

Politecnico di Milano
ING V - Scuola di Ingegneria dell'Informazione
Corso di Laurea Specialistica in Ingegneria Informatica



Enterprise Content Management Systems for Multinational Engineering Companies

Study Case: ABB - S.p.A - PA Division

Tesina di laurea di:

Rodrigo TERNI

Matricola:

748843

Relatore:

Fabio SCHREIBER

Correlatore:

Maurizio NOVELLI

Anno Accademico 2010/2011

Abstract

The thesis is based on the study of the importance and practical problems related to enterprise content management (ECM) for multinational companies. Since ECM has become a key for companies to succeed today, this thesis is meant as a guide to point out which are the challenges related to a successful ECM system, both in infrastructure as well as in the application itself. Both of these areas are broken down into smaller (yet significant) problems within a real ECM system.

A study case of ABB is provided, with actual specifications of ABB's infrastructure, to exemplify solutions for the aforementioned problems. Since there are many different problems that must be analyzed, we provide ideas for partial solutions that could solve each single problem separately, and also an example of suppliers that are capable of providing complete solutions to solve all problems with a single application. It is intended to serve as a model for multinationals that need to design a global process of content control, and also to stress the importance of managing and controlling the flow of information within a multinational company.

Contents

I	General design	4
1	Introduction	4
2	What is Enterprise Content Management?	5
3	Market Analysis	7
3.1	Overview	7
3.2	Gartner's Magic Quadrant	8
4	Problem statement	11
4.1	Infrastructural challenges	11
4.1.1	Storage	11
4.1.2	Data distribution on server farm	12
4.1.3	Bi-directional server farm replication	13
4.2	Application challenges	14
4.2.1	Offline data management	14
4.2.2	Massive import/export operations	15
4.2.3	Automating frequent time-consuming operations	16
4.2.4	Compliance and electronic signature	17
4.2.5	Transmittals	18
5	Conclusions	19
II	ABB Study Case	21
6	Project Requirements	21
6.1	Objectives	21
6.2	Internal typologies	21
6.3	Internal personnel responsibilities	22
6.4	Technical specification	24
7	Official ABB ECM Provider - SharePoint	25
7.1	Advantages	25
7.2	Limits and scalability	25

8	Solutions	26
8.1	Complete solutions	27
8.1.1	Cadac Organice	27
8.1.2	Sword CTSpace - Fusion Enterprise	29
8.1.3	EMC Documentum	31
8.1.4	Customized Microsoft Solution	32
8.2	Partial Infrastructure Solutions	33
8.2.1	Storage	33
8.2.2	Data distribution on server farm	36
8.2.3	Bi-directional server farm replication	39
8.3	Partial Application Solutions	43
8.3.1	Offline data management	43
8.3.2	Massive import/export operations	46
8.3.3	Automating frequent time-consuming operations	48
8.3.4	Compliance and electronic signature	55
8.3.5	Transmittals	56
9	Final decision	59
9.1	Considerations	61
9.2	Chosen Solution	62
	References	62

List of Tables

1	Ability to Execute Weighting	9
2	Completeness of Vision Weighting	10
3	BLOB Offloading - Solutions comparison	35
4	Summary Table - Problems and solutions	60

List of Figures

1	Gartner's Magic Quadrant: ECM, Q4 '10	8
2	Microsoft SharePoint - Global Distribution Deployment . . .	13
3	Microsoft - Rights Management Service	18
4	SharePoint 2010 - Application Limits	26
5	DocAve Replicator - Structure	40
6	DocAve Replicator - Topology	40
7	DocAve Replicator - Conflict	41
8	Colligo Contributor Client - Desktop Interface	43
9	Colligo Contributor Client - Conflict Resolution	44
10	Example of Access 2010 connected to SharePoint	47
11	Managed Metadata Service - Structure	50
12	Content Type Hub - Web application Creation	51
13	Content Type Hub - Web application Configuration	52
14	Content Type Hub - Web application Managed Metadata . .	52
15	Content Type Hub - Site Collection Layout	52
16	Content Type Hub - Site Collection Configuration	53
17	Content Type Hub - Site Collection Administration	53
18	Content Type Hub - Site Collection Publisher	54
19	Content Type Hub - Normal Project Information	54
20	Transmittal Alternative - Page Layout (General Info TAB) .	57
21	Transmittal Alternative - Page Layout (Related Documents TAB)	57
22	Transmittal Alternative - Standard Metadata Editing	58
23	Transmittal Alternative - Related Metadata Editing	58
24	Transmittal Alternative - Document Selection Screen	59
25	Transmittal Alternative - Content Type Transmittal Page Group	59

Part I

General design

1 Introduction

The main focus of this thesis is to study enterprise content management systems and point out the main difficulties related to its implementation and usage. The main objective of such a system is to manage content, including documentation, audio or video files and images, generated by the company or for the company. This management must comprehend all phases of the company's projects, including creation, development and archiving, and must follow a provided workflow. Additionally, the system needs to provide collaboration between users in all phases of a project.

The thesis began with an analysis of the current content management system of the process automation segment in ABB Italy. After studying the system, it became clear that there were several parts that weren't present or that could be significantly improved. ABB created this custom solution based on Microsoft SharePoint 2007, which is, as we will see further on, one of the main enterprise content management vendor currently available in the market. However, this custom application didn't predict the success that it could achieve, and thus didn't account for the scalability that the system would require. Enterprises today are realizing how managing their electronic content is essential for guaranteeing high performance and data availability for all employees, specially the managerial staff. Particularly, if we consider multinational companies, this task has an even greater necessity, since their content is distributed globally, so if a content management system isn't very well designed, companies face extremely difficult communication and performance problems.

There are two separate areas that a content management system should focus, firstly, the infrastructure and secondly, the application. The infrastructure should be the company's first concern, since a well implemented infrastructure will guarantee the system's scalability and performance, which is currently ABB's biggest problem. For example, every main software that provides the basis of content management, like Microsoft SharePoint, have their own particular form of maintaining and storing data, and they have well established limits on the size of the data that can be stored. This means that a company must find alternative ways of storing data if they are to store a larger volume of data than that predicted by the software's limit, otherwise, it might have a huge impact on the software's performance. And

this is only one of the problems that the company must address; there are still many others that will be mentioned further on.

After the infrastructure has been laid out, there is still a long way to go before the system can fulfill every requirement of the company, and thus comes in the second part of the implementation, which is the application part. Evidently, software like Microsoft SharePoint try to provide every possible requirement that a company may have, however, due to the specificity of different company areas and peculiarity of each one of them, this is quite a difficult task. And as a result, there are several requirements that aren't met by these software for large companies. We tried to address these main requests further on.

Finally, after the difficulties and objectives have been stated, we provide a thorough analysis of possible solutions following the requirements of ABB. And explain how a decision needs to be taken in order to take into account the most important factors for the company. It is important to realize that the solutions that are depicted as best suited are being considered for the case of ABB. They may not be the best suited solution for other companies.

2 What is Enterprise Content Management?

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. (AIIM, 2010)

ECM is unique in that it was developed to manage the creation and consumption of growing amounts of non-numeric content such as documents, Web pages, spreadsheets, diagrams and images, largely affected by the rise in popularity of the internet. (Glazer, Jenkins and Schaper, 2005)

The best ECM solutions deliver applications at the department level that integrate content management invisibly within the very act of collaboration. The transparent combination of content and collaboration benefits organizations tremendously by providing a place where simple ideas take root, are nurtured and finally mature into market-leading innovations. (Glazer, Jenkins and Schaper, 2005)

According to Kampffmeyer(2006), the five major ECM components and technologies can be categorized as:

1. *Capture* - This category contains functionalities and components for generating, capturing, preparing and processing analog and electronic

information. In this particular category, the most important applications are of recognition of human writing from scanned documents, in order to automate the process of transferring a paper document to an electronic copy.

2. *Manage* - The Manage components are for the management, processing, and use of information. They include the following traditional application areas:
 - DM or Document Management;
 - Collaboration (of supporting systems, groupware);
 - WCM or Web Content Management (including portals);
 - RM or Records Management (archive and filing management systems);
 - Workflow / BPM or Business Process Management.
3. *Store* - The components in this category are used for the temporary storage of information which it is not required or desired to archive. Even if it uses media that are suitable for long-term archiving, "Store" is still separate from "Preserve".
4. *Deliver* - The "Deliver" components of ECM are used to present information from the "Manage", "Store" and "Preserve" components. They also contain functions used to enter information in systems (such as information transfer to media or generation of formatted output files) or for readying (for example converting or compressing) information for the "Store" and "Preserve" components.
5. *Preserve* - The "Preserve" components of ECM handle the long-term, safe storage and backup of static, unchanging information, as well as temporary storage of information that it is not desired or required to archive.

On this thesis we will focus on the Manage, Deliver, Preserve and Store categories. However, since we only depicted the most important and relevant problems found in a real ECM system for a multinational company, we have made a different categorization of these problems. As mentioned earlier, the problems were divided into two different major areas, infrastructural problems, and application problems. Inside infrastructure we find problems related both to the Store and Preserve categories from Kampffmeyer's definition. And in the application section we find problems related to the Manage

and Deliver categories. The only category that was left out in the thesis was the Capture category, since ABB has no automated application to scan and transform paper documents in electronic copies.

3 Market Analysis

For a thorough market analysis, we tried to find information by leading experts in the market that perform such analysis. One of these experts, known for its state-of-the-art market analysis and for its cutting edge expertise and access to information, is called Gartner. This company has released, in the fourth quarter of 2010, a very interesting analysis on Enterprise Content Management. On the following sections we provide an excerpt from the Gartner report, which was allowed by the company, containing a short description of the ECM market, and an analysis on the most significant vendors available. The entire report is available for purchase from Gartner, and it is highly recommended for companies before actually deciding between ECM vendors.

3.1 Overview

ECM is becoming as much an essential part of an enterprise information infrastructure as it is an investment in strategies, suites and solutions. New delivery models, content as infrastructure, and process enablement can all help companies realize a worthwhile return on investment (ROI). But that's only part of the measurement of business value that buyers now expect from content investments.

The ECM market registered 4.8% growth in 2009, primarily due to global economic conditions. According to Gartner's statistics, worldwide ECM software license and maintenance revenue came to \$3.5 billion in 2009. Subsets of the market continued to generate higher revenue. For example, in Europe, the Middle East and Africa (EMEA), WCM revenue grew to claim almost 30% of this region's content management market.

Gartner forecasts that total software revenue in the ECM market will grow at a compound annual rate of 10.1% through 2014. The market's return to stronger growth is expected to start in 2010, with rates climbing to double digits and worldwide ECM software revenue exceeding \$5.7 billion by 2014. This represents a strong market, but it is also one that is undergoing transformation.

The ECM market has seen more mergers and acquisitions in 2010. Lexmark purchased Perceptive Software and Adobe acquired Day Software. We

see more vendors augmenting broader strategies in relation to information management by expanding their CCA foundations or promoting integration with Microsoft SharePoint. Underlying buyer concerns about technological dependency on vendors with multiple partners in capture, search, analytics and so on have spurred mergers and acquisitions. IBM's announced acquisition of Datacap further underscores the need to deliver a stable set of components on a single contract from a single vendor - though still rarely on a single stock-keeping unit.

3.2 Gartner's Magic Quadrant

Gartner's conclusions, displayed on Figure 1, are shown in a chart that has very specific evaluation criteria which are explained below.

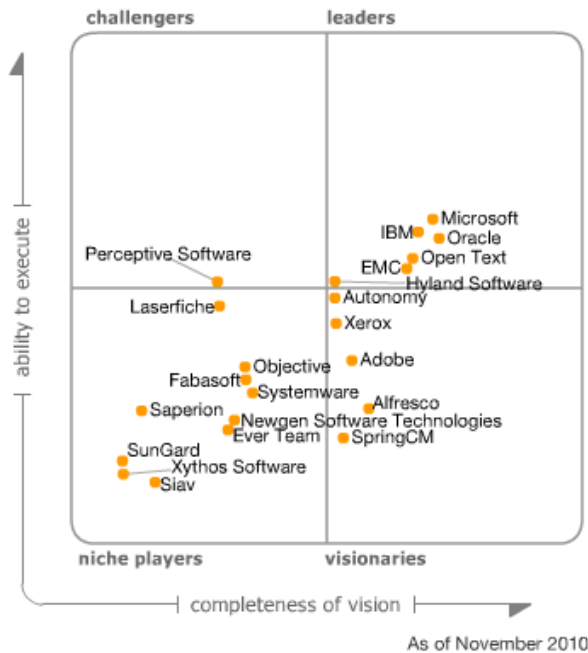


Figure 1: Gartner's Magic Quadrant: ECM, Q4 '10

The Magic Quadrant is copyrighted 2010 by Gartner, Inc. and is reused with permission. The Magic Quadrant is a graphical representation of a marketplace at and for a specific time period. It depicts Gartner's analysis of how certain vendors measure against criteria for that marketplace, as

defined by Gartner. Gartner does not endorse any vendor, product or service depicted in the Magic Quadrant, and does not advise technology users to select only those vendors placed in the "Leaders" quadrant. The Magic Quadrant is intended solely as a research tool, and is not meant to be a specific guide to action. Gartner disclaims all warranties, express or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose. The axis and their significance are:

- *Ability to Execute* measures how well a vendor sells and supports its ECM products and services on a global basis. In addition to rating product capabilities, we evaluate each vendor's viability, installed base, pricing, customer support and satisfaction, and product migrations from one major release to another.

