual Studio tools found most of the differences for the implementation team automatically. Any other questions were quickly resolved through email and Twitter.

One of the significant complications came from each vendor supporting different WS-Security authentication models, including one implementation that had no security

defined. This required some work when implementing the Web Services calls, but they were independent of the actual CMIS standard.

At approximately 7pm the night before the demonstration was presented at the conference, the iECM committee received an email from Nuxeo that their system was ready to be added to the search federator. It took only one hour to incorporate Nuxeo's repository to the demonstration system, including the time it took to make modifications for the uniqueness of the object model and security implementation. On a mature CMIS system, it is estimated that it would have taken just 15 minutes.

That quick turn-around is the power that the new CMIS standard gives the solution provider; the ability to modify a content application to use an additional, or different, CM repository. After mapping the metadata elements and the security model to the application, the CMIS implementation is minor in nature.

Working Towards a Complete Standard

Building on 2009's success, the iECM committee began designing the 2010 iECM-CMIS Demonstration in October 2009. The 2010 demo will be in a specific business domain, Content Management for Healthcare. Classification vocabularies already defined and commonly used in the Healthcare industry will be used to classify content including Logical Observation Identifiers Names and Codes (LOINC) and Systematized Nomenclature of Medicine (SNOMED).

This year's demonstration will show two CMIS scenarios. The Federated Repository scenario will once again be leveraged. In addition, the Application to Repository scenario will be used as the users of the system will be able to direct their inquiries to specific repositories on demand.

It is hoped that more vendors will be able to deploy instances of their product with content and related metadata that exemplify the content that would be found at Healthcare organizations such as hospitals, clinics, physician practices, medical libraries, etc. There are already commitments from more vendors than the 2009 demonstration.

The People behind the Demo

It is important to call-out the people that were critical to the success of the demonstration. These people gave time and effort to the success of the demonstration. All gave their time outside of their efforts on their day jobs.

- Thomas Pole: Thomas is the chair of the iECM committee for AIIM. He was in charge of the demonstration and was able to identify a host platform for the system for no charge. In addition, Thomas and his team built the User Interface for the demonstration.
- Laurence Hart: Laurence is a member of the iECM committee. He and his team were in charge of writing the federator and working with the vendors on their content and CMIS interfaces.

- Betsy Fanning: Betsy is the Director of Standards at AIIM. She helped keep the effort on track and coordinated with the various vendors that participated in the effort.
- The iECM Committee: The entire committee helped make sure that what was being done made sense from a business perspective. They worked with Thomas and Laurence to develop the requirements and design.
- The Vendors: This is more than the ones that participated in the final demonstration. All the vendors involved worked hard in this effort.
- Harris Corporation and Washington Consulting, Inc.: Thomas's and Laurence's companies, respectively, helped by allowing time and additional resources to build the demonstration. Thomas and Laurence both have full-time jobs and only support from their companies made this possible.

Interested in Doing More than Reading?

If you are interested in participating in the 2010 iECM CMIS demo project, please contact Thomas Pole (tpole@harris.com). If you wish to participate as a vendor, please contact Betsy Fanning (bfanning@aiim.org). Many skills and resources are needed, including:

- Testers
- Content classifiers
- Software developers
- Coordinators to help keep the team productive, including working with the vendors
- Vendors
- Healthcare professionals to act as Subject Matter Experts (SME)