OpenText Application Governance & Archiving™ for Microsoft SharePoint™

RIMtech Technical Assessment Report

Version 1.0

Bruce Miller

Where Software Meets Recordkeeping

www.rimtech.ca

For Report Updates go to http://www.rimtech.ca

OpenText has provided information and reviewed this report for factual accuracy.

Contents

Introduction

In this report we examine how OpenText solutions[1] integrate with Microsoft SharePoint 2010, via the OpenText product offering known as **OpenText Application Governance and Archiving for Microsoft SharePoint** (**AGA**)™. AGA does many things, however the exclusive focus of this report is to show how it can be deployed in order to achieve recordkeeping compliance – just one of the facets of AGA.

Many organizations with OpenText Content Server 10 have deployed SharePoint 2010. SharePoint 2010 is typically used to create and store documents, and is often used for extensive document collaboration, among other things. In this case there are two repositories in place, each storing documents, and each managing and handling documents in their own uniquely different ways. But OpenText Content Server 10 is compliant with modern corporate recordkeeping requirements – SharePoint 2010 (by itself) is not. This report will show how to use AGA to achieve recordkeeping compliance for documents that originate in SharePoint. This report is therefore not a comprehensive examination of all of AGA's capabilities.

AGA Overview

Content Server and SharePoint are each very sophisticated **ECM** (**E**nterprise **C**ontent **M**anagement) platforms. AGA is a comprehensive integration of the two products that offers five broad capabilities, as shown:

SharePoint Site Provisioning

Site Provisioning allows a SharePoint Administrator to design SharePoint Sites with greater overall structure, and to ensure consistency of this structure among different sites. This has a downstream beneficial effect on recordkeeping, in that records are easier to identify and process.

The SharePoint environment consists of Sites. Sites are places in SharePoint where people go to work on projects or places where information of a common purpose is organized and stored. Not unlike a website on the worldwide web, a SharePoint site will often consist of documents and folders, and the associated metadata required to manage this information. Each site can be configured with the particular users, security, tools and information needed for the goals of that particular site. Sites are the lifeblood of SharePoint, and it is not unusual for large organizations to have many thousands of sites, with no two alike.

Creating new sites, and managing sites so that they have some consistency is an enormous administrative effort. AGA's **Site Provisioning** capability greatly reduces that effort, by providing a means whereby a SharePoint administrator can define sites and replicate them to different locations throughout SharePoint. All the site features, including site structure (organization), folders, security, metadata, etc. can be easily deployed or replicated (provisioned) throughout SharePoint. Even Content Server capabilities such as archiving, recordkeeping policies, and search can be a part of provisioned SharePoint sites. This greatly reduces the effort of initial site creation, as well as ongoing maintenance of sites in SharePoint.

Cross-Platform Document Retrieval

With AGA in place, you can search across both Content Server and SharePoint repositories to find needed documents. A SharePoint user can run a single query, and both repositories will be seamlessly searched. Security control of both repositories is applied in accordance with the SharePoint users' privileges. Search results are uniformly presented in a consistent result list.

[1] OpenText Solutions refers to Content Server, Records Management, Storage Services, and Enterprise Library. We will refer to OpenText solutions as "**Content Server**".

And the SharePoint user can retrieve documents from Content Server into the SharePoint workspace.

SharePoint Storage Optimization

SharePoint stores its documents inside its underlying database. This means the database grows extremely large as the quantity of stored documents increases. This can negatively affect the performance of the system when searching and retrieving documents. It can also make the underlying IT infrastructure more difficult to maintain. With AGA, you can re-locate stored SharePoint documents outside the SharePoint database, transparently to SharePoint. This preserves maximum SharePoint performance as its document load rises over time. The documents are stored outside the database in a high-performance data store managed by OpenText Archive Server.

Recordkeeping

OpenText Content Server offers full compliance with industry's two leading recordkeeping standards – **US DoD 5015.2** and **VERS**. **US DoD 5015.2** is a US government standard and industry-wide de facto standard for electronic recordkeeping. Most ECM vendors comply with US DoD 5015.2. **VERS** is an Australian standard which goes well beyond US DoD 5015.2. Fewer ECM vendors are compliant with VERS. SharePoint is compliant with neither of these important standards. With AGA, a SharePoint site can gain the full benefit of recordkeeping compliance for documents archived from SharePoint to Content Server. Content Server also delivers a full **PRM** (**P**hysical **R**ecords **M**anagement) capability that can be accessed within SharePoint. With AGA in place, true "federated records management" can be achieved, in that recordkeeping compliance can be achieved across both the Content Server and SharePoint platforms[2].

Multi-Tier Storage

Multi-tiered storage refers to the ability of a system to store data on a storage medium suitable for its value and retrieval need. Newer data in need of speedy retrieval is stored on closer, faster (and usually more costly) media. Older data with a reduced need for fast retrieval is stored on slower, more remote (and usually less costly) media. Content Server, through its Enterprise Library Server facility, automatically manages the storage across multi-tiered storage, to reduce overall storage cost while maximizing retrieval performance. SharePoint does not have any such multi-tier storage capability. AGA however can bring the full value of multi-tier storage to SharePoint.

This report will deal only with the recordkeeping capability.

IT vs RM

AGA is essentially a bridge between two complex and different products. Simply installing it will not deliver recordkeeping compliance. It only gives you the *means* of achieving compliance. You (the organization) will have to create, and then maintain an environment that supports good recordkeeping. You'll need to configure AGA accordingly to support that environment. If you have an inadequate recordkeeping environment in SharePoint, and/or in Content Server, AGA cannot be expected to help you achieve compliance. To achieve success, you should:

> **Organize SharePoint**. Organize SharePoint documents, folders, metadata and security such that documents are more readily able to be archived as records at some later point.

[2] SharePoint documents have to be archived to Content Server in order to have recordkeeping applied to them.

Have strong Content Server recordkeeping. Have a well-defined and highly consistent records environment in Content Server to receive the archived records.

Motivate end users. Motivate end users, particularly SharePoint users, to participate in the extra effort that recordkeeping requires.

Build strong archive processes. Define and build AGA archive processes that are simple to use, consistently implemented across different user scenarios, and achieve a high degree of classification accuracy (more on this later).

Remember that AGA does a lot more than just recordkeeping. It's entirely possible it will be used with little or possibly no recordkeeping taking place, which can be perfectly normal. To a RIM professional, AGA will appear rather technical. This is by necessity, for two reasons:

1. AGA can do a lot more than just recordkeeping.
2. It has to bridge two systems that were obviously never designed to work together. It therefore has to overcome a myriad of technical differences between the two products.

AGA very much comes across as an IT product, not as a recordkeeping product. That's because it is. AGA has a great deal of technical heavy lifting to do just to archive a document in the first place. Within this process, recordkeeping is but an element – just a part of the process. But there's more to recordkeeping than just a technical component of archiving. You need to have a good recordkeeping environment in place that supports recordkeeping across the full document lifecycle, from initial creation in SharePoint, to archiving in Content Server as a managed record, and to final destruction in Content Server. The following actors will all have an important role to play:

IT Both Content Server and SharePoint system administrators will need to be involved. They will have to configure their systems to best suit archiving of records. Content Server and SharePoint should also be configured to maximize consistencies between two systems, to make archiving via AGA simpler and smoother. The Content Server administrator can also expect some degree of automatic creation of important elements such as user accounts, folders, and security settings. Not to mention that there will be a steady and substantial flow of incoming records into Content Server, which has to be carefully monitored for recordkeeping compliance.

RIM The RIM practitioner will have to influence how each system is configured, to achieve as much consistency as possible between the two. The minimum metadata required for recordkeeping should be defined and enshrined in each system. And all the recordkeeping mechanics, such as which SharePoint documents to declare as records, and when/how to best move them to Content Server, will have to be planned in advance, then AGA can be configured accordingly.

End Users AGA provides many ways to hide the technical mechanics of archiving from end users, by automating required processes where possible. A record presents different challenges from documents (non-records) however. When archiving a record, there is likely more work to be done by the contributing end user than would be the case for a non-record.. There are a minimum of (4) crucial metadata fields that must be completed for every document being treated as a record. Some of these fields cannot always be hidden, or taken care of by SharePoint (which we assume to be the source of the record). In a recordkeeping compliance setting, users will have to participate in the recordkeeping process at some point. You'll see that this participation can be measured as successful or unsuccessful. The organization will have to secure the (successful) cooperation and participation of the end users in the recordkeeping processes you define.

Records and AGA

With these two different systems deployed, how can the organization achieve full recordkeeping compliance of all documents consistently? There are two possible choices:

Apply recordkeeping independently to each system. Managing records within Content Server is straightforward, as it is an out-of-the-box capability. With SharePoint however it is necessary to custom-modify the product to meet recordkeeping requirements, or more likely, to use a specialized third party plug-in product to fill the recordkeeping gaps. The corporate retention schedule would have to be duplicated within both Content Server and SharePoint, and records would have to be separately managed and administered in two completely different platforms. This means a great deal of duplicated effort, not to mention the administrative overhead of managing records in two different platforms, each with a different approach, features and capabilities.

Move documents from SharePoint to Content Server. When ready, move documents from the non-compliant SharePoint platform (some might consider this to be a collaboration platform) to the compliant Content Server platform (some might consider this to be a compliance platform), where full recordkeeping control can be applied to them. Use SharePoint for document creation and collaboration, but not for long-term document storage or for recordkeeping.

AGA facilitates the second scenario, i.e. move documents from SharePoint to Content Server. We would do this to achieve recordkeeping compliance, i.e. we now have a means of achieving compliance of the documents we create in SharePoint, by archiving them to Content Server via AGA. AGA does not in and of itself deliver recordkeeping. If you move documents from SharePoint to Content Server, you now have the *means* of applying recordkeeping control to the documents you moved to Content Server. AGA does not apply recordkeeping control to any documents that are stored within SharePoint, nor does Content Server itself. In other words, to apply records control to documents in SharePoint, you need to *move* them from SharePoint to Content Server, and to move them you'll require AGA.

Using AGA does not mean you *must* apply recordkeeping. With AGA you can elect to simply move documents from SharePoint to Content Server without any regard whatsoever for recordkeeping. In this scenario, you would be using Content Server as a document archive only (as discussed below), without recordkeeping compliance. Recordkeeping is applied to any given document only to the extent that you choose to do so – AGA will only enforce recordkeeping if you configure it to do so. This is an extremely important point to note – AGA can get your documents moved from SharePoint to Content Server, where you can apply recordkeeping to them, but the recordkeeping is up to you. You'll see later in this report how AGA can facilitate recordkeeping during the move operation, but you have to plan and enforce the recordkeeping as part of the move operation.

OpenText uses the word "Archive" to mean "Move". The implication here is that documents start life in SharePoint, but at some point end up "archived" in Content Server, where formal recordkeeping control can be applied to them. **Be especially careful to note that "Archive" does not necessarily mean "archived as a record", it simply means the document has been moved to Content Server.** It may or may not be a record. And it may or may not be *managed* as a record (declared as a record), even if the document is in fact a record. Again, the extent of recordkeeping that occurs is determined by you, not AGA itself.

AGA can move (archive) a range of content from SharePoint to Content Server, from a single document all the way up to a complete SharePoint site containing documents, lists, emails, folders, and more. If we refer to all that content as "documents", it is up to you to specify to AGA which documents to manage as records, when, and how.

Abbreviations and Nomenclature

The following abbreviations will be used throughout this report:

AGA	**A**pplication **G**overnance & **A**rchiving, the supplied by OpenText to enable the integration of Content Server 10 and SharePoint 2010[3].
Archive	To **move** a document from SharePoint to Content Server. A copy or link of the original document may be optionally left behind in SharePoint.

[3] AGA can also be used with OpenText LiveLink 9.71.

Move and Archive will be used interchangeably throughout this report, and will mean the same.

Content Server **C**ontent **Server 10**. Refers to the overall **E**nterprise **C**ontent **M**anagement (ECM) offering from OpenText.

EL **Enterprise Library.** Often used to describe Content Server 10. For the purpose of this report, EL will be used to refer to the *repository* of Content Server 10, i.e. all documents stored in Content Server reside somewhere in an Enterprise Library. "*Storing* into Content Server" and "*Storing into* EL" in practice mean the same thing.

SharePoint SharePoint 2010

The term ***Document*** will refer to any form of recorded information, regardless of form or origin. A document means any and all of the following:

- A revisable text document (e.g. Microsoft Word)
- A fixed image (e.g. scanned TIFF, PDF, etc.)
- A revisable vector drawing (e.g. AutoCad)
- Created at outside the organization but stored within and utilized by the organization (.e.g downloaded from the web)
- An email, with or without attachments
- Multi-media files (e.g. video, audio)
- Social media (e.g. text messages, tweets, etc.)

The term **Record** will be used to refer to a document (per above definition) that is ***managed*** as a record, per the following:

- Is identified as a document of business value
- Is subject to a formal retention and disposition schedule
- Is locked down, i.e. cannot be deleted or revised once declared as a record.
- Will be destroyed only via a formal disposition process, run by the organization's designated **R**ecords and **I**nformation **M**anagement (RIM) professional.

Important AGA Tenets to Know

On the surface, the principle that underlies AGA is quite simple – move documents from SharePoint to Content Server. In reality however it's anything but simple. Each of the two platforms is a sophisticated content storage and management system that asserts quite a bit of control over documents. Each assumes they have exclusive control of every document stored within them. Each system manages the document in its own unique manner, different from the other. As a result, there are many technical "gotchas", or differences that have to be taken into account when moving a document between the two systems. Many square pegs have to be inserted into many round holes, so to speak.

AGA has to take each of these differences into account, resulting in some rather significant complexities. AGA is a technically sophisticated piece of middleware that inserts itself between Content Server and SharePoint, by inserting new capabilities directly into SharePoint, and extending the Content Server platform itself.

Following are a number of important tenets about AGA you should know before getting into the nuts and bolts of the product:

Entirely OpenText Product
No SharePoint Recordkeeping
Declaration Timing
Definition of Archive
Document Structure
Document Type
Document Links
Extensibility

Governance

Entirely OpenText Product

This product was developed entirely by OpenText, and not by Microsoft. The idea behind it is to make the use of Content Server as seamless as possible inside the SharePoint environment, and to provide a means of moving documents out of SharePoint into Content Server for long-term storage in a records-compliant environment. All the design and functionality are from the perspective of OpenText, i.e. Content Server is where the documents ultimately have to be stored in order to apply recordkeeping to them. The product design, feature/function, documentation, and support are provided exclusively by OpenText.

Limited SharePoint Recordkeeping

SharePoint has very limited recordkeeping capabilities, such as the ability to declare a document to be a record, and a limited ability to classify a document against a formal retention schedule. Furthermore, it has a designated site for the storage of records called the **Records Centre**. These capabilities however fall far short of the requirements of the defacto industry standard US DoD 5015.2, or the ICA Module 2 standard, adopted by the Government of Canada (among others). Content Server on the other hand is fully DoD 5015.2-compliant, nearly compliant with the ICA Module 2 requirements, and also meets the additional requirements of the Canadian Government. The underlying operating assumption with AGA is that Content Server will be used to manage records, not SharePoint. And in order for Content Server to apply compliant recordkeeping control, the documents must be stored in Content Server. Content Server does not apply recordkeeping to documents in SharePoint. To achieve recordkeeping compliance with Content Server and SharePoint, you must therefore:

1) Move documents from SharePoint to Content Server, via AGA
2) Use Content Server for recordkeeping, not SharePoint

If SharePoint has been configured for any degree of recordkeeping (such as metadata fields related to records), AGA can easily be configured to take full advantage of that. If you have applied any recordkeeping to documents in SharePoint, it will be far easier to carry that through to Content Server when you later move (archive) them to Content Server, by configuring AGA appropriately. If you happen to be declaring and classifying records in SharePoint before moving them to Content Server, you'll have to configure AGA to recognize the declaration state (a metadata field attribute), and the classification (a metadata field value), and ensure they line up with the corresponding values in Content Server.

OpenText makes the case that with AGA, it is possible to some extent to apply recordkeeping to documents in SharePoint by presenting the retention schedule (from Content Server) to the SharePoint user interface so that it can be used to declare and classify documents in SharePoint. Given that the SharePoint user can also access records in Content Server, SharePoint would in theory be a single point of access for records in both SharePoint and Content Server. This would require the construction of a custom declaration and classification process in SharePoint that uses both SharePoint and AGA components. This custom development is beyond the scope of this report.