To be considered for the Leaders quadrant, a vendor must provide most components natively, though they may be loosely coupled as a suite.

Although not explicitly identified as a core component, information access or search technology has always been a critical component of an ECM suite, and it plays a big role in helping companies sift through structured and unstructured information. All ECM products ship with a search engine embedded as a core component, so that users can create a full-text index and search the content stored in repositories.

Some vendors have added extended components, such as DAM for handling rich media, e-forms, and document and e-mail archiving, and document composition for high-volume generation of customized documents.

Table 1: Ability to Execute Weighting

Evaluation Criteria	Weighting
Product/Service	High
Overall Viability (Business Unit, Financial, Strategy, Organization)	High
Sales Execution/Pricing	Standard
Market Responsiveness and Track Record	Standard
Marketing Execution	Standard
Customer Experience	High
Operations	Standard

- *Completeness of Vision* focuses on potential. A vendor might succeed financially in the short term without a clearly defined vision or strategic plan, but it won't become a Leader.

A vendor with average vision anticipates change by accurately perceiving market trends and exploiting technology.

A vendor with superior vision anticipates, directs and initiates market trends, particularly if it integrates its vision for a broad range of areas, and capitalizes on product and service development.

Part of our assessment involves looking at how well each vendor understands changing requirements and market trends. We evaluate vendors on their awareness and adoption of emerging functionality, their technical architecture (for example, standards support, Web services and Web 2.0 capabilities), and their focus and abilities in federating and integrating with other content repositories and applications.

Table 2: Completeness of Vision Weighting

Evaluation Criteria	Weighting
Market Understanding	Standard
Marketing Strategy	Standard
Sales Strategy	Standard
Offering (Product) Strategy	High
Business Model	Standard
Vertical/Industry Strategy	Standard
Innovation	Standard
Geographic Strategy	Low

Thus, it is possible, by analyzing the conclusions provided by Gartner, to point out which are the company leaders on ECM and the fastest-growing companies. It is interesting to notice that Microsoft was depicted as the best rated company in this latest report by Gartner, which was the first time since its insertion in the market. This shows the rapid evolution and growth of the Microsoft SharePoint product in the ECM market. And it is also one of the reasons why it is the main focus of this thesis.

4 Problem statement

Since our study case is completely based on Microsoft SharePoint, due to the fact that it is ABB's official collaboration tool and, according to Gartner, one of the leaders in ECM, the problems are often focused in peculiarities of Microsoft's product, but may, however, not be present in other ECM vendors. For example, the storage problem is due to the particular form in which Microsoft SharePoint stores its data, but it is not present in EMC Documentum since it uses a different form of storage.

4.1 Infrastructural challenges

4.1.1 Storage

Microsoft SharePoint 2010 has a method of memory storage that is based on BLOBs(Binary Large Object), and they are stored in a SQL server. A BLOB is a collection of binary data stored as a single entity in a database management system. BLOBs are typically images, audio or other multimedia objects, though sometimes binary executable code is stored as a blob. Since BLOBs don't hold document structure, it is very difficult to divide them into different types of storage. As a result, all information regarding a same BLOB have to be stored together, which can be very expensive if it is not being entirely used. For example, if a same BLOB holds information about a closed project, and information about a current one, then the information on the closed project isn't being used, and could be stored at a cheaper form of storage. Therefore, companies tend to spend large amounts of money in memory storage because of the SharePoint method of storage.

The only solution to this problem, considering that the solution has to be deployed in SharePoint, is to use a method called BLOB offloading. This method is based on the division and storage of a BLOB in different types of memories, and it is not a simple process. All files contains two types of information, main data and metadata. Main data is all of the content of the file, for example, if the file is a document, the main data would be its body. Metadata, however, can be defined as data about data. It is data related to the specifics of the file. Using our document example, it would be the date of creation, document privileges, comments, and so on. Therefore, the process of BLOB division, is to capture the metadata assigned to each file and use it as a pointer to the main data that can be stored somewhere else. There are several different types of storage, each with its unique advantages and disadvantages, however, all of them less expensive then storing data on the SQL servers directly.

Using the BLOB offloading technique, the main server database would contain only the pointers to the file's metadata, while the files themselves would be stored somewhere else. This process would bring two great advantages. Firstly, it would save a considerable amount of expenses on cheaper memory storage options. And secondly, it would drastically decrease the size of the server database with a much smaller relative loss of performance, based on the fact that the search of files wouldn't be as vast as it was due to the size of the database. This technique could also be triggered based on file size basis, meaning that it could store larger files on cheaper memory storage and the smaller files in the server. These specifications on performance improvement and size shrinkage will be reviewed further on by each possible supplier.

4.1.2 Data distribution on server farm

When designing a database which will contain a huge amount of information, it is extremely important to consider how this data will be distributed on each server farm, keeping in mind the limits established by the Microsoft corporation on the SharePoint software as described further on.

Before analyzing the challenges of this problem, we must make some considerations. Firstly, we are considering that each server farm is related to a single country or region where the company is located, so as to guarantee that all projects are from the respective country or region. Next, regarding project allocation, we must keep in mind that SharePoint's main recommended limit is about the size of a single content database, which is 200GB per content DB, and that the company database will be considerably larger than that. And so, we know that it will be necessary to have several content databases. It is important to remember that this is a recommended software limit instead of a hard-wired one, which means that the software could run with larger content DBs, but it is not guaranteed that it would run on optimal performance. Therefore, the company would have to solve the problem of how to allocate its projects and site collections on each of these content databases.

The main challenges here would be to find the correct ratio between projects and site collections, as to optimize memory usage and performance. In addition, it would also be necessary to have different types of distribution of site collections among content databases, and a possible automated solution, so that there are differences of allocation for each project size. Lastly, there would have to be the possibility to transfer projects between databases, in case a project is aging and needs to be in a less expensive

file system. This last challenge is also connected with the storage problem, since a method of less expensive memory storage would have to be available to have a place to actually move older projects.

4.1.3 Bi-directional server farm replication

Since we are considering multinational companies, they hold several different server farms throughout the world. Most companies require a form of data replication between server farms to provide greater document availability. The optimal solution would be a completely distributed environment, in which every server farm would contain all information of all other farms. This is called a bi-directional replication between server farms, and involves high costs in deployment and storage and the company would have to possess a very fast network to avoid delays on data replication.

According to the requirements made by ABB, which are described further on, the biggest problem in a non distributed system is the latency time required for accessing information in foreign farms. And since usually old projects are used as a template for new projects, sometimes it is highly important for users to access different server farms that hold the project template that they need.

Microsoft currently recommends three different deployment configurations to accommodate geographically dispersed sites that are built with Microsoft Office SharePoint Server 2007 and Windows SharePoint Services 3.0. The three recommended deployment solutions are based on a central site and one or more regional sites. In each of the three, the central site hosts the majority of the SharePoint Products and Technologies solution for an organization.



Figure 2: Microsoft SharePoint - Global Distribution Deployment

Out of the three solution, the one that most resembles a distributed en-

vironment would be the distributed solution. This solution is depicted on Figure 2, where the central site is hosted in Madrid, regional sites are hosted at various cities around the world and each regional site is connected only to the central site. In this solution each regional site operates autonomously from other regional sites. And because it is not optimized for collaboration across regional sites, there is no bi-directional replication among all server farms. It is however a reasonable solution due to the fact that it has no implementation costs, considering that it is an innate functionality of SharePoint.

There are other solutions that actually implement a fully distributed environment using bi-directional replication, which will be reviewed further on at the ABB study case section.

4.2 Application challenges

4.2.1 Offline data management

An important functionality of a document control system is for users to be able to use the system even while offline. Needless to say that a multinational company has field workers and executives that are constantly traveling. These employees can't always be connected to the system, and often require to view and modify documents while disconnected. Microsoft didn't provide a solution for this problem in the old SharePoint version (Microsoft SharePoint 2007), however, in the new version it does. SharePoint Workspace 2010 is the name of this solution. It is a client application that provides fast, any-time interactive access to document libraries and lists on Microsoft SharePoint Server 2010 and Microsoft SharePoint Foundation 2010. It also provides options for creating Groove peer workspaces and Shared Folder workspaces.

Microsoft SharePoint Workspace 2010 provides a client for Microsoft SharePoint Server 2010 and Microsoft SharePoint Foundation 2010 that enables real-time synchronization of desktop content with SharePoint documents and lists. In addition, it provides options for creating Groove collaboration workspaces and synchronized shared folders. By using SharePoint Workspace 2010, information workers can synchronize online and offline content with a designated SharePoint site or collaborate with external partners and offsite team members through shared workspaces. Evidently, there are other solutions to be considered, however, since this is a functionality provided by Microsoft itself, it is a great leverage against all others.

It is important to remember that the biggest problem when users are

accessing content offline, is the conflict solving algorithm that a program will use to prevent loss of information. For example, if two users make changes to a certain document, and the synchronizing system simply merge these two different versions, bits of information could get lost. There are software that provide versioning of documents that were edited offline, so in our example, the first user's copy, would have a different version from the second one. This solution clearly avoids losing information, however, it will consume a large amount of memory, not to mention time for the project manager to put together all the bits of information contained on each copy. The optimal solution would be for a system to copy only the parts of a document that were altered and request that the user perform the merge of these parts, which will be condensed in a single file. In spite of still consuming the time of the user, it wouldn't consume a lot of memory and would not be subject to information getting lost. Further on, when different parts of the document were altered, and had no connection among themselves, the user could choose to merge automatically the new version of the document.

4.2.2 Massive import/export operations

An important operation when handling many projects simultaneously, is to be able to import and export large libraries to and from a server farm. For example, when a project is starting, and a new site collection will be created to hold all documents related to said project, a massive import operation would be required. This operation is based on the import of an initial library that holds all documents up to the point of the creation of the site collection, and all documents that will serve as reference to the completion of the project. Commonly, this initial library is very large, and currently, Microsoft SharePoint 2010 doesn't provide a form of automatically performing the import of a large library, so a document controller has to import each single document separately.

In addition, there are also moments in which an entire library, or even an entire site collection, has to be moved from one server farm to another. And if we consider that a single site collection can hold tens of thousands of documents, this operation can take several hours to be completed. So, even though this may seem like a simple problem, it can cause significant delays to a project.

There are many solutions that are applicable for these types of operations, many of them offered by highly recommended software houses, however, they are usually very expensive since they are often part of much broader software than one that just performs these operations. So we should

also consider, at this point, a simple, in-house solution that fulfills this requirement. We will see further on solutions that were considered by ABB to solve this particular problem.

4.2.3 Automating frequent time-consuming operations

There are many frequent operations that are usually performed manually for lack of support or out-of-the-box functionalities, which provide these desired operations. This results in a considerable waste of time and decreased productivity for the document controllers. Since these operations are usually repetition of previously performed tasks with very similar parameters, they can be automated and personnel can save a large amount of time.

The most important among these operations, is the provisioning operation. It is the task of creating and configuring a new project. Every time a project is created, it must undertake many configurations before it can actually be used. For example, we must associate the access levels of each person that will be connected to the project; allocate the correct memory space for the assumed size of the project; define the layout for the site collection; define the areas that will be present in the site collections; define the workflows of the project and so on. Many of these configurations are copied from other projects and slightly altered for new ones. Thus, it is a logical conclusion that old projects can serve as templates for new projects. By using projects as templates, the document controller can simply choose projects that most resembles the new one he is creating and configure it based on the chosen old project. Even though this is a very important tool for the system, software don't tend to focus on this part alone, and so, most of the solutions for this particular problem are custom ones.

Evidently, after a project is superseded, another operation is made necessary, the de-provisioning of the project. In this case, however, the requirement is very different. De-provisioning should operate in such a way as to provide the possibility for placing a project in "Read Only" mode after its archiving. Thus, after a project is stored, it cannot be altered in any way, guaranteeing its integrity. Another key aspect of this operation is that it must be able to move superseded projects to less expensive storage tiers, such as cloud storage or hard-disk storage.

