Scientific
Research
Publishing

# Evaluating Enterprise Content Management Tools in a Real Context

**M. J. Escalona\*, F. J. Domínguez-Mayo, J. A. García-García, N. Sánchez, J. Ponce**

Web Engineering and Early Testing Research Group, IWT2, University of Seville, Seville, Spain
Email: *mjescalona@us.es, fjdominguez@us.es, josepg@us.es, julian.garcia@iwt2.org,
nicolas.sanchez@iwt2.org

## Abstract

**Managing documentation in a suitable way has become a critical issue for any organization. Organizations depend on the information they store and they are required to have appropriate mechanisms to support the functional needs of information storage, management and retrieval. Currently, there are several tools in the market, both free software and proprietary license, normally named Enterprise Content Management (ECM) tools, which offer relevant solutions in this context. This paper presents a comparative study among several of the most commonly used ECM tools. It starts with a systematic review of the literature to analyze possible solutions and then it defines a characterization schema instantiated in a particular case, the Regional Government of Andalusia.**

## Keywords

## 1. Introduction

Nowadays, one of the most important assets for any organization is the information available at the staff's disposal. A suitable management of information is essential, therefore, document management is a key aspect [1]. Besides, in the current digital era, the right administration of digital documentation is crucial and handling the vast amount of digital documentation involves a hard task for many organizations, since lacking the proper mechanisms to manage documentation provide timely results.

Enterprise Content Management (ECM) [2] tools are tools solutions offering support mechanisms to assist

---

*Corresponding author.

this task. However, today's variability makes it difficult to know which of them the best for a concrete environment is.

This paper presents a comparative study among five of these tools as part of three main objectives:

- Find out which ECM tools the market offers at present.
- Compare them under homogenous criteria.
- Illustrate how this study can be adapted to a concrete environment.

A systematic literature review will be conducted in order to cope with the first aim. The paper shows that this line of work is not enough to analyze the current situation, thus it must be complemented with some other resources, which together with the former, will provide five tools for our study. The second aim is to define a characterization schema that will be instantiated in the paper for each approach under study.

Once finished, this work illustrates how the global comparative study can be customized in a concrete environment. For the third aim, the practical case of THOT Project will be used [3]. The Agencia de Obra Pública of Junta de Andalucía, Spain, is developing the project in liaison with the University of Seville, Spain, as it is further detailed.

The paper is structured as follows so as to cover the three aforementioned objectives. Firstly, Section 2 presents a related work section with a global view of ECM. Then, Section 3 introduces the mechanism used to get the two first goals, based on SLR and characterization schema. Later, Section 4 explains the detailed application of this mechanism and Section 5 presents THOT project and the study performance in its context. The paper finishes stating relevant conclusions and future lines of work.

## 2. Related Work

As regards reviews related to ECM systems, Scott [4] assesses the factors that lead to the user acceptance of an ECM system. Findings reveal the importance of cognitive engagement with technology. It mainly shows how a document perspective provides insights on the surprising results and highlights the importance of including the cognitive engagement construct in technology acceptance studies. Alalwan and Weistroffer [2] providea comprehensive literature review of ECM research. It proposes a conceptual framework of areas of concern regarding ECM and shows an agenda for future research on ECM, based on review and conceptual frameworks. After revising and classifying 91 ECM publications, it was concluded that ECM involved several sophisticated and interacting technical, social, organizational and business aspects. The authors suggested that today's literature concerning ECM could be grouped into three main pillars: the first pillar consists of the four ECM component dimensions (tools, strategy, process and people). The second pillar deals with the enterprise system lifecycle (adoption, acquisition, evolution and evaluation) and the final pillar is the strategic managerial aspect (change management and management commitment). An agenda for future research around the aforementioned pillars, in terms of review and suggested conceptual framework, is recommended.

As far as ECM implementations are concerned, Haug [5] includes a definition of a process model for ECM implementation in SMEs. This author identifies success factors related to ECM system implementation and proposes a definition of a new pattern for ECM technology development, compared to existing case studies. Thus, it contributes to the sparse literature on ECM implementation. In fact, the case seems to be the first longitudinal study regarding ECM implementations in SMEs.

In [6], Van Rooij explains that lessons learned from the implementation of ERP reveal that such implementations may be compromised by a large number of legacy issues. The author argues that the same issues may similarly affect the implementation of ECMs. Therefore, it is advised, with due adaptation, to take these issues into account when devising implementation strategies for ECMs.

In relation to ISO standards, many industry associations publish their own lists of particular document control standards that are used in their particular fields. ISO 2709:2008 [7] specifies the requirements for a generalized exchange format that will hold records describing all forms of material capable of providing bibliographic description as well as other types of records. It neither defines the length or the content of individual records nor assigns any meaning to tags, indicators or identifiers, being these specifications the functions of an implementation format. This ISO standard describes a generalized structure, a framework especially designed to enhance communication between data processing systems and not to be used as a processing format within systems. ISO 15836:2009 [8] establishes a standard for cross-domain resource description, known as Dublin Core Metadata Element Set. Similarly to RFC 3986, ISO 15836:2009 does not limit what might be a resource. This ISO stan-

dard defines the elements typically used in the context of an application profile, which constrains or specifies their use in accordance with local or community-based requirements and policies. However, it does not add implementation details, since it is out of its scope.

Another ISO standard to consider is ISO 15489 part 1 and part 2 [9]. This standard focuses on the principles of records management and establishes the basic requirements that allow organizations to establish a best practices framework that improves, in a systematic and effective manner, the creation and maintenance of its records, supporting in this way the organizational policy and objectives.

As regards capabilities about information retrieval services, ISO 23950 [10] defines the Information Retrieval Application Service and specifies the Information Retrieval Application Protocol. The service definition describes services that support capabilities within an application; these services are in turn supported by Z39.50 protocol. This description neither specifies nor constrains the implementation within a computer system. The protocol specification includes the definition of the protocol control information, the rules for exchanging this information and the conformance requirements to be met by the implementation of this protocol. This standard addresses connection-oriented and program-to-program communication intended for systems supporting information retrieval services and organizations such as information services, universities, libraries and union catalog centers. It does not address information exchange among terminals or via other physical media.

ISO 10244:2010 [11] gives detailed information associated with the activities organizations perform when documenting existing work or business processes (business process baselining), defining the level of information required to be gathered, methods of documentation for work or business processes, and procedures used when evaluating or analyzing work-business processes. This ISO standard provides organizations with tools to identify relevant aspects of work-business processes and document them in a standardized format, thus permitting a detailed analysis and identification of relevant technology, so that the processes can improve.

## 3. Our Mechanism for the Study

As it was introduced, our general study is divided in two different parts. The first one deals with looking for approaches and tools that will be considered in the study and the second one will compare them under a concrete set of criteria. For the former, we will use a Systematic Literature Review (SLR) and for the latter, we will define a characterization schema.

A Systematic Literature Review (SLR) is mainly carried out in order to find and develop innovative ideas for further research. In [12], authors conceive SLRs as the means of completing processes based on identifying, evaluating and interpreting all available documents focused on particular research questions or a specific investigation area. However, this process is not only associated with scientific environments, but also with any domain or environment (such as research, enterprise or engineering, among others), as it is not exclusively related to research work. In addition, it is normally used as a method for carrying out comparative studies on software tools or technology proposals.

Therefore, SLR aims to provide an exhaustive summary of the literature relevant to a research, technological or technical question.

The use of SLRs is relatively recent in the Software Engineering (SE) context, but it has gained significant importance in this area as a means to identify, evaluate and interpret all available data to answer research questions on a particular topic in SE. It has been growing in importance as a systematic and structured approach regarding literature reviews since 2004, when Barbara Kitchenham [13] proposed special guidelines that were adapted to cope with specific problems in the SE area. These guidelines have been used and evaluated in many contexts [14]-[17]. Last year, Kitchenham's proposal was updated again [18] in order to be implemented, taking into account recent results published by software engineering researchers concerning their experiences when performing SLRs, as well as their advices for improving the SLR process.

Moreover, there are other ideas or views to conduct systematic reviews. For instance, in [19], authors introduce different perspectives of SLRs. They issue their proposals after systematically selecting and analytically studying a large number of papers (SLRs) in order to understand the state-of-the-practice of search strategies in evidence-based SE.

In consequence, these SLRs proposals are highly directed towards answering research questions on some scientific knowledge. Nevertheless, SLR is not enough for a study led to compare technologies or tools solutions. For this reason, we will complement SRL with the two points below to carry out our study:

1) We will try to look for responses to our research questions not only in the research environment, but also in the context of a common search engine for enterprises, like Google.
2) We will analyze enterprise tendencies, like Gartner.

Apart from offering a SLR, we aim to provide a comparative study. For this purpose, we propose a characterization schema, based on SEG approaches [20]. This schema will allow us to offer a homogenous evaluation of each approach under study that, at the end of the process, will be essential for stating final conclusions and learned lessons.

## 4. Applying the Study

### 4.1. Finding the Approaches

As introduced, the guidelines for the systematic review stated in this work follow the protocol defined by Kitchenmham [15], which is one of the most acknowledged in SE. In addition, we take into consideration Wohlin and Prikladnicki's conclusions [21] about SLRs in SE. They consider that the search strategy is key to ensure a good starting point for the identification of studies and ultimately for the actual outcome of a particular study.

Nevertheless, this proposal initially centers on systematic reviews of research studies. In light of this, we have adapted this proposal to focus on ECM systems studies and all those related fields. A SLR essentially involves three phases: 1) planning the review; 2) conducting the review and 3) reporting the review.