For the purpose of this report, RIMtech assumes you'll be using SharePoint for document creation and collaboration, and only applying recordkeeping in Content Server. We further assume you have not yet implemented a sound recordkeeping process in SharePoint. This means you start with nothing and have to use AGA to declare the records as part of archiving. If however you are storing usable records-related metadata in SharePoint, you'll be further along the process and will need much less configuration of AGA. This report will propose specific SharePoint configuration suggestions to facilitate easier AGA implementation.

Assuming SharePoint documents are in no way declared as records, you will have to deal with the important issue of *timing*, as RM best practices encourage you to declare a document to be a record as early in its lifecycle as possible.

Declaration Timing

It's human nature for users to want to hold onto "their" documents as long as possible. Put another way, users may be reluctant to move their documents to another location. Presumably, SharePoint (and not Content Server) was used for good reason to create the documents in the first place. From their point of view, archiving from SharePoint to Content Server may appear tantamount to "giving up" their documents.

Their goal may be to keep them in SharePoint as long as possible and defer archiving to Content Server as late as possible.

Good RIM practice however states that a document of business value should be managed as a record (declared) as early in its lifecycle as possible, for the following reasons:

1) It is safely locked down as a record and cannot be altered
2) It is subject to the retention schedule
3) It is logically grouped with records of the same topic within the hierarchy of the retention schedule
4) It can be located by RIM personnel for legal holds, FOI requests, etc.

With AGA, documents that have been moved to Content Server are still fully retrievable within SharePoint. To a SharePoint user they are “still there”, even after archiving. This will greatly help to overcome user resistance to archiving documents earlier in their lifecycle.

AGA cannot help you decide when would be the right time to archive a document (or set of documents) into Content Server as a record. That will be a practical business decision. Some would say a document should be archived “when it is no longer needed”. The following are some considerations to take into account:

1) A document of business value that is not declared as a record is an **unmanaged record**. It is not under corporate control. The organization cannot be as accountable for it as compared to a declared record. It can be deleted, renamed, or edited. It may not be as easily searchable. Access control (security) may not be appropriate. The sooner it is declared to be a record, the better.
2) There may be pressure to reduce storage space in SharePoint.
3) Some exceptional documents may have a need for a higher accountability. If for any reason a particular document, or set of documents, is subject to higher than usual scrutiny, there may be pressure to move them to Content Server sooner, in order to apply recordkeeping accountability to them.

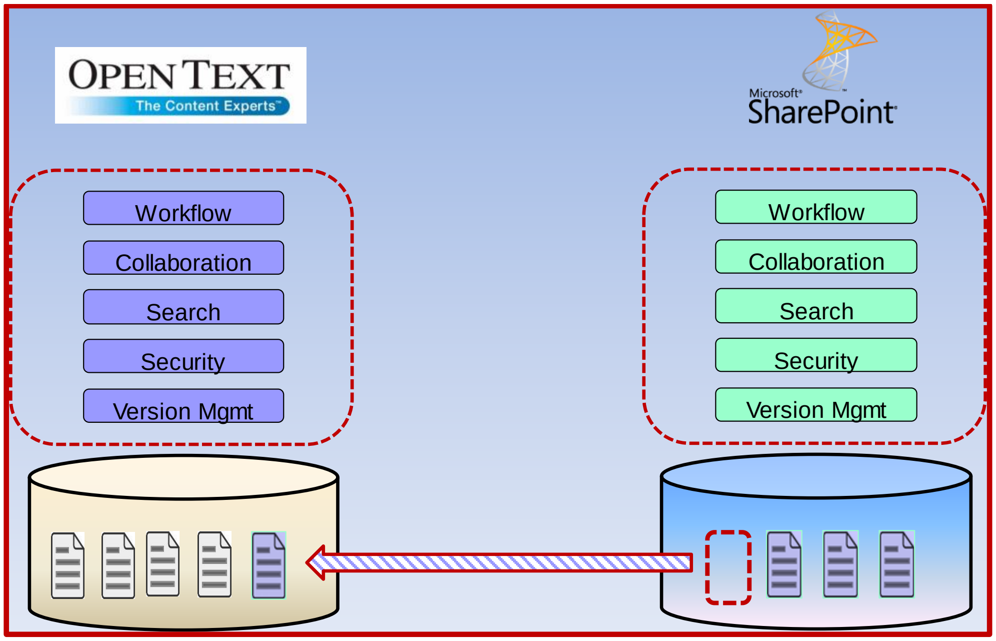
For good information governance, the goal therefore should be to archive them as records as soon as possible. The longer it waits, the more difficult it can be to declare it with the correct classification (dealt with later in this report). And because it can be retrieved even after archiving, there is little legitimate reason to defer it.

Definition of Archive

AGA allows you to “**archive**” documents from SharePoint to Content Server. The term “archive” really means **move**. AGA is a sophisticated tool to move documents from SharePoint to Content Server. There are many aspects of move that can be configured to suit your needs:

1) The original SharePoint document can be left behind (a *copy*) or deleted (a *move*).
2) If the original is deleted, a hyperlink[4] (link) can optionally be left behind as a substitute for the original document.
3) The document can be moved to Content Server as a declared record, or not.

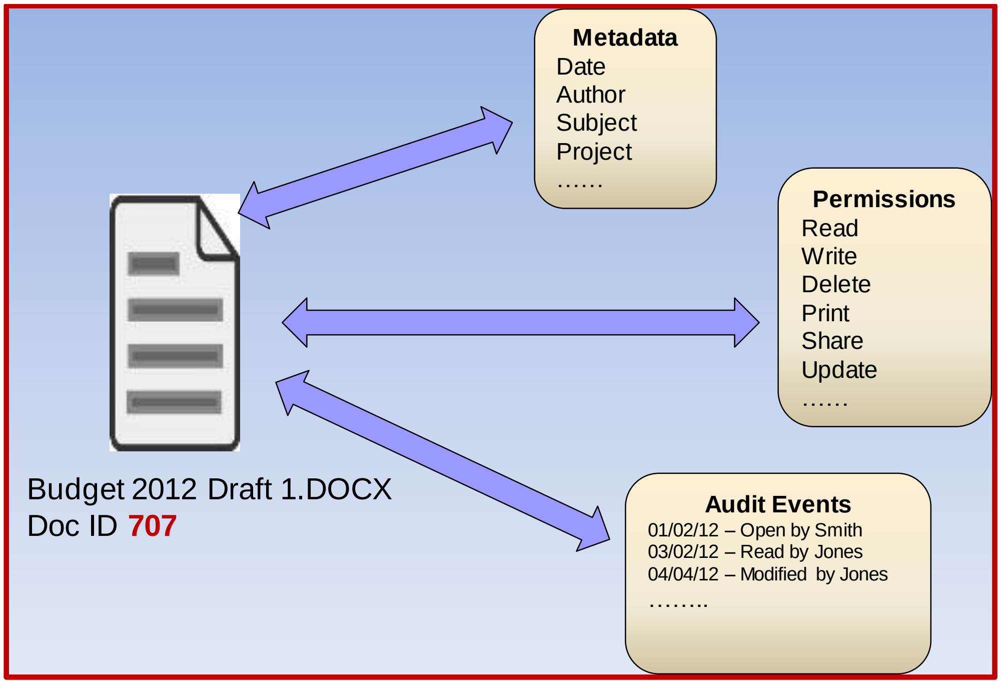
[4] Often shortened to simply “Link”, as used in the OpenText AGA documentation.



In a recordkeeping setting, “archive” has an implied meaning that the documents being archived are records, and that the archives consist only of records. Put another way, the classic meaning of the word assumes that an archive contains records (i.e. documents that are managed as records). This simply does not apply to AGA. AGA is a tool to move electronic documents from one system to another. Whether the moved documents are records or not is up to you – they can all be records, none of them can be records, or anything in between. AGA cannot control whether the source documents are managed as records, or even if the documents are records after being moved to Content Server. If your source (SharePoint) documents are not declared as records, and if you do not declare them as records as part of the archiving process, you are simply moving undeclared documents from SharePoint to Content Server, and not achieving any recordkeeping goals. There is some “middle ground” however. If you are utilizing some document structure in SharePoint, such as metadata or folder location to indicate either *record status* (record or not) or *classification* (which retention rule to apply to a document), AGA can take full advantage of that structure during archiving. AGA can act on this structure to know which documents are records, and how to classify them. The greater the recordkeeping structure you utilize in SharePoint, the simpler it will be to configure AGA to archive the target SharePoint documents into Content Server as records.

Document Structure

Content Server and SharePoint are both **E**nterprise **C**ontent **M**anagement (**ECM**) systems. In any ECM system, a document is more than just the document itself. Each document has several elements, as shown in the diagram below:



Document elements are:

Document ID. Each document is assigned a unique Document ID for life, to distinguish it from all other documents in the system. The original filename (e.g. Budget 2012 Draft 1) is stored as a metadata field that can be searched, however the document is now identified by its ID, not the original filename.

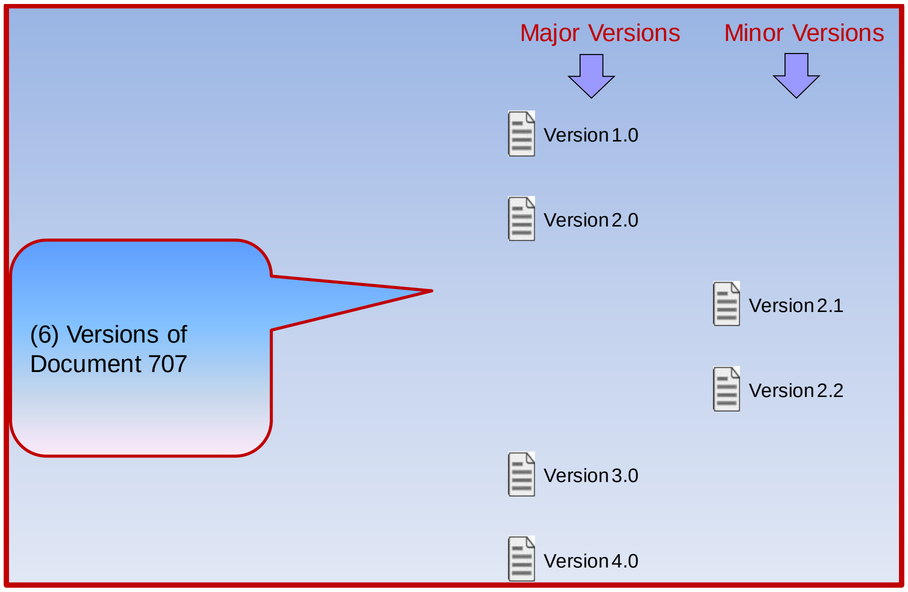
Metadata. Fields of descriptive data are assigned to the document, such as *Author, Date, Subject, Project Name*, etc. Some metadata fields are assigned internally by the system and are not normally seen by the user. Metadata fields can vary among different documents. Metadata can be searched by users.

Permissions. Defines what users can do to a particular document, such as Read it, Edit it, Delete it, etc.

Audit Events. A detailed recorded history of activities carried out on a document. For example, who read the document and when, who edited it and when, etc.

To understand AGA and how to utilize it properly, you need to understand each of these elements. This is because each time you archive a document from SharePoint to Content Server, you'll have to decide what to do with each of these elements.

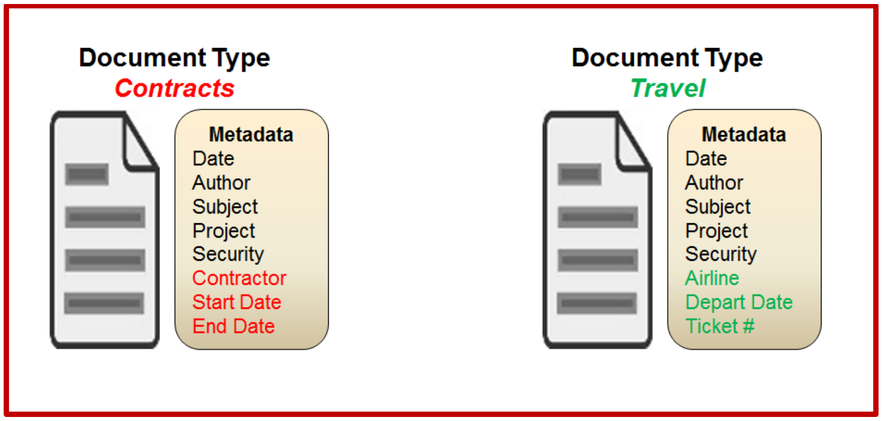
A document can also have multiple **versions**, as shown in the diagram below:



Again, you'll have to deal with versions when archiving documents to Content Server. Which versions do you wish to archive? Which versions should be declared as records? AGA will even allow you to specify that once you've archived a given version of a specified SharePoint document, all future versions created in SharePoint will be automatically archived to Content Server.

Document Type

There is an important element of ECM systems that you'll need to understand in order to deploy AGA – **Document Type**. The document type determines which metadata fields a document will have. A document type therefore specifies the metadata fields that all documents of that particular type will have. The diagram below shows document types *Contracts* and *Travel*, each with a specified set of metadata fields.



Notice in the example that each document type has a set of five fields common to each type, and three fields unique to each type. The common plus the unique fields together form the metadata for a given document type. All documents of type *Contracts* for instance will have the 8 fields shown.

Documents of a given document type therefore all have the same metadata fields in common with each other. All ECM systems, SharePoint and Content Server included, require that each document must have a document type defined. Put another way, all incoming documents must be a member of some document type. The document type is a mandatory, internal, system-defined attribute of all documents in both Content Server and SharePoint.

As you archive documents from SharePoint to Content Server, you will always have to align the metadata of the SharePoint document with that of the Content Server document, which in each system is driven by the document type[5] of that system.

Extensibility

The documents in SharePoint can, and likely will be, organized and structured very differently from those in Content Server. The metadata fields, security permissions, audit data, and versions will be quite different in each system. The manner and timing of records declaration can be vastly different in each, or non-existent in SharePoint. Yet you will have to account for all these differences when you move the SharePoint document to Content Server as a declared record.

Furthermore, *what* documents you archive, *when* you archive them, and *how* you archive them will be highly situational. Different business circumstances will drive these three factors. You can safely assume that you'll sometimes be presented with a unique situation that falls outside the normal situations you planned for, and for which you configured AGA.

You should be able to use OpenText AGA for just about any conceivable situation you encounter, thanks to the following two capabilities:

> **Configurability**. AGA is very highly configurable. OpenText has carefully and exhaustively accounted for most all possible combinations of document structure and methods of archiving. It allows you to explicitly specify the *what*, the *when*, and the *how* for each archiving operation. This means there's a lot of detail you'll have to specify for each archive operation, in order to account for all the possible combinations. But this is an understandable price to pay for configurability, which is critical for a product that has to deal with so many possible circumstances.
>
> **Extensibility**. Extensibility refers to the ability to customize a product to do something it does not do out of the box. It means that if you are presented with a set of circumstances that you cannot deal with through everyday configuration, you can customize the product as needed to deal with the circumstances. For example, you may wish to archive a SharePoint site that has thousands of documents within a folder structure that groups documents of a similar subject in a particular manner. You can design a one-time archiving process that classifies them correctly, declares, then archives them in a single operation, based on their unique grouping. To do this you'd have to add some custom code to AGA. AGA provides a number of ways by which you can extend (customize) the product. The more ways to customize the better, and AGA is highly extensible, by providing many different way to customize it.

Document Links

When you archive a document from SharePoint to Content Server, you have the option of deleting the original SharePoint document, and leaving a link behind in place of the original SharePoint document. This has the advantage that the SharePoint user can search for and find the document that was archived, even though it is no longer actually stored in SharePoint. In the SharePoint context menu (seen by clicking over a document's drop-down) the user can display its usual actions such as "*Open*", "*Send To*", "*Information*", etc. All of these context menu items will work for a linked document just as if it were the actual document.