Lastly, another common operation is to replicate global cross project metadata. Since there wasn't functionality in SharePoint 2007 that addressed this problem, it was necessary to manually replicate the desired metadata for all projects that would be accessed by different server farms. However, we will see later on that SharePoint 2010 has included a new func-

tionality called Managed Metadata that could solve this problem

4.2.4 Compliance and electronic signature

Organizations today must navigate a complex landscape of regulatory and legal compliance obligations. Companies must develop cost-effective and resource-efficient strategies to index, audit, secure, and manage their Microsoft SharePoint assets and users. Without such strategies in place, companies face excessive exposure to legal risk, financial shortfall, and productivity loss. (AvePoint, SharePoint Compliance Products)

These organizations around the world and across industries have invested in Microsoft SharePoint for automating their business processes. This investment has resulted in streamlined processes, reduced organizational costs, and improved collaboration and efficiency. However, when the need for a signature approval arises, organizations that haven't implemented a digital signature solution are forced to resort back to paper for obtaining approvals. This disruption in the electronic workflow minimizes the benefits of the investment in automation, while increasing organizational costs and bringing business to a halt. (Arx, CoSign Product)

Extending SharePoint with digital signature capabilities enables organizations to maximize their return on investment in automation while keeping business moving smoothly. Digital signatures keep your approval-based workflows electronic from start to finish; they streamline processes, improve collaboration and organizational efficiency, and cut costs. (Arx, CoSign Product)

Microsoft offers a product called Rights Management Service (RMS), which is focused on implementing security for documents and insuring compliance throughout an organization, see Figure 3. These are the main features of the product:

- *Application Support.* Support for AD RMS is already included within Windows Vista. Internet Explorer 7 and the 2007 Microsoft Office system already have support for AD RMS. The AD RMS client can also be installed on other Windows operating systems.
- *Persistent Protection.* Your content can be protected on the go. You specify who can open, modify, print, or manage the content, and the rights stay with the content—even after it has been transferred outside of your organization.
- *Usage Policy Templates.* If you have a common set of rights that you use to control access to information, a Usage Policy Template can be

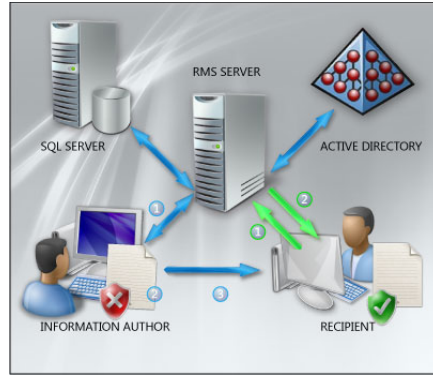


Figure 3: Microsoft - Rights Management Service

created and applied to content. This alleviates the need to recreate the usage rights settings for every file you want to protect.

- *AD RMS Software Development Kit.* The AD RMS Software Development Kit (SDK) can be used by independent software vendors (ISVs) to rights-enable their applications, meaning the application investments you've already made may be (or will become) compatible with AD RMS.

Unfortunately, this product does not predict the implementation of electronic signatures on Microsoft SharePoint, which is the main point in this problem. Thus it is necessary to resort to other suppliers that can fulfill this requirement.

4.2.5 Transmittals

A transmittal is, essentially, an official document sent from or to the company to provide confirmation of requests and deliveries. It is an important part in controlling and managing the progress of the projects and guaranteeing compliance by outsourced jobs. In addition, since transmittals usually carry signatures from project managers, they can be used legally for protection against delays and malpractice. A single transmittal can carry several attached documents, which can use a large amount of memory. Thus an important requirement on this phase is that documents should have the possibility to be physically attached to the transmittal or linked directed to the SharePoint portal, providing clients with the capacity to download directly from said portal.

In the end, the task of sending and receiving transmittals requires a particular workflow that involves the transmittal creation up to its issuing. Due to the fact that a transmittal will represent an official document, it is subject to several revisions before it can actually be sent. And each area of the company usually have a slightly different workflow for sending transmittals, so a solution for this particular problem should provide a flexible workflow for transmittals.

Sending and receiving transmittals are the most important functions for an engineering company, since they provide the official documentation between the company and clients or vendors. And because these functionalities are not present on SharePoint, they must be developed outside of the software's innate capabilities.

5 Conclusions

We conclude that even though software such as Microsoft SharePoint are very complete and powerful, they are usually not enough to fulfill every requirement of large corporations. And thus, every company that generates considerable content needs to invest in designing and implementing a better suited solution, be that a custom solution or an out-of-the-shelf one.

Enterprise content management systems are essential to guarantee the company's performance and productivity. It is a key aspect for companies to reduce bureaucracy and get things done with much more speed and simplicity. A well built system can even guarantee compliance from users, avoiding legal issues that may arise from loss of documents or signature fraud. Furthermore, it can provide a space for collaboration on projects, were users can share documents easily and work simultaneously in them knowing that they are always working in the final version. Managers can also insure that all users are subject to the defined workflow for a given project, which also reduces his workload, giving him time to focus on more important tasks. ECM can also provide companies with a different way of controlling projects. Every process that a company is undergoing is present on their ECM platform. This means that managers can have complete control of their employees and their progress.

It is important to understand that ECM is a term that is only beginning to be studied by corporations. There is still a lot to be developed and a lot to be learned about the immense possibilities that ECM can offer. One important function that can completely revolutionize the industry is the social aspect within the ECM portal. This functionality can improve significantly

5 Conclusions

connection inside a corporation and facilitate relationships and social gatherings. Thus, diminishing vertical structures and allowing employees to feel like they have a more complete understanding of the corporation and of the people who work in it.

In a nut shell, large corporations today cannot ignore content management, since it can improve substantially every phase of project development including communication with clients, suppliers and even employees.

Part II

ABB Study Case

6 Project Requirements

6.1 Objectives

The intended system should implement a centralized multinational document database capable of regulating informatics archiving. It should provide an optimal performance in communication during the development of ABB's projects and a correct documentation archiving after the completion of such projects.

All ABB project documentation should be managed in such a way as to guarantee immediate replication of content and necessary visibility and circulation among those involved during project execution.

The system should be developed in Microsoft SharePoint, which among many other advantages that are mentioned further on, allows:

1. Grouping of all project documentation;
2. Full control of document versioning and status;
3. Scheduling and registering document emission date;
4. Calculating document advancement progress by personalized filters;
5. Issuing transmittal of distributed documents;
6. Control distributed documents.

6.2 Internal typologies

There are three types of documents present in the current engineering document control system in ABB:

1. Client documents
 - (a) *Bid phase documents* contains all documentation officially provided by the client to ABB during the contract preparation phase.
 - (b) *General documents* contains the contract and all documents related to the development of the project. These documents are provided by the client to ABB under the terms of the contract or by organizations having contractual relationships with ABB.

2. *Project documents* contains all information, written or illustrated, regarding a specific project, which describes, document or certify activities, prescriptions, procedures and results.
3. *Vendor documents* contains all technical documentation(eg. blueprints, specifics) regarding suppliers services.

The document structure and typologies must continue to exist in the same form in the new document management system.

6.3 Internal personnel responsibilities

ABB currently holds six different profiles that access the document database. Each of them have different responsibilities and security privileges as follows:

1. Project Manager
 - Request IS for a project site in SharePoint;
 - Archive relevant documentation in its own directory;
 - Archive emails of relevant contractual and technical interest in SharePoint;
 - Indicate conservation period based on the agreements with the client and applicable laws;
 - Provide instructions to different units so that documents are collected after the acceptance of the blueprint and kept at the general archive;
 - Authorize document archiving, paper and electronic copy, on ABB general archive;
 - Monitor project's progress through EDC custom project functions.
2. Project Engineer
 - Manage all technical discipline responsible;
 - Manage all engineering archiving. Including documents related to the planning activities and control activities;
 - Respect SharePoint structure defined for the project, not tempering with it in any way;
3. Technical Discipline

- Manage all project team members;
- Manage archiving related to technical documents of the respective discipline.
- Respect SharePoint structure defined for the project, not tempering with it in any way;

4. Project Team Member

- Respect SharePoint structure defined for the project, not tempering with it in any way;
- Ensure proper archiving and tracking of documents within its competence;
- Archive emails of relevant contractual and technical interest in SharePoint;
- Approve discipline documents official revisions;
- Provide project's document collection as directed by the PM and PE.

5. Information System

- Monitor the integrity and security of stored data (viruses, unauthorized access, hardware and software malfunctions, etc.);
- Collect electronic documentation of active projects and make specified copies;
- Maintain copies of electronic documentation in suitable places in order to preserve its integrity;
- Create electronic archive structure, according to the official structure, and provide all accesses as specified by the PM(*Only for emergency/backup*);
- Manage any changes of access as requested by the PM, the only authorized reference(*Only for emergency/backup*);
- Deal promptly with applications bugs;
- Analyze, specify and implement all the improvements deemed necessary for the application software.

6. Document Controller

- Manage active project's set-up in SharePoint, with WBS data and Activity ID provided by project control;

- Create and maintains all the access requests for the projects in which is involved;
- Manage all sender/receiver list;
- Insert all document steps;
- Predispose the templates of transmittals broken down by recipient source (Internal, Vendor, Client);
- Issue transmittals(Project, Vendor, Client);
- Generate revisions of the technical documentation, verifying the consistency between the document in source and that in pdf format, and specially verify signature fields;
- Manage document review(Project, Vendor, Client);

The profiles and their respective responsibilities should remain the same in the new document management system.

6.4 Technical specification

ABB's current system was completely developed in house based on Microsoft SharePoint 2007, and has been used by ABB Process Automation division since June 2005. These are the technical specification of the system taken on 2009:

- Over 50 Web intranet portals and growing fast;
- Over 30 Web internet portals (for customers and suppliers);
- Used for large projects;
- 800 GB of space allocation;
- More than 650 thousand documents archived;
- More than 80 thousand transmittals issued;
- Approximately 300 ABB named users in write mode and 200 in read mode;
- Approximately 100 internet users (customers and suppliers).

The solutions for the new system have to take into account the size and specifications of the company today, and in the future, since the expectations of growth are substantial.

7 Official ABB ECM Provider - SharePoint

Among all of the possible ECM providers that we have seen on Gartner's report, ABB's official tool is Microsoft SharePoint. Thus, in this study case, it will not be considered solutions involving other providers, except if they integrate directly with SharePoint.

7.1 Advantages

The main advantage of Microsoft SharePoint 2010 is that it provides a state of the art software that implements the basic layer of a document control system. The software has a simple installation and configuration process, including a smooth upgrade from SharePoint 2007.

SharePoint 2010 also provides a single, scalable infrastructure that allows better control over server resources. Additionally scalable services architecture helps you centralize services like Search, My Sites, and Taxonomy. These services can be managed through Central Administration and scripted by Windows PowerShell. Also, because the architecture is extensible, third-party companies can build and add services to the platform. Lastly, it helps to reduce costs by consolidating intranet, extranet, and Internet sites on a single platform-on-premises or in the cloud.

Since an out-of-the-box application wouldn't be enough to fulfill all of ABB's needs, an important advantage of SharePoint is that it provides a platform for customized solutions. One can use the features of SharePoint 2010 just as they are or quickly create secure and easy-to-use solutions for specific business needs.

7.2 Limits and scalability

SharePoint's 2010 test results and guidelines provided are based on a single SharePoint Server 2010 farm. There are three types of limit definitions:

- *Boundaries* are absolute limits that cannot be exceeded by design. It is important to understand these limits to ensure that you do not make incorrect assumptions when you design your farm.
- *Thresholds* are those that have a default value that cannot be exceeded unless the value is modified. Thresholds can, in certain circumstances, be exceeded to accommodate variances in your farm design, but it is important to understand that doing this may affect the performance of the farm in addition to the effective value of other limits.

- *Supported limits* define the tested value for a given parameter. The default values for these limits were defined by testing, and represent the known limitations of the product. Exceeding supported limits may cause unexpected results, significant decrease in performance, or other harmful effects.

All of SharePoint's limits, as defined by Microsoft, can be found on Microsoft's tech support site (<http://technet.microsoft.com/en-us/library/cc262787.aspx>), however a quick reference to them follows on Figure 4.

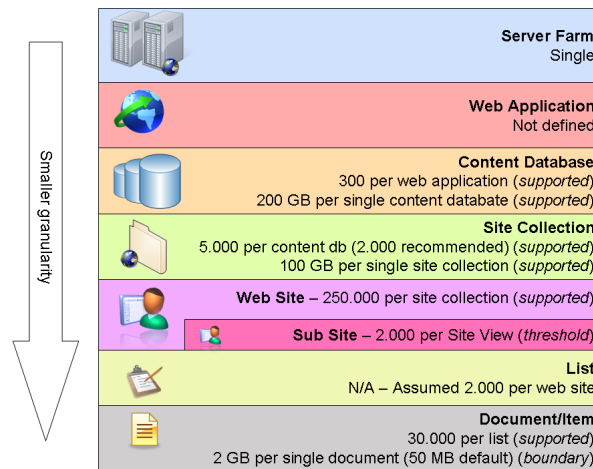


Figure 4: SharePoint 2010 - Application Limits

8 Solutions

The solutions were analyzed in two different forms, solutions that delivered one single product to supply every or most of ABB's requests and partial solutions to supply each individual request. Evidently, due to the complexity of ABB's needs the final choice became constricted to products that would supply most of ABB's requirements rather than many products for each need.