- The stages associated with planning the review are: identification of reviewer needs, specification of the research question(s), development a review protocol and evaluation of that protocol.
- The stages associated with conducting the review are: identification of research, selection of primary studies, study of quality assessment, data extraction and monitoring, and data synthesis.
- The stages associated with reporting the review are: specification of dissemination mechanisms, formatting the main report and evaluation of that report.

The planning phase has two main goals; on the one hand, deciding which method will be used to conduct the review and on the other hand, identifying and formulating the thesis that the systematic review will prove. Regarding the first goal, this work aims to answer the next questions:

- Question 1: What ECM systems currently exist in the market and what do they offer?
- Question 2: How ECM systems can be adapted to the general guidelines of the Andalusian Public Administration?
- Question 3: What is the most appropriate ECM system the Andalusian Public Administration, and more specifically, the Contracting Services for Transport and Infrastructure Constructions must use?
- Question 4: What areas of improvement are needed for the selected ECM system?

A large number of identified search keywords picked up from these questions will be used in the review process, as **Table 1** summarizes.

The following databases have been considered in the systematic literature review: ACM Digital Library, EI Compendex, IEEE Xplore, ISI Web of Knowledge, Science Direct, SCOPUS, Springer Link and Wiley Inter Science Journal Finder. **Table 2** shows the fields where the defined search keywords have been applied in each database. Besides, this table represents the logical relationship among these fields.

Once all planning phase goals have been achieved, the review process enters in the review phase, which consists in finding and evaluating the adequacy and relevance of many primary studies associated with the research question as possible sources for further analysis. Primary studies are searched through the aforementioned databases by means of the keywords represented in **Table 1** together with the search fields shown in **Table 2**. Then, a strategic definition for evaluating the adequacy and relevance of the studies is needed after the search.

Firstly, keywords of **Table 1** are searched for each logical criterion in the search field included in **Table 2**. Secondly, the set with the previous primary studies is reduced according to the following inclusion criteria:

1) The primary study must have been published in the last four years, that is, from 2010 to 2013 (both included). This exclusion criterion is considered realistic and acceptable in the context of this work, because the number of ECM systems and versions has increased in the last years. Therefore, the considered primary study must be recent in order to infer useful conclusions.
2) The paper must focus on Computing Science.
3) The paper must have been published in any influential magazine (for instance, Journal Citation Reports indexed).

**Table 1.** Search keywords.

| Abbreviation | Search Keywords |
|---|---|
| SMEC | Solutions for Managing Enterprise Contents |
| SMDP | Software for Managing Document Processes |
| EMCRE | ECM in Real Environment |

**Table 2.** Logical criteria for fields used in each database.

| Database | Logical Relationship of Search Fields |
|---|---|
| ACM Digital Library | "Title & Abstract" |
| Ei Compendex | "Title & Abstract" |
| IEEE Xplore | "Abstract" or "Publication Title" or "Index Terms" |
| ISI Web of Knowledge | "Title & Abstract" |
| Science Direct | "Abstract, Title, Keywords" |
| SCOPUS | "Title & Abstract" |
| Springer Link | "Title & Abstract" |
| Wiley Inter Science Journal Finder | "Title & Abstract" |

Thirdly, a new discrimination is conducted through a fast reading of each primary study. First of all, the title theme of the primary study must be linked to the topic of this work. For example: "ECM", "Enterprise Contents Management" or "Document Process". The introduction and abstract must mention the goals of the research questions posed in this section, once that condition is satisfied and this primary study is catalogued as promising.

Finally, after carrying out this review, we neither find concrete solutions for ECM systems nor studies on them. We have found some work associated with theoretical proposals in the context of ECM systems. Therefore, classic search engines and those running on the Internet have been selected in order to conduct a specific survey of this type of systems. In addition, we have considered the last Gartner's analysis concerning Enterprise Content Management [22], which is popularly known as Magic Quadrant and was presented in October 2012. Gartner [23] is an international research and consulting company dealing with Information Technology (esta última frase quizá debería ser una nota al pie, una aclaración).

It follows the criteria below among the pre-selection ones:
- Basic functionality (core components). It takes into account the capabilities and/or applications to manage the content life cycle:
o Document Management. It refers to basic skills such as locking and unlocking documents (check-in/check-out), change control and versioning, full-text indexing, security, library cataloging and document types, and advanced capabilities such as support for document composition, association life cycles, taxonomy and content replication.
o Records Management. It is used for long-term retention of content through automation and implementing policies that ensure legal compliance and regulatory framework. The minimum requirement is the ability to enforce retention of business critical documents, based on a records retention program. Higher ratings are given for compliance with certification standards and the requirements model for electronic document management.
o BPM (Workflow/Business Process Management). It deals with supporting business processes, routing content, assigning work tasks and states, and creating audit trails. The minimum requirement is the document review and approval workflow. Higher scores are given to graphics-capable solutions to define routing processes in series and in parallel.
o Document Imaging/Image-processing. They are applications for capturing, processing and managing images of paper documents. There are two possibilities for these component solutions: 1) Document Capture (scanning hardware and software, character recognition technologies and forms processing technology), either using native capabilities or through a formal partnership with a third provider, for example Knowledge Lake,

Kofax, EMC (Captiva) or IBM (Datacap) and 2) the ability to store images of scanned documents as "other" type of content in the repository, both in a folder or through an electronic process route.

o Interoperability functions/extended components. They deal with the ability to share data and enable the exchange of information and/or digital assets and the ability to generate electronic forms and integrate them into email and packaged applications.

o WCM (Web Content Management). It controls the content and interactions with Web solutions. This includes content creation features, such as templates, workflows and change management, and content distribution functions that offer packaged or on-demand content to Web servers. The minimum requirement is a formal partnership with a WCM provider.

o Social content/collaboration. It refers to collaboration, knowledge management and shared documents, such as blogs, wikis and online support between users. Social content, including videos, is the fastest growing category of new content in organizations. This feature is becoming increasingly important due to the incorporation of social networks in today's society.

• Market positioning of ECM solution taking a strategic approach, that is, how they can help these solutions for businesses and organizations to take control of their content and, consequently, increase efficiency, enhance collaboration, and facilitate information sharing.

According to this criterion, firstly the analysis conducted in 2012 by Gartner ® on Enterprise Content Management tools is taken into account and secondly, the study of global Internet searches trendlines conducted on certain key words is considered.

On the one hand, we integrate Gartner's analysis into ECM based on the evaluation of different weighted criteria whose overall results are represented by means of a two-dimensional matrix that assesses suppliers in terms of their expectation for future performance and tools execution capacity.

The weighting established by Gartner for basic functionalities is listed below:

• Document Management: 15%
• Document Imaging/Image-processing Applications: 18%
• Workflow/BPM: 22%
• Records Management: 13%
• Web Content Management (WCM): 7%
• Social Content/Collaboration: 15%
• Interoperability/Extended components: 10%

**Figure 1** shows the result obtained through Gartner Magic Quadrant for ECM solutions ®.

On the other hand, a study on the Internet searching habits regarding ECM solutions has been recently conducted in Spain. This study has been carried out through Google Inc. Insights for Search. This tool generates trendlines in Google from certain keywords historical search.

**Figure 2** shows the result of this study:

These trendlines have allowed us to corroborate the importance of various ECM solutions, showing high activity in Alfresco solution.

Therefore, the following ECM solutions (in alphabetical order) have been pre-selected attending to preliminary analyses: Alfresco [24], Docu Ware [25], EMC Documentum [26], IBM File Net [27], Knowledge Tree [28] Nuxeo [29], Open Text [30] and MS Sharepoint [31].

## 4.2. Defining the Characterization Schema

The characterization schema checklist is composed of a set of attributes and qualities assessed in each ECM solution.

This schema enables applying a brief and homogeneouse valuation of each approach under study. It is described in parts to facilitate reading and initial presentation, although, as noticed, the instantiation of the evaluation of each tool is presented as a whole.

Functional modules: The following functionalities are included natively and minimally in this group of items. They are described in the previous section so, in this case, **Table 3** only enumerates the possible values that each of them can take. ✓ represents that the feature is completely covered by the approach, ✗ means that the feature is not supported by the approach and ⊘ represents that the feature is partially supported by the approach.

