ECMs and Institutional Repositories: The Case for a Unified Enterprise Approach to Content Management

Malcolm Wolski, Natasha Simons and Joanna Richardson
Griffith University

Universities are currently developing responses to manage the explosion of research content. There is an expectation by these institutions as well as governments, funding agencies and other stakeholders that research data will be well managed, available and accessible to users as appropriate.

The large enterprise content management (ECM) platform vendors are evolving into "information management frameworks". The ECM solutions being marketed by these vendors are underpinned by content repositories, promising to manage all of the enterprise's digital assets. One might logically question whether a university actually needs separate institutional repositories (IR) systems and infrastructure such as DSpace, for example, to manage research data. If these new enterprise solutions overcome the historical shortcomings traditionally associated with research content, then what is the future of the IR? The implementation of SharePoint along with new research data services at Griffith University has been a catalyst for beginning to question some of the fundamental paradigms which have underpinned the current thinking about an enterprise approach to research infrastructure and the role of research repositories.

Having conducted a literature review, the authors outline the roles of enterprise content management systems and institutional repositories in the context of strategies, processes, and technologies rather than as single products. The focus is on architecture and a management approach rather than technological solutions.

This paper explores the synergies between institutional repositories and enterprise content management systems and how research content would fit within the traditional enterprise content management system model. It concludes that there are major benefits in taking a unified enterprise approach to managing research content within a university.

Introduction

In Australia, government funding and policy guidelines are placing pressure on universities to increase the accessibility of their research output. There is a clear focus in the literature on the need to improve research data management, sharing, and accessibility to help meet these objectives (Bolton, 2010; Borgman, 2010; O'Brien, 2010). Increasingly this focus has also highlighted the need to improve data capture and methods of collaboration.

In developing and supporting the research infrastructure to help achieve these objectives, it is clear that the content—and content is used here to encompass all research output—will not achieve critical mass by virtue of individual voluntary effort or small siloed solutions. Historically in most universities, libraries have taken on the role of stewardship in relation to preserving research content in what is commonly referred to as the institutional repository (IR). However one needs to consider the scale, volume and breadth of the solution required to capture, process and preserve data as well as how to make it shareable with easy to use interfaces that lower the burden on researchers to better manage their outputs. The infrastructure and investment required is beyond most library organisations.

On the other hand, universities like all organisations have always been required to manage their structured data and documents to ensure that regulatory and policy requirements are being met. However the explosion of digital content replacing the traditional print-based documents and the increasing volume of digital content underpinning day-to-day business processes (e.g. web site content, marketing videos, videoed lectures) have resulted in many organisations looking for new enterprise solutions to manage this content. Because of the level of investment and the demand to share content (structured and unstructured), there has been a move away from departmental siloed solutions towards what is now referred to as enterprise content management systems (ECMs).

At the systems and infrastructure layer, the ECM and IR environments have become more complex with solutions provided through Software as a Service (e.g. SharePoint in Office365) and Infrastructure as a Service (e.g. Research Data Storage Infrastructure (RDSI)). Discovery tools can potentially harvest data from a range of content management systems such as internal repositories (e.g. DSpace, SharePoint, Squiz Matrix), hosted repositories (e.g. Digital Commons) and third party repositories (e.g. ASSDA) and databases (e.g. Peoplesoft HR, Research Administration Database).

In addition functionality is rapidly developing to provide richer environments where relationships between content are exposed or information about usage is published to further opportunities for collaboration between researchers. In this new content ecosystem, functional and architectural issues at the content management layer need to be addressed to meet new demands.

However given the relatively recent focus on the evolution of enterprise content management systems into information management frameworks, one might logically question whether a university actually needs systems and infrastructure such as DSpace, for example. If these new technologies overcome the major obstacles traditionally associated with siloed research content, then what is the future of the content management layer for research content? The implementation of SharePoint at Griffith University has been a catalyst for beginning to question some of the fundamental paradigms which have underpinned the current thinking about an enterprise approach to research infrastructure and the role of research repositories.

This paper explores the relationship of research content management to enterprise architecture and where research content fits within the traditional enterprise content management (ECM) system model. The discussion is based on an analysis of the drivers for change, the roles of ECMs and IRs, and the incentives for an integrated enterpise approach.

Drivers for research content management

Universities are developing a range of approaches for dealing with the information explosion. Within these organisations there is an expectation that an ever-increasing volume of data is required to be available to users at all times and that the user's tolerance for data which cannot be easily discovered is continuing to decrease.