However, since the objective of the thesis is to perform a thorough analysis of the best possible answer when designing a document management system, all of the partial solutions that were found are also documented. In this manner, even if these products are not the best solution for ABB, they could be the answer to smaller companies that have less complex needs and

require only a couple of these products. It is also important to mention that the partial solutions section contains one or two examples of software or custom programming that could be used to solve each particular problem, but there are a large number of alternative solutions available that aren't mentioned. The objective of this section is not to point the best software, but rather possible solutions and how they would function and improve the system.

8.1 Complete solutions

8.1.1 Cadac Organice

The Cadac Organice Product Suite provides solutions for engineering document management and document control and is based entirely on Microsoft SharePoint, Microsoft's platform for Enterprise Content Management and Project Collaboration. It provides project-driven organizations a flexible and efficient environment to manage (project) documents and to collaborate with internal and external project partners. The Cadac Organice Product Suite leverages these benefits and enhances SharePoint with a user friendly interface and the engineering document management and document control functionality customers need.

Cadac Organice offers a practical yet comprehensive solution at an affordable price. It provides high return on investment due to higher quality, stricter compliance and increased efficiency in (engineering) document management and project collaboration. It provides a comprehensive document management solution for engineering industries, including solutions for input and creation, management and control, output and distribution of (engineering) documents.

The Product Suite is divided into several modules that fulfill needs separately. These are the available modules, and a brief summary of its objectives:

- *Cadac Organice Explorer* is a smart client for SharePoint providing a user friendly interface and enhancing SharePoint with engineering document management functionality. Main functionalities include navigation (users can create, store, search, retrieve, view, check in, check out, edit and revise documents from the interface), office integration, email integration, CAD integration, site provisioning tool and document numbering tool.
- *Cadac Organice Transmit* helps Document Controllers distribute and

transmit multiple SharePoint documents efficiently to multiple recipients at once and offers management information to Document Controllers who need to track transmittals that have been sent. Main functionalities include creating, sending and tracking transmittals.

- *Cadac Organice Workbox* builds productive serial and parallel state workflows in SharePoint to automate business processes. Main Functionalities include creating and executing workflows and following workflow history.
- *Cadac Organice Publish* publishes documents that are stored in SharePoint to other SharePoint document libraries or sites. You can publish documents in their native format, but you can also convert documents into an independent file format, so they can be opened, viewed and annotated by others without the native applications. Main functionalities include publishing and converting documents.
- *Cadac Organice Scan* incoming correspondence (postal registration) and other hardcopy documents can be scanned and stored in the SharePoint document environment, making all documents centrally and digitally available to users. Hardcopy documents can first be registered and then scanned into the SharePoint environment, or first scanned and then registered and moved in SharePoint. Scanned documents are interpreted using Optical Character Recognition (OCR) and converted into searchable PDF documents.
- *Cadac Organice Connect* connects SharePoint data with data in ERP, PSA and PDM/PLM systems, and vice versa. This concept leads to a connection layer and a family of connectors that can be used to integrate the SharePoint environment with 3rd party business applications in the customer's environment.
- *Cadac Organice Gatekeeper* helps companies that collaborate with external partners in the design process, that want to standardize use of CAD within their engineering teams, or that outsource design work to external partners. It helps by allowing them to define and automatically enforce design standards, thereby ensuring the quality of their CAD files. Using design standards ensures interoperability between CAD Engineers and prevents errors caused by misunderstanding or misinterpretation.

8.1.2 Sword CTSpace - Fusion Enterprise

Sword CTSpace provides document management and collaboration workflow applications for the construction and engineering industry that increase the efficiency and reduce the risks in managing a built asset throughout its entire lifecycle.

Fusion Enterprise products provide engineering and drafting teams the functionality to successfully leverage an organization's generic document management system into a configuration management system for engineering projects. The provided functionality allows for:

- Robust version control;
- The management of relationships between engineering documents;
- The management of compound drafting files;
- Engineering document collaboration;
- Document viewing and markup.

Sword CTSpace's online offerings allow engineering development consortiums to quickly deploy an engineering content management solution without any IT investment. Engineering teams from multiple organizations can come together via the extranet to collaborate on engineering content in a secure and controlled environment.

Fusion Enterprise customers can leverage the on-line offerings to extend engineering content collaboration to external partners without impacting the integrity of the internal enterprise document management system.

Policies, processes and procedures can be quickly translated into paperless forms and workflows that are integrated across the multiple organizations normally involved in a construction project. The logging, creation, distribution, tracking, and response to: bid packages, RFIs, change requests, change orders, issues logs, etc. are performed on-line.

The Paragon Cost Management System allows owners and program / project managers the ability to track actual costs against a budget. Reporting tools help keep projects on track financially by comparing actual results to predicted results for a given project milestone.

The content generated and used by Fusion Enterprise is stored in your ECM platform and can be searched using your Enterprise Search and Business Intelligence tools. Fusion Enterprise adds a set of Engineering Content functionality that applies across the whole spectrum of your asset lifecycle.

Fusion Enterprise has been specifically designed to manage engineering documentation in the context of a construction project or operating plant or facility. It enables organizations to automate everyday document control operations and document-centric business processes using a standard web browser.

Here is a summary the software's advantages:

- Allows organizations to quickly deploy document control and automate business process to enforce compliance to corporate and governmental standards;
- Enables anybody within an entire organization or supply chain to quickly perform complex design and engineering tasks;
- Supplements an organization's chosen ECM system with modules that optimize the business processes of an enterprise from top-to-bottom, from project management to workflow management. These modules integrate seamlessly with the native ECM platform, allowing project participants to work in a familiar environment while lowering implementation costs and eliminating the need for extra training time;
- Project Management Integration;
- Document Lifecycle;
- Transmittals;
- Document Distribution;
- Document Relationships;
- CAD Drawing Management;
- Alerts Warnings;
- EAM Integration;
- Concurrent Engineering;
- Document Viewing, Markup Approval;
- Workflow Management;
- Change Control;
- Batch Import Export;

- PDF Generation.

As we can see, many of ABB's requirements are theoretically fulfilled in the specifications of the software, however there are still a couple of them that weren't mentioned, such as server-farm replication or storage possibilities.

8.1.3 EMC Documentum

Even though EMC is an ECM vendor, and it was decided that Microsoft SharePoint 2010 will be the official vendor for ABB, EMC was still considered since the company provides an integration tool between their product and SharePoint. Thus, the solution would be to implement the EMC Documentum product and use the EMC integration tool to continue using SharePoint. This is an important and very possible solution since EMC is considered one of the main authorities in ECM as we have seen from the market analysis.

EMC Documentum Platform provides essential capabilities for managing enterprise content and is the foundation for EMC Enterprise Content Management and Intelligent Case Management product offerings. Service-oriented and based on industry standards, the Documentum Platform adheres to Content Management Interoperability Services (CMIS) and supports a broad range of operating systems, databases, application servers, and enterprise applications. It adds intelligence to the organization by offering text analytics as well as advanced and federated search. EMC Documentum's main benefits are:

- *Industry-leading content management;*
- *Unified platform and code base;*
- *Open, standards-based architecture;*
- *Content analytics and federated search;*
- *Compliance and security;*
- *Global distributed deployments;*
- *Enterprise class;*
- *XML component management.*

As regards to the integration between EMC and SharePoint, the provided solutions help companies enhance and control their Microsoft Office SharePoint Server deployments. EMC can provide a proven, enterprise-class content management infrastructure for Microsoft SharePoint and Microsoft Office desktop applications. End users access a single content platform that addresses business process, compliance, and archiving requirements. These are the key benefits:

- *Seamless integration with Microsoft Office* - Access EMC Documentum services such as browse, search, check-in/check-out, and workflow directly from the Microsoft Office Desktop, Outlook, and SharePoint applications.
- *Reduced compliance risks* - Avoid penalties during litigation or audits by automatically capturing and classifying records content across Microsoft Office applications and documents.
- *Storage efficiency* - Lower storage costs by migrating older or infrequently accessed content from primary storage systems to more cost-effective archival storage.
- *Automated business processes* - Leverage business process management capabilities within the EMC Documentum repository from the SharePoint interface.
- *Easy browsing and search* - Use federated search services to search SharePoint content as well as EMC Documentum and other third-party content.

8.1.4 Customized Microsoft Solution

Another possibility for a product that could suppress every request from ABB was to design another application from scratch. However, since ABB doesn't currently possess the availability to develop this application in house, it was decided to search third party suppliers that could implement the necessary improvements to SharePoint 2010, and to perform the migration of the current Document Control System to the new platform.

From the several possibilities, the chosen one was Microsoft itself. The company has an area strictly for custom solutions development for multinational companies, and due to its guaranteed complete support with the SharePoint tool, they were the natural first choice.

However a customized solution can indeed fulfill each single request from ABB, there are still downsides when choosing this option, which are shown in the Conclusions section of the study case.

8.2 Partial Infrastructure Solutions

8.2.1 Storage

RBS FILESTREAM Provider This solution is provided by Microsoft as a free download and is supported in SharePoint 2010. It essentially uses the RBS interface supported by SharePoint 2010 to wrap SQL 2008's FILESTREAM capability.

RBS uses a provider to connect to any dedicated BLOB store that uses the RBS APIs. Storage solution vendors can implement providers that work with RBS APIs. The RBS FILESTREAM provider uses the SQL Server FILESTREAM feature to store BLOBs in an additional resource that is attached to the same database and stored locally on the server. The FILESTREAM feature manages BLOBs in a SQL database by using the underlying NTFS file system. The data is not stored in the MDF file, but in another file that is associated with the database.

This implementation of the FILESTREAM provider is known as the local FILESTREAM provider. You can conserve resources by using the local RBS FILESTREAM provider to place the extracted BLOB data on a different (cheaper) local disk. Since SharePoint Foundation 2010 does not currently support remote RBS FILESTREAM provider, this implementation can only be used on local hard disk drives, thus excluding remote storage devices as a possible storage solution.

Given that the main purpose at this point is to store content remotely in different types of file systems, this solution is not feasible.

DocAve Extender (AVEPOINT) This solution essentially implements a remote RBS/EBS provider. Therefore, it is capable of allowing BLOB off-loading not only to local hard disks, but to any network addressable file share, File Transfer Protocol system, or cloud-based storage system (including Rackspace, Amazon, and Microsoft). With this technique, there can be shrinkage of up to 95% of SQLServer's memory size, thus increasing performance and decreasing memory overload.

The DocAve Extender is a free software, however to achieve the desired functionalities, it is required to purchase the DocAve Backup and Recovery module as well, which has license costs included. This module offers full

fidelity backup and recovery, from an individual content item to an entire SharePoint environment and all of its farm-level components, maintaining all metadata, securities, and version histories. The implementation is based on a Business Criticality Matrix, which classifies all SharePoint content by both business criticality and usage activity, and automatically associates it to appropriate backup schedules. It also automatically detects new SharePoint content and performs scheduled or on-demand full, incremental, and differential backups.

StoragePoint (Metalogix) This third party software leverages the SharePoint EBS (WSS3/MOSS 2007/ SFS/MSS 2010) and SQL RBS (SFS/MSS2010) provider interfaces to externalize SharePoint content BLOBs in real-time. StoragePoint storage profiles determine when content is externalized and are configured by SharePoint Web Application, Content Database, or Site Collection. It also delivers an Intelligent Archiving capability that allows a SharePoint or site collection administrator to define rules for moving BLOBs to less expensive tiers of storage as they age or become less relevant to the day-to-day operation of the business

Performance wise, the product features an asynchronous operations capability. This means that lesser performing tiers of storage (i.e. the Cloud) can be performed without impacting the end user experience. A SharePoint administrator simply sets up a system cache where BLOBs will be temporarily written until they are moved to their final destination.

SharePoint content BLOBs can be sent to virtually any on-premise (i.e. SAN, NAS, CAS) storage platform. There is also support for several Cloud storage platforms. Additionally, BLOBs can be externalized using the same folder structure and filename used in SharePoint instead of an arbitrary date/time folder structure and GUID for the filename.

Best suited solution In this case the chosen solution would be Metalogix's product StoragePoint. It contains every desired functionality for this particular problem, including the preservation of folder structure. A more detailed comparison of each requirement is studied on table 3.

The only downside of this solution would be the license costs, however, given the amount of expenses that would be saved by cheaper file storage it is a highly recommended solution.