For the next groups of items, features can take a value between 1 and 4. 1 represents that the approach does not cover the feature and 4 is the best value given. Values 2 and 3 are obviously intermediate values depending
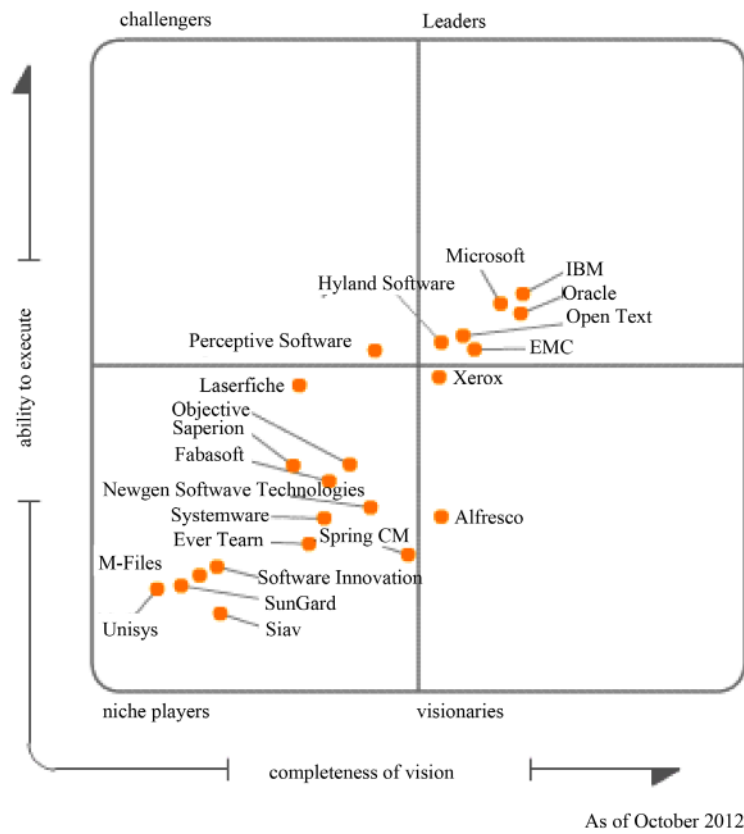
**Figure 1.** Gartner comparative results.
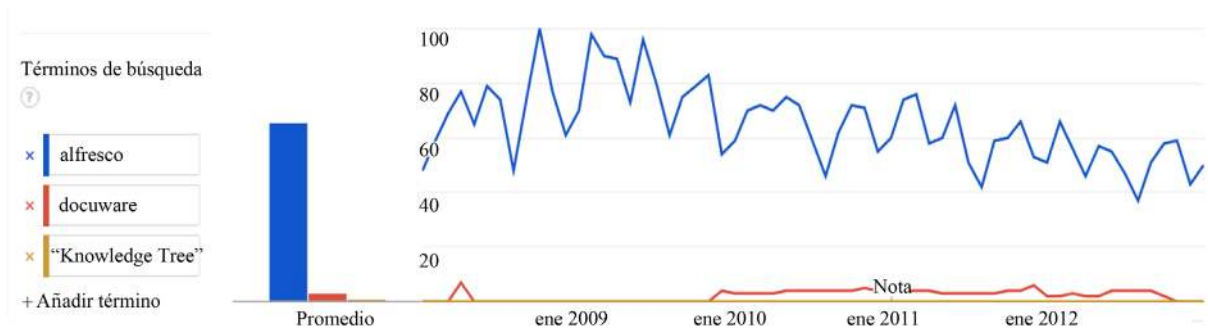


**Figure 2.** Google Inc. results.

**Table 3.** Functional modules.

| Group | Features | Values |
|---|---|---|
| | Document Management | {✓, ×, ⊘} |
| | Records Management | {✓, ×, ⊘} |
| | BPM (Workflow/Business Process Management) | {✓, ×, ⊘} |
| Functional Modules | Document Imaging/Image-processing | {✓, ×, ⊘} |
| | Interoperability functions/extended components. | {✓, ×, ⊘} |
| | WCM (Web Content Management) | {✓, ×, ⊘} |
| | Social content/collaboration. | {✓, ×, ⊘} |

on their coverage.

a) User Orientation. Although ECM systems offer standard solutions on its orientation towards the end user, many companies need using easy and versatile systems because not all their employees have the same user profile to handle computer tools. Next, the Sub-features of this feature are described below:

▪ Usability compliance. It measures whether the system is user-friendly for non-expert employees in ICT tools.

▪ Accessibility. It analyses the accessibility level of the system based on the UNE 139803:2012, which is based, in turn, on WCAG 2.0.

▪ Document preview. It evaluates whether the system provides a suitable interface for previewing and resizing documents.

▪ Drag & Drop. It reviews whether the system provides visual utilities of drag & drop.

▪ Bulk uploads. It facilitates massive uploads of documents and evaluates whether the system allows users to upload a large amount of documents in a single operation.

▪ Undo. It assesses whether the system allows the user to undo an operation.

▪ WYSIWYG Editor (What You See Is What You Get). It evaluates whether the system has a user-friendly WYSIWYG editor for creating workflows or forms, for example.

▪ Customization. It measures the customization degree of GUI (Graphical User Interface) the ECM system provides.

▪ Groups and Social networks. They help a business gain contacts, clients and increase public conscience. Even entrepreneurs who run small businesses from their homes can take advantage of this resource to set up a global presence. Consequently, we have considered this aspect to be essential. This Sub-feature evaluates whether the system provides an integration mechanism for sharing documents and information through different social networks such as Facebook, Twitter and LinkedIn, among others.

▪ Multilanguage. This Sub-feature evaluates whether the system provides a GUI multilanguage and allows both, managing information and documents in different languages and integrating the ECM system with translation software.

b) Functionality to capture, access, retrieve and view documents. The ability to capture, access, retrieve and display, within the group, includes those features that let anyone customize the system according to the preferences of users or the organization being implemented. Next, we describe the Sub-features of this feature:

▪ Degree of Cataloging. It assesses whether the system allows complete and flexible cataloging folders and documents, *i.e.*, the system provides default metadata and allows creating new metadata. In addition, this Sub-feature evaluates the ability to index metadata and create semantic relationships among documents.

▪ Agrupation. This Sub-feature evaluates whether the system allows grouping documents according to their content type in order to present them as a single content.

▪ Thesaurus Support. It appraises whether the system can be integrated with some kind of thesaurus.

▪ Digitalization. It controls that the system can scan documents without using third-party software component.

▪ Bulk Upload. It reviews whether the system provides functionality to upload documents massively within digitization or importation processes.

▪ Content Generation. It monitors that the system can generate new content (in different formats) from a document.

▪ Office Integration. It assesses whether the system allows its integration with office IT such as Microsoft Office, Google Docs and Open Office, among others.

▪ Forms and Templates. It tests whether the system can create forms and templates in order to store content homogeneously.

▪ Integration Forms Managers. It evaluates whether the system allows its integration with forms manager (e.g. formul@) in order to store content homogeneously.

▪ Advanced Search Methods. It measures whether the system provides agile methods to find documents, e.g., using search patterns or describing relevance criteria of results and metadata patterns, among others.

▪ Search Algorithms. It evaluates whether the system can configure the search algorithm, *i.e.*, places to look for, customizations of search syntaxes, accents tolerance and case sensitivity, among others.

▪ Display Formats. It guarantees that the system allows displaying contents of documents in different formats electronically.

c) Documental Life Cycle. This feature enables us to assess the level or degree of support the system offers to

the document cycle. Next, we describe the Sub-features of this feature:

- Check-in Check-out. It evaluates whether the system allows locking (using check-in and check-out operations) a document when a user is modifying it.
- Life cycle. It evaluates whether the system supports the document life cycle: creation, revision, classification, search, management, distribution, archiving and destruction.
- Versioning Support. It assesses whether the ECM system provides mechanism to control versions.
- Actions traceability. It tests whether the system allows auditing and monitoring actions on documents.
- Inconsistencies management. It evaluates whether the ECM system provides mechanism to control inconsistencies and conflicts between related documents.
- Dissemination management. It controls that the system can support the content dissemination based on user subscriptions.
- Conservation. It evaluates whether the ECM system allows defining retention periods on documents. This aspect ensures the integrity of electronic documents throughout their life cycle.
- Destruction. It lets users know whether the ECM system allows physical destruction of documentation. The destruction operation must be registered by the ECM system.
- Physical actions. It evaluates whether the ECM system allows carrying out physical operations on documents, for instance, to compose/decompose documents, to seal, to add notes or to append text, among others operations.

d) Workflows. This feature, can assess whether the tool supports management with business processes. Below, we describe the Sub-features of this feature:

- Supported standards. It evaluates the supporting degree of standards such as XML-RDF, Wf-XML and WARIA.
- Management support. It controls that the system can support the definition and maintenance of processes and procedures.
- Available options. It tests whether the system provides additional functionality to create workflows (for instance, alternative flows, sequential flows or collaborative flows).
- Motorization. It evaluates whether the ECM system allows process monitoring.
- Simulation and modeling. It lets us know whether the system provides simulation and modeling utilities.
- Graphical utilities. It points out whether the system provides graphical utilities to design workflows.
- Task management. It evaluates whether the system provides mechanism to manage tasks, deadlines and alerts.

e) E-Government. Measuring the support and features involving e-Government is essential, given the context in which this project is developed. This group of features evaluates the degree of support this context offers. Therefore, we can evaluate whether the system allows accessing and using valid electronic documents according to the guidelines of the Spanish National Interoperability Scheme. The Sub-features of this feature are described below:

- Electronic documents. It evaluates whether accessing to services, data and electronic documents and using them is admitted (according to the guidelines of the Spanish National Interoperability Scheme).
- Digital signature. It enables us to test whether the system allows signing online documents, as well as validating electronic signatures according to a signature policy shown in the electronic document signature.
- Accreditation and representation. It confirms whether the ECM system is able to use accreditation mechanisms and identify every citizen by recognizing his/her electronic signature.
- Indexes support. It lets us evaluate whether the system generates signed electronic indexes for each electronic record that have been processed.
- Unique identification. It confirms whether the system identifies each document exclusively.
- Minimum metadata. It assesses whether the system allows allocating mandatory minimum metadata to every electronic document according to the guidelines of the Spanish National Interoperability Scheme.
- Classification plan support. It evaluates whether the ECM system allows classifying electronic documents in accordance with the classification schema of the Spanish Public Administration.
- Official time synchronization. It lets us know whether the system allows synchronization with the official time.