Coupled with the demands of dealing with traditional corporate data, there is the added complexity of support for research data. A key driver—along with collaboration and building upon the work of others—is reproducibility of scientific results. "The scientific record ... should make data available, and contain enough information about methods and practices, that another scientist could reproduce the same results starting from the same data" (Lynch in Hey, 2009). Big Data presents its own challenges. In his plea for the creation of better tools to support the whole research cycle, Gray (Hey, 2009) asserts that "we now have terrible data management tools for most of the science disciplines". The scale of individual experiments in some disciplines—and particularly their data rates—makes this issue a formidable challenge (Bell in Hey, 2009). According to Thwaites (2012), "the days of researchers finding or generating data, downloading it to their desktop PC and analysing it there are rapidly disappearing".

Research data is not limited to just the traditional datasets generated in the sciences and social sciences. Investigations in the creative and performing arts are producing research outputs in a wide range of non-text-based formats. In Australia and New Zealand, research outputs reported as part of their respective national research assessment exercises have included:
- Published scholarly (academic) work such as books, book chapters, journal articles and conference papers
- Work presented in non-print media such as films, videos and recordings
- Other types of outputs such as products, performance and exhibitions

New solutions are appearing on the scene to deal with the large scale volume of unstructured content (e.g. Islandora). Organisational units managing institutional repositories (IRs) are seeking solutions to deal with infrastructure issues associated with scale and storage of non-text based content; discoverability and accessibility issues; and workflow and collaboration tools.

From the commercial world large vendors such as Microsoft are marketing their "enterprise products" to universities as providing solutions to these types of problems.

Role of enterprise content management systems

Universities—like other major organisations—are grappling with new sources of content that are contributing to the growth in the amount of enterprise content. In addition they are handling increased volumes of paper-based documents, a higher percentage of which is being digitised. In response the large enterprise content management (ECM) platform vendors are evolving into "information management frameworks" (Clarke et al, 2012). While the literature tends to focus on an ECM in terms of administrative functions such as digitising services, document and records management, and—in more recent times—email and web content management, there has always been a digital asset function, which is underpinned by repositories.

Enterprise Content Management systems first appeared in the literature in the early 2000s in terms of ECM being recognised as a separate concept within the field of Information Systems (Grahlmann et al, 2012). In the literature and practice, a notable difference in definitions of an ECM is that whereas some authors restrict ECM content to semi- or unstructured data (Reimer in Grahlmann et al. 2002, 272), other authors such as Smith & McKeen (2003) include all content (Grahlmann et al). As an example of a narrow scope the AIIM website (2012) states:

> Enterprise Content Management (ECM) is the strategies, methods and tools used to capture, manage, store, preserve, and deliver content and documents related to organizational processes. ECM tools and strategies allow the management of an organization's unstructured information, wherever that information exists. http://www.aiim.org/What-is-ECM-Enterprise-Content-Management

Davis et al (2012) assert that while much of an organisation's relevant information assets are not defined as structured data, it is not always correct to think of them as "unstructured". They argue that information created, captured, and stored in enterprise communications platforms, desktop file applications, and content management systems all have a "level of structure in their native state and a significant amount of associated metadata that enables both categorization and analysis".

Another aspect highlighted in the literature has been the role of ECMs in content life cycle from capture through to management as they relate to organisational processes (Grahmann et al, 2012). For example, Smith & McKeen's (2003, 647) define ECMs as: "... the strategies, tools, processes and skills an organization needs to manage all its information assets (regardless of type) over their lifecycle."

For this paper the definition of an ECM by Grahlmann et al (2012, 272) is believed the more suitable for a university as it covers both learning and teaching content and research content:

> Enterprise Content Management comprises the strategies, processes, methods, systems, and technologies that are necessary for capturing, creating, managing, using, publishing, storing, preserving, and disposing content within and between organizations.

The business case for an ECM has been based on the economical savings from moving siloed departmental systems into single enterprise wide solutions, which then assists with regulatory compliance as well as connecting people with the content (the "digital assets") they need to do their jobs (deCathelineau, 2012). Vom Brocke et al (2011) cite a number of organisational benefits of an ECM including rationing of content, enhanced consistency and timeliness of content, improved collaboration, cost saving in information processing, and enabling sustainable knowledge management. The latter is generally driven by the need for compliance because of government and industry regulations. Korsvik and Munkvold (2010) highlight that the explosive growth of content, predominantly unstructured, means that more efficient processes for managing content have economic benefits as well. For these reasons ECMs have high relevance across all sectors.

ERPs, Research Administration Systems and Records Management Systems were developed in a print-centric environment which is now moving to a multichannel web-centric approach underpinned by digital

content. An enterprise approach is required to manage this content to avoid the siloed approach. In addition Vom Brocke et al (2011) conclude that the lines between business process management and ECM are becoming increasingly blurred.

Both Rosziewicz (2010) and Korsvik and Munkvold (2010) point out that because data has multiple uses specific to units of operation within the enterprise and because the enterprise must also deal with legacy data, the first challenge is to develop a core taxonomy to enable systems to interact and workflows to be developed that run across boundaries.