Table 3: BLOB Offloading - Solutions comparison

Requirement	RBS FILESTREAM provider	DocAve Extender	StoragePoint
Manage Tera bytes of Content in a Single Content Database	NOT recommended	YES	YES (2T+)
BLOB Encryption	NO	YES (256 bit)	YES (256 bits)
BLOB Compression	NO	YES	YES (Zip64/Deflate OOB)
Implement Tiered Storage/HSM	NO	NO	YES
Move content to less expensive tiers of storage as it ages	NO	(No implementation of HSM)	YES
Move content to less expensive tiers of storage as its state changes	NO	NO	(Intelligent BLOB archiving)
Move content to less expensive tiers of storage as it retains versions	NO	NO	YES (Intelligent BLOB archiving)
Leverage Less Expensive NAS Storage	NO	YES	YES
Leverage Inexpensive Cloud Storage	NO	YES	YES
Store BLOBs on WORM-Compliant Devices	NO	YES	YES
BLOB/File Shred on Delete	NO	YES	YES
Use Windows DFS for BLOB Replication	NO	NO	YES
Hold folder structure on external storage	NO	NO	YES
Use cache for slower performing tier of storage	NO	NO	YES

8.2.2 Data distribution on server farm

In order to list the possible solutions on this point, it is important to enumerate the necessary steps, or rationalizations to find a suitable answer.

Firstly, we must define the correct content DB/Site collection ratio, so that we can define ways of spreading the content according to this ratio. To do so, we consider that a single content DB can contain several site collections, and that a single site collection can contain as much as 25.000 sites. A single site collection can easily hold many projects. However, unlike shifting site collections between content DBs, which is a fairly simple task, dividing projects that are in the same site collection is a very time consuming operation. And since projects have very different deadlines, size and priority, a change between the databases in which they are stored is a very common operation. Thus reaching the logical conclusion that a site collection must hold a single project. Consequently, with this solution, it is achievable:

- Self consistency;
- Scalability;
- Portability(Shifting between content databases);
- Replicability;
- Cleanness in archiving and historicizing;
- Creation of site collection can be based on a custom module on a specific content DB, and based on a provided topology of the project(small, medium and large).

The second task is to effectively discuss how the distribution of content will be made among content DBs. There are two possible solutions for this problem, the first one is to create a battery of preexisting content databases and assign projects accordingly; and the second one is to dynamically create content databases and assign projects. An analysis of both possibilities follows:

Battery of preexisting content DB The only advantage that can be assumed from this solution would be the possibility to schedule back ups prior to the insertion of new projects, since the content DBs would already be created. However, there are several disadvantages in this scenario, for example: manual creation and management of projects; loss of control in

content DB assignment and naming convention determined prior to implementation. As a further analysis, we consider the three possible topologies for projects as described earlier. The following is the intended manner in which the application should handle each of them.

1. *SMALL* project

- Create content DB denominated prior to implementation: *SMALL_1*, *SMALL_2* ... *SMALL_n*;
- The application, provided with a new small project, has to know in which content DB to insert the project by interrogating each content DB to check availability;
- In case of saturation, the application needs to block the operation and notify the administrator;
- The administrator needs to manually switch off a saturated content DB to avoid creation of new site collections from other sources.

2. *MEDIUM* project

- Create content DB denominated prior to implementation: *MEDIUM_1*, *MEDIUM_2* ... *MEDIUM_n*;
- The application, provided with a new medium project, has to know in which content DB to insert the project by interrogating each content DB to check availability;
- In case of saturation, the application needs to block the operation and notify the administrator;
- The administrator needs to manually switch off a saturated content DB to avoid creation of new site collections from other sources.

3. *LARGE* project

- Manual creation of content DB for large projects;
- The application cannot dynamically create a content DB given that the administrator has to firstly create a special content DB. He also has to instruct the tool, or the responsible for the project creation, to deposit the project in the content DB.

Dynamically create content DB Differently from the previous solution, by dynamically creating content DBs, most of the disadvantages of the static creation are turned into advantages. There is no need for administrator management in project creation, since it is an automated operation; There is full control in content DB assignment and naming convention is defined at run time. The only disadvantage of this solution is, ironically, the only advantage from static creation, which is backup. In this case, content DB back up is postponed after its creation. The application will send an email to the administrator warning him of the need for backing up the new content DB. Like before we also analyze the response of the application for each project topology:

1. *SMALL* project

- No content DB created in advance;
- The application, provided with a new small project, creates the site collection in the only present and active *SMALL_x* content DB;
- If the new project exceeds the small project quota for the current content DB, the application creates a new content DB (respecting the correct *Naming* convention), switches off the old one, creates the site collection in the newly created content DB and sends an email to the administrator to request backup.

2. *MEDIUM* project

- No content DB created in advance;
- The application, provided with a new medium project, creates the site collection in the only present and active *MEDIUM_x* content DB;
- If the new project exceeds the medium project quota for the current content DB, the application creates a new content DB (respecting the correct *Naming* convention), switches off the old one, creates the site collection in the newly created content DB and sends an email to the administrator to request backup.

3. *LARGE* project

- No content DB created in advance;

- The application, provided with a new large project, creates a new content DB(respecting the correct *Naming* convention), creates the site collection in the newly created DB, switches off the content DB and sends an email to the administrator to request backup.

As it can be easily inferred, the most suitable choice in this matter would be the dynamical creation of content DBs, providing more freedom in project creation management.

The last, but not least important task, is to define how projects can be transferred among content DBs, but since a site collection has been defined to contain a single project, the transferring of projects is fairly easy since site collections can be easily moved between content DBs.

8.2.3 Bi-directional server farm replication

No document replication Because of the extent of this particular problem, it is important to consider the alternative of remaining with a separate structure and abandoning the idea of replicating all material to every server farm. In this case, the only implementation necessary would be the creation of a single web application to every server farm, providing considerable savings in material replication and license acquirement for third party software. The downside of this solution would be the latency when accessing foreign information. Additionally, since one of ABB's requests is for users to be able to access their documents offline, if we consider that this request will be fulfilled, than the problem of replication is solved in a different manner. If users need to have content from different server farms, they can simply set an automatic update of this content to be viewed offline, so that their computer will always contain the necessary documents, even if they are offline, removing the latency time, and remaining always updated.

Content deployment with Microsoft SharePoint Server Like it was said before, Microsoft provides a possible solution to this problem using their own technology. However, this solution does not implement a completely bi-directional replication, instead, it provides a unilateral one, since there would be a primary server farm that would contain all information about other server farms.

Even though this is not exactly what is requested by ABB, it is still a solution worth mentioning. Specially since it has no implementation costs, and is fairly simple to be conceived.

DocAve Replicator (AvePoint) DocAve Replicator performs live, event-driven, as well as scheduled, or off line replication, as shown on Figure 5. Synchronization and management of all content, configurations and secu-

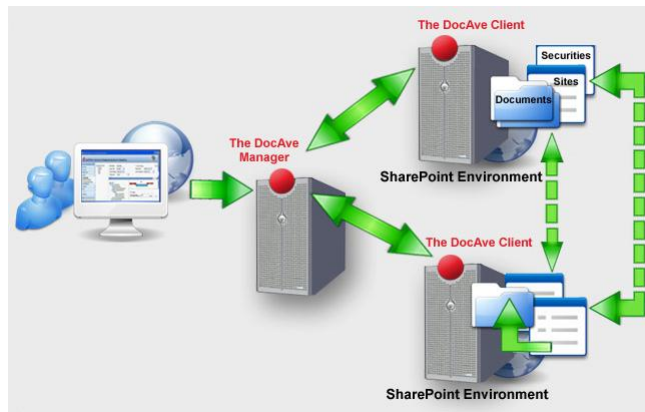


Figure 5: DocAve Replicator - Structure

rities is performed with full fidelity. It provides administrators with the ability to replicate all elements one-way, two-way, or one-to-many, within a single SharePoint environment, or across multiple SharePoint farms - even over an extranet. The topology as desired by ABB is shown on Figure 6,

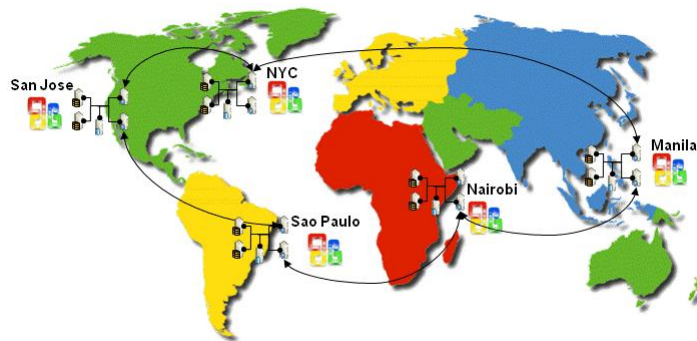


Figure 6: DocAve Replicator - Topology

where regional locations would maintain their own local SharePoint farms, but - directly through this SharePoint local farm - regional users can access both:

1. Content that is available to them through their local farm;

2. Content replicated from the other farms throughout the deployment.

Utilizing this two-way (bi-directional) replication poses unique challenges, however. Depending on the frequency of content and configuration updates, it is possible that a conflict may occur between the actions of multiple parties on a given content item or configuration. DocAve Replicator provides an elegant means to handle such incidences by providing administrators with an interface to define precedence rules to resolve conflicts, as it can be seen on Figure 7. Rules can be based on modification time, document versioning,

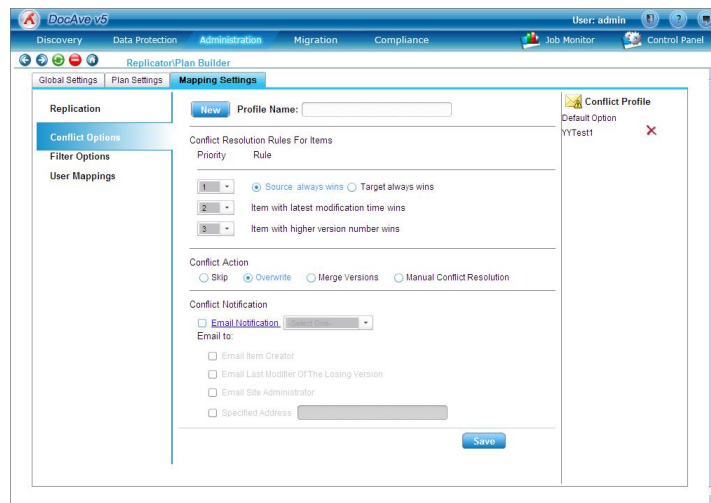


Figure 7: DocAve Replicator - Conflict

or source/target precedence. Additionally, administrators can provide end users with greater control by allowing manual conflict resolution. In this scenario, notifications alerting end users about conflicts are sent, so they may take the appropriate course of action to resolve the issue. Lastly, DocAve replicator also provides integration with DocAve Extender, so it provides replication of file system storage. However, it is still unknown whether it would provide this replication with software such as StoragePoint.

Replicator (Syntergy) Syntergy's Replicator, like DocAve Replicator, provides bi-directional replication and its a very recommended solution by IT professionals. It is not a caching technology, but its instead true replication of the content to another SharePoint site, providing inter-farm replication services for MOSS 2007 and WWW v3.0. No client software is required to

be installed and user interface matches SharePoint's look and feel to help reduce the learning curve of administrators.

Replicator is capable of keeping information current across geographic barriers, collaborating in virtual teams across the country and the world. Syntergy's Replicator for SharePoint supports the ability to define other SharePoint servers and create a partnership between them. At the list and site level you can then select specific areas that you wish to replicate and the target location for the replication. By setting up two servers talking to each other you have bi-directional replication which includes conflict resolution.

As provided by Syntergy's website(www.syntergy.com), here is a list of Replicator's key features:

- Integrated into SharePoint - no additional training of external applications;
- Remote Differential Compression - Only transfer the blocks that change, not the whole file;
- Keep User and Group information synchronized between sites;
- Replicate permissions assigned on the list;
- Rules Engine to selectively replicate based on meta-data;
- Supports replication of check-in/check-out state and all meta-data, including content types and site columns;
- Conflict resolution support;
- Map library to an alternate named site;
- HTTP/HTTPS protocol only, no infrastructure changes;
- No crawlers that put a load on the server, all event based replication.

Best suited solution According to professional reviews, one of the best suited software available for fulfilling ABB's requirement of bidirectional replication is the Syntergy Replicator. In spite of being the most expensive solution in terms of license fees, it fulfills completely the task of global availability of content. However, it remains to be decided, if the advantages of having global content available to all server farms supersedes the disadvantages of spending considerable investment in license costs. In addition, it is important to consider that ABB's network has significantly increased

its speed in the last few years, which means that the latency when accessing foreign information has diminished from what it used to be. Thus, the solution of no document replication is also to be considered if ABB is to continue increasing its network speed, and even more so if ABB is immersed in an offline data availability solution.

8.3 Partial Application Solutions

8.3.1 Offline data management

Colligo Contributor Client for SharePoint This is a desktop software that increases SharePoint adoption by providing a seamless user experience online and offline. It enables information workers to instantly access, modify, and create SharePoint documents, list items, and metadata - anywhere. Mobile and remote workers remain productive, regardless of network latency or unavailability, while working in the field or from branch office, and while traveling.