f) Interoperability Compliance. A specific section dealing with interoperability has been included: Integration with tools. This sub-feature evaluates whether the ECM system provides mechanisms (e.g. APIs) to integrate

with third-party tools. The Sub-features of this feature are described as follows:

- Connection. It examines whether and to what degree APIs that enable integration with third-party systems are provided.
- ERPs. It confirms whether the system supports direct access to documents from ERP systems and allows creating transactions and exchanging data in both directions.
- Capture tools. It lets us know whether the system supports integration with capture tools.
- Email. It evaluates whether the ECM system supports integration with email solutions.
- CMIS repository. It controls that the ECM system can be integrated with any repository (for instance, web portals and virtual offices) that implements the CMIS standard.
- Web services. It confirms whether the system provides integration mechanisms based on Web services using XML (eXtensible Markup Language) and MetaData.
- Single window. It guarantees that the ECM system can be integrated with the main single window platforms (for instance, Solicit@ and Present@).
- Electronic record. It confirms whether the ECM system can be integrated with the main electronic record platforms (for instance, @riesand SIGEM).
- Records manager. It tests whether the ECM system can be integrated with the main records management platforms (for instance, Trew@).
- Electronic files. It evaluates whether the ECM system can be integrated with the main electronic file platforms and custody of documents (for instance, Archiv@).
- Digital signature. It assesses whether the ECM system can be integrated with the main digital signature platforms (for instance, @firma, viafirm@).
- EAI. It asserts whether the ECM system can be integrated with EAI tools (Enterprise Application Integration).
- Streaming. It analyzes whether the system can be natively integrated with streaming servers in order to see and/or listen to online audiovisual content.

g) Security and Control. One of the major objectives of document management solutions is to ensure information security, by controlling accessing the system from inside and outside the organization and managing the documents including such information either to archive or destruct them. As a result, these solutions must provide services that ensure that the information stored is secure. It evaluates whether the system is functional enough to analyze data, or otherwise, whether the system allows using third-party tools. Next, we describe the Sub-features of this feature:

- Data Analysis. It studies whether the data analysis system is allowed, or otherwise, third-party tools to perform data extraction system can be used.
- Exportation. It evaluates whether the system is functional enough to export data, or otherwise, it allows using third-party tools.
- Activity indicators. It confirms whether the system provides indicators to measure each user's activity in the system.
- Granularity. It assesses whether the system provides security utilities to control the access to a specific document or some parts of it.
- LOPD. It evaluates that the system allows switching between HTTPS and HTTP protocols according to the system module to which the user is accessing or encrypting contents and documents traceability, among others aspects.
- Logs. It controls whether the system allows audits (activity log) and issues reports on all actions taken.
- SSO. It reviews whether the system supports different kind of authentication: Single Sign ON, LDAP or Kerberos.
- Notifications. It specifies whether the system reports any security problem to the administrators.

h) Architecture. Open architecture. It evaluates whether the system has an open architecture (*i.e*., the system allows adding, upgrading and changing its components) or a closed architecture (*i.e*., the software manufacturer chooses the components, and the end user does not intend to upgrade them). Next, we describe the Sub-features of this feature:

- Open architecture. It evaluates whether the system has an open (to add, upgrade and change its components) or a proprietary architecture.
- Browsers. It assesses whether the system interacts successfully with popular Web browsers such as Internet

Explorer, Mozilla or Chrome, among others.

- Mobility. It lets us know whether the system provides a mobile interface for Smart Phone and Tablets.
- Development kit. It reviews whether the system offers a self-development kit (SDK, API or Web services).
- Cloud solution. It confirms that the system can provide a cloud solution.
- Administrative capabilities. It analyzes whether the system allows managing, monitoring and optimizing its architectural resources.
- Programming language. It studies the implementation degree of the programming language.
- Version. It revises the policy deployment and version management of the ECM solution, *i.e.*, update procedure and pre-upgrade tasks.
- Multiplatform. It evaluates whether the system supports multiple technological platforms (operating systems, Web servers, application servers and DBMS, among others).
- Extensibility. It studies whether the system has a modular platform and is easily extended using plugins, modules or extensions.
- Volumes. It tests whether the system successfully manages a large amount of information.
- High availability. It evaluates whether the system allows high availability configurations, *i.e.*, active-active or active-passive clusters and automatic recovery, among others.
- Scalability. It evaluates whether the ECM system can be automatically adapted (without losing quality) to increase workflows and system requests.

i) Cost. Cost (both initial and long-term by maintenance) is one of the most important factors any organization must take into account when choosing an ECM solution. The Sub-features of this feature are described below:

- Licenses. It evaluates the type of license and its cost.
- Infrastructure. It analyses the total cost (software and hardware) of the minimum necessary infrastructure.
- Open source. It assesses the existence of an active community development.
- Maintenance and support. It calculates the cost of multi-year support and maintenance of the solution.

j) Assistance and RM (Roadmap) Support. This last feature listed in the latter group includes aspects for the evaluation of the Characteristics support, assistance and roadmap provided by the ECM solution. Next, we describe the Sub-features of this feature:

- Certification program. It evaluates whether the ECM solution designer provides a recognized and complete certification program.
- User Manuals. It lets us know whether the manufacturer provides enough online manuals regarding the system operation.
- Service support. It confirms whether the designer offers a 24-hour support.
- Formation service. It revises whether the designer promotes classroom training.
- Roadmap. It assesses the roadmap of the system, *i.e.* a plan with short-term and long-term goals.

## 4.3. Applying the Characterization Schema

Once characterization schema is defined, it is time to instantiate every approach under study and evaluate them according to our criteria. The first analysis excludes four of our initial approaches: Docu Ware, Knowledge Tree, Nuxeo and MS SharePoint. They are discarded due to the fact that basic functional features are not supported by these approaches. Thus, in our view, these approaches are not considered suitable ECM tools. However, Nuxeo is an exception. ECM alone does not cover the basic functional features, therefore it is supported by a commercial solution, named Athento [32] that is a proprietary solution that enriches Nuxeo. In terms of these considerations, we will instantiate our characterization schema only in Alfresco, Athento + Nuxeo, ECM Documentum, IBM Filenet and Opentext.

Besides, it is necessary to consider three important aspects before presenting the final instantiation:

1) Multiple evaluation: The numerical value of each feature was calculated according to three evaluations:

a) We asked each company that works with these approaches for its own evaluation of each approach.

b) The group of researchers of this paper makes our own evaluation, based on our expertise. As we worked in different projects related to ECM, we installed each tool to test it.

c) We selected a group of 3 students pursuing a PhD in computer sciences without any expertise in ECM tools

and asked them to assess each feature individually.

The results shown were merged and discussed, and the points on which the three groups disagreed were studied in detail. Thus, although our evaluation was subjective, it mixed a set of tests from different levels.

2) Multiobjective optimization: Although it is out of the scope of this paper, we have to mention multiobjective optimization [33]. After revising the results obtained and shown below, the problem that has arisen in this comparative study can be classified as a multiobjective optimization problem. The comparative study that has been executed comprises a large number of desirable characteristics for the tools to develop (this is detailed in the classification schema) and the relative weight of each of these multiple features depends largely on the particular work environment.

Therefore, and attending to the principles established in multiobjective optimization problems [34], this section will cope with the instantiation of the classification chart according to each characteristic specific values, that means that each studied characteristic is posed as a simple optimization problem. That is, each feature is independently studied without taking into account other characteristics. Thus, tables presented in Section 5 do not take into account the multiobjective approach.

3) Preference composition: The preference theory is framed in the theory of specific methods of assessing preferences to quantify changes in people's welfare [35], caused by a change in the quantity or quality of a good or service, which, in turn, is grounded in neoclassical welfare theory methods.

If we considered the comparative proposal under this theory, it would be necessary to assess the preferences of every feature marked in each of the three groups of evaluation given (companies, researcher and students). Thus, the assessment issues an appraisal which varies from 1 to 4, where 4 represents the highest score and 1 the worst solution. It is reminded that the allowable values shown in the classification could range from 1.99, therefore at the end, it was decided to reduce this range to facilitate results understanding. This will allow us to have an order of preference for each of the characteristics of the classification. With regard to subjective criteria, proposals will be sorted from lowest to highest in terms of group provisions, leaving the possibility of being ranked only in the case of factors that can be objectively measured.

After these remarks, we present the results of our study.

**Table 4** shows the first set of concepts: Functional modules. As the coverage of these functional characteristics is an obligatory constraint for our approaches, all of them are well valued in this group of items. However, we have to stick out Athento/Nuxeo solution, which currently does not support the record management module completely. Designers promise to include it in the new version of the solution.

