Content Management Interoperability Services

Matt Moore*

Content Management Interoperability Services (CMIS) is an emerging standard that allows compliant Enterprise Content Management (ECM) systems to share documents. This article discusses the reasons for the emergence of CMIS and outlines some of the key elements of the standard. It outlines the response by one vendor (Nuxeo) and the use of the CMIS by Dutch local government. The article concludes with some comments on the future direction of CMIS.

INTRODUCTION: ENTERPRISE DISCONTENT MANAGEMENT

Does your organisation use a content management system? Or more to the point for this article, does it use more than one? Perhaps the records management team use an Electronic Document and Records Management System (EDRMS). However the web team loathe the EDRMS with a passion and have two systems of their own: one a workflow-heavy publishing tool, and the other an open-source online community tool. Marketing have their digital asset management system for pictures and videos – and they do not trust anyone else with that. Meanwhile the IT department have a bunch of SharePoint licences that they got free with their Exchange server purchase and they think everyone else should be using that. Everyone else agrees with the IT department – provided they can keep on using their systems as well.

Enterprise Content Management (ECM) is a commonly used term but a rarely occurring organisational state.[1] Enterprises would like to have a single sleek and flawless system for managing their content in the same way that most of us would like the sleek, flawless body of a supermodel. However, many things get in the way of this desired ideal:

- organisations have invested time and money in historical point solutions (eg records management) that they are reluctant to throw out;
- some vendors claim to offer "ECM", but it is unclear as to whether they do (or indeed can) offer full-spectrum solutions for everything;
- new functionality (eg social media, mobile) is emerging. The market is far from fully mature;
- even if organisations do manage to manage their content and then commit to one vendor (and, realistically, the players in the vendor's ecosystem that will fill the gaps) – what happens if you need to share content with other organisations? Does sharing end at the firewall? What happens if your organisation is merged with another one that has made different choices?

Historically, organisations have dealt with these issues by obtaining point-to-point connectors between different platforms. Some of these need to be built from scratch, some are supplied by one or other of the vendors and others are created by third parties. They can be expensive and will often require upgrading when one of the underlying products is upgraded.

Both vendors and users within the industry recognise that this is not a sustainable way to operate – which is why a common standard for content management is so important. Content Management Interoperability Services (CMIS) is a recently developed standard that attempts to meet this need.[2]

* Matt Moore is a director of Innotecture, an occasional lecturer at the University of Technology Sydney and a former chair of the New South Wales Knowledge Management Forum. He has spent over a decade working in knowledge and information management, learning and development and internal communications with organisations such as IBM, Oracle and the Australian Securities and Investment Commission. Many thanks to Jane Zupan, Gershon Janssen, David Choy and Ryan McVeigh for their assistance in the writing of this article.

All webpages cited were viewed 22 May 2012.

[1] McKinnon C, "ECM in an Interoperable World" (2011), http://www.slideshare.net/CherylMcKinnon/aiim-new-england-ecm-in-an-interoperable-world.

[2] "Content Management Interoperability Services (CMIS) Version 1.0. OASIS Standard" (1 May 2010),

A SHORT HISTORY OF CMIS

There are some pre-existing standards in the content management space but they tend to be tied to a particular technology or user case. The Open Document Management API (ODMA) allows user access to a repository but not a system-to-system. Content Repository for Java Technology API (JSR-283) requires a Java interface. HTTP Extensions for Web Distributed Authoring and Versioning (WebDAV) focuses on extending HTTP protocols. None of these options can be applied to a broad enough selection of ECM vendors.[3]

In 2006, the US-based Association for Information and Image Management (AIIM) formed a committee of vendors, users and academics to work on iECM: Interoperable Enterprise Content Management. Out of these discussions, three vendors (Microsoft, IBM and EMC) created their own project. This then grew to seven vendors and a draft CMIS specification was submitted to the Organization for the Advancement of Structured Information Standards (OASIS) in 2008. It ended up with 19 vendors collaborating on the final version that was ratified in May 2010. The standard is currently under the stewardship of the CMIS Technical Committee which consists of employees of key vendors (eg IBM) and independent experts.

WHAT IS CMIS?

CMIS does not cover all functions of all ECM products. It focuses on the core functionality that these systems tend to have.[4] In a sense, it describes the baseline functionality that a content management system should have, then provides a framework for such systems to interact. It is independent of a particular platform (eg Documentum), a specific programming language (eg Java), or a set of protocols (eg REST, SOAP).

What makes up the CMIS standard? CMIS outlines a data model that covers four common objects managed by a repository and then describes the services that can be applied to them. The objects are "things" in the repository and the services are what can be done with or to those things.