Roszkiewicz (2010) has identified poor planning and poor workflow development as a common factor in the underutilisation of ECMs. Therefore he concludes that implementing best practice ECMs needs the same attention as the deployment of Enterprise Resource Planning systems (ERPs). Grahlmann et al (2012) have concluded that because of the dissimilar characteristics and the diversity of functionalities within the scope of an ECM system, any implementation of an ECMS will be an integration of multiple software products. For example, on the one hand a records management archive requires that archived records remain unaltered. In another part of the organisation, on the other hand, the records management system may be required to allow versioning.

All the definitions of an ECM refer to it in terms of strategies, processes, and technologies rather than as a single product.

Role of institutional repositories

The purpose for hosting an institutional repository has changed over the last ten years. Institutional repositories evolved from a need to archive and preserve scholarly materials (Johnson, 2002), specifically research publications (Jain, 2011; Crow, 2002). Crow (2002) described it as an application scholarly in scope, cumulative and perpetual, open and interoperable. Lynch (2003) expanded the scope of institutional repository beyond just an archive application, describing it as a set of services providing a stewardship role.

Recent developments in Australia have seen the emerging trend of enhanced discovery and data sharing services (Lougee, 2009; Wong et al, 2009) and more focus on the content lifecycle - from data capture through to publication and preservation. Reznik-Zellen et al (2012) describe the emerging role of an institutional repository as providing a richer environment supporting virtual communities and promoting community engagement.

The contemporary institutional repository is now a rich ecosystem of data stores, content management functions, access management, discovery, and collaboration services. Jain (2011) concludes that (a) IRs are an ongoing service, not a project, as needs and technologies keep evolving and (b) they need a clear articulation of vision, strategy and tactics whether it is an institution-centred, researcher-centred or a general public centred vision.

Research data is characterised by the need to link to external tools and services, the importance of external discovery, and the diversity of formats both within—and between—research projects. The demand to manage research data through its entire lifecycle to meet a variety of specific discipline needs has resulted in an IR landscape that involves multiple technologies and services. In some cases specific research communities have developed their own customised solutions. Institutions are seeking ways to integrate these products into a single environment to move content and share data. The increasing need for scholarly communication and collaboration and the related dynamic technological environment demands that the institutional repository be flexible and adaptable to new trends.

Role of content management systems for learning and teaching

In recent times, managers of IRs have been turning their attention to the content used for learning and teaching. The literature shows that authors have recognised the relationships between repositories for research content and learning as far back as the early 2000s. In the following diagram (Figure 1), Lyon (2004) describes how she envisages the integration of research into the learning lifecycle. As research is undertaken and results are written up in publications, the latter are self-deposited by their authors into the institutional repository. This is then linked to by the learning management system (LMS). In addition to

the IR, the LMS also links to content in other organisational repositories: data generated by research as well as traditional learning objects. In this detailed model it clearly demonstrates the critical role of repositories in learning and teaching workflows.
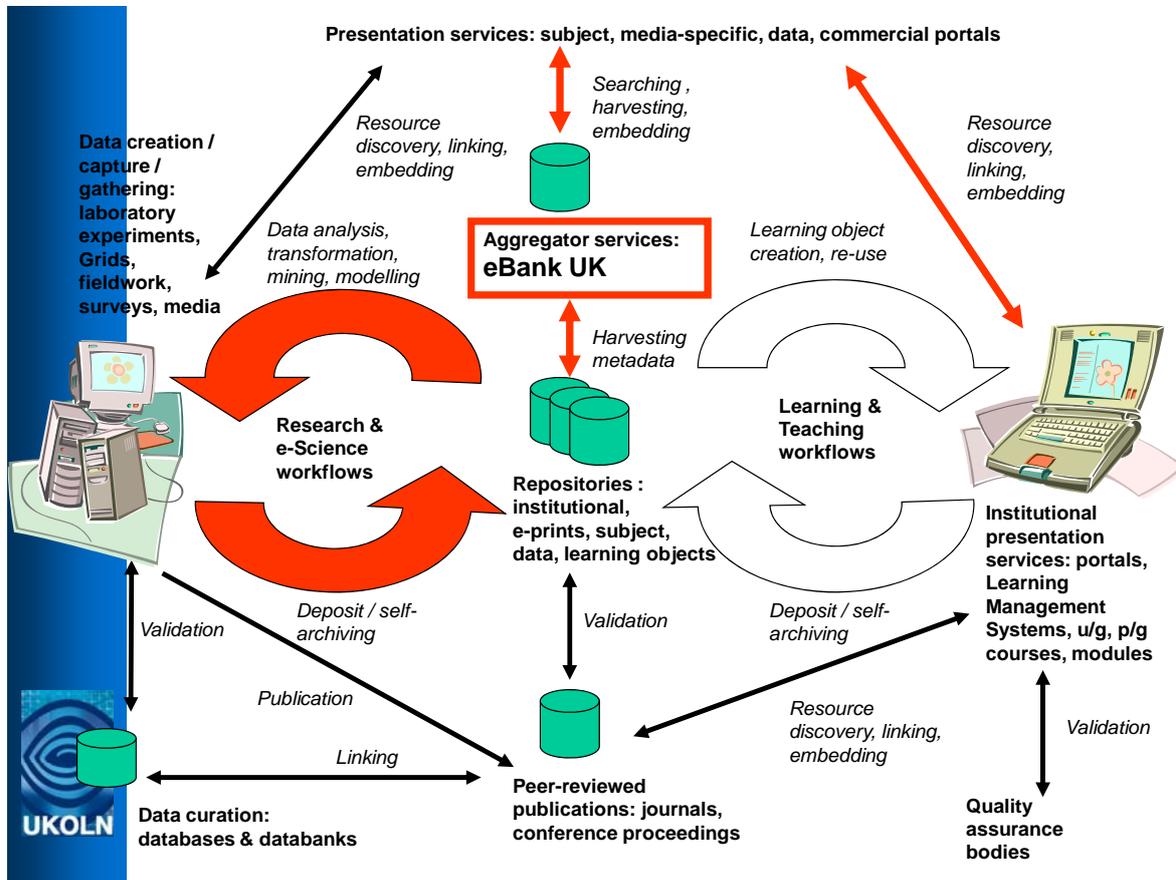
Figure 1: Integration of Research Content in the Learning Lifecycle