**Table 5** values user orientation characteristics. This section highlights Alfresco side for its native ability to preview a lot of formats. It also includes capabilities for streaming video and audio (the display is supported on a Flash Document previewer). An additional advantage is how easily Alfresco customize user interface using

**Table 4.** Characterization schema instantiation: Functional modules.

| Group | Features | Alfresco | Athento / Nuxeo | EMC Documentum | IBM Filenet | OpenText |
|---|---|---|---|---|---|---|
| Functional Modules | Document Management | ✓ | ✓ | ✓ | ✓ | ✓ |
| | Records Management | ✓ | ⊘ | ✓ | ✓ | ✓ |
| | BPM (Workflow/Business Process Management) | ✓ | ✓ | ✓ | ✓ | ✓ |
| | Document Imaging/Image-processing | ✓ | ✓ | ✓ | ✓ | ✓ |
| | Interoperability functions/extended components | ✓ | ✓ | ✓ | ✓ | ✓ |
| | WCM (Web Content Management) | ✓ | ✓ | ✓ | ✓ | ✓ |
| | Social content/collaboration | ✓ | ✓ | ✓ | ✓ | ✓ |
| | Features | ✓ | ✓ | ✓ | ✓ | ✓ |

**Table 5.** Characterization schema instantiation: User orientation.

| Group | Features | Alfresco | Athento/ Nuxeo | EMD Documentum | IBM Filenet | Open Text |
|---|---|---|---|---|---|---|
| User orientation | Usability compliance | 1 | 1 | 1 | 1 | 1 |
| | Accessibility | 1 | 1 | 1 | 1 | 1 |
| | Document preview | 4 | 1 | 2 | 2 | 2 |
| | Drag & Drop | 1 | 1 | 1 | 1 | 1 |
| | Bulk uploads | 1 | 2 | 1 | 1 | 1 |
| | Undo | 1 | 1 | 1 | 1 | 1 |
| | WYSIWYG Editor | 2 | 1 | 4 | 4 | 3 |
| | Customization | 3 | 2 | 2 | 2 | 1 |
| | Groups and Social networks | 3 | 1 | 2 | 2 | 2 |
| | Multilanguage | 2 | 1 | 4 | 3 | 3 |
| | TOTAL | 19 | 13 | 19 | 18 | 16 |

Spring Surf and languages, like Javascript, as well as SDK to design components dashlets.

In contrast, EMC Documentum platform manages users and geographically distributed locations contents adapted to every language, culture and currency, as well as stores multilingual content in shared repositories.

This platform can provide a single virtual repository that encompasses multiple geographic locations under the shape of either a single distributed repository or federations of repositories, which are groups of repositories that work together. The virtual repository allows users to access content regardless of language or geography. Multilingual presentation management is the unique feature that binds versions of the same content in different languages so that users can choose the language they will use for communication.

In addition, EMC Documentum provides a platform for developing and deploying advanced business solutions and case management, cloud-based optimized, including EMC on Demand deployments. This solution, called xCP elaborates applications working on four different levels: data model, processes/services, user interface and reporting (BAM).

Tabla 6 values functionality to capture, access, retrieve and view documents. IBM Filenet, EMC Documentum and Opentext provide good answers to this set of features. On the one hand, IBM Filenet is notable for the addition of two search engines, Content Search Services, based on Lucene, and Content Search Engine, based on Autonomy K2.On the other hand, EMC Documentum platform incorporates FAST index server, an enterprise search technology leading the industry. The search capability is modular and entails alternative engines for Documentum specific market offerings. For example, the publication of the original manufacture of Documentum is developed for software suppliers incorporate it into their products. Documentum platform offers open-source Lucene alternative as the default engine. However, the FAST search engine for all the standard offerings to business customers is built into the repository.

Table 7 shows the instances for the treatment of documental life cycle in each tool. There are no tangible differences between EMC Documentum, IBM FiletNet and Opentext solutions in this set of features, highlighting very slightly that both, EMC Documentum and IBM FileNet incorporate tools that facilitate screening, traceability and possible inconsistencies and conflicts solutions, as well as implement automatic content rules.

Table 8 instances aspects related with workflows. This section points out EMC Documentum and IBM FiletNet to provide extended functionality in the workflow engine. However, IBM Filen Net is vaguely better positioned to deliver a powerful console for process simulation.

In Table 9, aspects related with e-Government are presented. Here, the case of Nuxeo/Athento solution is different since it is a management solution in charge of adding more electronic records to document management utilities.

**Table 6.** Characterization schema instantiation: Functionality to capture, access, retrieve and view documents.

| Group | Features | Alfresco | Athento/ Nuxeo | EMC Documentum | IBM Filenet | Open Text |
|---|---|---|---|---|---|---|
| Functionality to capture, access, retrieve and view documents | Degree of Cataloging | 2 | 1 | 3 | 3 | 2 |
| | Agrupation | 2 | 1 | 3 | 3 | 3 |
| | Thesaurus Support | 1 | 1 | 2 | 2 | 2 |
| | Digitalization | 1 | 3 | 1 | 1 | 1 |
| | Bulk Upload | 1 | 2 | 1 | 1 | 1 |
| | Content Generation | 2 | 1 | 3 | 3 | 3 |
| | Office Integration | 2 | 1 | 3 | 3 | 3 |
| | Forms and Templates | 2 | 1 | 3 | 4 | 3 |
| | Integration Forms Managers | 1 | 1 | 2 | 2 | 1 |
| | Advanced Search Methods | 2 | 1 | 4 | 3 | 4 |
| | Search Algorithms | 2 | 1 | 4 | 4 | 3 |
| | Display Formats | 3 | 1 | 2 | 2 | 2 |
| | TOTAL | 21 | 15 | 31 | 31 | 28 |

**Table 7.** Characterization schema instantiation: Documental life cycle.

| Group | Features | Alfresco | Athento/ Nuxeo | EMC Documentum | IBM Filenet | Open Text |
|---|---|---|---|---|---|---|
| Documental Life Cycle | Check-in/check-out | 1 | 1 | 1 | 1 | 1 |
| | Life cycle | 2 | 1 | 3 | 3 | 3 |
| | Versioning Support | 2 | 1 | 3 | 3 | 3 |
| | Actions traceability | 2 | 1 | 4 | 4 | 3 |
| | Inconsistencies management | 2 | 1 | 4 | 4 | 3 |
| | Dissemination management | 2 | 1 | 3 | 3 | 4 |
| | Conservation | 2 | 1 | 3 | 3 | 3 |
| | Destruction | 2 | 1 | 3 | 3 | 3 |
| | Physical actions | 1 | 1 | 1 | 1 | 1 |
| | TOTAL | 16 | 9 | 25 | 25 | 24 |

Aspects related with interoperability are presented in **Table 10**. These features are well supported in general by the different ECM solutions. However, Alfresco Enterprise vaguely stands because of its large implementation in Public Administration. This has led to integration works developed in liaison with the staff responsible for the electronic issue of e-Government in Andalusia and SIGEM [36] in Extremadura.

**Table 11** includes the three solutions that longer have stayed in the market: Dcoumentum EMC, IBM File Net and Open Text. Nevertheless, IBM Filenet receives a higher score because it has proper and specific tools to analyze data and identify activity indicators: File Net Business Activity Monitor.

In characteristics related with architecture refer in **Table 12**, we would like to highlight once again the three solutions that longer have stayed in the market. There is a relevant difference especially with Athento/Nuxeo,

**Table 8.** Characterization schema instantiation: Workflows.

| Group | Features | Alfresco | Athento/Nuxeo | EMC Documentum | IBM Filenet | Open Text |
|---|---|---|---|---|---|---|
| Workflows | Supported standards | 4 | 1 | 3 | 3 | 2 |
| | Management support | 2 | 1 | 3 | 3 | 3 |
| | Available options | 2 | 1 | 4 | 4 | 3 |
| | Motorization | 2 | 1 | 4 | 4 | 3 |
| | Simulation and modeling | 2 | 1 | 3 | 4 | 3 |
| | Graphical utilities | 2 | 1 | 3 | 3 | 3 |
| | Task management | 2 | 1 | 4 | 4 | 3 |
| | TOTAL | 15 | 7 | 24 | 25 | 20 |

**Table 9.** Characterization schema instantiation: e-Government.

| Group | Features | Alfresco | Athento/Nuxeo | EMC Documentum | IBM Filenet | Open Text |
|---|---|---|---|---|---|---|
| e-Goverment | Electronic documents | 1 | 3 | 1 | 1 | 1 |
| | Digital signature | 1 | 2 | 1 | 1 | 1 |
| | Accreditation and representation | 1 | 2 | 1 | 1 | 1 |
| | Indexes support | 1 | 1 | 1 | 1 | 1 |
| | Unique identification | 1 | 1 | 1 | 1 | 1 |
| | Minimum metadata | 1 | 3 | 1 | 1 | 1 |
| | Classification plan support | 1 | 2 | 1 | 1 | 1 |
| | Official time synchronization | 1 | 1 | 1 | 1 | 1 |
| | TOTAL | 8 | 15 | 8 | 8 | 8 |

**Table 10.** Characterization schema instantiation: Interoperability compliance.

| Group | Features | Alfresco | Athento/Nuxeo | EMC Documentum | IBM Filenet | Open Text |
|---|---|---|---|---|---|---|
| Interoperability Compliance | Connection | 2 | 2 | 1 | 1 | 1 |
| | ERPs | 2 | 1 | 3 | 3 | 4 |
| | Capture tools | 1 | 3 | 1 | 1 | 1 |
| | Email | 1 | 1 | 3 | 3 | 4 |
| | CMIS repository | 4 | 3 | 2 | 2 | 2 |
| | Web services | 1 | 1 | 1 | 1 | 1 |
| | Single window | 2 | 1 | 1 | 1 | 1 |
| | Electronic record | 1 | 1 | 1 | 1 | 1 |
| | Records manager | 2 | 1 | 2 | 1 | 1 |
| | Electronic files | 2 | 1 | 1 | 1 | 1 |
| | Digital signature | 1 | 2 | 1 | 1 | 1 |
| | EAI | 1 | 1 | 1 | 1 | 1 |
| | Streaming | 2 | 1 | 1 | 1 | 1 |
| | TOTAL | 22 | 19 | 19 | 18 | 20 |