As you plan your archiving strategy for AGA, you'll have to decide when to leave links behind. You need to be aware of the following limitations of a link:

1) The integration with the **Microsoft Office** suite is different. There is no property information panel. There is no task pane in Microsoft Office with information about the SharePoint site or library where the link is stored.
2) SharePoint **Workflows** cannot be used with links. This means a custom workflow cannot include the linked document.
3) Microsoft **SharePoint Workspace** and **Outlook** do not work with links if you should go offline with the document. You cannot take the linked document offline and work on it remotely, and return it to SharePoint.

[5] SharePoint refers to its Document Type as ***Content Type***. For the purpose of this report, we will assume they are equivalent.

Should these link restrictions be too limiting, you'll see later in this report that there are alternatives to linking.

Governance

How do you achieve the "governance" in AGA's name? Governance is a broad term that can mean a lot of different things, but the implication is that using AGA will increase your degree of information governance (or accountability). We will define good governance as including a minimum of the following:

> **Good recordkeeping**. Identifying those documents that should be managed as records, and applying recordkeeping to them accordingly. Keeping and destroying records in accordance with the official approved retention schedule. Measuring and confirming that documents are consistently and correctly classified. Destroying documents properly and accountably via formal disposition process, governed by the retention schedule.
>
> **Supporting Retrieval**. Providing timely, effective retrieval against mandated special requests such as Freedom of Information, court-ordered, or litigation requests. Providing retrieval of records and non-records, of all documents in all locations.
>
> **Applying Legal Holds**. Protecting specified documents from deletion in the event of court-ordered legal holds.
>
> **Minimizing Holdings**. Reducing or eliminating redundant, transient, or otherwise un-needed documents, to reduce overall volume.

How does AGA help you achieve good governance? Good governance comes not from any product, but by your application of the product, along with the underlying corporate policies and procedures you put in place to back it. Content Server is a tool that, if used appropriately, can support good governance practices. Compared to Content Server, SharePoint does not have the compliant recordkeeping capabilities to support good governance. Presumably OpenText is implying that by moving documents from SharePoint, a platform that does not support most of these good governance practices, to Content Server, which has features that support good governance, you will more readily be able to deliver good governance. RIMtech has no argument with this position, but remember that Content Server and AGA will not themselves deliver good governance. The governance comes from your corporate policies and procedures, and how you follow these using Content Server. Good governance starts with good recordkeeping, which AGA does not deliver. AGA will only *facilitate* the recordkeeping you are already practicing in SharePoint and/or Content Server.

Recordkeeping Principles and Objectives

At the highest level, there are (3) fundamental electronic recordkeeping practices or elements that must be carefully implemented and judiciously tracked in any successful electronic recordkeeping system, including your combination of Content Server and SharePoint. Each element is independently quantifiable and therefore measurable. Each is an end user practice, and each is manifested either in specific features of the software, or the manner in which the software is used. They are:

Qualification
Declaration
Classification

Qualification

Qualification refers to the process by which a user determines if a document is a record or not. A document would typically qualify as a record if it has business value, records a decision, or some other criteria appropriate to the business. RIMtech defines the **Qualification Rate** as the percentage of a group of documents that qualify as a record. For example, about 5% of a day's worth of email might qualify as a record, whereas 95% of a day's worth of written MS Word documents might qualify.

Declaration

Declaration is a state of a document. A document declared as a record is locked down and cannot be deleted except via the official disposition process, and has been assigned an appropriate retention rule and access controls. Declaration also refers to the physical act of following through on the qualification of a document, i.e. submitting into a host system as a record, where it is managed as a record. Theoretically, 100% of all documents that are qualified as records should be declared as records. Hence, RIMtech defines the **Declaration Rate** as the percentage of qualified records that have been declared as records in the target system in which they are stored. For instance, a declaration rate of 95% means that 95% of those documents that *should* be declared (qualified as records) *are* declared.

A document may be a record, but has not been declared as such. If not declared, it is considered an **unmanaged record**, as the system software cannot apply recordkeeping to it. It is, in essence, a “lost record”.

Classification

Classification is the process whereby a record has been assigned an appropriate retention rule from the official retention schedule. It will either be the right (appropriate) rule, or the wrong rule. RIMtech defines the **Classification Accuracy Rate** as the percentage of a group of records that are known to have the correct retention rule assigned. The classification accuracy rate must be above a minimum threshold, typically in the range of 80-90%, in order for Disposition to take place.

Before we examine the details of using AGA to achieve these recordkeeping goals, it is important to understand the following general electronic recordkeeping principles:

Document Status
Case vs. Admin
Classification Accuracy
Declaration and Classification
Declaration Timing
Document Duplication
Retention Schedule

Document Status

In modern electronic recordkeeping systems a document is not simply a record (or not). RIMtech defines **document status** as denoting whether a document is a *record*, a *work in progress*, or a *reference document*. Instead of simply asking “*What is a record?*” we instead ask users to choose one of the following three options that best applies to the document they are placing into a system such as SharePoint:

Record The user wants this document preserved as a record. Put another way, they user does not want it to be auto-deleted! It is “finished” and in play (it has been circulated among one or more user). It will be placed into SharePoint as a declared record.

Work In Progress Placed into SharePoint, but not yet complete. Safe to assume that because it is something being worked on, it is a business record that eventually will be declared as a record.

Reference Not a record, but a document that is needed to conduct work. Probably a document received from outside (e.g. downloaded from the web).

Below are some of the practical, real-world guidance we can consider giving users to help them make the appropriate choice. This guidance is general only, and would have to be tailored to a particular organization:

Document Status	Choose This When
Record	• You do not want this document to be automatically deleted • You (or others) will need this in the future • This appears to be a record • Not transitory/temporary in nature
Work In Progress	• The document is not yet finished • You are not ready for others to see it • It is "work" -- everyday business
Reference	• You (or others) need this to do your work • OK for this to be deleted in 2 years (or whatever the corporate auto-delete period may be)

Note that these choices deal with the *need* for the document, from the user's perspective. We do not ask them to deal with the esoteric meaning of records versus non-records. Our goal is to make all users feel comfortable selecting record status as close to 100% of the time as possible.

What to do with the document, based on the Status selection? Again that will vary widely among different organizations, but below is a typical general-purpose answer that can act as a basis for your particular circumstances:

Document Status	Store in SharePoint As	Retention
Record	Declared Record, classified appropriately	Determined by CATEGORY
In-Progress	Classified, declared as non-record Will be declared in the future	1. Notify user after 6 months (if not declared) 2. Auto-convert after 1 year
Reference	Non-record, not classified	Auto-delete after 2 years

Case vs Admin

The principle of a case versus admin record type is a hugely important principle in recordkeeping that you'll need to understand. It drives the mechanics of recordkeeping in any system, including SharePoint and Content Server. SharePoint does not natively understand or support the principle of case records and their proper treatment. This has a profound effect on our ability to use SharePoint`s native capabilities to manage records. However if we do not correctly implement this principle, we will fail to achieve proper recordkeeping in Content Server, let alone in SharePoint. Let's first examine the theory behind case vs. admin.

Also known as a **File Plan**, a sample Retention Schedule fragment for a hypothetical organization is show in Appendix 2. It is a three-level hierarchical classification scheme based on business activities. Records are "classified" against this scheme by assigning a rule from this schedule to the document. For example a document entitled "*Insurance Policy 2009*", bearing filename Insurance Policy 2009 Final Approved.docx should be classified against Administration/Insurance, whereupon it will be destroyed after 7 years. Each category in the file plan has an identifying number, composed of numbers and/or letters representing that category, and each level separated by a dash. For instance, the category **Insurance** has file number 10-20, 10 representing the Primary ADMINISTRATION, and 20 representing the secondary level category INSURANCE.

The retention rule for Safe-Tee Security is "*2 years after contract end, then Destroy*", the same rule for all 100 contractors. Because each contract bears a different contract end date, the rule must be applied to each contractor separately. All records related to Safe-Tee Security must be treated as one group of records, separate from all other contracts.

It is important to note that all categories fall into one of the following two “file types[6]”:

Admin Ongoing admin or business activity that never ends, i.e. “Budgeting”, “Travel”, etc.

Case Person, Place, Event, Thing, Project. e.g. “Peggy’s Cove Plane Crash”, “Taxpayer Smith, Fred”, Employee “Jenkins, Ruth”. A Case has a defined beginning and end. Disposition is triggered by an event date.

Typically 50-75% or more of all records are members of *case* files. “File” in this case is the old-fashioned term used for a collection of records, however to avoid confusion with the modern computer term “file”, where possible we will refer to a *category* instead of a *file*. We will refer to a computer file (i.e. something stored on a disk) as a *document*, instead of a file. Disposition is triggered very differently on each category type:

Admin ***Document age***. Documents are treated on an individual basis, per-document basis. Only documents where age > retention period are qualified. This means that older records can be destroyed and younger ones remain behind.

Case ***Event date***. All documents in this category are qualified for disposition or none at all. They are always treated as a group – never separated.

In our file plan fragment, Admin categories are shown in blue – there are 5 (*Travel, Insurance*, *Accounts Receivable*, *Accounts Payable*, *Accounts Receivable*, and *Budgeting*). These categories bear a simple time period expiration as their retention rule. Note that all remaining categories are case categories, wherein the retention rules are triggered by an event date (which is usually unknown at creation time).

It is important to understand that more often than not, a declared record will be classified against a case category, which will have downstream implications on the mechanics of AGA archiving.

Classification Accuracy

Any electronic recordkeeping system is only as good as the accuracy of its classification. This is because if the classification accuracy drops below a certain level, the records manager will not have sufficient confidence to run Disposition. Most organizations want 100% accuracy, however that is just not possible, due to the human element – some users do not care, some make mistakes, etc. One thing is absolutely certain – every records manager will draw a line below which they will simply not feel comfortable running disposition. In our view, if the classification accuracy rate drops below this threshold, the project has failed, as records disposition cannot be conducted.

> **BEST PRACTICE**
>
> Measure ***Classification Accuracy***, by department, each day. Measure ***Qualification Rate*** and D***eclaration Rate*** at least once per month. Take immediate action if either number falls out of range.

Classification accuracy of a large (millions) document collection is easy to measure, via statistical sampling. Suppose 1,000 records were declared on a given day. Sample 10 from each department (as an example) at random, and if 7 of the 10 (or 70 of the 100, assuming 10 departments) are correctly classified, we can safely extrapolate that to mean 70% of the 1,000 records are accurately classified. Put another way, we have 300 misclassified documents on that day that will ultimately be destroyed too early or too late. Whatever your acceptable accuracy threshold may be, you must measure it each and every day.

If users consistently **qualify** appropriately, **declare** the documents as records, and properly **classify** them, we can proceed with ***Disposition***, where we will destroy records in accordance with the official

[6] “File” is the traditional language used to define a category within the file plan

approved retention schedule. Disposition is the "end game", where everything comes together in a way that allows us to accurately destroy massive quantities of recorded business records with full legal confidence. If declared records are not classified correctly, too many would be destroyed too early or too late, which could prevent any disposition from taking place at all.

Each of these three factors (*qualify*, *declare*, *classification*) is in fact a **critical project measurement metric**[7]. ***Everything*** in any electronic recordkeeping endeavour contributes in some way, however indirectly, to one of these three factors. Each factor can, and must, be measured against an acceptable threshold. Any organization can readily agree upon a minimum acceptable threshold for each of these three factors. The following acceptability thresholds are very common, based on RIMtech's experience across many electronic recordkeeping projects:

Qualification	**5%** of all incoming email, **95%** of all documents created and downloaded.
Declaration	**95%** of all business records should be declared into SharePoint.
Classification Accuracy	A minimum **85%** of all declared records must be properly classified

Each of these three factors is easily measured, and **must** be regularly measured, by department, on a regular basis. Further, we need to "hit" all three to succeed. We cannot fail on any of the three, or the entire project will fail. Put another way, if we cannot run Disposition with complete confidence, which is based on the accuracy of the classification, the project simply is not successful.

Declaration and Classification

Whenever we declare a record, it must be *classified*, or assigned a retention rule. For each different method of declaration, there are many different means whereby the user can classify it. This makes for a virtually unlimited number of possible declare and classification permutations. In some declaration methods, the classification is highly implicit – the user may not even be aware they are classifying the document. In other declaration methods the classification is highly explicit, in that the user has to manually tell the system where to classify the document – it is an obvious effort. And there are many variations in between implicit and explicit. In some circumstances we will be able to hide the classification entirely from the user so they are not even aware of it.

No matter how implicit or hidden the classification may be, the following two classification realities are absolutely undeniable:

Classification is driven by document content. The correct classification is driven entirely based on what the document is about – its *subject matter*. It does not matter if the document is a single-sentence email message or a rambling 500-page report – the content determines the classification. And the user has to somehow identify to the host system what that content is about, no matter how oblique a method may be in use.

Classification is entirely user-driven. Something from the user will always drive the classification, no matter how oblique the method. It may be their selection in a CONTRACTOR field "Safe-Tee Security". It may be that they are storing the document in a working folder called ACTIVE MATERIALS, which may have a retention policy linked to "Safe-Tee Security". Perhaps the SharePoint Content Type is TRAVEL REQUEST, which routes the document to a folder that applies the appropriate retention rule.

Classification is therefore only going to be as accurate as the user's ability to tell SharePoint what the document is about. Hence, in implementing good declare methods, we have to make it as easy and reliable as possible for the user to tell us what this document is about. It boils down to the act of choosing the appropriate (matching) category from the file plan. Whatever the document subject matter may be, there has to be a file plan category that matches.

As is the case with all modern electronic recordkeeping solutions, there are three "stages" that every user will have to pass through every time they manually declare a document into SharePoint as a record, otherwise known by RIMtech as the "**Three Cs**":

[7] See details in RIMtech's forthcoming book "*Managing Electronic Records*"

1) They will have to **Categorize** the document as a *Record*, *Work In-Progress*, or *Reference*
2) They will have to **Classify** the document against the file plan (if categorized as a record, or in-progress).
3) They will have to **Complete** a minimum of 4 metadata fields

    a. *Content Type* (can be auto-completed in some cases)
    b. *Admin* (What is this document about?)
    c. *Category* (If a record or in-progress)
    d. True Document Date

These three things take time and effort to complete, and represent a barrier to successful adoption. The criticality of this cannot be overstated – a good declaration capability is a make-it-or-break-it feature for any electronic recordkeeping solution.

In the case of an AGA deployment we are dealing with documents that are created in SharePoint, then there is some unknown time delay, and then the documents are archived to Content Server. It is probably safe to assume that most often, this archiving will be done by a third party. Users are simply unlikely to return to a document long after they completed it, then invest the effort in archiving it to Content Server.

If the documents in SharePoint are already declared as records, and they are correctly classified, then the archiving operation is simply a move of declared records from SharePoint to Content Server. If however the documents in SharePoint are not declared as records, then we must declare them as part of the archiving operation. Unless the archive operation is conducted by the original document author (unlikely, as discussed), then that means we will have to declare and classify the document as part of the archive operation.

Declaration Timing

For this report, we assume Content Server will be used for recordkeeping, and SharePoint is in essence a “source” of records for Content Server. For successful recordkeeping therefore, you must:

1) **Ensure SharePoint documents are Qualified**. Ensure SharePoint documents are appropriately qualified, i.e. the end users, or the AGA archiving process you define, are aware of which documents should be identified as records.
2) **Declare Records During Archiving**. Documents that are qualified as records are declared as records by the time they reach Content Server. Either you declare them inside SharePoint and follow through by archiving them as records in Content Server, or you declare the documents as records at the time you archive them in Content Server.
3) **Achieve Classification Accuracy**. If you establish for example a 90% minimum accuracy threshold, 90 out of every 100 documents you move to Content Server must be verified to be correctly classified.

RIMtech assumes there is no practical way to correct classification errors after they occur, due to the sheer level of effort. Therefore classification accuracy has to be achieved during (or before) declaration.

In practical terms, like any other system, your users are responsible for identifying which SharePoint documents are records (qualification). You must ensure that by the time they are archived into Content Server as records, it has the correct classification (90% of the time). There are therefore three possible scenarios to get your SharePoint documents declared as records into Content Server:

Declare in SharePoint Before Archiving. If the Qualify, Classification, and Declaration have already been done in SharePoint, then you can simply configure AGA to carry it through to Content Server, by preserving the document’s records-related metadata fields.

Declare During Archiving. Identify those SharePoint documents that you need to declare as records, classify them, then move them to Content Server as declared records.