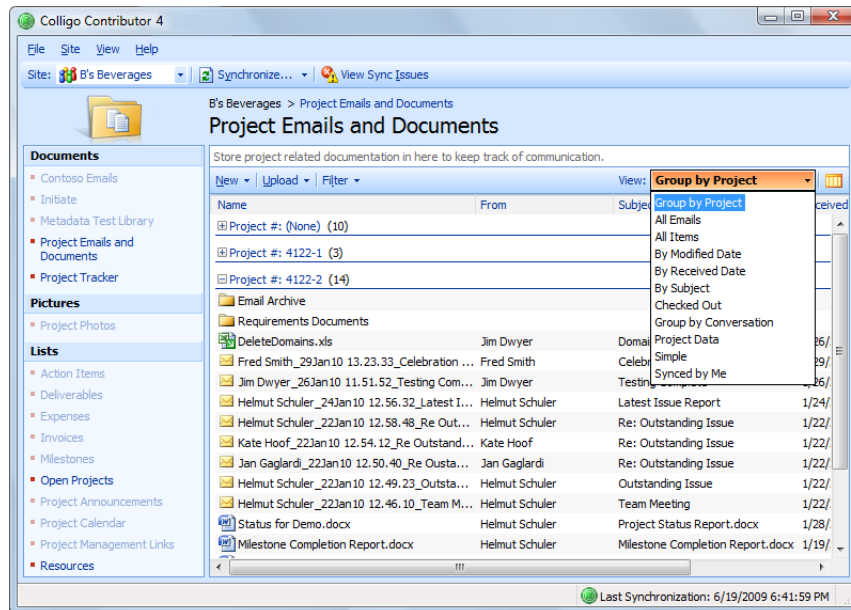


Figure 8: Colligo Contributor Client - Desktop Interface

Contributor Client synchronizes SharePoint content to a local encrypted data store for instant access - improving user productivity over slow connections or when working offline. Its intuitive drag-and-drop interface is

tightly integrated with Microsoft Office desktop applications so it's easy to use. 2-way sync, data encryption, document check-in / check-out, and elegant conflict resolution enable authorized users to keep content safe and up-to-date. Desktop interface can be seen on Figure 8.

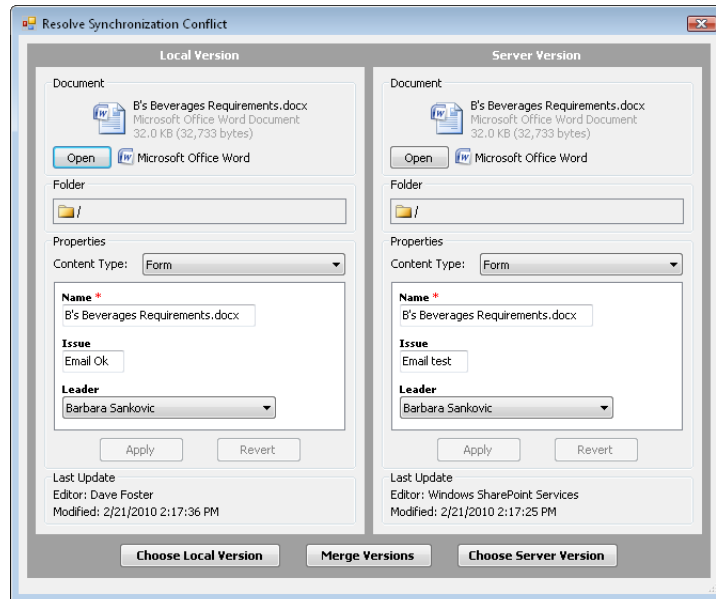


Figure 9: Colligo Contributor Client - Conflict Resolution

Users can select specific documents, document sets, libraries, or entire SharePoint sites for sync to their laptops.

Check-In and Check-Out is supported when users are working online. Elegant conflict resolution, integrated with Microsoft Word compare and merge, notifies the user if conflicting edits have been made by multiple users while offline. Figure 9 illustrates this conflict resolution interface.

This solution covers all desired aspects of ABB's desire, thus solving the problem of offline data availability. The only downside of this solution are the license costs.

SharePoint Workspace SharePoint Workspace 2010 is the succession of Microsoft Office Groove 2007. SharePoint Workspace 2010 is a client application that provides fast, any-time interactive access to document libraries and lists on Microsoft SharePoint Server 2010 and Microsoft SharePoint Foundation 2010. SharePoint Workspace 2010 also provides options for cre-

ating Groove peer workspaces and Shared Folder workspaces. SharePoint Workspace 2010 is more versatile than Microsoft Office Groove 2007 and can be integrated with Microsoft SharePoint Server 2010 or can run independently.

Microsoft SharePoint Workspace 2010 provides a client for Microsoft SharePoint Server 2010 and Microsoft SharePoint Foundation 2010 that enables real-time synchronization of desktop content with SharePoint documents and lists. SharePoint Workspace 2010 also provides options for creating Groove collaboration workspaces and synchronized shared folders. By using SharePoint Workspace 2010, information workers can easily synchronize online and offline content with a designated SharePoint site or collaborate with external partners and offsite team members through shared workspaces. SharePoint Workspace 2010 is included with Microsoft Office Professional Plus 2010.

Some of Microsoft's Workspace 2010 key features follows:

- Choice of workspace types;
- Easy setup;
- Offline and online collaboration. *Information workers can easily synchronize online and offline work, through a SharePoint workspace, Groove workspace, or Shared Folder workspace. Content is synchronized dynamically among online collaboration points and updates are transmitted immediately when an offline client comes back online. When an offline client reconnects, SharePoint Workspace automatically adds offline contributions to the workspace and applies workspace updates to the previously offline client;*
- Immediate access to latest document versions;
- Bandwidth usage optimization;
- Integration with Windows logon;
- Microsoft Communicator integration and built-in workspace messaging;
- Familiar user interface;
- Searchable content;
- Common file dialog boxes;

- Safe Mode operation;

We then conclude that this software, as well as Colligo's software, also fulfills ABB's requirements. And Microsoft's Workspace presents the advantage of being distributed by Microsoft itself, which guarantees full compatibility.

8.3.2 Massive import/export operations

In this case, as it was said before, using third party software can implicate high licensing costs, even though they can solve the problem. Firstly, we analyzed one of these solutions, and then provide alternative possibilities that could be customized in house.

Metalogix SharePoint Site Migration Manager 2010 SharePoint Site Migration Manager migrate sites, lists and libraries between SharePoint servers while retaining views, versions, user-info, web parts, permissions and metadata. Migrate from SPS 2003, WSS v2, MOSS 2007, WSS v3.

The software's key features are:

- *Reorganize or re-template sites during migration* Re-define site structure while migrating. Split Sites and Site collections into multiple targets and re-template Sites. Administrators and content owners can define what content is migrated.
- *Migrate all list types* Migrate Document Libraries, Issues, Tasks, Contacts, Announcements, Discussions and Custom Lists while preserving views, metadata, and user-edit information.
- *Retain your valuable data* Preserve views, version chains, metadata, and user-edit information.
- *Migrate Permissions* Map and migrate site permissions from WSSv2/SPS 2003 and WSSv3/MOSS 2007 to SharePoint 2010 easily.
- *Migrate Web Parts* Migrate all out-of-the-box and custom web part content from WSSv2/SPS 2003 and WSSv3/MOSS 2007 sites to SharePoint 2010.
- *Connect to multiple sites* Connect to multiple SharePoint sites, across any number of servers, for easy consolidation or distribution of your SharePoint data.

- *Work remotely* Do all the heavy lifting for your SharePoint migration from the convenience of your own machine. SharePoint Site Migration Manager connects to any SharePoint site that you can access with your browser – even those outside your organization’s intranet.
- *Compliant* Run migration activity through the supported SharePoint APIs. There are no un-supported, direct writes to the SharePoint database.
- *PowerShell Enabled* Access and extend all SharePoint Site Migration Manager functionality via PowerShell commandlets, to script and automate all migration functionality.

Custom solutions A possibility could be to make use of SharePoint’s Edit Datasheet functionality, which has significantly improved in SharePoint’s new release. Ultimately, the document controller would have the possibility of editing data with an excel-like interface but directly online.

A valid alternative could also be using Microsoft Office Access 2010 with data connection. In this case the document controller would have the possibility to use the Access interface connected to data to modify documents, and also particular fields. An example of how Microsoft Access 2010 connected to SharePoint would look like is provided on figure 10.

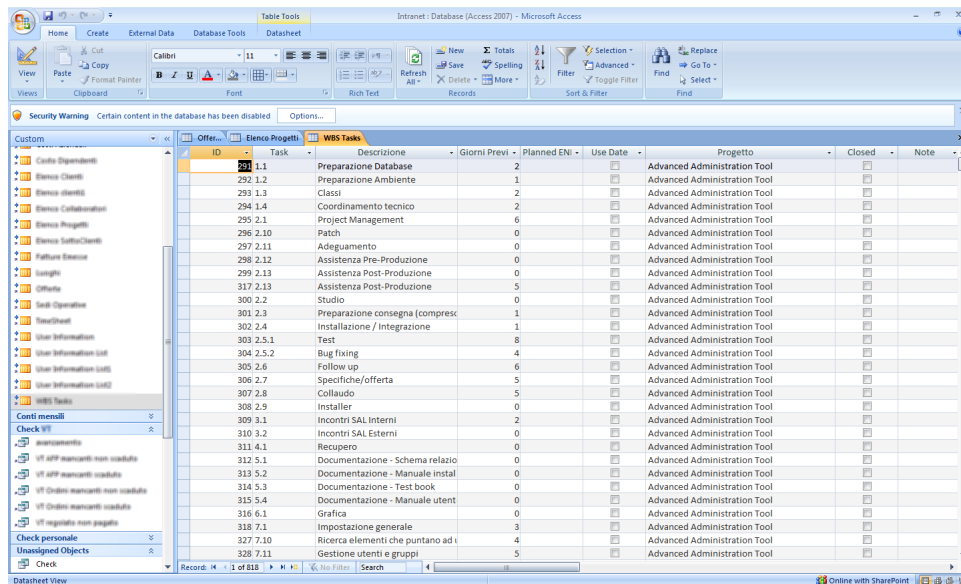


Figure 10: Example of Access 2010 connected to SharePoint

8.3.3 Automating frequent time-consuming operations

Provisioning

Template from old projects The best solution that we could find for provisioning, is to create templates to be used by newly formed projects. A template can be considered a Site Definition, in which an interface will be presented to the user containing all required data for the creation of a new project, as well as fields for selection of users and access levels. After the user has compiled these information they will be sent to approving users and in the final stage, according to the defined work flow, a job will be scheduled for the creation of a new Site Collection. It is important to think of a Timer Job running on the system with high priority that allows performing Power Shell operations such as creating Content Database.

A Site Definition will be constructed for each project typology. Upon the creation of a new Site Collection(equivalently to a project), a Site Definition will be chosen from among a prepared basket which will support the entire Site Collection structure. These Site Definitions are to be considered molds of project, which lose their relation with a project as soon as it is created. Massive changes to the structure or to the pages will be managed in the future with the implementation of custom programming.

It remains to be studied the exact number of Site Definitions to be designed and the data structures and data itself to be exposed in the Homepage.

De-Provisioning

As we have seen before, the data distribution of a single project per site collection allows for a simple transfer between content BDs. In such a way that if a project is no longer being used, it can simply be moved to a content DBs in a less expensive memory tier and be set to "Read Only" to avoid users from altering consolidated information.

Managing global metadata

Microsoft Business Connectivity Services is included in Microsoft SharePoint Server 2010 and the Microsoft Office 2010 suites. This module grants a set of services and features that provide a way to connect SharePoint solutions to sources of external data and to define external content

types that are based on that external data. External content types resemble content types and allow the presentation of and interaction with external data in SharePoint lists (known as external lists), Web Parts, Microsoft Outlook 2010, Microsoft SharePoint Workspace 2010, and Microsoft Word 2010 clients.

By using Microsoft Business Connectivity Services, one can design and build solutions that extend SharePoint collaboration capabilities and the Office user experience to include external business data and the processes that are associated with that data. This evolution of the BDC (Business Data Catalog) would help in case there were large lists to handle which were already recorded on external data sources.

Managed Metadata is a hierarchical collection of centrally managed terms that you can define, and then use as attributes for items in Microsoft SharePoint Server 2010.

A *term* is a word or a phrase that can be associated with an item in SharePoint Server 2010. A term set is a collection of related terms. You can specify that a Microsoft SharePoint Server column must contain a term from a specific term set. Managed metadata is a way of referring to the fact that terms and term sets can be created and managed independently from the columns themselves. There are two types of term sets, those that are local to a site collection context and those that are global, which are created outside a site collection context. Users can see only global term sets and term sets that are local to the user's site collection. Additionally, terms can be divided into two types:

- *Managed terms*, which are usually pre-defined, can only be created by users with the appropriate permissions, and are often organized into a hierarchy.
- *Enterprise keywords*, which are simply words or phrases that have been added to SharePoint Server 2010 items. All enterprise keywords are part of a single, non-hierarchical term set called the keyword set.

