

# OPEN-SOURCE ENTERPRISE CONTENT MANAGEMENT USING WORKFLOWS: AN IMPLEMENTATION CASE-STUDY FOR HIGHER EDUCATION INSTITUTIONS

C. MAICAN<sup>1</sup> R. LIXĂNDROIU<sup>1</sup>

**Abstract:** *Organizations are continuously challenged by the management of increasing amounts and varieties of digital information types and formats. In the form of a case-study, this paper presents a small part of the implementation process of an open-source Enterprise Content Management software and a workflow for routing documents within a university. The first part of the paper makes an introduction to the concepts surrounding the Enterprise Content Management software, the second part of the paper briefly presents the Business Process Management (BPM) and the third part focuses on the workflow process as part of the BPM within an organization.*

**Key words:** *enterprise content management, open source, university.*

## 1. Introduction

Like other major organizations, universities nowadays cope with new sources of content that influence the growth in volume of enterprise content. Additionally, they control increased volumes of paper-based documents, from which higher and higher percentages are digitized. As a reaction, large enterprise content management (ECM) platform vendors change into “information management frameworks”[2]. While the literature tends to focus on an ECM in the relationship with administrative purposes such as digitizing services, document and records management, and, lately, email and web content management, there has always been a digital asset function, which is underpinned by repositories [14].

Attempting to state what ECM is, some authors (e.g. [1], [11]) rely on the definition instituted by the Association for Information and Image Management (AIIM): “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”. In addition, other researchers have presented wider explanations of ECM: for example, [3] explains ECM as “the strategies, processes, methods, systems, and technologies that are necessary for capturing, creating, managing, using, publishing, storing, preserving, and disposing content within and between

---

<sup>1</sup> Centre “Advanced Research on Mechatronics”, *Transilvania* University of Braşov.

organizations”. Another characteristic underlined in the literature has been the functions of ECMs in content life cycle starting from capture until its management, as they relate to organizational processes [3]. For example, [10] 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 [3] is believed more suitable for a university as it covers learning, teaching 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 [16].

According to [5] the scope of ECM spans from suitable preparation, presentation, and personalization of information for end users to the development of algorithms for the purpose of information processing and information transformation. The concept of ECM thus outspreads the original notions described in [3], [8], [11].

Forrester [16] classifies the ECM software based on the way content is used: to support customer-facing business activities or to support internal company activities. The applications for managing persuasive content or researched-backed content for front-end business activities is now a separate type within the customer experience management environment, while content management technologies to be used inside organizations is detailed into three categories: foundational (check-in/checkout, permissions, archiving, and basic workflow, records and retention management, taxonomy, search and eDiscovery), business (document management and team collaboration), and transactional (multichannel capture,

enterprise report management and e-forms).

From a technical point of view, according to Gartner [16], the key components of an ECM suite are:

- Document management : check-in/check-out, version control, security and library services for business documents, compound document support and content replication;
- Image-processing applications: document capture (scanning hardware and software, optical and intelligent character recognition technologies and form-processing technology), the ability to store images of scanned documents in the repository as a regular content type in a folder, and the ability to route them through an electronic process;
- Workflow/business process management (BPM): document review and approval workflow, often including graphical process builders, and serial and parallel routing;
- Records management: long-term retention of content through automation and policies, ensuring legal, regulatory and industry compliance, enforcing retention of critical business documents, based on a records retention schedule;
- Web content management: includes content creation functions, such as templating, workflow and change management, and content deployment functions that deliver prepackaged or on-demand content to Web servers;
- Social content: document sharing, collaboration, knowledge management and project team support. Blogs, wikis and support for other online interactions are also important components. Social content, including video, is the fastest-growing category of new content in the enterprise;

- Extended components: digital asset management, document composition, e-forms, search, content and analytics, email and information archiving, email management and packaged application integration.

The studies conducted by both Gartner

and Forrester, presented in Figure 1, underline that commercial software leads the market while only one ECM open-source provider is presented. From this point of view, we believe that the biggest advantages of open-source software are the access to the application source code and low initial licensing costs.