Declare After Archiving. Move from SharePoint to Content Server as non-records. Ignore recordkeeping, and leave it up to a Content Server user or process to later identify, classify, and declare the documents as records within Content Server.

RIMtech is of the opinion that qualification, classification, and declaration are more likely to happen earlier in the document's lifecycle than later, most often when the user either creates the document (initial creation), or when they finish it (complete it and transmit it to others, via email or print). The classification also is more likely to be more accurate earlier in the lifecycle, as the user is engaged in the subject matter of the document, which ultimately determines the classification. The more time elapses after the document has been completed, the less likely the user is still engaged in the matter and willing to invest effort into it. In general, a user remembers less, and cares less, the more time has passed since creating the document.

If you accept that sooner is better, then your goal should be to get the documents declared in SharePoint, long before they are archived into Content Server via AGA. The archiving process then becomes a simple matter of moving an already-declared record from SharePoint to Content Server (assuming the classification is correct). If a given document or set of documents is not declared as a record in SharePoint, you will have to do the following as part of the AGA Archive process:

1) Determine if it is a record (qualify it)
2) Classify it correctly
3) Declare it as a record in Content Server upon moving it.

In essence, you are applying recordkeeping after the fact. This can be extremely challenging, for the following reasons:

1) You are not familiar with the subject matter, which makes qualification and classification difficult
2) You have to assess (read and understand) the target documents in advance of archiving them to Content Server
3) A given set of documents, such as a folder of documents or an entire SharePoint site, can contain documents of many different possible subjects, each of which would have to be independently classified

If you are not concerned with recordkeeping, then none of this matters – AGA is a highly effective tool for moving documents from SharePoint to Content Server. In a recordkeeping setting however, you need to treat the document accordingly, and that means getting them *qualified*, *classified*, and *declared* into Content Server as records. In this report we assume a full recordkeeping environment where most all documents being moved from SharePoint to Content Server are records, and need to be properly managed as such.

AGA cannot help you identify which SharePoint documents are records (qualification). It will however facilitate (assist, not do) classification, and it can see to it that the document is moved to Content Server as a declared record. It is up to you to qualify the SharePoint documents, classify them, and tell AGA to declare them as records.

Obviously, the more of this that can be done in SharePoint the better. If all records are declared, and properly classified in SharePoint, archiving via AGA becomes a very simple exercise of physically moving the documents to Content Server.

Document Duplication

Recordkeeping best practices are unequivocal when it comes to copies of records – copies are undesirable and must be avoided. Duplicates of records constitute an unmanaged instance of a managed record, often cannot be accounted for, diminish the trustworthiness of a record, are exposed to unexpected discovery, and can "compete" with the official record.

AGA is all about moving (archiving) documents from SharePoint to Content Server. It is quite flexible, and allows you to archive documents as a **move** (delete original in SharePoint) or as a **copy** (keep original in SharePoint). Keeping copies behind is bad recordkeeping practice, and is to be avoided except in exceptional circumstances. Fortunately, AGA is as configurable as it is flexible. You can easily configure each archive operation for move or copy. In a normal recordkeeping setting, it is imperative that you always configure it for a move, not a copy. Leave the copy scenario for exceptional circumstances.

Retention Schedule

You have two systems that are used to create and store documents – Content Server and SharePoint. But you have only one corporate retention schedule. Because Content Server is being used as the records system and SharePoint only as a source of records, you've entered your retention schedule into Content Server, not in SharePoint. In order for an end user to classify a document, the user must have access to the retention schedule. However, it "lives" in Content Server, not in SharePoint. In SharePoint, you might want to have a user classify a document long before it gets archived to Content Server. And you'll certainly need to classify documents as part of an archive operation, so they can be archived into Content Server as declared records. You have one retention schedule, but two systems that need access to it.

AGA solves this dilemma by allowing you to present the retention schedule to the SharePoint user in either of the two following scenarios:

> **A web part in a SharePoint Page**. The SharePoint user can bring up the retention schedule from Content Server while inside a normal SharePoint page. The SharePoint user cannot select a value from the schedule, but they can consult the full retention schedule to help them properly classify the document during any SharePoint declaration process you might have. It can also be used during searching to browse the various corporate topics and case files.
>
> **During Archiving**. You can define an archiving process that can either present the retention schedule to the user for manual selection of the retention category, or use the retention schedule within an automated process to assign a valid category based on some known document metadata or folder location.

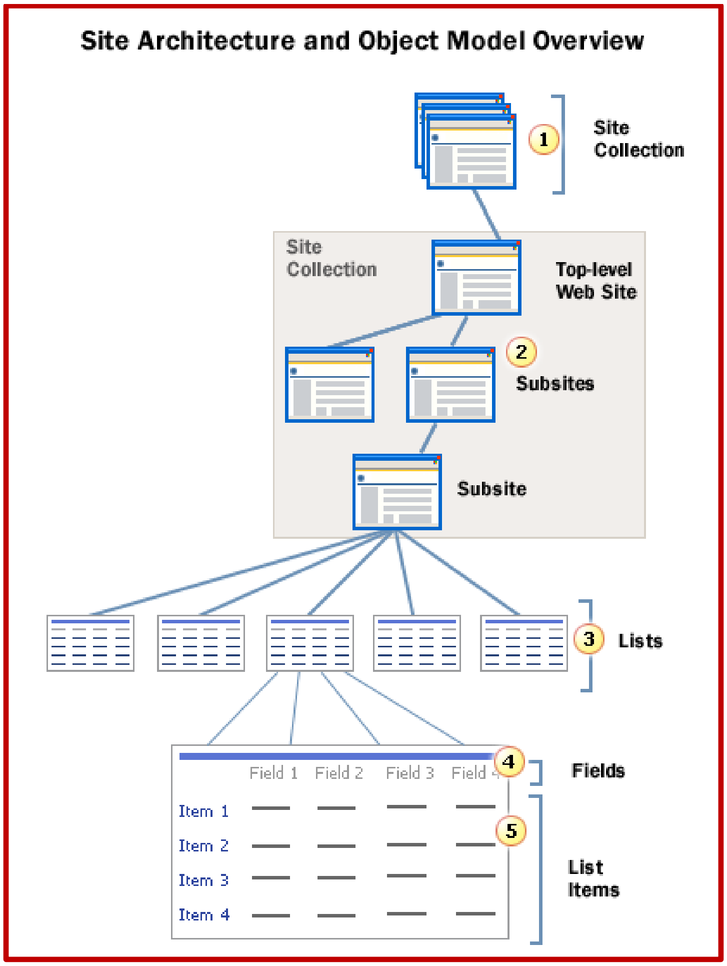
Key SharePoint Concepts

In order to fully understand and deploy AGA, you'll need to know the following essential high-level concepts about SharePoint:

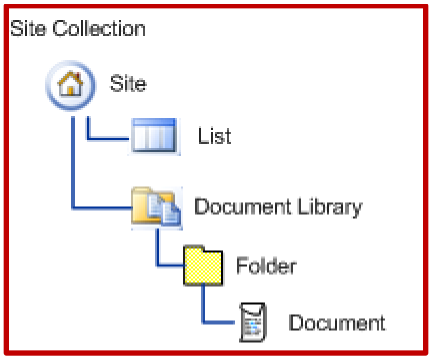
SharePoint Structure
Site
Page
Library
Web Part

SharePoint Structure

To utilize AGA you'll need to fully understand how SharePoint is organized (structured). The principal building block of SharePoint is the **site**. As shown in the diagram that follows, sites are hierarchically structured into subsites, with subsites inheriting properties from parent sites.



Each site contains **pages** (not shown on the diagram). These sites (or more specifically, the pages within the sites) can contain **lists**. Lists in turn contain fields of data (metadata) about the items in the lists. These lists can be lists of documents, in which case the list is called a **library**. The diagram below shows the hierarchy within an individual site:



You need to know that a document library contains documents and or folders. Documents can reside within a folder, or directly within a library, without being in a folder. This is an important point. Within an ordinary Windows shared server, all documents must be in a folder. In SharePoint, documents do not need to be in a folder – they can reside in a library with no folders at all. With AGA, you can define an archive operation that archives an individual document, a folder (and all the documents it contains), or a library (and all the folders and/or documents it contains), or an entire site (and all the libraries and documents it contains).

Site

A SharePoint **site** is a collection of pages, lists, and libraries configured for the purpose of achieving an express goal. A site may contain sub-sites, and those sites may contain further sub-sites. Examples of SharePoint sites would include collaboration (team) sites, document workspaces, group work sites, and meeting workspaces. Sites have navigation, themes, custom permissions, workflows, and have the ability to be configured or customized in a number of ways. In order to achieve a greater degree of maintainability, sites typically inherit site-level settings from their parent sites.

A site contains one or more document libraries, which in turn contain documents. With AGA you can archive an entire site, a library within a site, a folder within a library, or a document within a library. If you archive an entire site, you are archiving all of the site components, including themes, permission, workflows, etc. If however you archive a library, you are archiving only the documents within that specific library.

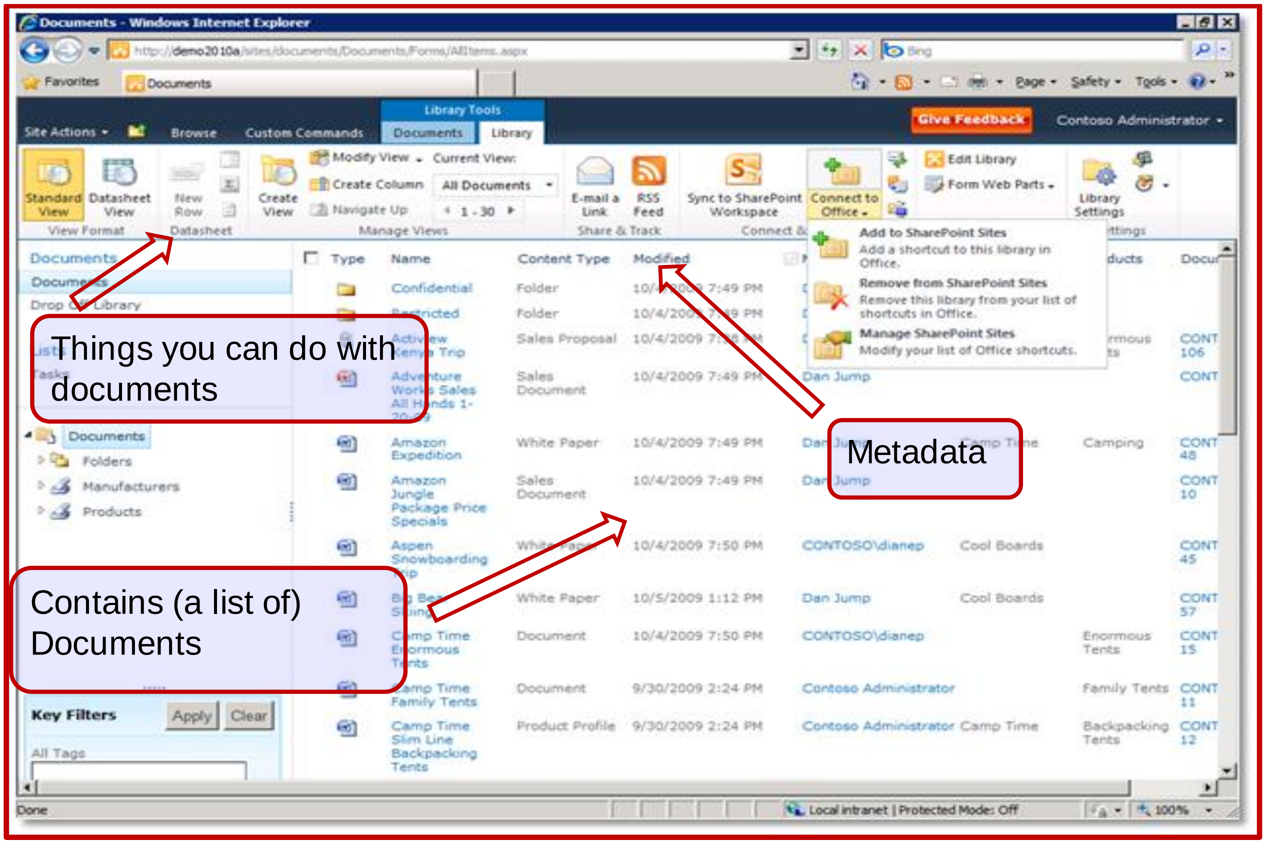
Pages

Unlike most ECM products, SharePoint does not really have a pre-defined, fixed user interface. Instead, users navigate **pages** when they use SharePoint. All the pages together form the overall user interface. It works very similar to that of a website – users navigate forward by clicking on links or menu buttons, and back using the BACK button in their web browser. A page can be customized to do pretty much anything you want within SharePoint. Each page has a unique address, or URL, again similar to that you see when browsing a website on the worldwide web. The address is shown in the browser’s address bar.

A SharePoint site is comprised of **pages**. Pages may be duplicated across multiple sites, or they may be unique to a given site. Pages can also inherit properties from a parent site. You’ll see that AGA provides custom pages to SharePoint, as well as adding special sections to existing SharePoint pages.

Library

A SharePoint **library** is a location on a site where you create, store, update, and collaborate on documents with team members. Each type of library — document, picture, form, wiki page, asset, report, record, data connections, and slide — displays a list of documents and key information about the documents, such as who was the last person to modify the document. Libraries are typically custom-built for specific projects for a specific purpose and feature their own unique set of behaviors and features. A typical SharePoint document library screen is shown below:



With AGA you can specify which documents from which libraries can be archived, or you can archive an entire library in a single operation. Libraries are the areas where SharePoint users will most often be working with their documents, so they are an important point of interface with AGA.

Web Part

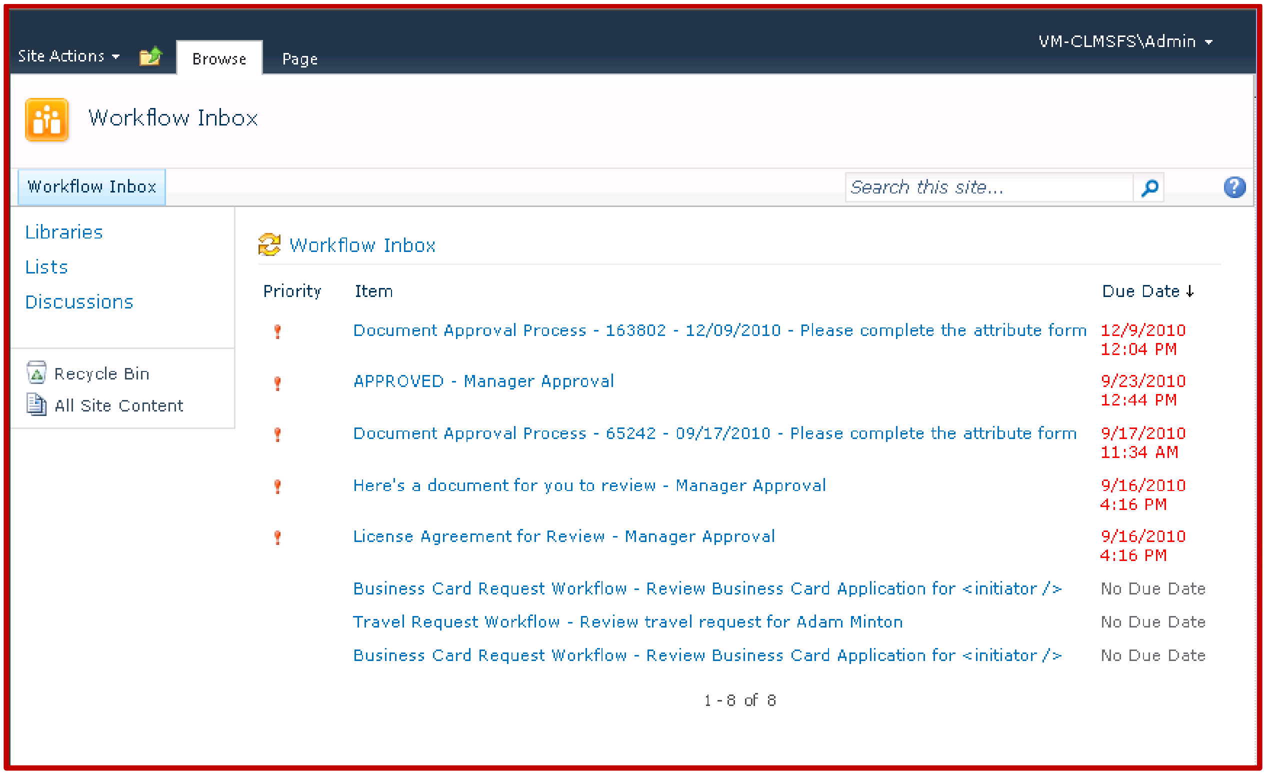
The SharePoint user interface is made up of web pages. Most web pages are customized for some intended purpose to suit the business at hand. A **Web Part** is a piece of a SharePoint page added by a third party. AGA adds several web parts inside SharePoint, including:

> **EL Web Part**. A fully functional Content Server user interface, but inside SharePoint. This is essentially the OpenText Content server Web User Interface client, but presented within SharePoint. The SharePoint user can access, position to, and navigate Content Server folders, display and work with documents in Content Server.
>
> **Search**. Search and retrieve from Content Server. Consists of a *Search* part and a *Results* part. Search can include not only documents stored directly in Content Server, but documents available from external systems but accessed via Content Server, such as email systems, SAP business management systems, etc. Search Results are highly configurable so as to present only needed data, and in any display format.
>
> **WorkFlow**. Access and view Content Server work flows.
>
> **Physical Records**. Borrow physical records (e.g. boxes of records or other artifacts), and request records.