A review of some of the seminal early writings about the potential of institutional repositories reveals that they were envisaged in fact as containing potentially all of an institution's digital assets. Boundaries would logically be determined by institutional strategic goals. Lynch (2003), for example, suggests that "a mature and fully realized institutional repository will contain the intellectual works of faculty and students—both research and teaching materials..." According to JISC (2010) the fact that most institutional repositories have ultimately focused on research outputs can be attributed primarily to the lack of incentives for creators to share and the lack of recognition by institutions of the value of sharing learning and teaching content.

In the case of the JISC-funded CLIF (Content Lifecycle Integration Framework) Project, which concluded in March 2011, it investigated the integration work required to create a digital content lifecycle based on the Fedora Commons repository software, Microsoft SharePoint and the virtual learning environment, Sakai. The authors (Green et al, 2012) concluded that "the lack of the most up-to-date standards in the interfaces for content management presented by both Sakai and SharePoint ... does not make the task of getting these systems to work together any easier. It is concluded from this experience that all content management systems should be encouraged to make it as easy to get content out as to get content into them in order to facilitate seamless flow and enable the digital content lifecycle across systems".

This is not to say that all content should be stored in a single repository. Repositories should be considered as integrated components of larger systems and distributed infrastructure (Payette, 2010). This reinforces the definitions described earlier of IRs and ECMs in which they resemble an ecosystem of products with an overall governance structure relying on standards and workflows for integration.

Is integration the future of the IR and the traditional ECM?

From the discussion above it is evident that the definitions of an ECM and an IR are very similar. It could be said that the IR is evolving to become the ECM of the research community and increasingly the learning and teaching community. This is because the traditional ECM approach has not addressed the specific needs of these communities.

In the following discussion about the future of the IR in the ECM, the reference to the ECM is in the context of it as an enterprise framework and architecture rather than a discrete number of products such as TRIM or SharePoint or Fedora.