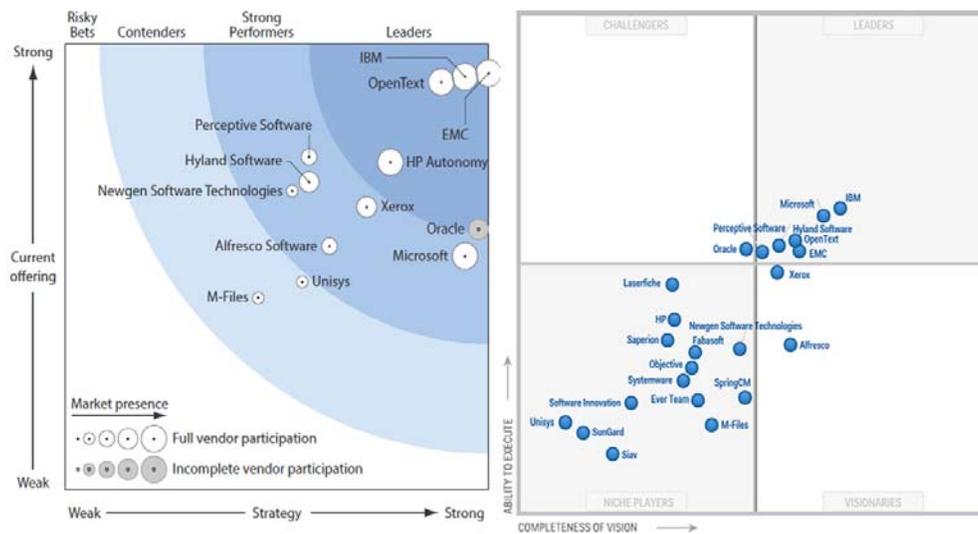


Fig. 1. *Forrester's Wave (left) and Gartner's Magic quadrant (right) for ECM.*

## 2. Flow management inside organizations

ERPs and Records Management Systems were initially created in print-centric situations that move to a multichannel web-centric approach at present, underpinned by digital content [16]. For avoiding the siloed approach (the lack of collaboration and standardization between business units) for managing this content, an enterprise approach is required. In addition, [12] conclude that the lines between business process management and ECM are becoming increasingly blurred because at present major ECM vendors include BPM as a component of their product. Both [7] and [4] indicate that, because data is used by multiple organizational units within the business,

each one with their type of usage, and because the business must also deal with legacy data and legacy systems, the first challenge is to develop a core taxonomy to enable systems to interact; furthermore workflows should be developed that run across boundaries [16].

In a world where there was only hardcopy information both on print and legacy systems, there used to be a physical limit to the extent of data that could be stored and retrieved. It is believed that Microsoft, with the introduction of their office-suite and OS, allowed users to overcome this physical boundary. As a result there was an exponential increase in information. Thus, document management software has become an essential component of many organizations as they

try to deal with the vast volumes of data they have gathered over time. At the simplest level, all users who create folders into which they store content produce in fact a basic document folder structure that permits them to store, retrieve and manage document content. The difference between this local-based type of document management and the one offered by ECM applications is only the scale of the content to be managed. ECMs are intended to support the whole organization aiming at the management, the creation, the storage, the retrieval and the retirement of information stored as documents. Contrasting to a file structure on a regular computer, an ECM entails a centralized repository that is used to manage the storage of any type of data that could be of value to an organization while protecting the same against loss.

According to [17], as content deposited inside an ECM is usually self-contained (for a document it cannot be implicitly assumed that it has any relationship with any other stored information), a well-designed document management repository encourages finding and sharing information easily. It does this via search tools while the documents being stored are classified according to classification schemes or taxonomies.

For implementing a document workflow inside an organization, a simple DMS is not enough, because it cannot handle the flow of the documents across the various structures of the implementing organization. In this case, for handling not only the document flow, but also other data and information flows inside organizations, a BPM application could be implemented. At their basic level, business processes are distinct events or collections of events accomplished by individuals within an organization with some results in a desired business goal. The goal can be characterized by any metrics and it may

overlap any organizational units, but at the same time it is typically considered necessary or desirable to a company's functioning.

At the more advanced end of the BPM scale, business processes can be seen not just for managing the current state of the company but also for proposing strategies and producing hypothetical business processes to be tested before deployment [17]. This type of business modelling still falls under the banner of BPM, although it is less about the organization as such and more about testing the current environment for identifying improvements that could be made.

The management of business processes should not be seen as isolated events because in many scenarios they are linked to any number of related business processes and in many cases business processes can be layered into larger, more inclusive processes that may or may not be co-dependent [17]. This is predominantly the situation when we regard the operational business processes such as those in the manufacturing field, where processes are often inter-related and co-dependent.

From a content management point of view, BPM has historically been associated with (Enterprise) Document Management applications to a certain extent, given that the execution of the processes in these applications often results in documents or records for storage.