The four objects are documents, folders, relationships and policies. This list of objects explicitly excludes certain entities such as programming interface objects, user profiles, compound or virtual documents, workflows and business processes, events and subscriptions. All objects have a unique ID and a set of named properties. All objects may also have a content stream (ie the raw data in a video file) with one or more renditions (eg a thumbnail). All objects may also have an access control list (ACL) that defines who or what has permission to do something with the object in question.

A *document* object represents a standalone information asset. Document objects are the elementary entities managed by a CMIS repository. Documents can be version-able (eg they can be checked in and out), file-able within one, many or, indeed, no folders. Documents can be located by discovery services (ie search).

A *folder* object represents a logical container for a collection of "file-able" objects, which include documents or sub-folders. Unlike documents, folders can only sit in one folder – not several or none.

A *relationship* object represents an instance of directional relationship between two objects ("source" and "target"). Relationships are "non-invasive" in that creating or removing one should not affect the source or target object. The support for relationship objects is optional.

A *policy* object represents an administrative policy, such as a retention management policy, which may be "applied" to one or more "controllable policy" objects. Whether or not an object is controllable is specified in its object-type definition.

CMIS services focus on "CRUD" operations: Create, Retrieve, Update and Delete. It also includes services that deal with navigation, discovery, versioning and access control lists.

http://www.docs.oasis-open.org/cmis/CMIS/v1.0/os/cmis-spec-v1.0.html.

[3] Hart L, "The Challenge of CMIS" (2009), http://www.wordofpie.com/2009/04/20/the-challenge-of-cmis.

[4] Choy D, "CMIS Technical Overview, Part 1" (2008), http://www.youtube.com/watch?v=83tt0LgZudA; Choy D, "CMIS Technical Overview, Part 2" (2008), http://www.youtube.com/watch?v=yhoTOTl6phc.

THE INDUSTRY RESPONSE

The point of a standard is that many people use it, otherwise it is not really a standard. So what has been the uptake among organisations with ECM-related systems? The AIIM's *State of the ECM Industry 2011* comments:

> The development of CMIS (Content Management Interoperability Services) promises to improve connectivity between ECM systems, and between ECM and other enterprise applications. In our 2009 survey, 69% of our end user respondents were "not sure what it is". This fell to 54% in 2010 and 42% this year. The number making a commitment to adopt it as a standard in their organization has doubled since 2009 to 8%, although this is still a small proportion of the overall potential.[5]

Awareness of CMIS has steadily increased over the last two years and has now crossed the 50% midpoint mark. The AIIM survey does not rigorously test respondents on their CMIS knowledge so exactly what "awareness" means is ambiguous. The number of ECM users committed to adopt it is still low, therefore examples of organisations that are using CMIS are of interest. Indeed, the main driver behind this article was to identify these actual uses. The following two sections discuss the activities of a vendor and then look at a group of ECM product users.

VENDOR PERSPECTIVE: NUXEO

Nuxeo is an open source content management platform that is used by organisations as diverse as the BBC, the Jeppesen Boeing and the French Atomic Energy Commission. As a content management software vendor, Nuxeo has frequent customer requests to integrate their technology with other enterprise software solutions across the spectrum – web content management and web portals, business intelligence and reporting, scanning and capture etc. Historically, these integrations were not always easy. Project scope, feasibility and maintenance had to be carefully defined to determine if benefits outweigh the costs. Version upgrades for either side could cause side effects for the integration and require additional maintenance work.

CMIS is a standard, vendor-neutral protocol that enables one technology to exchange information with another. Even if either side changes or upgrades, the standard CMIS protocol remains the same, so that the integration remains stable. In the past year, Nuxeo has leveraged CMIS to provide connectors with other CMIS-compliant applications, offering customers the opportunity to quickly and easily align Nuxeo content management and document management applications with existing information sources and other systems.

- Nuxeo and Hippo: Nuxeo and Hippo, a leading vendor of commercial Java Open Source Web Content Management (WCM), have built a connector that allows for content from a Nuxeo-based ECM application to be accessible on a Hippo-based website, eliminating the need to duplicate information or move ECM content out of its managed context;
- Nuxeo and Ephesoft: Nuxeo and Ephesoft, provider of an intelligent document capture system, have built a CMIS-based connector between their platforms. The connector between the Ephesoft content capture platform and the Nuxeo content management platform is a fully open-source, end-to-end solution for managing the content lifecycle from capture to publish to archive;
- Nuxeo and Liferay: Liferay has certified the Nuxeo Platform as a content source via CMIS for Liferay Portal 6.1 Enterprise Edition (EE). Liferay Portal can serve as a front end for the Nuxeo document management platform.