**Table 11.** Characterization schema instantiation: security and control.

| Group | Features | Alfresco | Athento/ Nuxeo | EMC Documentm | IBM Filenet | Open Text |
|---|---|---|---|---|---|---|
| Security and Control | Data Analysis | 2 | 1 | 3 | 4 | 3 |
| | Exportation | 1 | 1 | 2 | 2 | 2 |
| | Activity indicators | 2 | 1 | 3 | 4 | 3 |
| | Granularity | 2 | 1 | 3 | 3 | 3 |
| | LOPD | 1 | 1 | 2 | 2 | 2 |
| | Logs | 2 | 1 | 3 | 3 | 3 |
| | SSO | 2 | 1 | 3 | 3 | 3 |
| | Notifications | 1 | 1 | 2 | 3 | 2 |
| | TOTAL | 13 | 8 | 21 | 23 | 21 |

**Table 12.** Characterization schema instantiation: Architecture.

| Group | Features | Alfresco | Athento/ Nuxeo | EMC Documenum | IBM Filenet | Open Text |
|---|---|---|---|---|---|---|
| Architecture | Open architecture | 1 | 1 | 1 | 1 | 1 |
| | Browsers | 2 | 1 | 3 | 3 | 3 |
| | Mobility | 2 | 1 | 2 | 2 | 2 |
| | Development kit | 2 | 1 | 4 | 4 | 3 |
| | Cloud solution | 3 | 1 | 3 | 2 | 2 |
| | Administrative capabilities | 1 | 1 | 2 | 2 | 2 |
| | Programming language | 2 | 1 | 2 | 2 | 2 |
| | Version | 2 | 1 | 3 | 3 | 3 |
| | Multiplatform | 3 | 2 | 3 | 3 | 3 |
| | Extensibility | 2 | 1 | 3 | 3 | 3 |
| | Volumes | 2 | 1 | 3 | 3 | 3 |
| | High availability | 2 | 1 | 4 | 4 | 4 |
| | Scalability | 3 | 1 | 4 | 4 | 4 |
| | TOTAL | 27 | 14 | 37 | 36 | 35 |

because this solution neither focuses on providing cloud service, nor a high-availability good solution for large volumes of information.

**Table 13** analyses aspects related with cost. Athento/Nuxeo is very well valued in this comparative solution as the cost of licensing and support is low in comparison to the other standard solutions. It is followed by Alfresco Enterprise thanks to the large community of developers who currently covers.

In the area of software licensing, it should be indicated that suppliers' offer for non-free tools concerns named users (or named users bags, such as IBM Filenet). This would imply a high cost for large deployments, if compared to Alfresco Enterprise or Athento/Nuxeocosts. Besides, IBM FileNet licensing costs include DB2 and WebSphere Application Server licenses.

Finally, in **Table 14**, assistance and RM support are instantiated. EMC and IBM FiletNet Documentum area little better rated in this group of features since they have an extensive program of certification, a lengthy road-map, as well as a broad presence in the market, apart from providing a wide range of training and online support options.

**Table 13.** Characterization schema instantiation: Cost.

| Group | Features | Alfresco | Athento/ Nuxeo | EMC Documenutm | IBM Filenet | Open Text |
|-------|----------|----------|----------------|----------------|-------------|-----------|
| Cost | Licenses | 2 | 4 | 1 | 1 | 1 |
| | Infrastructure | 3 | 3 | 1 | 2 | 1 |
| | Open source | 4 | 2 | 1 | 1 | 1 |
| | Maintenance and support | 1 | 3 | 2 | 2 | 2 |
| | TOTAL | 10 | 12 | 5 | 6 | 5 |

**Table 14.** Characterization schema instantiation: Assistance and RM support.

| Group | Features | Alfresco | Athento/ Nuxeo | EMC Documentm | IBM Filenet | Open Text |
|-------|----------|----------|----------------|---------------|-------------|-----------|
| Assistance and RM Support | Certification program | 3 | 1 | 4 | 4 | 3 |
| | User Manuals | 2 | 1 | 3 | 3 | 3 |
| | Service support | 3 | 1 | 3 | 3 | 3 |
| | Formation service | 2 | 1 | 4 | 4 | 3 |
| | Roadmap | 2 | 1 | 3 | 3 | 3 |
| | TOTAL | 12 | 5 | 17 | 17 | 15 |

## 4.4. Conclusions from the General Evaluation

In order to obtain conclusions to develop the Document Management System from these ECM tools within the THOT Project, it is very important to take into account not only the value that all these ECM tools offer to us but also other important factors when these tools are used like the costs, risks or incertitude of these tools.

As regards the value, the most valuable ECM tool depends on the project scope. So, each ECM tool has its strength and weakness points and the relevance of each tool feature depends on the project scope. So, just with the context we could say which ECM tool is the most suitable one to be used in the development of the final Document Management System.

As regards costs, there are several ECM tools in the market, both free software and proprietary license, which offer appropriate solutions in this context. We must consider that we just going to pay a fee for the licenses of these ECM tools if it is sure that we are going to receive the value we need for the THOT Project but with minimal costs, risks and incertitude in the development of the final solution and its deployment in the Andalusian agency. Thus, we try to make the Andalusian agency more competitive using the Document Management system.

But not only is important the costs of a license but the risks and incertitude that organizations assume with the implementation of this Document Management system. For instance, these can be measured by different aspects like an existing community of developers about the tool, the quantity of existing documentation, existing forums and tutorials about the tool among other things.

So, we have evaluated these tools by solving a multiobjective optimization problem. Multiobjective optimization is an area of multiple criteria decision making, which is concerned with mathematical optimization problems involving more than one objective function to be optimized simultaneously. For a nontrivial multiobjective optimization problem, there does not exist a single solution that simultaneously optimizes each objective. The multiobjective study has taken into account the cost of 5 years, risks and competitiveness based on previous values to state the following conclusions:

- According to cost

Establishing a 5 year period as the framework for this comparison, Alfresco Enterprise solution has involved

higher costs than other solutions evaluated, although the discharge, as in the case of Athento/Nuxeo package does not increase significantly in the functional aspect as it depends on the named user number.

Furthermore, ECM Documentum, IBM Filenet and Opentext tools are aligned in terms of cost/value where IBM File Net and Documentum ECM are offering greater functional scope than Nuxeo/Athento package which is on the other side of the scale.

Regarding cost, Nuxeo/Athentois located just below ECM Documentum, IBM Filenet or Opentext solution in the framework of this comparison, therefore this multiobjective relationship Nuxeo/Athento should be discarded for it offers a lower functional scope. **Figure 3** graphically represents the results:

- According to risk

IBM File Net, EMC Documentum, Open Text and Alfresco Enterprise are distinguished by their position in the market (see Gartner Inc. Google and outcome) and therefore they assume less risk and uncertainty. Regarding value, IBM File Net is best positioned at a reasonable cost compared to other solutions. **Figure 4** shows the relationship between Risk and Value whereas **Figure 5** represents the relationship between Risk and Cost.

- Finally, according to competitive (**Figure 6**).

IBM File Net has a slightly better position than EMC Documentum and Open Text. Despite the three solutions are highly competitive in the context of this comparison, they are far from other solutions evaluated.

## 5. Customizing Our Study in a Real Project

This section presents a concrete use of the study, after analysing our work in general. We introduce a view of a real project named THOT to cover the three objectives of our work. We ponder on each characteristic for the concrete necessity of the project.