Seven reasons for integrating the IR and ECM

Common technology architecture: IR and ECM development have common technology drivers including workflow processes to handle the huge increase in unstructured content during its lifecycle, management of digital assets, and efficient access to material and technology usability issues. They share common architectural components in services, processes, access, repositories and storage. The following architectural diagram (Figure 2) is based on an ECM framework developed by Grahlmann et al (2012) and has been extended to encompass IRs, library, and learning and teaching content. For example, in the framework below an additional functional grouping (Transform/Enhance) has been added to align with key processes in learning and research environments.
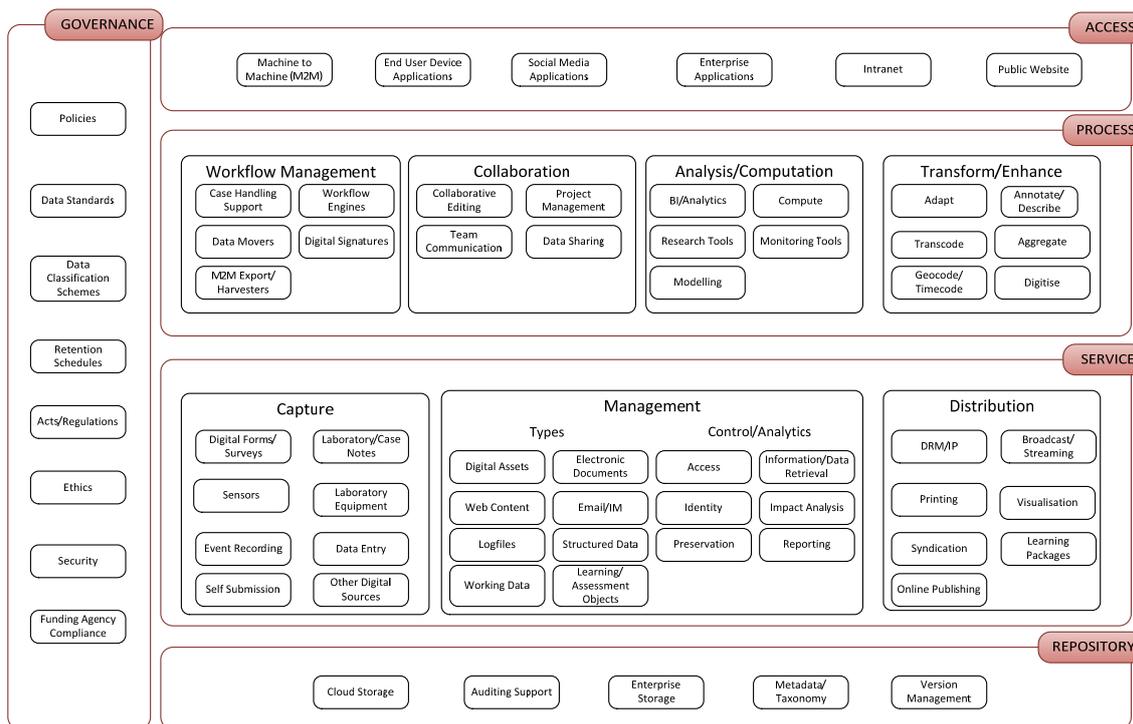
Figure 2: Integrated IR/ECM Architectural Components

Additionally in the Service component there are requirements for distributing content through broadcasting and streaming channels. These are used for learning and teaching (e.g. lectures), research (e.g. videotaped interviews) and administrative purposes (e.g. marketing campaigns). There is also a requirement to manage retrieval of information/data that is an integral function of both IRs and ECMs. New technologies and their applications are not only used for research but also are being used

increasingly in administrative services (e.g. sensors used for monitoring building environments for optimising energy use).

Both IRs and ECMs are progressively dealing with more unstructured data and non-text-based content. Many of the services and tools required to deal with unstructured data are common to both, e.g. streaming services, workflows and services for moving large objects. While IR and eresearch environments have traditionally been dominated by open source products, commercial products are now being used (e.g. SharePoint). While the ECM environment has traditionally been dominated by commercial products, open source products are now being used (e.g. ELMS Drupal based learning management system, Confluence, MySource Matrix). IRs and ECMs could benefit by developing single scalable enterprise solutions for these services.

Common reporting requirements: Both ECM/IRs require statistics, analytics and reporting for regulatory purposes. In both ECMs and IRs there is a need for audit trails to comply with specific acts and regulations or to meet standalone agreements on data access. Both need reporting for disposing of content or reporting out-of-date formats. For IRs statistics, analytics and reporting are required for measuring impact and usage. This can range from citation statistics to file downloads. Emerging needs around tracking data citation and managing DOI services also have automated reporting tools. Common technology approaches can be used to meet these requirements and where necessary built into the underlying core infrastructure.

Common data standards: Both ECMs and IRs are trying to address the whole content lifecycle. Naimoli and Fari (2008) observe that an ECM must support the whole lifecycle with the ability to move information from creation and capture to sharing and collaboration to publication then archival or disposal. Grahlman et al (2012) have concluded there are limitations of a standardised IT solution in a large heterogeneous enterprise. IRs are facing the same situation as solutions are being sought to manage research content during its lifecycle. The specific requirements of discipline groups make it difficult to rely on a single product (Payette, 2010). A critical success factor common to ECMs and IRs is the development of a general set of enterprise data standards (Roszkiewicz, 2010), which should be based on open access standards (Day and Ball, 2009). Any enterprise-wide or bespoke solution must adopt both mandated and open access standards and models. An integrated enterprise-wide approach to data standards is required to provide a consistent set of source content at the appropriate level of access in a timely manner.

Common content classes: All enterprise content is subject to the same compliance requirements with new and existing regulations and guidelines. Some content is shared or duplicated within IRs and ECMs. For example a copy of a data sharing agreement or copy of a grant may be retained with the research data as well as preserved in the central registry. The retention periods for preservation and re-use may differ from the regulatory requirements. Content classified under other legislation (e.g. privacy or copyright constraints) needs to be classified and managed consistently across the enterprise. One of the impacts of implementing these services and complying with regulations and guidelines in relation to research content and business content is that a common content classification scheme is required (e.g. Queensland Government, 2011).