As Jane Zupan, Nuxeo Product Marketing Manager says:

> Often our clients and prospects don't ask us about CMIS and the promise of interoperability. Their requests are more pragmatic; they have multiple systems and sources of information, and if these technologies talk to each other, they are far more valuable than if they are isolated, or require manual intervention. They want their scanned documents to be automatically ported to their content

[5] AIIM, *State of the ECM Industry 2011* (2012) p 12, http://www.aiim.org/Research/Industry-Watch/State-of-the-ECM-Industry-2011.

management platform, or their document management software to interact with their intranet, so that their documents have a single source, rather than copying the same information from one software solution to the other.[6]

CASE STUDY: THE DUTCH GOVERNMENT

Since 2005, the Dutch government has been developing the NORA, which is the Dutch acronym for "Dutch Government Reference Architecture". The NORA-family consists of various technical reference architectures: NORA, GEMMA, PETRA, WILMA and MARIJ. Each reference architecture focuses on a different tier of the Dutch government and its constituent government organisations. These architectures are maintained and further developed by organisations that are closely related to that tier of the government and understand their specific needs and requirements. They are all also Dutch women's names.

One very active tier of the Dutch government is the distributed local governments (municipalities). The Netherlands has just over 400 municipalities. They are the most common point of contact for citizens for many of their government interactions, eg requesting a passport or drivers licence, or for inquiries about property taxes. In fact, for the last two years, municipalities have been positioned as the key gateway to government services.

Municipalities are required to adhere to the mandatory GEMMA reference architecture. The GEMMA reference architecture is maintained by KING, an independent institute which provides advice, assists, encourages and supports municipalities in their organisation (development) and implementation of tasks.

The GEMMA reference architecture has a few key aspects, the most significant of which is the concept of case management, together with its specific information models. To follow this architecture, one has to implement the core components such as a workflow system, a case management system, a document/records management system and broad integration with national registries, such as GBA (Basic Registration of Natural Persons), BAG (Basic Registration for Addresses and Buildings), and NHR (National Register of Companies). The StUF standard, which in Dutch stands for "Standard Exchange Format", must be used for information exchanges between these components. A critical concept in this reference architecture is interoperability: if all components speak the same language, then broad interoperability can be achieved.

When the municipalities started to design their new architectures, it was thought that systems compliant with this architecture would have their own internal document repositories. This resulted in limited product offerings from only a few vendors, undesired migration scenarios and interoperability issues. And as most municipalities already used one or multiple ECM systems to store most or all of their business process related content, it was very desirable to keep these ECM systems in place, if only for cost reasons. So rather than replacing these systems with new solutions, they had to be included in the new architecture. However, for these commercial ECM systems to work and integrate within the GEMMA architecture, they should be compatible with both the architecture and the StUF information exchange standards and this was not generally the case.

As a result of this gap, the municipality of Woerden started an initiative in late 2010 to develop an additional standard. Woerden contacted KING to help with the standards process and invited other willing municipalities and IT vendors to join. A first meeting was convened in December 2010. It was agreed to create a specification that translates between the StUF standards and the OASIS CMIS standard. This would allow commercial ECM systems to comply with the reference architecture.

During the course of 2011 more municipalities and IT vendors joined the effort which resulted in a first draft specification by September 2011. As the group felt the best way to prove and then improve the usability of the specification was by using it, the members agreed on a pilot phase in which a range of vendors and system integrators would work on implementing the specification for this CMIS Bridge and then engage in testing together with the municipalities.

[6] Personal communication with Jane Zupan.

So far, a CMIS Bridge has been implemented by NovoGov (a vendor). The CMIS Bridge is a generic "bridge" between the StUF information exchange standard and any CMIS-compliant implementation that supports the specifications' change log functionality. A working setup has been also created with Alfresco.

The impact of these efforts will be profound.

- This effort promotes the use of standards-based ECM systems in favour of creating specific (add-on) solutions. CMIS is the big differentiator in this equation: instead of building custom codes on each and every ECM system, the standards-based CMIS interface can be used to allow for interoperability.
- CMIS allows for 400+ municipalities to continue to use their existing ECM solutions or giving them the freedom to select their own (new) ECM solution, while still adhering to the mandatory government architecture and information exchange standards – in effect allowing many of the 400+ municipalities' ECM solutions to coordinate with each other on a large scale.
- The CMIS standard also becomes a criteria for ECM solution purchases. DIMPACT, a cooperative association by and for municipalities with 30 members, has the sole purpose of selecting and purchasing solutions for its members in order to implement the parts of the GEMMA architecture. In its recent tender it specifically referred to the specification of this CMIS Bridge as its future direction.