### 5.1. A Global View of Thot Project

Nowadays, the Andalusian Public Administration has already implemented or is currently implementing different initiatives, which are driving the need for a deep change in document management systems:
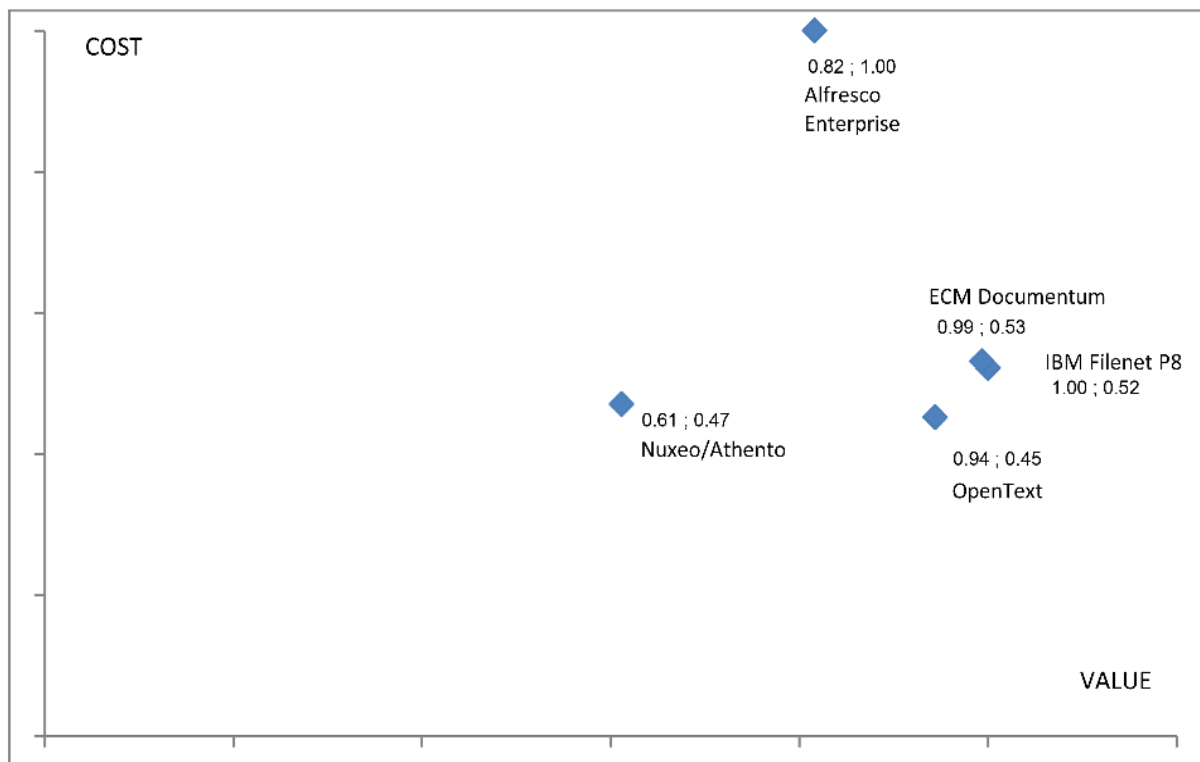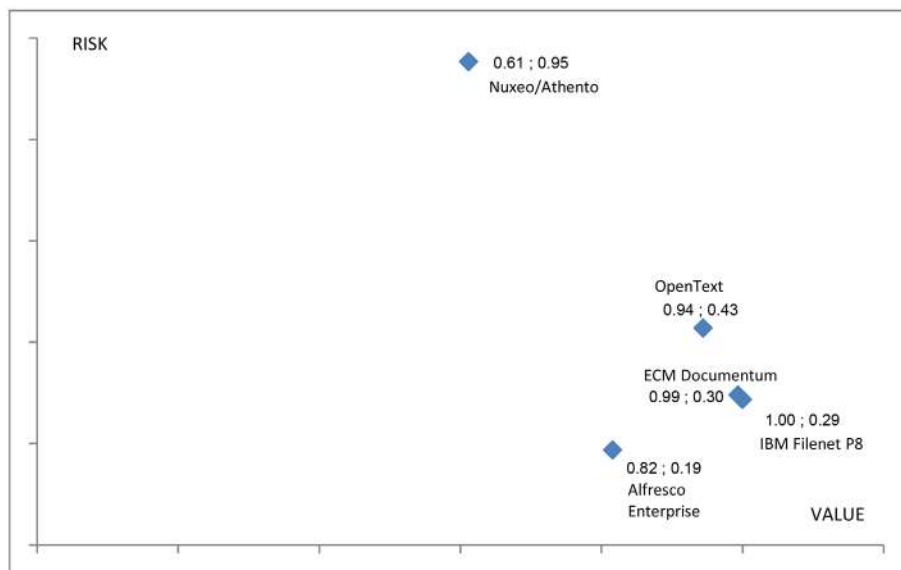


**Figure 3.** Diagram according to cost.

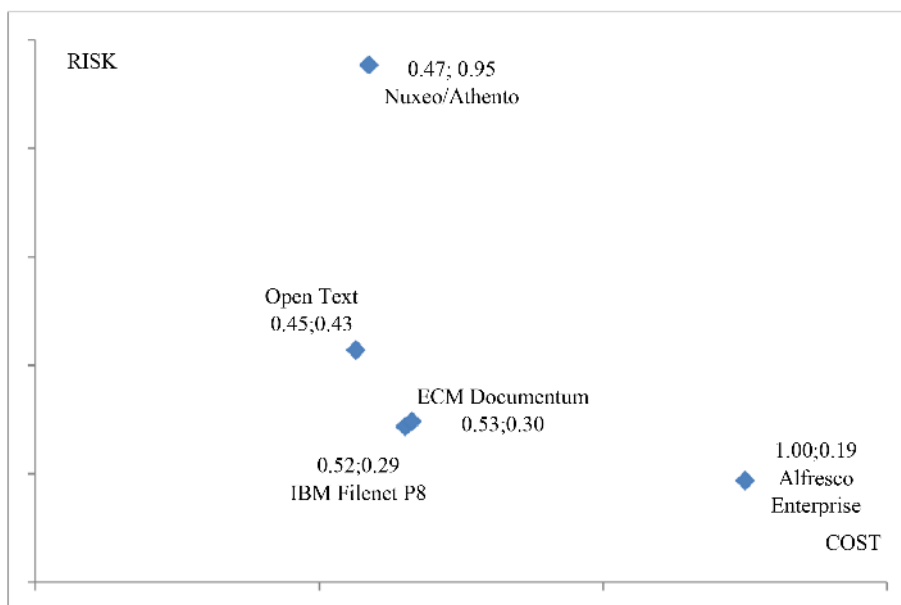**Figure 4.** Risk vs Value.



**Figure 5.** Risk vs Cost.

- The financial management of records is now being supported by JUPITER [37] Information System, although this situation will change because the Ministry of Finance and Public Administration of the Regional Government of Andalusia has launched an initiative to migrate the current JUPITER System to an ERP technology platform. In fact, there are already some institutions in Andalusia where this technology has been implemented, such as the Public Agency of Contracting Services for Transport and Infrastructure Constructions.
- In July 2011, the Andalusian Parliament approved a new law concerning establishing ERIS-G3 [38] system, which is an e-Government system to manage electronic procedures to process public records in each public agency.
- @rchivA [39] Information System that facilitates proper management of the Patrimony Documentation of Andalusia is providing researchers and the general public with the necessary tools to consult and distribute such Patrimony through new technologies.
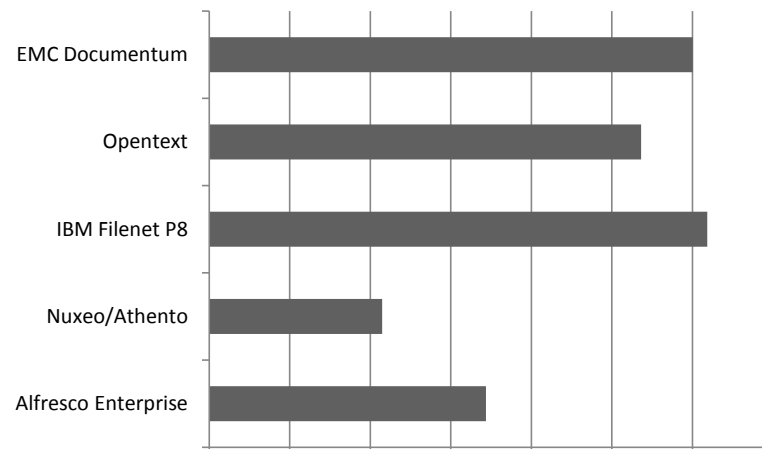
**Figure 6.** Competitive.

Therefore, in the near future we will witness a general scenario in the Regional Government of Andalusia Administration compiling:

- A clear objective to improve the efficiency and effectiveness of strategic, tactical and financial resources,
- A high level of automation of management and business processes, and
- A technology that integrates data and processes and also enables integrating document management.

This increases the opportunity (and also the need) to propose innovations in document management, in this case related to public infrastructure and services. Then, we propose the implementation of a document management model to be applied to this context. THOT project is an e-Government project aiming at implementing an ECM system in the Public Administration of the Regional Government of Andalusia (Spain). The project has been granted for a total of 621,250 Euros. It tends to define an environment for managing agile documentary records, framed within the existing real world, to be implemented by appropriate technological solution aligned with guidelines and the current context. The specific objectives of this research project are:

- Objective 1: Create a document management solutions model integrated with ERP (Enterprise Resource Planning).
- Objective 2: Optimize resource investment in Andalusia for the development of the new ERP technology platform, providing new adjacent functionality as document management model.
- Objective 3: Develop such a model applied to records of procurement of services and transport infrastructure projects of the Public Works Agency of Andalusia, responding to specific casuistry project technical documents and favoring the existing need to exploit the spatial reference information (such as support for planning and sectoral processes).
- Objective 4: Apply European recommendations and e-Government principles to promote the format conversion, from printed papers to online versions, in written communications between management and contractors as well as facilitate compliance with current regulations related to e-Government.
- Objective 5: Achieve a large number of digital documents instead of paper in order to be able to save and keep them.
- Objective 6: Standardize and structure information in order to encourage the development and applicability of information technology and communication in the public works sector.
- Objective 7: Apply standards in the document management system that allows a secure interoperability with other systems.

This project aims to make a qualitative leap covering different disciplines of research and innovation, such as document management policies, e-Government policies, dissemination and integration policies in Web environments (aspects that are treated and profusely investigated by different groups and research fora, both national and international). That enables organizations to provide a common framework for document management and cover the need to have a comprehensive management to complement business processes from beginning to end. The project offers, with innovation and research jobs, a solution that not only permits organizations to manage documents intelligently, but also to distribute, maintain and custody them.

## 5.2. Customization of Our Study

We held a set of meetings with functional users in order to adjust our study to the concrete case of THOT Project. At theses meeting, we studied each group of features depending on its relevance in THOT context. Results are presented in **Table 15**.

## 5.3. Conclusions from THOT Project

As can be realized along this paper, non-free comparative tools have clear advantages in opposition to other tools. Particularly, IBM Filenet has obtained best weighted overall result just for a minimum difference. This result can be justified from the fact that both IBM Filenet and EMC Documentum are solutions with extensive experience and worldwide deep implementation.

They have a better performance in more general characteristics of ECM tools such as ability to capture, access, retrieve and display document management life cycle or workflow, security, access control and activity or support issues, manual and roadmap.