New services such as minting Digital Object Identifiers (DOIs) for specific content may require some content to be handled differently than what is on current retention schedules. New discovery and data sharing services require a consistent approach to what can be shared and with whom. The discoverability of content will benefit from the capture and exposure of the latent relationships between content classes, e.g. storing the relationship between a publication, its underlying data and the researcher who produced it utilising semantic web technologies.

Common content creators: Both IR and ECM have common content creators. Grahlman et al (2012) and Korsvik and Munkvold (2010) have noted the important role of the unit (departmental) information manager. In most universities a central records function already exists to perform a centralised service although content collection may be decentralised to business units. Most universities have administrative functions which encompass processing, verifying and reporting research publications but require researchers or their nominees to input details about publications. These researchers are now also expected to upload their research content into IRs. In the IR landscape the role of data librarian and unit (research

group) information manager is becoming increasing important (Choi and Rasmussen, 2009; Delserone, 2008). Even where laboratory equipment is generating large volumes of data a person/s is usually responsible for managing that data. In more and more cases the people being asked to enter and manage content in the ECM are undertaking a similar role in the IR.

Common record quality issues: As administrative, academic and researcher communities seek to have content more discoverable and shareable the need for accurate descriptions and metadata is required. Typically the best person to describe and upload data is located at the point of capture or close to the source. Universities tend to have academic and research units in which their administrative staff are processing both business and academic/research content. There are advantages to developing a common skill set to deal with both IR and ECM content and to have a common systems approach. This becomes even more critical as more traditional business content needs to be preserved with the research data to make it reusable sometime in the future. For example, central record systems hold licenses, contracts, access agreements, ethics clearances that need to be preserved with the related research data. Other content required to manage and preserve research data is held within human resource and research administration systems. A core component of both the IR and ECM is common content ranging from structured data from enterprise systems such as human resources, research administration systems to unstructured data held in central record repositories. Content from these sources needs to meet defined quality standards for compliance, discovery and reuse. Given the synergies between the records management function, IR and other groups who manage structured and unstructured content, it makes sense to adopt a common approach.

Common issues on resistance to use and low uptake: There are shortcomings in both the current IR and ECM environment that are well covered in the literature that have led to low uptake of systems and resistance to mandated use. Seamless, integrated and easy-to-use tools for project management and collaboration are typically not readily available in either environment. Missing also is the availability of easy-to-use workflow engines and import/export tools to capture, process and move large amounts of data, especially to minimize data entry. Both environments need to respond to social media and other rapidly emerging web 2.0 tools in common use by the same end-user audiences. There are opportunities to economise on development and change management programs to deploy common solutions to both IR and ECM.

First Steps to integrate IRs with the ECM

From a design point of view both ECMs and IRs are faced with a similar challenge: how to design solutions that minimise the IT burden and make it easy to get content in and out at the appropriate levels of access (Naimoli and Fari, 2008; Roszkiewrcz, 2010). While both ECM and IRs suffer from psychological resistance to use (Grahlmann et al, Quinn, 2010) there are differences between the communities involved. Quinn (2010) noted that faculty members typically have a greater degree of autonomy when it comes to mandating use of such systems and the resistance to using IRs can be specific to the faculty group such as concerns about copyright and plagiarisms. Quinn also noted that the infrequent use of systems makes them appear difficult to use and in these cases inaction becomes a form of resistance. Andersen (2008) also noted that successful deployments of ECMs were dependent on end-user uptake and that passive resistance was an issue. In both the ECM and IR community, end-users need to be convinced of the benefits or rewards of using such systems. Both IRs and ECMs need to improve the ease of use to lower the burden on the end-user. Developing enterprise-wide solutions will require different solutions to target the different groups to maximise user benefits and require comprehensive change management and communication planning (Naomi and Fari, 2008).

IRs were originally focussed on the public user to provide open access to research material but increasingly they need to address both researcher and institutional needs. Conversely ECMs originally focused on institutional needs for compliance and to share documents but increasingly need to address individual and public requirements (e.g. web content). This difference in approach has led to the development of specific services, access mechanisms etc. For example, a key driver in the research community is reproducibility; the creation of persistent identifiers (PIDs) is a key requirement. As part of a global effort to improve access to research data, there is growing impetus for an international culture of data citation using the Digital Object Identifier (DOI) system.

This new integrated ECM architecture will need to address some critical requirements of IRs. The underlying theme of shareability and discoverability is a priority which is reflected in the use of open access standards, standard formats, shared taxonomies and ontologies. Examples of these can be seen in standard OAI-PMH, green and gold open access models. The new ECM will need to take into account its place in national and international infrastructure such as exporting and importing data to and from other systems, harvesting by Google, machine-to-machine connections to the National Library of Australia (NLA) and Research Data Australia (RDA).