AGA web parts[8] are selected, pre-packaged pieces of Content Server application functionality that appear inside (on) SharePoint pages. The SharePoint user is unaware that a web part is on the page, the web part is simply a part of the page. The Content Server functionality looks as if it was built inside

[8] AGA offers web parts for other records sources such as SAP or email integration.

SharePoint. The screenshot below shows a web part from AGA that allows a SharePoint user to see workflow status from Content Server:



Priority	Item	Due Date ↓
!	Document Approval Process - 163802 - 12/09/2010 - Please complete the attribute form	12/9/2010 12:04 PM
!	APPROVED - Manager Approval	9/23/2010 12:44 PM
!	Document Approval Process - 65242 - 09/17/2010 - Please complete the attribute form	9/17/2010 11:34 AM
!	Here's a document for you to review - Manager Approval	9/16/2010 4:16 PM
!	License Agreement for Review - Manager Approval	9/16/2010 4:16 PM
	Business Card Request Workflow - Review Business Card Application for <initiator />	No Due Date
	Travel Request Workflow - Review travel request for Adam Minton	No Due Date
	Business Card Request Workflow - Review Business Card Application for <initiator />	No Due Date

1 - 8 of 8

Underlying Principles of AGA

To effectively and confidently deploy AGA, you'll need to have a deep understanding of **document metadata** and **permissions**.

Document Metadata

For every document you move from SharePoint to Content Server, you'll need to ensure the metadata fields are carried over to Content Server. The goal is simple – for every SharePoint document field, you need to preserve that field 100% in Content Server. For each different Content Type in SharePoint, you will encounter a different set of metadata fields. You'll need to explicitly tell AGA which SharePoint fields you have, and create an identical set of fields in Content Server. Essentially, you have to mirror each of the SharePoint content types and metadata fields in Content Server. That way, when you archive a document to Content Server, there will be a matching document metadata structure to receive the incoming field values.

Note the difference in terminology:

<u>Content Server</u>		**<u>SharePoint</u>**
Category	*Equates to*	**Content Type**
Attribute	*Equates to*	**Column**

In various Content Server documentation, you'll often see the Content Server term *Attribute* referred to as a *Property*, and a SharePoint *Column* as a *Field*. *Attribute* and *Property* (Content Server) equate to *Column* and *Field* (SharePoint), all of which refer to document metadata fields.

To ensure metadata fields are carried over to Content Server properly, you'll need to accomplish three things for each document type in SharePoint:

1) Map SharePoint fields to equivalent fields in Content Server

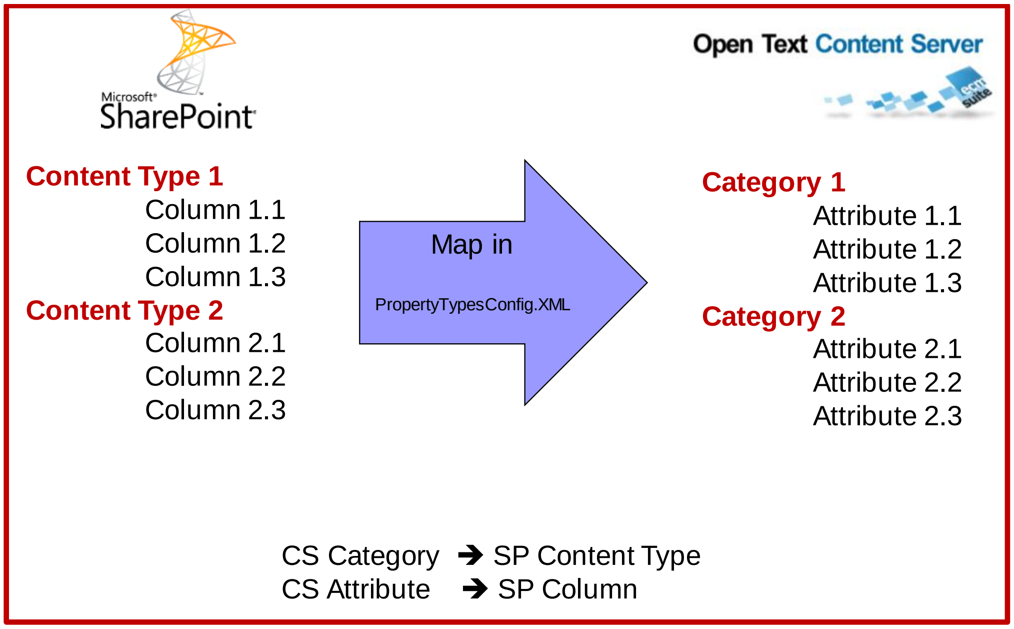
2) Ensure field data types are compatible
3) Set field values in Content Server where needed

You will have to insure that the mapping, data types, and values are all 100% compatible and fully configured (and tested) for each document you archive, before the archiving is carried out. There could be hundreds, even thousands, of content types in a large system (potentially times 2 for Content Server and SharePoint). SharePoint content types feature inheritance, which AGA can take advantage of to help reduce the total number of mapping combinations required. Regardless, prepare for a significant effort of field mapping in advance of AGA deployment. After deployment, you'll have to maintain these mapping as new content types are created and modified in both systems. We will detail each of these three steps in order.

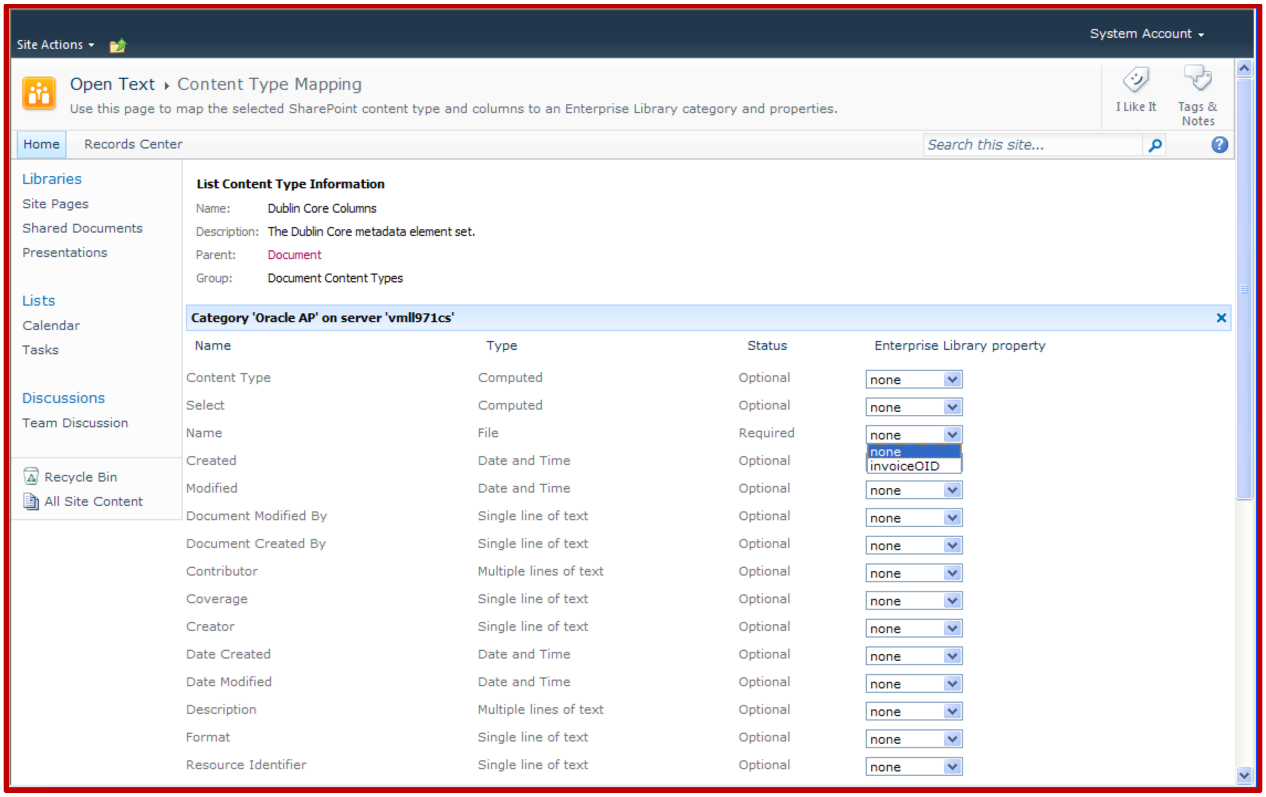
If you archive a document that does not have the field mapping completed, AGA dumps all SharePoint metadata fields into a single Content Server field called *AllProperties.* While this transfers the metadata over to Content Server, it is not recommended, as the fields are not separated for effective searching. This may be perfectly acceptable in a non-recordkeeping setting, where you are simply doing a "mass archiving". In a setting with strong information governance requirements however, this represents poor information management practice, and should be avoided. Document metadata is an important part of all records, and must be visible, and as retrievable as possible. In a non recordkeeping setting, this feature might actually be considered a benefit – i.e. a means of stripping unwanted metadata from SharePoint documents.

Map Metadata Fields

To map metadata fields, you need to explicitly define in AGA each field for each SharePoint content type, then specify a matching content type (Category) in Content Server, along with a field for that category that matches each field in the SharePoint document. The example below illustrates how the mapping is achieved:



The following screencap shows how you map SharePoint fields to Content Server fields, using the AGA Administrative module in SharePoint:



Use this page to map the selected SharePoint content type and columns to an Enterprise Library category and properties.

List Content Type Information

Name: Dublin Core Columns
Description: The Dublin Core metadata element set.
Parent: Document
Group: Document Content Types

Category 'Oracle AP' on server 'vmll971cs'

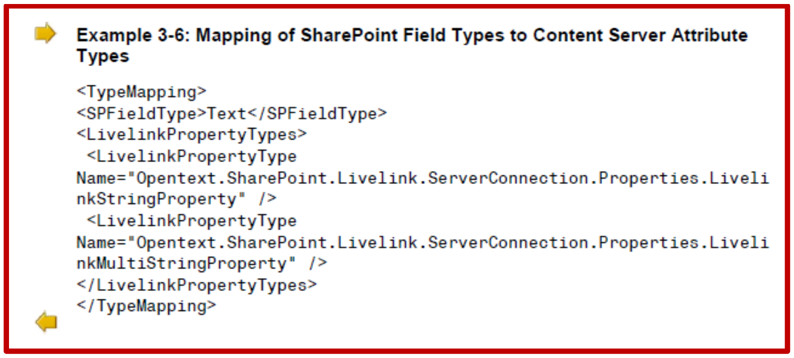
Name	Type	Status	Enterprise Library property
Content Type	Computed	Optional	none
Select	Computed	Optional	none
Name	File	Required	none none invoiceOID
Created	Date and Time	Optional	
Modified	Date and Time	Optional	none
Document Modified By	Single line of text	Optional	none
Document Created By	Single line of text	Optional	none
Contributor	Multiple lines of text	Optional	none
Coverage	Single line of text	Optional	none
Creator	Single line of text	Optional	none
Date Created	Date and Time	Optional	none
Date Modified	Date and Time	Optional	none
Description	Multiple lines of text	Optional	none
Format	Single line of text	Optional	none
Resource Identifier	Single line of text	Optional	none

If you have a SharePoint field for example called *Project Name*, you will need an equivalent field in Content Server of the same data type that equates to the SharePoint field *Project Name*. If a suitable target field does not already exist in Content Server you will have to first create the target field in Content Server, then run the AGA mapping module to match up the equivalent SharePoint fields. You can use any existing Content Server target field that might be suitable for your use – it does not have to bear the exact same field name.

You can gain greater control over the mapping, and customize it, by mapping fields manually. You define the mapping by editing the following AGA XML file:

PropertyTypesConfig.XML

Below is a fragment from the XML file that illustrates the syntax (rules of entering the data) you would use to record a field mapping:



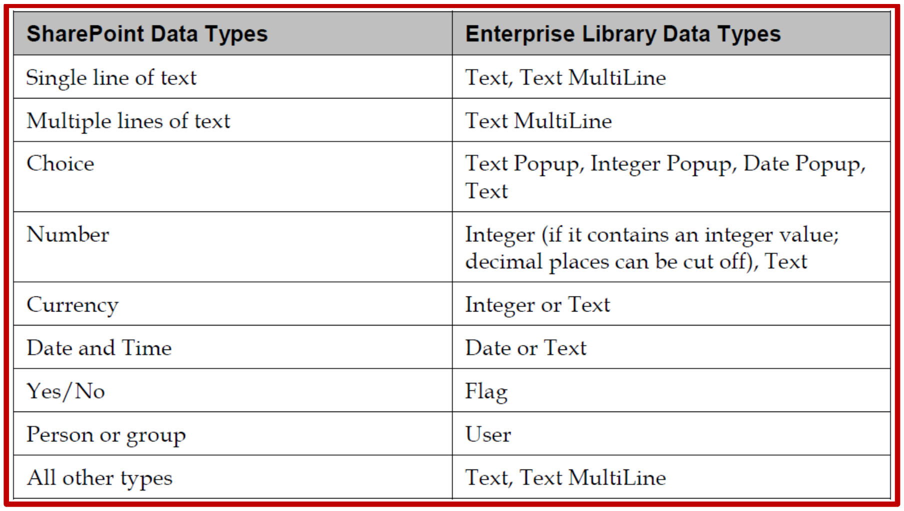
Example 3-6: Mapping of SharePoint Field Types to Content Server Attribute Types

```
<TypeMapping>
<SPFieldType>Text</SPFieldType>
<LivelinkPropertyTypes>
 <LivelinkPropertyType
Name="Opentext.SharePoint.Livelink.ServerConnection.Properties.Liveli
nkStringProperty" />
 <LivelinkPropertyType
Name="Opentext.SharePoint.Livelink.ServerConnection.Properties.Liveli
nkMultiStringProperty" />
</LivelinkPropertyTypes>
</TypeMapping>
```

Data Types

For every field you map, you must ensure the **data type** is compatible. For instance, you cannot map a SharePoint multiline text field such as *Document Description*, which stores textual data, to a Content

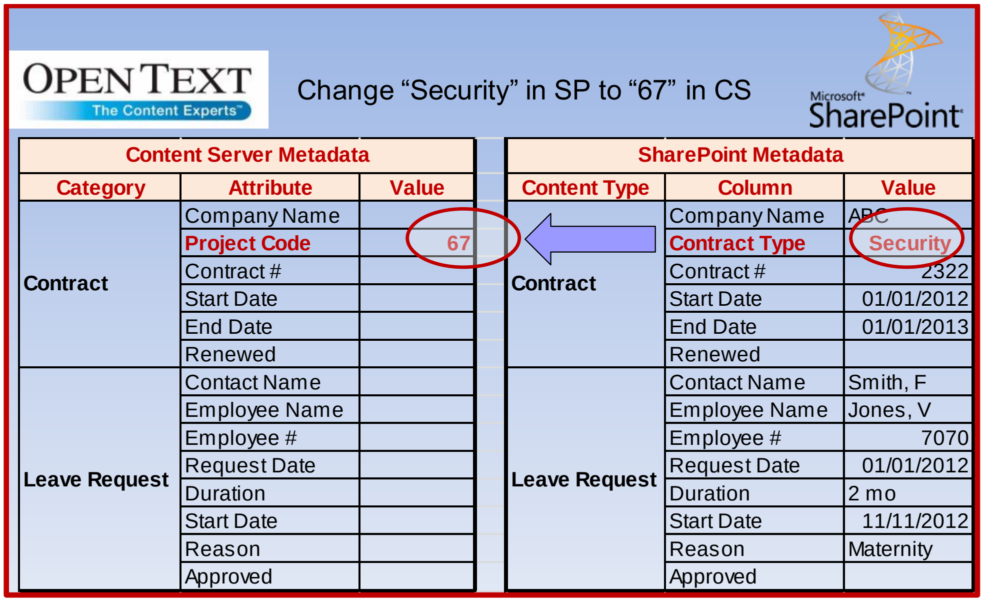
Server field *Price* which stores only integer (numeric) data. AGA automatically checks the data type of each field you map, and it will only display onscreen the fields that are compatible types. The table below shows the data types for Content Server (referred to as Enterprise Library), and SharePoint.