In ECM, we can frequently see discussions about workflows as a substitute way of representing a business process [13]. Perhaps the essential difference is that, whereas business processes are able to exist independently of software, for ECM applications, workflows are normally embedded in those products or applications and for them, the business processes represent specific tasks within the application rather than the organization

as a whole. For example a workflow designed to achieve the progress of building a web page, starting from the author creating the page, passing through the editors approving it and ending with an automated release by the WCMS, is a business process even if this one does not usually exist outside the WCMS.

A BPM cycle can be clustered into six essential components that develop from each other sequentially, starting with: process vision, process design, model of the process, process implementation, monitoring of its success of failure and optimization, if required.

The process is not based on a single life-cycle because as soon as the process has been improved, it often returns to the vision stage for taking into account the fact that the world is continuously fluctuating and the processes have to progress for incorporating these fluctuations [17]. Therefore, the designed processes are continually tested for effectiveness.

Among other tools which are often necessary for management processes and are somehow related to BPM, and often part of ECM, records management allows an organization to store a given number of information items for later evidential retrieval. Records could be anything from contracts to invoices and other useful documents and thus, anything that an organization may need to keep for later remembrance or endorsement of certain events. Policies and procedures have to be created for managing the life cycle of the records, from their creation to their final retirement and destruction and for ensuring that a full information trail is available for every aspect of the process [12]. The policies and procedures govern the approach in which a record is captured and preserved, and the way users may access the record during its entire life cycle. According to [18], efficient management of records allows organizations to:

- Know what records they have, and locate them easily;
- Increase efficiency and effectiveness;
- Make savings in administration costs, both in staff time and storage;
- Support decision making;
- Be accountable;
- Achieve business objectives and targets;
- Provide continuity in the event of a disaster;
- Meet legislative and regulatory requirements;
- Protect the interests of employees, clients and stakeholders.

Records management offers solid return on investments to organizations, starting from economic best practices in reducing storage costs for documents, and up to putting into practice legislative requirements. An unmanaged record system makes the execution of responsibilities more demanding, costs organizations time, money and resources, and makes them vulnerable to security breaches, prosecution and embarrassment. In an unmanaged records environment, up to 10% of staff time is spent looking for information [18].

### **3. Implementation of an open-source ECM for document registration**

The single ECM open source software present in both Gartner and Forrester diagrams is Alfresco and consequently we have decided to implement it. Among the requirements for an ECM, we could mention: file organization, document management, taxonomies, workflows, document versioning, file and folder-level security, indexing and searching - all these features are already present in the chosen ECM.

Regarding the document flow inside the university, we have considered the implicit workflows and, because they do not

comply with our business process management, we have had to redesign them. The most important flow regards the administrative document flow which we have described and implemented using *BPM Activiti* and which is depicted in Figure 2. In this flow, all the documents have to pass through the Registration Office and they direct the documents according to their required recipient. The electronic workflow is always started by the Registration Office, being available only in their user-interface. All the documents from the University are addressed to the Management, therefore the electronic documents follow this route. The Management may approve or reject directly the document or, as it is most often the case, may ask for additional information from various administrative offices – this is electronically accomplished by using the Exclusive

Gateway (multiple IFs) that directs the document back to the Registration Office for being transmitted to another office (depicted as an User). The user has to fill-in a form or may attach additional documents for fulfilling the Management's requirements for more information. These steps could be repeated infinitely or until the Management decides that the initial document is either approved or rejected.

In order to easily find the documents inside the systems, these can be characterized using tags. Also all the documents passing through the Registration Office are attached a registration number and are automatically added to a DataList accessible only to the users inside this office, containing all the required data for identifying what is the office or department where the electronic and the physical documents are processed at that moment of time.

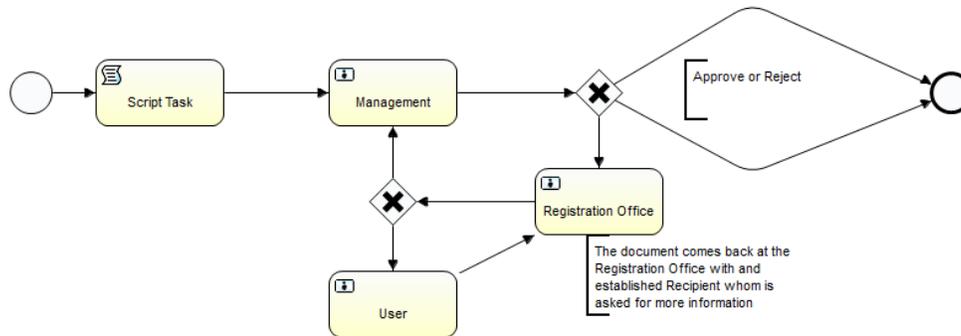


Fig. 2. Workflow for document registration and tracking.