Those involved so far have included:

- municipalities of Woerden, Almere, Amstelveen, Apeldoorn, Breda, Ede;
- IT Vendors: NovoGov, BCT, Centric, Circle Software, Decos, Dimpact, Exxellence, Interaccess, InteractionNext, PinkRoccade;
- KING (maintenance and support body of the standard and standards process);
- Gershon Janssen: voting member of the CMIS TC, one of the editors of the Dutch "CMIS Bridge Specification" and architect at NovoGov, a Dutch software company who has successfully created one of the first working prototypes.

The specification will be formalised in mid-2012. The Dutch government has a "comply-or-explain policy" with respect to agreed upon standards. For this reason it is expected that vendors will quickly use this specification for the CMIS Bridge in order to create this very much desired interoperability with commercial ECM systems that support CMIS.

THE FUTURE OF CMIS

CMIS will ultimately succeed or fail based on its uptake. Several factors will influence the speed and extent to which this happens.

The first factor is the range of vendors that support CMIS. SharePoint is a growing part of the enterprise content marketplace and so Microsoft's support for this standard is critical. Likewise the support of OpenText, IBM, Oracle and EMC Documentum is critical for the Enterprise Content Management marketplace. However, the manner in which a vendor has implemented the standard will differ. For example, CMIS 1.0 does not specify how metadata is managed and different vendors have taken different approaches to metadata management.

Also important is the speed at which the functionality of CMIS expands. CMIS 1.1 is already in draft form and it will be submitted for public review and approval in mid-2012. It includes the following additions:

- browser binding will make it easier for developers to build browser-based applications that use CMIS content;
- records management capabilities will be increased around the retention and holding of objects;
- properties (secondary object types) that can be dynamically added to or removed from objects during their lifecycle;
- the capability to create, modify or delete primary or secondary object types.

Whether this will be enough remains to be seen. CMIS is very much focused on the management of documents rather than, say, the presentation of web content.

Thirdly, CMIS is only one of a number of content-related developments and how it relates to these other trends will be critical.

- Cloud computing and ECM-as-a-Service: The shift from on-premise technology to software services that are not only hosted but owned by someone else is a major talking point in the technology world – and ECM has not been immune to this. However, fully cloud-based corporate ECM environments are still relatively rare in Australia. One reason for this is that content management system integration is seen as daunting and cloud-based systems are perceived as difficult to integrate. For some vendors, CMIS is seen as a critical element of a move into the cloud because it decreases some of these integration pains. However, other cloud vendors have indicated that CMIS is not a priority for them.
- Social media is producing ever increasing amounts of content that ECM systems will need to manage in some way. Some ECM vendors are offering plug-ins to major enterprise social collaboration platforms such as Yammer and Jive.
- Mobile device proliferation also has impact on the ECM world – requiring a greater variety of ways to present content. The screen of a smart phone is different to a desktop monitor. There are already some vendors providing CMIS-compliant tools for Apple and Android devices.

In summary, the future of CMIS is far from assured. Nevertheless, it remains the only standard of its kind. I would advise Australian information management professionals who use a document-oriented content management system to understand the basics of CMIS and then to ask the vendor (or, more likely, vendors) about their level of CMIS compliance. This compliance level may vary across the product set. It may be premature to include CMIS compliance as an "essential" requirement in a request for tender, however, it is certainly "desirable" and, at the very least, there is discussion to be had with vendors around the platform's current integration capabilities and its future integration roadmap.[7]

In the medium term, our information environments will get more complicated and not less so. We need tools that will help us in making this complexity manageable. CMIS may prove to be one such tool and, at present, it is certainly an option worth exploring.

APPENDIX 1: OASIS MEMBERS WHO APPROVED CMIS 1.0

- Adobe
- Active Endpoints
- Alfresco
- ASG
- Booz Allen Hamilton
- Day Software
- dotCMS
- EMC
- FatWire
- fme AG,
- ISIS Papyrus
- IBM
- Liferay
- Microsoft
- Nuxeo
- Open Text
- Oracle
- SAP
- SAPERION
- WeWebU

[7] Waldhauser S, "5 Myths About the CMIS Standard" (2011), http://www.digitallandfill.org/2011/08/5-myths-about-the-cmis-standard.html; Choy D, "In Search of Common Ground" (2011), http://www.oasis-open.org/committees/download.php/42999.