Another implication of ECM operating as part of a national and international infrastructure is a much more complex security and access management regime. While identification and authentication services such as those offered by the Australian Access Federation (AAF) meet some of these requirements, the problem of multiple identifiers for individual researchers remains. Funding agencies, publication houses and government agencies (e.g. NLA, UK Names project) have their own identity services for researchers. If an institution is going to interact with external systems, methods are required to match identities especially if machine-to-machine transactions are to take place. To develop an integrated ECM this global issue needs to be addressed at the enterprise level.

Discoverability, shareability and the open access approach also need to address legal and regulatory complexities not normally found in the more traditional ECMs. The digital rights management (DRMs) frameworks required to preserve and share research material are much more complex not only to preserve the material but also to manage access. Enterprises will need to identify reporting, retention and auditing requirements in addition to the more traditional archival requirements.

The new integrated ECM architecture will need to address the requirement of persistence and long term sustainability. Some research data is required to be kept for long periods in cases where it underpins published research (e.g. drug discovery) or is data that can only be captured once (e.g. temperature reading at a site). A number of services have emerged to cater for this requirement. For example one of those is minting Digital Object Identifiers (DOI) issued when research is published. Within the institution any data file associated with a DOI will need to be given a special classification to ensure its long term preservation and link to the minted DOI. This has implications down through the architecture to storage management services.

There is also the need to capture, process and move data in a way that reduces manual input and labour-intensive intervention. This could range from providing easy to use tools for researchers through to writing complex workflows to capture, move and process data from laboratory or sensor equipment, e.g. process medical images in large volumes, handling and managing different file formats. Because of the nature of research and rapid technology development, scalability and adaptability have become critical to accommodate rapid growth and new developments in short timeframes. This has seen the increasing reliance on hosted and cloud services becoming part of the institutional architecture.

The biggest barrier to integration may well prove to be the organisational silos currently managing IRs and the various component systems of the current ECMs. The first step will be to develop an ECM architecture (both technical and information) that incorporates all of the current IR services and addresses emerging demands. As several authors have noted, it is not advisable to attempt to deploy an ECM solution across the enterprise in a single staged approach, but to target specific units or groups that could benefit from these services and approach. In an integrated ECM perhaps the target technologies should be those services that both ECM and IRs identify as a priority. These could include streaming services, storage and persistent identifiers, better access to key institutional data (e.g. HR data, central identity services, publication data), workflow technologies, project management and collaboration tools.

Conclusion

Both ECMs and IRs have many common elements and drivers which highlight the advantages of taking an enterprise-wide approach when planning further development of institutional repositories. There are economic benefits in developing a single enterprise-wide technical architecture to identify common technologies and processes to share development costs. It will involve utilising multiple technologies, processes, services and strategies. This should result in recognition that content management systems need to be treated as enterprise systems in their own right. This will help ensure the allocation of appropriate funding to build robust, scalable and sustainable solutions.

A consistent and coordinated approach to enterprise information management will help ensure that content is captured, managed, shared and preserved across the institution in such a way as to reduce the risk of breaching regulatory and policy guidelines. The use of semantic web technologies to expose the relationships among content in various sources will assist in breaking down information silos. An enterprise-wide approach to data standards and data classification schemes will allow the various staff with information management roles to better understand what data they need to preserve and what access requirements must be addressed.

An added benefit is that staff across the university will have a clearer view of what content is to be managed and how it is to be managed. More importantly source material should be more easily discoverable and shareable, thereby avoiding duplication and re-entry. An enterprise-wide approach aimed at reducing the effort required by end-users to capture and preserve content and to build information management expertise will improve uptake by the research community.

References

Andersen, R. (2008). The rhetoric of enterprise content management (ECM): confronting the assumptions driving ECM adoption and transforming technical communication. *Technical Communication Quarterly*,17(1), 61–87.

Bolton, S. (2010). The business case for the adoption of a UK standard for research information interchange. London: JISC.

Borgman, C. L. (2010). Research Data: Who will share what, with whom, when, and why? *Fifth China-North America Library Conference*. Beijing**:** 1-21 http://works.bepress.com/borgman/238/

Choi, Y. & Rasmussen, E. (2009). What qualifications and skills are important for digital librarian positions in academic libraries? A job advertisement analysis. *Journal of Academic Librarianship*, 35(5), 457-467.

Clarke, S., Davis, M., & Edwards, R. (2012). *Enterprise content management 2011/2012: Driving business value from content management*. London: Ovum.

Crow, R. (2002). The case for institutional repositories: a SPARC position paper. Washington, DC: SPARC.