SharePoint Data Types	Enterprise Library Data Types
Single line of text	Text, Text MultiLine
Multiple lines of text	Text MultiLine
Choice	Text Popup, Integer Popup, Date Popup, Text
Number	Integer (if it contains an integer value; decimal places can be cut off), Text
Currency	Integer or Text
Date and Time	Date or Text
Yes/No	Flag
Person or group	User
All other types	Text, Text MultiLine

Set Field Values

A field's **value** refers to the data within the field. For instance, field *Contract Type* could contain values *Security* or *Grounds Maintenance*. Occasionally you might need to force a SharePoint field value into a Content Server field. For example, you might be mapping a SharePoint field to a Content Server field that already exists, and you need to force a value from the incoming document field into that field, overwriting the usual default value in Content Server. The example below shows how you might map SharePoint field *Contract Type* to Content Server field *Project Code*, and insert value ***67*** into the Content Server field:


OPEN TEXT
The Content Experts
Microsoft
SharePoint

Change "Security" in SP to "67" in CS

Content Server Metadata			SharePoint Metadata		
Category	Attribute	Value	Content Type	Column	Value
Contract	Company Name		Contract	Company Name	ABC
	Project Code	67		Contract Type	Security
	Contract #			Contract #	2322
	Start Date			Start Date	01/01/2012
	End Date			End Date	01/01/2013
	Renewed			Renewed	
Leave Request	Contact Name		Leave Request	Contact Name	Smith, F
	Employee Name			Employee Name	Jones, V
	Employee #			Employee #	7070
	Request Date			Request Date	01/01/2012
	Duration			Duration	2 mo
	Start Date			Start Date	11/11/2012
	Reason			Reason	Maternity
	Approved			Approved	

To set field values in this manner, a small amount of custom programming[9] is needed. The custom mapping logic is programmed into the **ValueConverters** section of AGA configuration file PropertyTypesConfig.XML. Programming such as this should only be required for special circumstances.

Permissions

Permissions define *who* can do *what* to *which* documents. Permissions are comprised of the following elements:

Users Named User account, e.g. Smith, Fred

Groups Groups of users. Every member of the group has the group's specified permissions

Document Permissions The specific actions available to a user who has access to the document. Examples of actions would be *edit, delete, create new version*, etc.

A user's effective permission to a given document determines what they can do with that document. Their permissions are granted in one of the following two ways:

Direct The system administrator explicitly grants user *User1* permissions *Edit, Delete, Print* to document *707*.

Indirect Group *Group1* has *Edit, Delete* permissions to document *707*. Because User *User1* is a member of group *Group1*, *U1* has *Edit* and *Delete* permission to any documents that *Group1* has rights to see, in addition to any other permissions he may have been granted directly.

When archiving documents via AGA, your goal typically is to make the permissions of the destination (Content Server) document the same as they were in the source (SharePoint) system. This assumes you are moving the document, not leaving a copy behind in SharePoint. If however you are leaving a copy behind (not recommended), then obviously it becomes important to match the permissions of each system. This simplifies both the user experience in each system, as well as the system administration, not to mention that you would not want to inadvertently violate any access control policies by having a different security scheme in each system. Even if you are moving a document from SharePoint to Content Server, you would not normally want the security permissions to differ just because it was moved to a different system.

To align permissions in both systems, there are two things you'll need to accomplish:

1) Match SharePoint users and groups with Content Server users and groups
2) Map Document Permissions

We will look at each separately, then see how they work together to form the overall permissions scenario.

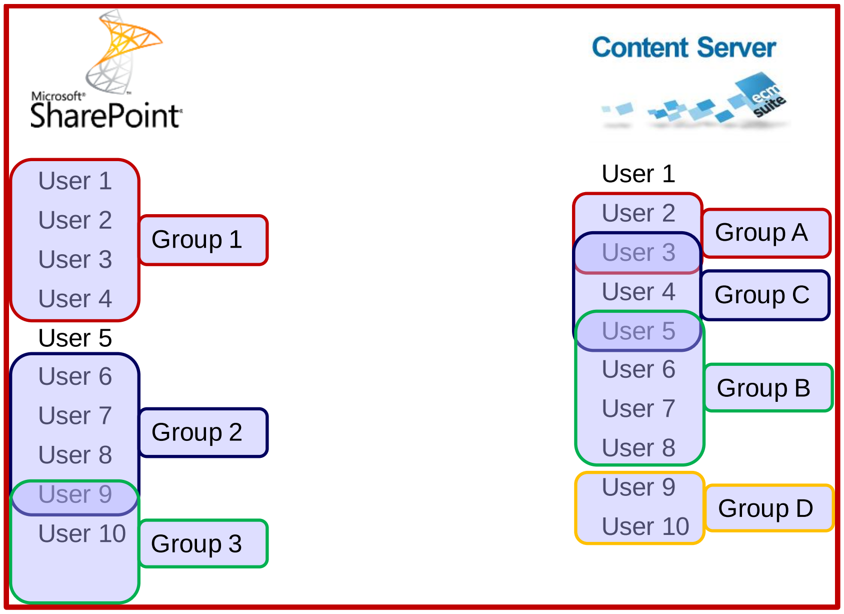
Users and Groups

Each system will have users and groups (of users) defined. Each defines users and groups somewhat differently[10], but the concept is the same. To align permissions in both systems, you should first ensure that the users and groups in each system are identical[11]. The diagram below shows how the two systems could hypothetically be running. They each have the same 10 users, however their groups are entirely different. Notice that SharePoint User 9 is a member of groups 2 and 3, whereas in Content Server, the same user is a member of only group D.

[9] Use the .NET Windows programming platform.

[10] Semantic and naming differences between user and group names in each system (such as sensitivity to case) can be custom-specified in an AGA configuration screen.

[11] Users must be matched in each system for user login authentication to work. This ensures that SharePoint users never have to manually log onto Content Server – they are automatically logged in with their (matching) account.



In order for document permissions to work in the same manner in both systems, the users and groups **must** be the same in each system. The diagram makes this obvious – if a document you are about to archive from SharePoint to Content Server provides permission to it in SharePoint via the user's membership in a group, the permission can only be replicated in Content Server if Content Server also has that same group. Similarly, if for example user 9 exists in SharePoint but not in Content Server, how can this user access the document once archived to Content Server?

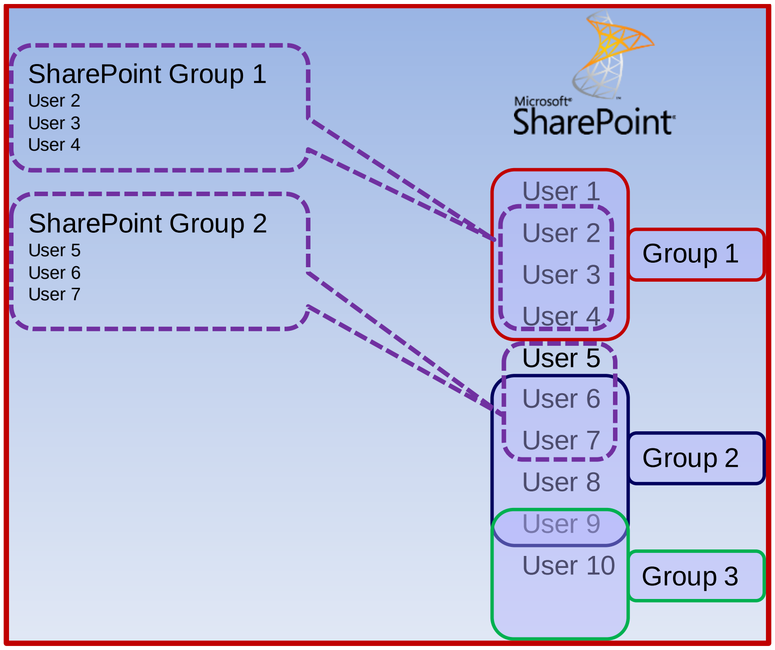
Clearly your first important step in aligning permissions is to **replicate users and groups in Content Server and SharePoint**. OpenText recommends using Microsoft Windows Active Directory services to define users and groups outside of SharePoint. This way, both SharePoint and Content Server are deriving permissions from a single common source, minimizing the need for duplication. Regardless of the approach you use, this is an ongoing, permanent operational decision you have to make, and preserve for the life of each system. Duplicating user accounts (or maintaining in Active Directory) is a recommended best practice, for the following reasons:

1) **Simplifies User Operations**. The more different the security models of each system, the more trouble for users to understand, Many users will often report what they believe to be a system failure ("*The system lost my document!*") when the system refuses to show them a document they believe they can access.
2) **Simplifies Administration**. The greater the differences, the more administrative time and effort will be required to understand, maintain, and explain the differences.
3) **Reduces Complexity**. Each difference between the two systems increases the complexity of the overall integrated solution. This complexity can potentially be difficult to manage.

The Administrators of both Content Server and SharePoint should ensure continual replication of users and groups, if at all possible. When using AGA to archive a document, AGA can be configured to automatically create users in Content Server that are contributing permissions to the document in SharePoint. While this is a very good thing for AGA to do, the end result is that users "pop up" in Content Server in an ad hoc manner over time, as documents are archived. This preserves the permissions integrity of the document once transferred, but can have a huge impact on the overall planning and architecture of the destination Content Server system, as users and groups emerge over time without any foresight or planning. Particularly in large ECM deployments of several thousand users, system architecture is very carefully planned, deployed, and maintained.

If however the users and groups of both systems are replicated, each document archive operation becomes far simpler, with no implications of foreign (to Content Server) groups or user accounts suddenly appearing unexpectedly. Where possible, try your best to replicate users and groups, and leave the automatic AGA creation for the exceptions.

SharePoint has a special type of user group called a **SharePoint Group**. These special user groups are tied to a SharePoint site, and are deleted when the site is removed. They are “temporary”, in that they only exist for the life of the site with which they are associated. The diagram below shows two SharePoint groups, in addition to the “Regular” groups[12].

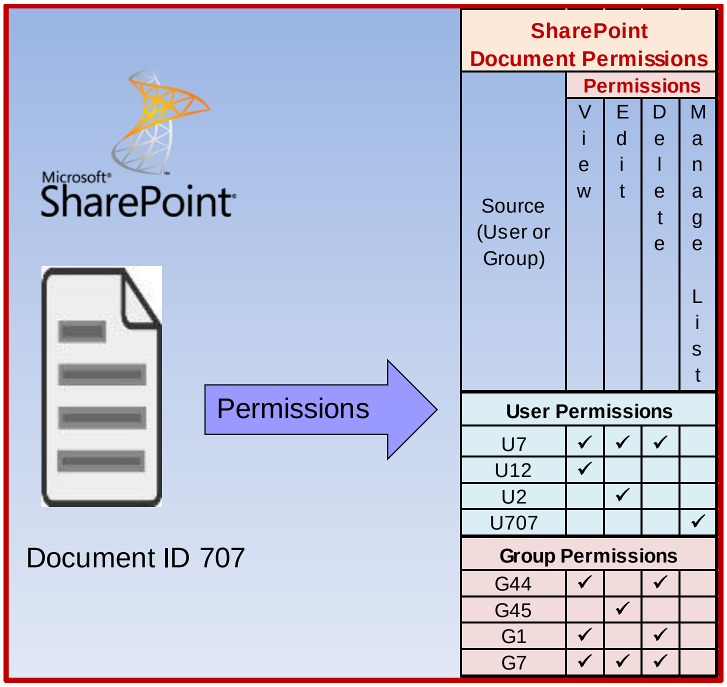


Content Server has no equivalent concept of a temporary group – all groups are permanent (until manually deleted). When you use AGA to archive a document to Content Server, you will lose these temporary SharePoint group settings. Consider not using these special groups in SharePoint if possible, or at least define them such that you are not dependent on the permission derived from them once the documents are archived to Content Server.

Document Permissions

Each document in an ECM system has its permissions that determine who can do what with the document. Each permission is an action that can be carried out on the document. Each permission is assigned directly from a user, or indirectly via a group. Think of a document’s permissions as a matrix of actions versus assignments (users or groups). The diagram below shows document permissions for SharePoint.

[12] So-called “regular groups” in SharePoint are created and maintained in the Microsoft Windows Active Directory. Users and groups are created outside SharePoint, but used internally by SharePoint. SharePoint Groups are created and stored exclusively within SharePoint itself.

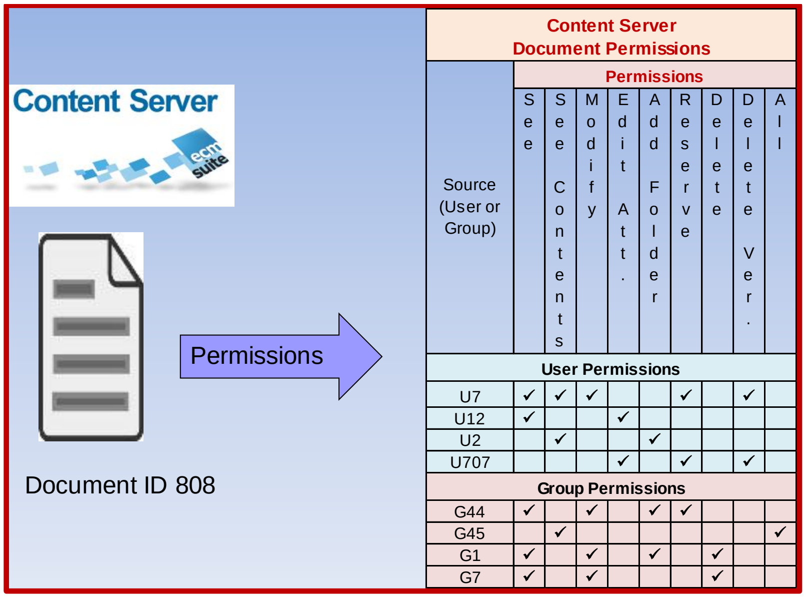


SharePoint Document Permissions

Source (User or Group)	Permissions			
	View	Edit	Delete	Manage List
User Permissions				
U7	✓	✓	✓	
U12	✓			
U2		✓		
U707				✓
Group Permissions				
G44	✓		✓	
G45		✓		
G1	✓		✓	
G7	✓	✓	✓	

SharePoint has four discreet permissions, or actions, that can be assigned to a document – *View, Edit, Delete*, and *Manage List*. For any given document, these actions are granted either from users or groups.