In addition, a due-date is attached to electronic documents, because of the regulations the organization has to respond to requests in maximum 30 days. By using this date we can easily find out what office is late with providing the management with the required information for finalizing the request in due time.

The documents inside the workflow and after it is finished could be part of the non-repudiation scenarios [6] via using digital

signatures for signing documents and adding comments to them besides the comment fields available directly in the workflow interface.

### 3. Conclusions

The growing volume, complexity and variety of digital content makes ECM applications a vital element of organizational IT. Additionally, the

necessity of making quick and correct decisions imposes changes to the organizational environment, positions ECM in an important role for decision support. Using an open-source ECM, we have implemented a system for tracking documents inside a university based on and using workflows rules. This system involves business rules that are enforced by applying BPM processes as workflow steps. All the generated records are kept in place based on records management rules.

## References

1. Alalwan, J. A., Weistroffer, H. R.: *Enterprise content management research: A comprehensive review*. In: *Journal of Enterprise Information Management* (2012), 25(5), 441–461.
2. Clarke, S., Davis, M., Edwards, R.: *Enterprise content management 2011/2012: Driving business value from content management*. London. Ovum (2012).
3. Grahlmann, K. R., Helms, R. W., Hilhorst, C., Brinkkemper, S., van Amerongen, S.: *Reviewing Enterprise Content Management: a functional framework*. In: *European Journal of Information Systems* (2012), 21, p. 268–286.
4. Korsvik, K., Munkvold, B.E.: *Enterprise content management in practice - one size does not fit all*. In T. Fallmyr (Ed.), *Nokobit 2010*. Gjøvik, Norway: Tapir Akademisk Forlag.
5. Laumer, S., Daniel Beimborn, Christian Maier, Christoph Weinert, *Enterprise Content Management* (2013), DOI 10.1007/s12599-013-0291-3.
6. McCullagh, A., Caelli, A.: *Non-Repudiation in the Digital Environment*, In: *First Monday*, Volume 5, Number 8 - 7 August 2000.
7. Roszkiewicz, R.: *Enterprise metadata management: How consolidation simplifies control*. In: *Journal of Digital Asset Management* (2010), 6, 291–297.
8. Scheepers R: *A conceptual framework for the implementation of enterprise information portals in large organizations*. In: *European Journal of Information Systems* (2006), 15(6):635–647.
9. Simons, A., et al.: *Decision support capabilities of enterprise content management systems: An empirical investigation, Lessons Learned from Implementing Enterprise Content Management at the National Public Administration in Liechtenstein*. In: *Enterprise Content Management in Information Systems Research, Progress in IS 2014*, pp 199–216, DOI 10.1007/978-3-642-39715-8\_12.
10. Smith, H. A., McKeen, J. D.: *Developments in Practice VIII: Enterprise Content Management*. In: *Communications of the Association for Information Systems* (2003), 11, p. 647–659.
11. Tyrväinen, P., Päivärinta, T., Salminen, A., Iivari, J.: *Characterizing the evolving research, on enterprise content management*. In: *European Journal of Information Systems* (2006), 15(6):627–634.
12. vom Brocke, J., Simons, A., Clevén, A.: *Towards a business process-oriented approach to enterprise content management: the ECM-blueprinting framework*. In: *Information Systems and e-Business Management* (2011), 9, 475–496.
13. von Brocke, J.: *BPM Research in Education – On the Role of Enterprise Content in Business Process Management*. Available at: <http://www.bptrends.com/class-notes-bpm-research-in-education-on-the->

- role-of-enterprise-content-in-business-process-management/
14. Wolski, M., Simons, N., Richardson, J.: *ECMs and Institutional Repositories. The Case for a Unified Enterprise Approach to Content Management*. In: THETA: The Higher Education Technology Agenda 2013 , 7-10 April 2013, Hobart, Tasmania.
  15. \*\*\* Forester Research, *The Forrester Wave™: Enterprise, Content Management*, Q3 2013, 2013.
  16. \*\*\* Gartner. *Magic Quadrant for Enterprise Content Management*, 2013.
  17. \*\*\* Cognidox. *A Buyers Guide for Document Management Systems*, 2008.
  18. \*\*\* The National Archives of Scotland, Records Management. Available at: <http://www.nas.gov.uk/recordkeeping/recordsmanagement.asp>.