There are several benefits that can be attained by using managed metadata. Among these it is important to mention:

1. Facilitates more consistent use of terms, as well as more consistent use of the enterprise keywords that are added to SharePoint Server items;
2. Higher degree of confidence in metadata due to consistency;

3. Simple searches can provide more relevant results because items have consistent attributes;
4. Easier insertion of new values for metadata, improving the dynamics of documentation control.

Enterprise Content Types with managed metadata service In addition to sharing managed metadata, you can also use the managed metadata service to share content types. By creating a new managed metadata service and specifying a site collection as the content type hub, you can share all content types in the site collection's content type gallery.

You can create multiple managed metadata services, and share multiple term stores and content types from multiple site collections. However, each managed metadata service must specify a different term store. When you specify a nonexistent database for the term store, a new database is created. On Figure 11 the corporate managed metadata service is the primary

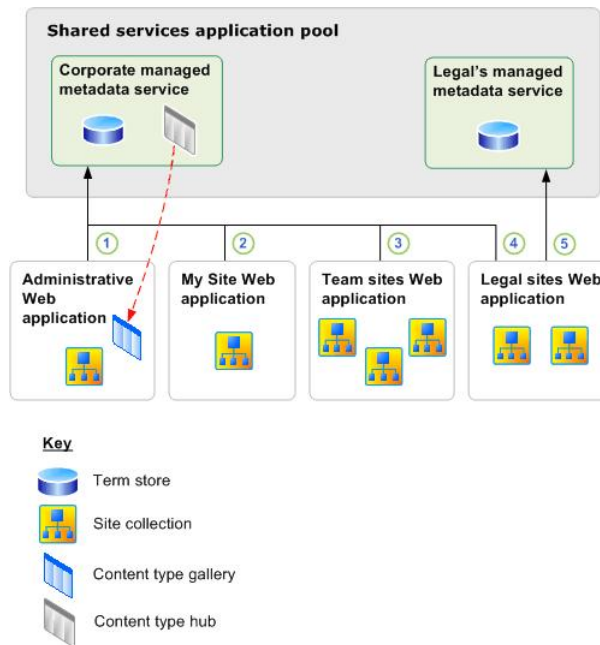


Figure 11: Managed Metadata Service - Structure

managed metadata service for all SharePoint Server 2010 sites in the company. The corporate taxonomy is represented by global term sets in the term store that is associated with the corporate managed metadata service. The

content type hub that is associated with the corporate managed metadata service makes shared content types available to users of all site collections.

Every Web application has a connection to the corporate managed metadata service. The connections from the My Site Web application, the team sites Web application, and the legal sites Web application, numbered 2, 3, and 4 in the figure, all have restricted access to the corporate managed metadata service. Restricted access lets users of the sites in these Web applications use the shared content types and global term sets, to add new enterprise keywords, and to create local term sets, but it prohibits them from modifying global term sets.

Best suited solution In order to achieve ABB’s goal for managing external data, the best suited solution would be to make use of managed metadata service, a SharePoint’s innate functionality. With this methodology, global metadata can be made available for all site collections and can be centralized in a single location. This solution intends the creation of a content type that would act as the hub for all other web applications. It would provide information to other web applications through a publish/subscribe structure provided by managed metadata. We will now describe how the Content Type Hub should be created to fulfill ABB’s request. It should first be created a single web application, with a single managed path for all child site collections (using *node sites*). For this example this path will be:

- <http://ABB.SharePoint.Portal>

Figures 12, 13 and 14 show the creation and activation of this site collection to become the content type hub. After configuring the web application, the HUB site collection would be configured, Figures 15 and 16 show this process.

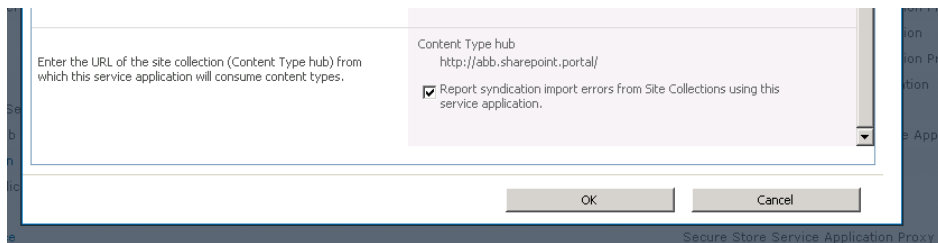


Figure 12: Content Type Hub - Web application Creation

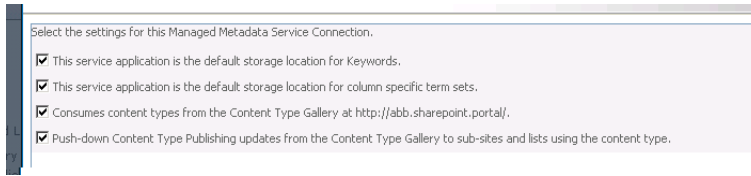


Figure 13: Content Type Hub - Web application Configuration

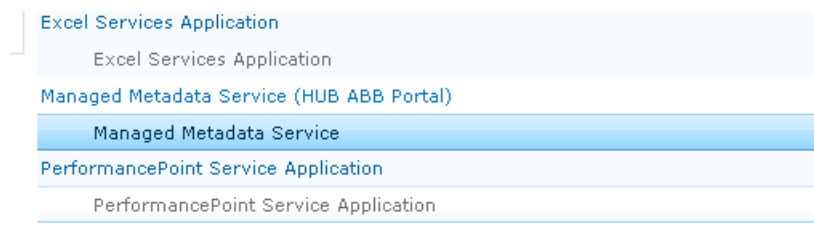


Figure 14: Content Type Hub - Web application Managed Metadata

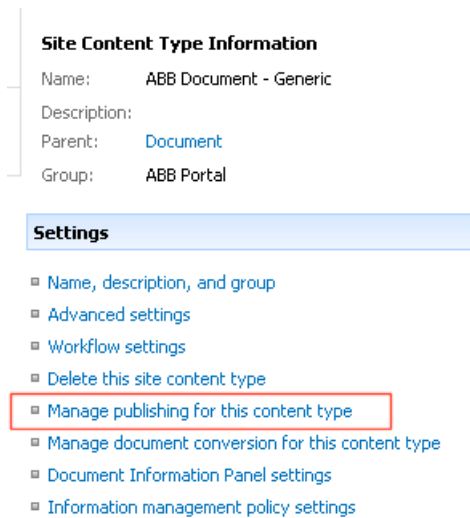


Figure 15: Content Type Hub - Site Collection Layout

It is important to mention that there are no defined quota for the site collection, and that this site collection would contain all general information, cross project information and pages for management and administration of projects (eg. Creation of new project with pre-approved workflow).

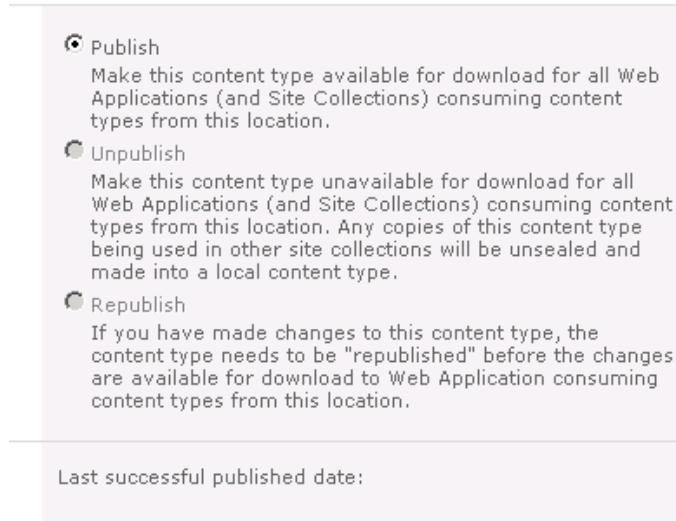


Figure 16: Content Type Hub - Site Collection Configuration

The propagation of a content type is managed through a Timer Job. An example of how to publish a content type can be seen on Figures 17 and 18.

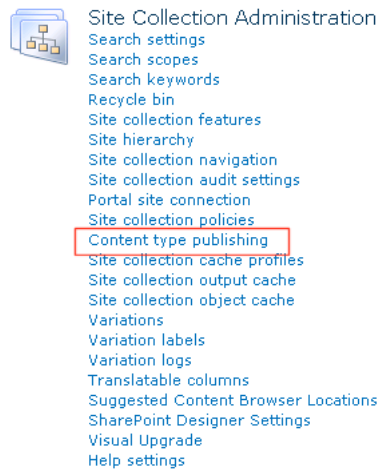


Figure 17: Content Type Hub - Site Collection Administration

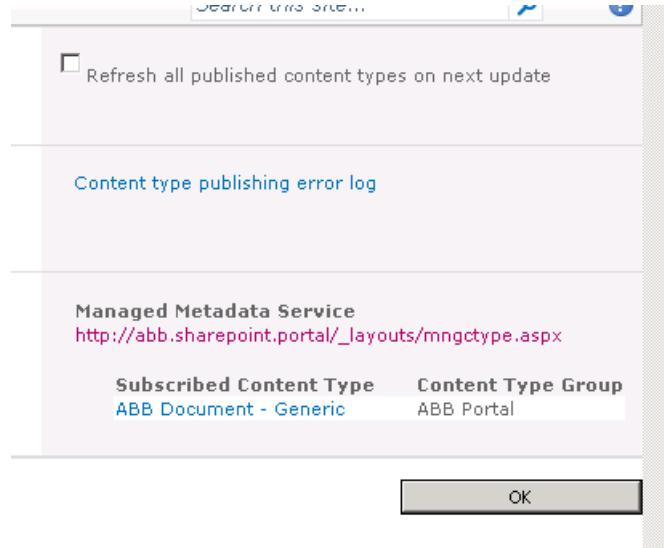


Figure 18: Content Type Hub - Site Collection Publisher

Lastly, an example of a project with path: `//ABB.SharePoint.Portal/sites/Progetto1`, which is a content type connected to the normal site collection, would contain the information provided on Figure 19.

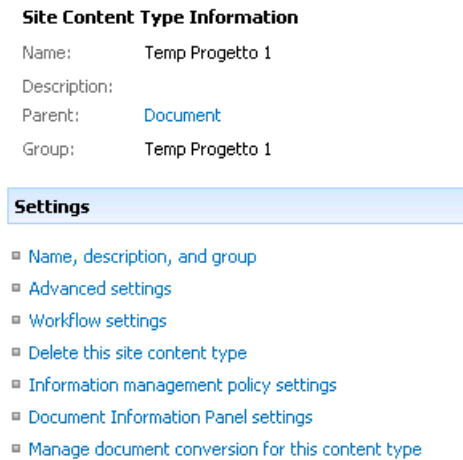


Figure 19: Content Type Hub - Normal Project Information

This node, like all others connected through the hub node, inherits the content types from its parent on the standard scheduled Timer Job mode. This means that the deployment of a content type in the root node will have effect over all site collections, thus centralizing the management of deployment and of data modifications.

8.3.4 Compliance and electronic signature

CoSign - Digital Signatures for Microsoft SharePoint CoSign is the a digital signature solution that is seamlessly integrated with Microsoft SharePoint 2007/2010, enabling users to approve, agree, accept, and audit legally enforceable documents, list items, and records that are centrally located on the SharePoint server. The application is compatible with all major document types including PDF, Microsoft Word, Excel, and InfoPath. Furthermore, it allows compliance with regulations worldwide, meeting the needs of organizations in the most tightly-regulated industries.

CoSign digital signature capability with Microsoft SharePoint is available in three structures suited to the company's needs, however for the considered solution, the only necessary structure would be the CoSign Standard Support for SharePoint. Which includes signature clearing rules, sectional signing, reason code management, configurable signature blocks, graphic-signature image management, and signatures that are backward compatible across different versions of Microsoft Office? Users may sign any document type using CoSign prior to checking it into SharePoint, and may also leverage CoSign's sign and save (WebDAV) support to sign PDF documents and save them back to SharePoint servers in one step. The only missing feature in this structure is to signing PDF directly, even though this requirement can also be fulfilled if desired by using also the CoSign Add-on. If it was a company's desire, there are more advanced features such as programming new compliance rules, but this isn't the case in ABB.

Infowise Signature Field The application implements a custom field that enables signing SharePoint items or document content and properties with user name and password to comply with regulatory requirements. The field stores, in an encrypted form, the user identity, time and date of signing, item version (if enabled) and a hash sum of signed fields' values.

Entering a valid user name and password pair during the data entry or editing phase will sign the item and a small icon will appear to indicate that the item has been signed and the signature is intact. If the item is edited

again, without re-signing, the signature is broken and the icon is replaced to indicate this fact.