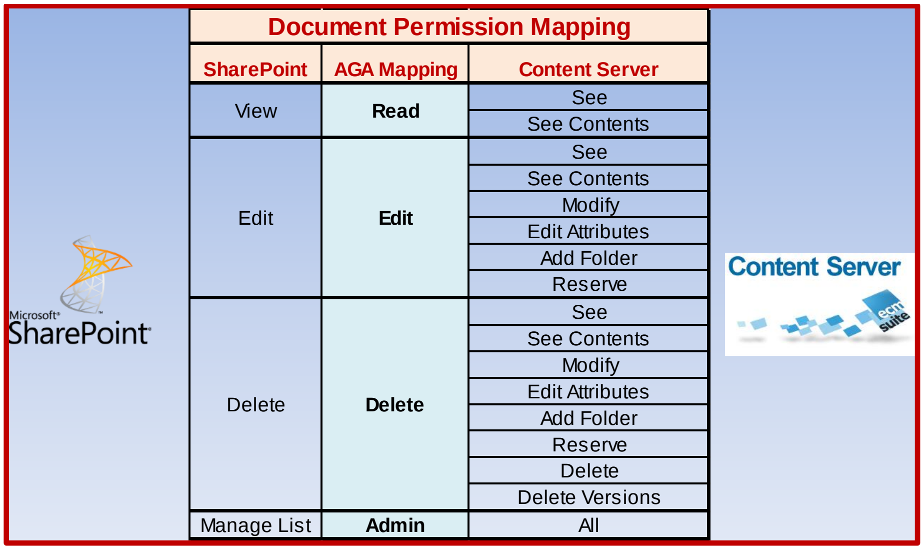
Content Server on the other hand has a very different set of document permissions, as shown:



Content Server Document Permissions

Source (User or Group)	Permissions								
	See	See Contents	Modify	Edit Att.	Add Folder	Reserve	Delete	Delete Ver.	All
User Permissions									
U7	✓	✓	✓			✓		✓	
U12	✓			✓					
U2		✓			✓				
U707				✓		✓		✓	
Group Permissions									
G44	✓		✓		✓	✓			
G45		✓							✓
G1	✓		✓		✓		✓		
G7	✓		✓				✓		

Content Server has nine permissions, each of which is completely different from SharePoint's four permissions. In order to archive a document from SharePoint to Content Server, and maintain equivalent security, there must be a way to rationalize these two different security models. Fortunately, AGA has done this for us, in the form of a **permission mapping**, as shown below:


Microsoft SharePoint
Content Server
ecm suite

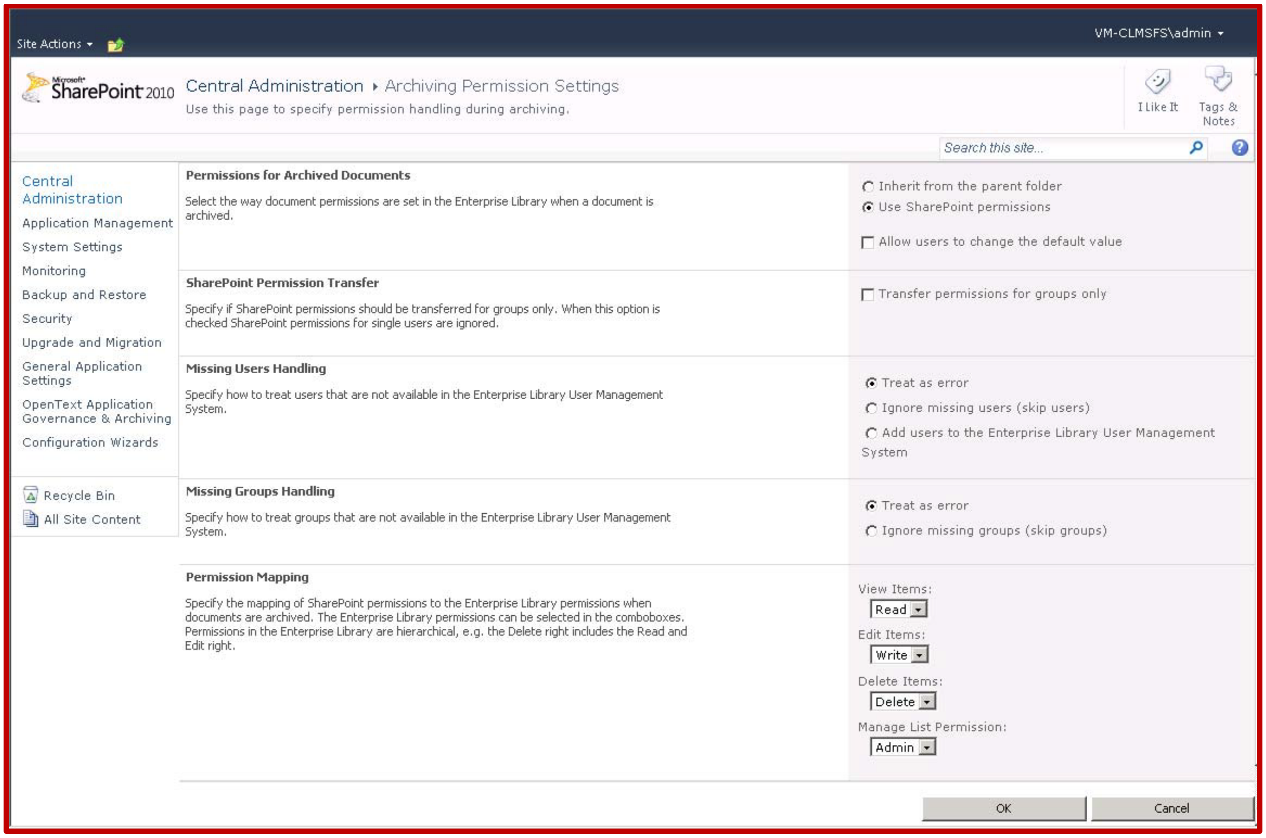
Document Permission Mapping

SharePoint	AGA Mapping	Content Server
View	**Read**	See
		See Contents
Edit	**Edit**	See
		See Contents
		Modify
		Edit Attributes
		Add Folder
		Reserve
Delete	**Delete**	See
		See Contents
		Modify
		Edit Attributes
		Add Folder
		Reserve
		Delete
		Delete Versions
Manage List	**Admin**	All

OpenText has effectively “bundled” multiple Content Server permissions into a group that will be switched on in Content Server when the mapped SharePoint permission is requested. As shown in the preceding diagram, the SharePoint permission *Edit* will switch on Content Server *See, See Contents, Modify, Edit Attributes, Add Folder,* and *Reserve* permissions. AGA refers to that bundle of Content Server permission as “Edit”. When you define an AGA archive operation, for each of the four SharePoint permissions, you can specify one of the four AGA Bundles, shown in the diagram as “AGA Mapping”.

AGA Permissions Configuration

You’ll have to define how permissions are to be handled for all archiving operations for each SharePoint site, i.e. it is a global site setting that will be applied to all archiving operations at all times for the specified SharePoint site. The screencap below shows AGA’s **Permission Configuration** form, found on the Administration console in SharePoint:



Central Administration ▸ Archiving Permission Settings

Use this page to specify permission handling during archiving.

Permissions for Archived Documents

Select the way document permissions are set in the Enterprise Library when a document is archived.

○ Inherit from the parent folder
● Use SharePoint permissions

□ Allow users to change the default value

SharePoint Permission Transfer

Specify if SharePoint permissions should be transferred for groups only. When this option is checked SharePoint permissions for single users are ignored.

□ Transfer permissions for groups only

Missing Users Handling

Specify how to treat users that are not available in the Enterprise Library User Management System.

● Treat as error
○ Ignore missing users (skip users)
○ Add users to the Enterprise Library User Management System

Missing Groups Handling

Specify how to treat groups that are not available in the Enterprise Library User Management System.

● Treat as error
○ Ignore missing groups (skip groups)

Permission Mapping

Specify the mapping of SharePoint permissions to the Enterprise Library permissions when documents are archived. The Enterprise Library permissions can be selected in the comboboxes. Permissions in the Enterprise Library are hierarchical, e.g. the Delete right includes the Read and Edit right.

View Items: Read
Edit Items: Write
Delete Items: Delete
Manage List Permission: Admin

OK | Cancel

There are five parameters you can specify:

Permissions Mode. There are two choices:

Inherit from Content Server Folder. SharePoint permissions are discarded, and permissions are inherited from the destination Content Server folder. Use this mode when the users are comfortable with the Content Server permission structure they are about to apply. For an archive operation of multiple folders and subfolders, the source folder structure should match that of the destination structure, so the inheritance will be compatible.

This would be useful if you wish to change the permissions of the newly-archived folder structure. For example, you may wish to change a complex SharePoint permissions arrangement on the original folders to a simple read-only on the destination structure. The original permissions arrangement may not be needed once the work is complete and archived.

Use SharePoint Permissions. Apply the existing SharePoint permissions to Content Server. This assumes the destination Content Server users and groups exist. Apply the SharePoint to Content Server permission mapping as defined.

Group Only. Ignore SharePoint permissions granted directly to users, and include only those granted via groups.

Missing Users Handling. Specifies how to proceed if a SharePoint user is missing in Content Server. Either proceed with the archive and ignore the missing user(s), or create the users automatically in Content Server.

Missing Groups Handling. Specifies how to proceed if a SharePoint group is missing in Content Server. Either proceed with the archive and ignore the missing group(s), or cancel the archive operation.

Permission Mapping. Specify which Content Server document permissions will map to which (bundles) of Content Server permissions, as shown earlier.

There are two special scenarios to consider when planning your permission strategy and mapping:

Permission Synchronization

You can synchronize the permissions during either of three archiving modes –*Create Link, Child Link*, and Synced *Copy*. Once you've had AGA deployed for a while, most of your documents will have been moved (hopefully), in line with good recordkeeping practices. Inevitably however you'll have some that were archived with a link left behind in SharePoint. SharePoint users can continue to search and read the archived document, by clicking on the link to it in Content Server.

But what if the permissions to the (link) have changed in SharePoint? The two systems are now out of sync, as the permissions in each system are different. Worse, you will not know – there is no mechanism by which you (or anyone for that matter) are notified when a permission has changed in SharePoint that ought to be reflected in the actual documents, stored in Content Server.

You'll need to have AGA update permissions (propagate SharePoint changes to Content Server) periodically. You run this update process as frequently or as infrequently as you wish. In an AGA configuration screen, you specify a schedule (date, time of day, repeat frequency) you'd like the update to run. Windows Scheduler will then automatically update permissions per the schedule.

For more advanced scenarios, you can programmatically trigger a permission update. Say for instance you wanted to run an update each time 1,000 documents were archived. To do this, you have to create a **PowerShell cmdlet** (command-let). This is a series of very high-level programming language commands that you assemble into a series of instructions called a **script**. It requires technical knowledge, and uses AGA's **S**oftware **D**evelopment **K**it (SDK). You can write a script that will examine all the archived document links in SharePoint, check for permission changes (or whatever triggering criteria you specify) then propagate these changes in Content Server. Regardless of how you trigger a permission update, OpenText recommends you update permissions on a regular basis, to make sure the permissions do not fall too far out of sync for too long.

Restoring Documents Back to SharePoint

AGA allows a SharePoint user to **restore** a document from Content Server back into the SharePoint workspace. Note the important difference between *restore* and *retrieve* below:

Restore Copy an archived document from Content Server back into the SharePoint workspace

Retrieve From within SharePoint, find and view a document that exists within Content Server

If you restore a document from Content Server to SharePoint, the permissions on that particular document may have been re-mapped from its original native SharePoint settings to different Content Server settings. If in fact the original SharePoint permission settings were re-mapped in Content Server, and the original SharePoint documents were not preserved as part of the archiving operation (good information governance practice), then the original SharePoint permission settings cannot be recovered, as they were never recorded anywhere (practically speaking, this is a too much to ask for). What permissions to assign the restored documents? Remember that the Content Server and SharePoint permissions differ from each other. AGA therefore has no choice but to apply a "best fit" – a reasonable translation between the two different permissions models.

When you restore an archived document back to SharePoint, AGA will automatically map the Content Server permissions to standard SharePoint permissions in accordance with the following mapping table:

SharePoint Restore Permission Mapping		
Minimum Content Server Permissions		SharePoint Permissions
See	➔	View
Modify	➔	Edit
Delete	➔	Delete
All	➔	Manage List

If the original SharePoint document's permissions were not re-mapped, the permissions of the restored document will be the same as those of the original.

AGA Overview

AGA consists of the following broadly-defined sections or components:

Administration
SharePoint Pages
SharePoint Web Parts
Archiving
Search
Extensibility
Site Deployment

Administration This is where you go in SharePoint to set up, configure, and maintain AGA's various features and functions. An AGA Administration page is located in the SharePoint Central Administration for global settings that affect all users, all SharePoint sites, or all archiving actions across all of SharePoint. All SharePoint sites are collectively known as the SharePoint "farm". Site collection and site-specific AGA administrative functions are located within the administration area of SharePoint sites.

AGA allows you configure a great many elements of the integration, such as settings that determine how and when archiving takes place, the metadata and permissions mappings, etc. These parameters are all recorded in various XML files, stored within SharePoint. If you wish, you can manually edit these XML configuration files directly, in addition to using the configuration pages that AGA supplies in SharePoint.

You can apply these settings to the farm, site collection, or site level, which allows you to determine the scope of libraries/documents the setting will affect. In most cases, these settings override any global setting that may be in effect for the location in question. This way, you can selectively decide which settings affect which documents.

SharePoint Pages AGA supplies a number of SharePoint **pages**. These allow SharePoint users to interact with Content Server from inside SharePoint in a number of ways, such as archive documents, search Content Server, display search results from Content Server, borrow physical records from Content Server, and more. AGA also implements many dialog boxes, property sheets, and similar page components within SharePoint. These pages, page components, and web parts (discussed later) define the SharePoint user interface/experience with AGA.

SharePoint Web Parts AGA has a number of web parts that appear on several existing SharePoint pages. They interact seamlessly with the existing SharePoint web pages, and allow the SharePoint user to work with Content Server. The most notable web part is the **EL** (**E**nterprise **L**ibrary) **Web Part**, which basically presents a comprehensive user interface to Content Server, inside SharePoint. There are seven AGA web parts:

EL Web part. Content Server user interface inside SharePoint.
LiveReport. Generate Content Server reports.
LiveReport Results. View Content Server LiveReports inside SharePoint.
Workflow Inbox. Display Content Server workflow tasks inside SharePoint.
Search Box. Simple search of Content Server, while inside SharePoint.
Advanced Search. Advanced Content Server search, from inside SharePoint.
Search Results. View and work with Content Server search results.

Archiving Allows documents to be moved from SharePoint to Content Server. There are manual (interactive) and process-driven (automatic) means of archiving documents. Documents can be archived as records, or as non-records. SharePoint Lists and Sites can also be archived.

Search The means whereby SharePoint users can search Content Server only, or search Content Server and SharePoint, to find documents that exist in either system.

Extensibility The means by which you can customize the integration via simple configuration (no development required), or extending the product by adding in your own custom features (software development required).

Site Deployment The means by which you can easily replicate sites throughout SharePoint without having to recreate them from scratch every time, including any AGA settings and features.

The heart of the AGA product is the archiving process itself. You can define how and when documents are archived, treat them as records if needed, account for the likely differences in metadata and permissions (among other differences), and even customize the archiving for one-off unique situations. As stated earlier, documents can be archived as records, or not. This report is concerned exclusively with archiving for the purpose of recordkeeping.

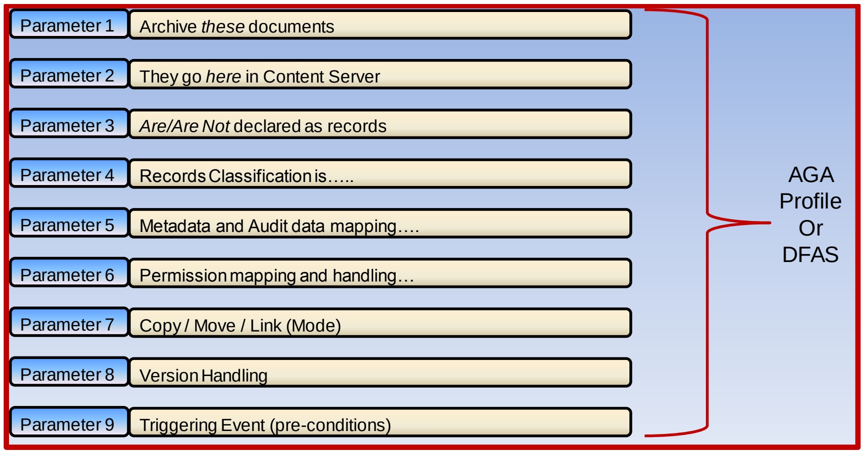
Archiving Documents

In theory, archiving a document from SharePoint to Content Server can be as simple as an end user manually archiving a single SharePoint document to Content Server. It can be a single archive of an entire SharePoint site, consisting of many thousands of folders and their documents. Or it can be anything in between. And the documents can be archived as records, which introduces another layer of information management considerations, which add to the complexity. On top of all this, the document metadata, permissions, and recordkeeping must be properly accounted for. Each time you contemplate an archive, the following questions **must** be answered:

1) What SharePoint documents do I want to archive?
2) Where do they go in Content Server?
3) Should the source documents be declared as records in Content Server?
4) If the documents are undeclared records in SharePoint, how can they be properly classified and declared during archiving?
5) How are the metadata fields and audit data to be mapped?
6) How are the document's original permissions in SharePoint to be transferred to (or translated to) Content Server?
7) Should the original stay behind (copy), or be deleted (move)?
8) Should a link to the Content Server document be left behind in SharePoint after archiving?
9) Archive just this document, or multiple versions (versions handling)?