However, as expected, the lowest score is framed in specific aspects of the context of this study such as e-Government, and aspects related to cost, such as license regarding number of named users.

Finally, we would like to remark the case of Athento/Nuxeo solution because, as previously commented, Nuxeo solution does not achieve an initial value in the functional coverage that justifies its introduction in this study (it has been incorporated due to the studied experiences aroused the interest of improving Athento solution). We have also observed that the resulting package is quite far from the highest values because this study is focused exclusively on document management tools, whereas Nuxeo solution is centered on aspects more related to administration files management.

## 6. Conclusions and Future Work

This paper presents a comparative study of five different ECM tools. The work uses an enriched systematic literature review in order to decide which tools are included. After carrying out this review, this paper presents a characterization schema that allows comparing each approach under study in a similar way.

The schema is applied to each approach and a set of global results are obtained and commented. Moreover,

**Table 15.** Characterization schema instantiation for THOT project.

| Group of items | Weigth | Alfresco | Athento/ Nuxeo | EMC Documenum | IBM Filenet | Open Text |
|---|---|---|---|---|---|---|
| 1. User Orientation | 0.5 | 19 | 12 | 19 | 18 | 16 |
| 2. Functionality to capture, access, retrieve and view documents | 1.0 | 21 | 15 | 31 | 31 | 28 |
| 3. Documental Life Cycle | 1.0 | 16 | 7 | 24 | 25 | 20 |
| 4. Workflows | 0.2 | 16 | 7 | 24 | 25 | 20 |
| 5. e-Government | 1.0 | 8 | 15 | 8 | 8 | 8 |
| 6. Interoperability Compliance | 1.1 | 22 | 19 | 19 | 18 | 20 |
| 7. Security and Control | 1.1 | 13 | 8 | 21 | 23 | 21 |
| 8. Architecture | 1.3 | 27 | 14 | 37 | 36 | 35 |
| 9. Cost | 1.8 | 10 | 12 | 5 | 6 | 5 |
| 10. Assistance and RM | 1.0 | 12 | 5 | 17 | 17 | 15 |
| **RESULT…** | | **164** | **116** | **206** | **207** | **192** |
| **Pondered Result…** | | **16.13** | **12.09** | **19.64** | **19.77** | **18.66** |

the paper illustrates how this study can be customized to a concrete environment with its real application to THOT Project. Thus, the three aims that are cited on the introduction are achieved.

As a final conclusion, we would like to add that it is a very active environment and there are continuous changes in tools and versions. This study was finished in February 2014, taking into consideration the current situation of the analyzed tools at this moment.

For this reason, in terms of future work, we would conclude that the evaluated tools presented a constant process of change and evolution. In fact, the lines that seem most promising and empower these tools can be clearly aligned with strategic Action Digital Economy and Society, specifically with the following subject priorities:

- Future Internet (networks, services, things or people).
- Cloud computing: development, innovation and selection of technologies and solutions.
- Mobility: technology-based services and mobility, networking and mobile systems products.
- Safety in the use of applications, especially in e-ID.
- Cyber-security and digital trust, protecting particularly vulnerable groups.
- Open/Linked/Big Data: re-use of public sector information and knowledge to create value.
- Social Media for their business-generating potential and provided services.
- Digital content: systems, platforms, services and processes that facilitate design, production and packaging.
- Systems, platforms, services and processes for new solutions and audio-visual broadcasting.

## Acknowledgements

## References

[1] AIIM (2013) The Global Community of Information Professionals. http://www.aiim.org/

[2] Alalwan, J.A. and Weistroffer, H.R. (2012) Enterprise Content Management Research: A Comprehensive Review. *Journal of Enterprise Information Management*, **25**, 441-461. http://dx.doi.org/10.1108/17410391211265133

[3] THOT Project. http://www.aopandalucia.es/inetfiles/agencia_innovacion_estructura/14122012112320.pdf

[4] Scott, J.E. (2011) User Perceptions of an Enterprise Content Management System. *Proceedings of the* 44*th Hawaii International Conference on Systems Science* (*HICSS*-44 2011), Koloa, 4-7 January 2011, 1, 9, 4-7.

[5] Haug, A. (2012) The Implementation of Enterprise Content Management Systems in SMEs. *Journal of Enterprise Information Management*, **25**, 349-372. http://dx.doi.org/10.1108/17410391211245838

[6] Van Rooij, J.C. (2013) Legacy Issues in the Implementation of Enterprise Content Management (ECM). *International Journal of Information*, **3**, 120-123.

[7] ISO 2709:2008, Information and Documentation—Format for Information Exchange. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=41319

[8] ISO 15836:2009, Information and Documentation—The Dublin Core Metadata Element Set. http://www.iso.org/iso/catalogue_detail.htm?csnumber=52142

[9] ISO 15489-1:2001, Information and Documentation—Records Management—Part 1: General. http://www.iso.org/iso/catalogue_detail?csnumber=31908

[10] ISO 23950:1998, Information and Documentation—Information Retrieval (Z39.50)—Application Service Definition and Protocol Specification. http://www.iso.org/iso/catalogue_detail.htm?csnumber=27446

[11] ISO 10244:2010, Document Management—Business Process Baselining and Analysis. http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=45935

[12] Zhang, H. and Babar, M.A. (2013) Systematic Reviews in Software Engineering: An Empirical Investigation. *Information and Software Technology*, **55**, 1341-1354.

[13] Kitchenham, B. (2004) Procedures for Performing Systematic Reviews. Keele University, Keele, 33.

[14] Brereton, P., Kitchenham, B.A., Budgen, D., Turner, M. and Khalil, M. (2007) Lessons from Applying the Systematic Literature Review Process within the Software Engineering Domain. *Journal of Systems and Software*, **80**, 571-583. http://dx.doi.org/10.1016/j.jss.2006.07.009

[15] Kitchenham, B. and Brereton, P. (2007) Introduction to Special Section on Evaluation and Assessment in Software Engineering EASE06. *Journal of Systems and Software*, **80**, 1423-1424. http://dx.doi.org/10.1016/j.jss.2006.10.031

[16] Staples, M. and Niazi, M. (2007) Experiences Using Systematic Review Guidelines. *Journal of Systems and Software*, **80**, 1425-1437. http://dx.doi.org/10.1016/j.jss.2006.09.046

[17] Kitchenham, B. and Charters, S. (2007) Guidelines for Performing Systematic Literature Reviews in Software Engineering, Technical Report EBSE 2007-001, Keele University and Durham University Joint Report.

[18] Kitchenham, B. and Brereton, P. (2013) A Systematic Review of Systematic Review Process Research in Software Engineering. *Information and Software Technology*, **55**, 2049-2075.

[19] Zhang, H., Babar, M.A. and Tell, P. (2011) Identifying Relevant Studies in Software Engineering. *Information and Software Technology*, **53**, 625-637.

[20] SEG (Software Enginnering Group) (2007) Guidelines for Performing Systematic Literature Reviews in Software Engineering Version 2.3. EBSE Technical Report. EBSE-2007-01. School of Computer Science and Mathematics. Keel University and Department of Computer Science. University of Durham. United Kingdom.

[21] Wohlin, C. and Prikladnicki, R. (2013) Systematic Literature Reviews in Software Engineering. *Information and Software Technology*, **55**, 919-920.

[22] Gilbert, M.R., Shegda, K.M., Chin, K., Tay, G. and Koehler-Kruener, H. (2012) Magic Quadrant for Enterprise Content Management. Article ID: G00237781.

[23] Gartner Company Is Based in Stamford, Connecticut, USA. http://www.gartner.com

[24] Alfresco. http://www.alfresco.com

[25] DocuWare. www.docuware.com

[26] Documentum. http://www.emc.com/domains/documentum/index.htm

[27] IBM FileNet. http://www-03.ibm.com/software/products/us/en/filecontmana/

[28] KnowledgeTree. www.knowledgetree.com

[29] Nuxeo/Athento. http://www.athento.com/nuxeo/

[30] OpenText. http://www.opentext.es/

[31] MS Sharepoint. http://office.microsoft.com/es-es/sharepoint-server-help/introduccion-a-sharepoint-HA102772778.aspx?CTT=5&origin=HA010378184

[32] Athento. http://www.athento.com/

[33] Afshin-Mansouri, S., Gallear, D. and Askariazad, M.H. (2012) Decision Support for Build-to-Order Supply Chain Management through Multiobjective Optimization. *International Journal of Production Economics*, **135**, 24-36. http://dx.doi.org/10.1016/j.ijpe.2010.11.016

[34] Renaud, J.E. (2012) Interactive Multiobjective Optimization Procedure.

[35] Espinosa, O. (2012) Método de Valoración de Preferencias' Análisis Conjunto': Una Revisión de Literatura (Method of Preference Valuation'Conjoint Analysis': A Review of the Literature). Econografos Escuela de Economía. http://dx.doi.org/10.2139/ssrn.2194963

[36] SIGEM. https://www.comparaiso.es/plan-avanza

[37] JUPITER Project, Junta de Andalucía. http://www.juntadeandalucia.es/repositorio/usuario/listado/fichacompleta.jsf?idProyecto=19

[38] ERIS-G3 Project, Junta de Andalucía. http://www.juntadeandalucia.es/repositorio/usuario/listado/fichacompleta.jsf?idProyecto=680

[39] Archiva (@rchiva), Junta de Andalucía. https://ws024.juntadeandalucia.es/ae/adminelec/areatecnica/archiva