Davis, M., Mukherjee, S., & Shelma, M. (2012). *Why BI has embraced enterprise search and retrieval (ESR)*. London: Ovum

Day, M., & Ball, A. (2009). Institutional repositories: the great debate. *Bulletin of the American Society for Information Science and Technology,* 35 (4*),* 22-25.
Retrieved from http://www.asis.org/Bulletin/Apr-09/Bulletin_AprMay09_Final.pdf

deCathelineau, S. (2012). Taking an enterprise view of content management. *Health Management Technology,* 33(3), 20-21.

Delserone, L. M. (2008). At the watershed: preparing for research data management and stewardship at the University of Minnesota Libraries. *Library Trends*, 57(2), 202-210.

Grahlmann, K. R., Helms, R. W., Hilhorst, C., Brinkkemper, S., & van Amerongen, S. (2012). Reviewing Enterprise Content Management: a functional framework. *European Journal of Information Systems, 21*, 268-286.

Green, R., Awre, C., & Waddington, S. (2012). The CLIF Project: The repository as part of a content lifecycle. *Ariadne, 68*.

Hey, T., Tansley, S., & Tolle, K. (Eds.). (2009). *The fourth paradigm: data-intensive scientific discovery*. Redmond, WA: Microsoft Research.

Jain, P. (2011). New trends and future applications/directions of institutional repositories in academic institutions. *Library Review,* 60(2), 125-141.
http://www.emeraldinsight.com/journals.htm?issn=0024-2535&volume=60&issue=2

Joint Information Systems Committee. (2010). *Digital Repositories InfoKit*. London: JISC.
http://www.jiscinfonet.ac.uk/infokits/repositories/types

Johnson, R. K. (2002). Institutional repositories: partnering with faculty to enhance scholarly communication. *D-Lib Magazine,* 8(11), 1-7.

Korsvik, K., & Munkvold, B. E. (2010). Enterprise content management in practice - one size does not fit all. In T. Fallmyr (Ed.), *Nokobit 2010*. Gjovik, Norway: Tapir Akademisk Forlag.

Lougee, W. (2009). The diffuse library revisited: aligning the library as strategic asset. *Library Hi Tech,* 27(4), 610-623.

Lynch, C. A. (2003). Institutional repositories: essential infrastructure for scholarship in the digital age. *Portal: Libraries and the Academy,* 3(2), 327-336.

Lyon, Liz (2004). Realising the scholarly knowledge cycle: the experience of eBank UK. *CNI Task Force Spring Meeting*. Alexandria, Virginia: CNI.

Naimoli, Michael & Fari, Gabor (2008). Breaking barriers to enterprise content management. *Applied Clinical Trials;* 17(9):58-64.

O'Brien, L. (2010). Innovation, university research and information infrastructure: making sound investments in information infrastructure. *EUNIS (European University Information Systems) 2010*. Warsaw.

Payette, S. (2010). Repositories and cloud services for data cyberinfrastructure. *2010 IASSIST Conference*, Ithaca, NY. http://duraspace.org/node/957

Queensland Government ICT Policy and Coordination Office (2011). *Information asset lifecycle*. Brisbane: Queensland Government.

Quinn, B. (2010).Reducing psychological resistance to digital repositories. *Information Technology and Libraries*, 29(2) 67-75.

Reznik-Zellen, R., & Adamick, J. (2012). Supporting virtual communities through disciplinary repository development. *Library Hi Tech,* 30(2), 275-290.

Roszkiewicz, R. (2010). Enterprise metadata management: How consolidation simplifies control. *Journal of Digital Asset Management,* 6, 291-297.

Smith, H. A, & McKeen, J. D. (2003). Developments in Practice VIII: Enterprise Content Management. *Communications of the Association for Information Systems*, 11, 647-659.

Thwaites, T. (2012). Creating a working environment for big data. *ANDS Newsletter*, 13:7.

vom Brocke, J., Simons, A., & Cleven, A. (2011). Towards a business process-oriented approach to enterprise content management: the ECM-blueprinting framework. *Information Systems and e-Business Management,* 9, 475-496.

Wolski, M., & Richardson, J. (2011). A framework for university research data management, *EDUCAUSE Australasia*. Melbourne.

Wong, W., Stelmaszewska, H., Barn, B., Bhimani, N., & Barn, S. (2009). *JISC user behaviour observational study: User behaviour in resource discovery: Final report*. London: JISC.

Authors: Malcolm Wolski (contact author)

Division of Information Services, Griffith University

Nathan, QLD Australia 4111

m.wolski@griffith.edu.au

Natasha Simons

Division of Information Services, Griffith University

Nathan, QLD Australia 4111

n.simons@griffith.edu.au

Joanna Richardson

Division of Information Services, Griffith University

Nathan, QLD Australia 4111

j.richardson@griffith.edu.au