The field can be set up to sign one or more of the other fields in the list, the signature will only affect the field that the field is attached to, changes to other fields of the item will not break the signature.

According to professionals, this is a simpler solution than the one displayed above, however it perfectly fits ABB's requirements and it has a significantly lower license fee cost.

Best Suited Solution In this case, both solutions completely solve ABB's requirement. The solution provided by CoSign appears to be more complex and complete, however it has a much higher license fee cost, and since it is not precisely necessary these advanced features, the best suited solution in this case is the one provided by Infowise.

8.3.5 Transmittals

Firstly, it was considered the use of Document Sets, a new SharePoint 2010 functionality. However, since the documents needed to be physically present in a document set, it would be impossible to add the same document to several document sets, which in this case, would represent Transmittals.

Thus, the second and best suited solution, would be to transfer the transmittal logic to the logic of publishing pages. In this case, when a given user needed to create a transmittal, he would basically create a New Page in SharePoint. This page would then be attached to the Page Layout EDC-Transmittal, which would contain all the metadata required for a transmittal. Also, by using the User Control custom, attached to the Page Layout as well, the user would view and define contacts and attached documents.

A practical demo is displayed on the following figures. Figure 20 shows the Page Layout, which is divided in tabs to provide a much faster view of the metadata. Figure 21 shows the related documents tab of the Page Layout. Figure 22 illustrates the standard metadata editing interface. Figure 23 shows the related metadata editing interface. Figure 24 illustrates the model dialog in SharePoint 2010, with an application page to provide document selection. And finally, Figure 25 shows the group of pages of the Content Type Transmittal.

CONTRATTO 1

The screenshot shows a software interface for 'CONTRATTO 1'. On the left, there is a navigation menu with sections 'Lists' and 'Libraries'. Under 'Lists', items include 'Related', 'Substituted', 'MacroAreas', 'Countries', 'Document Type', and 'Document SubType'. Under 'Libraries', items include 'Attachments' and 'CONTRATTO 1'. The main content area has three tabs: 'Generale', 'Allegati', and 'Documenti Relazionati'. The 'Generale' tab is active, displaying 'INFORMAZIONI GENERALI'. The content is organized into two columns. The left column contains 'Immagine relativa :', 'Macro area : BUSINESS DATA', and 'Document type : DOCUMENT TYPE ABC123'. The right column contains 'Descrizione documento : STEFANO PAGE CONTENT', a line of asterisks 'AA', 'Country : ITALIA', and 'Document subtype : DOCUMENT 1'.

Figure 20: Transmittal Alternative - Page Layout (General Info TAB)

CONTRATTO 1

The screenshot shows the same software interface for 'CONTRATTO 1', but with the 'Documenti Relazionati' tab selected. The main content area is titled 'DOCUMENTI RELAZIONATI e SOSTITUITI'. It contains two sections: 'Documenti correlati : CONTRATTO 1' and 'Documenti sostituiti :'. The left navigation menu remains the same as in Figure 20.

Figure 21: Transmittal Alternative - Page Layout (Related Documents TAB)

8 Solutions

The screenshot shows a SharePoint metadata editing interface for a document titled "CONTRATTO 1". The interface is divided into several sections:

- Top Bar:** Includes a ribbon with "Save & Close", "Check In", "Paste", "Copy", "Undo", "Font", "Paragraph", "Styles", "Spelling", "Markup Styles", and "HTML" options. A status bar below the ribbon indicates "Status: Checked out and editable. Publication Start Date: Immediately".
- Left Navigation:** A sidebar with "Lists" and "Libraries" sections. Under "Lists", "CONTRATTO 1" is selected.
- Main Content Area:**
 - Informazioni Generali:** Contains fields for "Immagine relativa" (with a "Page Image" label and a "Click here to insert a picture from SharePoint" link), "Descrizione documento" (with a "Page Content" label and the text "STEFANO PAGE CONTENT"), "Macro area" (with a "Macro" dropdown set to "BUSINESS DATA"), "Country" (with a "Paese" dropdown set to "ITALIA"), "Document type" (with a "DocType" dropdown set to "DOCUMENT TYPE ABC123"), and "Document subtype" (with a "DocSubType" dropdown set to "DOCUMENT 1").
 - Bottom Left and Right:** Each contains a button labeled "Add a Web Part".

Figure 22: Transmittal Alternative - Standard Metadata Editing

The screenshot shows a SharePoint metadata editing interface for a document titled "CONTRATTO 1". The interface is divided into several sections:

- Top Bar:** Includes a ribbon with "Save & Close", "Check In", "Paste", "Copy", "Undo", "Font", "Paragraph", "Styles", "Spelling", "Markup Styles", and "HTML" options. A status bar below the ribbon indicates "Status: Checked out and editable. Publication Start Date: Immediately".
- Left Navigation:** A sidebar with "Lists" and "Libraries" sections. Under "Lists", "CONTRATTO 1" is selected.
- Main Content Area:**
 - Documenti Relazionati e Sostituti:** Contains sections for "Documenti correlati" (with a "CONTRATTO 1" entry and a "Gestione Documenti Correlati" button) and "Documenti sostituiti" (with a "Gestione Documenti Sostituiti" button).
 - Bottom Left and Right:** Each contains a button labeled "Add a Web Part".

Figure 23: Transmittal Alternative - Related Metadata Editing

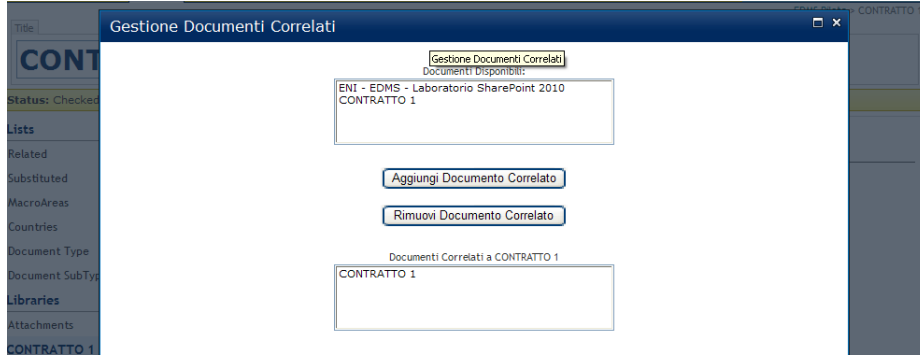


Figure 24: Transmittal Alternative - Document Selection Screen

Name	Document Date	Modified	Modified By	Blocco	DocType	DocSubType	Macro	Paese
Page Layout : EDMS Contracts (1)								
Documento-Prova-1	8/27/2010 11:00 AM	8/30/2010 1:10 PM	Stefano Magri		DOCUMENT TYPE ABC123	DOCUMENT 1	BUSINESS DATA	ITALIA

Figure 25: Transmittal Alternative - Content Type Transmittal Page Group

In the particular case of transmittals there are a large number of third party vendors that perform these types of operations. However, since they all involve high license cost fees, we decided to focus the solutions into simple custom ones that wouldn't involve these costs.

9 Final decision

The final decision made by ABB depended on several different points which are noted further on. We added in this section a summary table (Table 4), containing the main discussed points to serve as reference to all the problems and solutions discussed in the thesis, including which problems are solved by the considered providers. It is important to remember that this table is intended as a summary for the considered ABB study case, and the specifications which are written as "Not Applicable (N/A)" in no way indicate that the company doesn't provide the functionality in question, simply that it is not applicable for the case of ABB.

Table 4: Summary Table - Problems and solutions

Requirement	Cadac Organice	Sword	EMC Documentum	Custom Microsoft Solution	Other Supplier
Storage	N/A	N/A	YES	To be developed	StoragePoint (Metalogix) N/A
Data distribution on server farm	YES	N/A	YES	Embedded	
Bi-Directional server farm replication	N/A	N/A	YES	To be developed	Replicator (Synergy)
Offline data management	YES	N/A	N/A	YES (Microsoft Workspace) To be developed	Colligo
Massive import/export operations	N/A	YES	N/A	To be developed	SharePoint Site Migration Manager (Metalogix) N/A
Site Provisioning	YES	N/A	N/A	To be developed	N/A
Manage Global Metadata	YES	N/A	YES	Embedded (Managed Metadata)	N/A
Email Integration	YES (Microsoft Outlook)	N/A	N/A	Embedded (Microsoft Outlook)	N/A
Workflow Management	YES	YES	YES (Simple workflows)	Embedded (Simple workflows)	Nimtex Workflow
PDF Generation	N/A	YES	N/A	To be developed	N/A
Electronic signature and Compliance	N/A	N/A	N/A	To be developed	CoSign
Scan Documents	YES	N/A	N/A	To be developed	N/A
Transmittals	YES	YES	N/A	To be developed	N/A

9.1 Considerations

The final evaluation was based in three main parts. The first part was related to the considered company's reputation, capability and expertise. The following point were considered, each with a given weight:

1. Quality of references
2. Capability to support ABB
3. Viability of the solution provider
4. Commitment to SharePoint

Secondly, as it was discussed all throughout the thesis, the infrastructural challenges that were fulfilled by each company were considered. Evidently, each of the points that are abstracted in the thesis were more specified for ABB's needs, and are here pointed out:

1. Global Performance in ABB Distributed environment
2. Connectivity and Integration to other Tools (LN and ABB standards)
3. Leverage Native SharePoint Solution
4. Security & Role Base Access
5. Offline Capability for Installation and Commissioning
6. Deployment, Upgrade, Operation & Support
7. Capacity to handle a big volume of Documents (mandatory)
8. Project Archive Management

Lastly, the application requirements were discussed. These are the main points that were considered:

1. Document Life Cycle & Team Collaboration
2. User Friendly, Search, Navigation & Project Landing Capability
3. Internal and External Communication
4. Reporting, Tracking Progress and Dashboard
5. Workflow/Tasks management for handover of work packages, approval, release ... processes

6. Reuse Project Structure & Template from Previous Project
7. Rendition services
8. Synchronize Project Metadata to Files

Even though these are the main points that were considered, that are several other points that aren't mentioned here but that were also critical for deciding which was the best suited solution for ABB.

9.2 Chosen Solution

Unfortunately, due to compliance regulations and disclosure agreements, it is not allowed for ABB's notifications and grading for each of the suppliers to be written in this thesis. It is also not possible to be revealed which was the chosen supplier to implement the new ECM system for ABB.

References

- [1] Ulrich Kampffmeyer, *ECM - Enterprise Content Management*. Project Consult, Hamburg, Germany, 2006.
- [2] AIIM. Association for Information and Image Management. *What is Enterprise Content Management*. <http://www.aiim.org/What-is-ECM-Enterprise-Content-Management.aspx>. 2010.
- [3] David Glazer, Tom Genkis, Hartmut Schaper, *Enterprise Content Management Technology - What you need to know*. Ontario, Canada, 2005.
- [4] Gartner, *Magic Quadrant for Enterprise Content Management*. November 16th, 2010.

Suppliers - Partial Solutions

- [5] DocAve Extender, *AvePoint*.
Web site: <http://www.avepoint.com/sharepoint-storage-extender-docave/>
- [6] StoragePoint, *Metalogix*.
Web site: <http://www.storagepoint.com/>
- [7] DocAve Replicator, *AvePoint*.
Web site: <http://www.avepoint.com/sharepoint-replication-docave/>

References

- [8] Replicator, *Syntergy*.
Web site: <http://www.syntergy.com/products/sharepoint/replicator/enterprise/>
- [9] Colligo Contributor Client, *Colligo*.
Web site: http://www.colligo.com/products/sharepoint/contributor_client.asp
- [10] SharePoint Workspace 2010, *Microsoft*.
Web site: <http://technet.microsoft.com/en-us/library/ee649102.aspx>
- [11] Metalogix SharePoint Site Migration Manager 2010, *Metalogix*.
Web site: <http://www.metalogix.net/products/sharepoint-site-migration-manager/>
- [12] Business Connectivity Services in SharePoint Server 2010, *Microsoft*.
Web site: <http://technet.microsoft.com/en-us/sharepoint/ee518675>
- [13] CoSign, *Arx*.
Web site: <http://www.arx.com/products/digital-signature-for-SharePoint>
- [14] Infowise Signature Field, *Infowise*.
Web site: <http://www.infowisesolutions.com/product.aspx?id=SignatureField>

Suppliers - Complete Solutions

- [15] Cadac Organice, *Cadac Group*.
Web site: <http://www.organice.com/Pages/default.aspx>
- [16] Sword CTSpace - Fusion Enterprise, *Sword Group*.
<http://www.sword-ctspace.com/section/view/405>

Vendors

- [17] Microsoft SharePoint 2010, *Microsoft*.
Web site: <http://sharepoint.microsoft.com/en-us/pages/default.aspx>
- [18] EMC Documentum, *EMC*.
<http://www.emc.com/domains/documentum/index.htm>