These are a lot of questions to answer before you can archive a document. However they cannot be avoided – ever. You simply must not – cannot – archive a document without answering each and every question. There is no such thing as a “one size fits all” answer to these nine questions. Different documents, different users, and different sites in SharePoint will each present a totally unique set of circumstances that will force you to re-examine the answers to each of these nine questions.

For the purpose of this report, we will define an **archiving operation** as the act of archiving one or more documents from SharePoint to Content Server. This can be a single document, an entire site, an entire library, or any other possible combination. We assume that the operation, regardless of how many documents it archives, is carried out from start to finish, with no intermediary steps. An archiving operation can be initiated by either a manual *user request* (one-time ad-hoc manual operation), or by a *process*, such as when a pre-determined set of criteria have been met. Either way, each archiving operation will always require the answers to the nine questions. RIMtech will refer to these answers as **archive operation parameters** (**AOP**s). More technically stated, an archive operation “consumes” a set of AOPs in order to move the document(s) from SharePoint to Content Server. The archive operation simply cannot complete (at least not correctly) without these AOPs. Again, the parameters are (more or less) the answers to the 9 questions, as shown below:


Parameter 1
Archive *these* documents
Parameter 2
They go *here* in Content Server
Parameter 3
Are/Are Not declared as records
Parameter 4
Records Classification is…..
Parameter 5
Metadata and Audit data mapping….
Parameter 6
Permission mapping and handling…
Parameter 7
Copy / Move / Link (Mode)
Parameter 8
Version Handling
Parameter 9
Triggering Event (pre-conditions)
AGA
Profile
Or
DFAS

There are two ways to initiate (trigger) an archive operation with AGA:

Manual A SharePoint user, at a time of their choosing, arbitrarily selects document(s) from SharePoint and archives them to Content Server. The AOPs are specified by the user as part of the operation, or pre-defined in advance. AGA refers to this as **Interactive Archiving**. Either documents or folders may be archived using this approach.

Process-driven The AGA Administrator has pre-determined a set of criteria (rules) in SharePoint that must be satisfied to trigger the archive operation. The AOPs are defined in advance. The operation proceeds and completes without user interaction. A given operation can be repeated as many times as needed. AGA refers to this as **Profile-Based Archiving**. Only documents may be archived using this approach – not folders.

Interactive Archiving

Interactive archiving refers to the means whereby a user can manually, at any time, arbitrarily select a SharePoint document (or site or library) and archive the document(s) to Content Server. The user can simply select a document out of the blue and ask AGA to archive it. AGA will oblige, however it will need the AOPs in order to proceed. AGA can be “fed” the AOPs in one of the following two ways:

Prompt the User. Ask the user for the answers to each question. There are no pre-defined answers it can utilize.

Pull From Configuration File. Answers (all or some) have been pre-defined and recorded in a system configuration file.

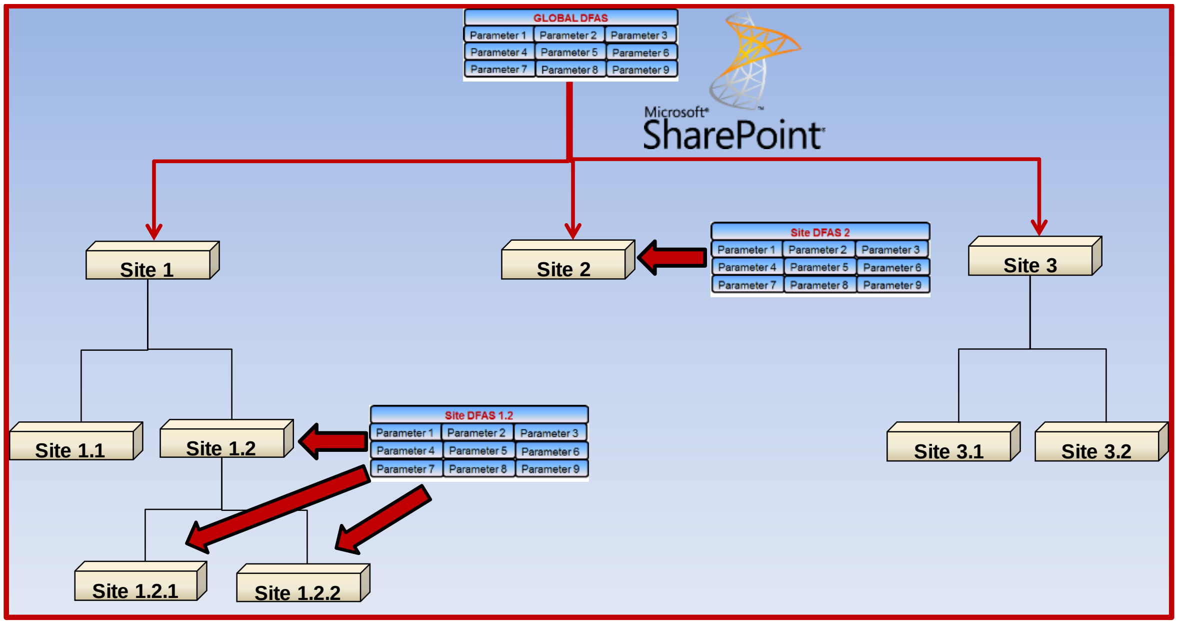
Let's assume for the moment there is no pre-defined configuration that contains the answers in advance. Is it realistic to expect a user to deal with these 9 questions in a reliable or consistent manner? A highly motivated user with plenty of time and patience, coupled with a good deal of knowledge about the target Content Server system and location, just might be willing and able to get the job done. And what if the user targets a SharePoint folder of multiple documents, or a SharePoint library of multiple folders? The technical and information management issues suddenly become even more complex. Which documents are records? Can they all be classified against the same retention rule? What if different documents have different metadata fields? What if permissions are different (different documents have different users and groups)?

To add one final layer of complexity, how much does the user know about the target location in Content Server? Is it a suitable location? How well will the document metadata and permissions fit in the target location? Is recordkeeping applied in the target location?

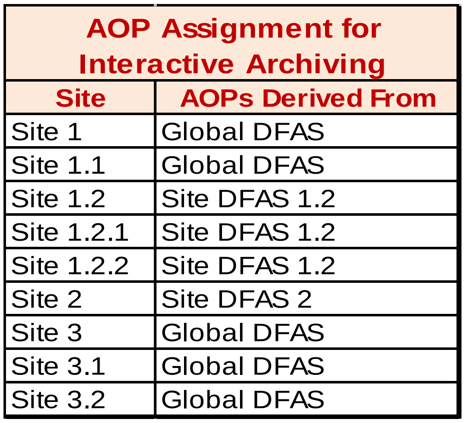
Generally speaking, for most users in most circumstances, it is not realistic to expect them to deal with these issues. Most of these questions have deep technical and information management implications. Few end users are qualified for this, and are not likely to be sufficiently motivated to invest sufficient time and trouble to figure it out. And to supply the AOPs effectively, they need to have a good understanding of the *IT configuration* (folder, metadata, permissions structure) and *IM configuration* (declaration status, classification accuracy) of both the source documents and destination location. Most likely, this will simply be too much to ask of the user.

This means a document really ought not to be (cannot realistically be) arbitrarily selected and archived by a SharePoint end user. A better approach is to pre-define the AOPs in advance for each document(s) you wish to archive, and have AGA automatically use the pre-determined AOPs during the archive operation.

AGA allows you to pre-define a set of AOPs in the form of a **DFAS** (**D**ocument/**F**older **A**rchive **S**ettings) file. There are two types of DFAS files – Global, and Site. The **Global DFAS** (there can only be one) specifies AOP settings for all SharePoint sites. **Site DFAS** files specify AOP settings for specified sites, which overrides the global DFAS settings. The AOP settings recorded in the global DFAS, and those in the various site DFAS files, are recorded in internal AGA XML files. The diagram below shows how the global and site DFAS files apply AOPs to various SharePoint sites.



In the diagram, note the global DFAS specifies AOP settings for all SharePoint sites, as they follow the hierarchical relationship of SharePoint sites. A site DFAS however has been applied to site 1.2 and site 2. Site 2 has AOP settings from Site DFAS 2. Site 1.2, and the subsites under it, have the AOP settings from Site DFAS 1.2. The following shows which sites have which settings:



AOP Assignment for Interactive Archiving	
Site	AOPs Derived From
Site 1	Global DFAS
Site 1.1	Global DFAS
Site 1.2	Site DFAS 1.2
Site 1.2.1	Site DFAS 1.2
Site 1.2.2	Site DFAS 1.2
Site 2	Site DFAS 2
Site 3	Global DFAS
Site 3.1	Global DFAS
Site 3.2	Global DFAS

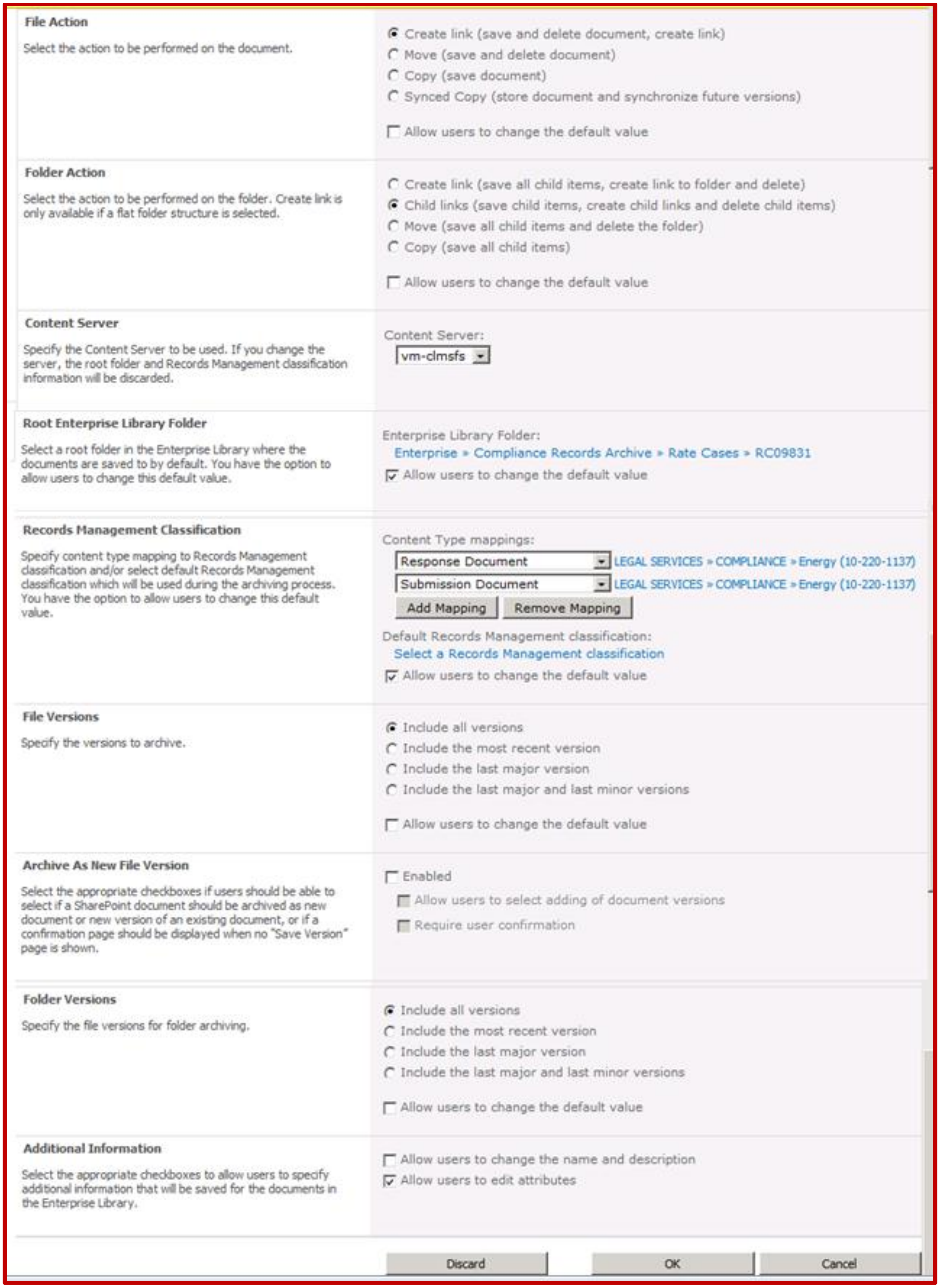
If you have defined a global DFAS, then the AOPs are known in advance, and an interactive archiving operation manually initiated by the SharePoint user becomes a single-click operation. The user can select a document(s) and trigger the operation in one of two ways:

Select from the document`s context menu **Send To → Enterprise Library**.

Click the **SEND** button on the Enterprise Library portion of the button bar.

If the user is in a SharePoint site with a site DFAS assigned, the global DFAS setting will be overridden by the local site DFAS settings.

The screencap on the following page shows the configuration settings for either the Global or Site DFAS. Note there are nine AOPs:



File Action

Select the action to be performed on the document.

● Create link (save and delete document, create link)
○ Move (save and delete document)
○ Copy (save document)
○ Synced Copy (store document and synchronize future versions)

☐ Allow users to change the default value

Folder Action

Select the action to be performed on the folder. Create link is only available if a flat folder structure is selected.

○ Create link (save all child items, create link to folder and delete)
● Child links (save child items, create child links and delete child items)
○ Move (save all child items and delete the folder)
○ Copy (save all child items)

☐ Allow users to change the default value

Content Server

Specify the Content Server to be used. If you change the server, the root folder and Records Management classification information will be discarded.

Content Server:
vm-clmsfs

Root Enterprise Library Folder

Select a root folder in the Enterprise Library where the documents are saved to by default. You have the option to allow users to change this default value.

Enterprise Library Folder:
Enterprise » Compliance Records Archive » Rate Cases » RC09831
☑ Allow users to change the default value

Records Management Classification

Specify content type mapping to Records Management classification and/or select default Records Management classification which will be used during the archiving process. You have the option to allow users to change this default value.

Content Type mappings:

Response Document	LEGAL SERVICES » COMPLIANCE » Energy (10-220-1137)
Submission Document	LEGAL SERVICES » COMPLIANCE » Energy (10-220-1137)

Default Records Management classification:
Select a Records Management classification
☑ Allow users to change the default value

File Versions

Specify the versions to archive.

● Include all versions
○ Include the most recent version
○ Include the last major version
○ Include the last major and last minor versions

☐ Allow users to change the default value

Archive As New File Version

Select the appropriate checkboxes if users should be able to select if a SharePoint document should be archived as new document or new version of an existing document, or if a confirmation page should be displayed when no "Save Version" page is shown.

☐ Enabled
☐ Allow users to select adding of document versions
☐ Require user confirmation

Folder Versions

Specify the file versions for folder archiving.

● Include all versions
○ Include the most recent version
○ Include the last major version
○ Include the last major and last minor versions

☐ Allow users to change the default value

Additional Information

Select the appropriate checkboxes to allow users to specify additional information that will be saved for the documents in the Enterprise Library.

☐ Allow users to change the name and description
☑ Allow users to edit attributes

The AOPs are:

File Action. This is the archiving mode that will be used when archiving a *document* (See Archiving Modes)

Folder Action. This is the archiving mode that will be used when archiving a *folder* (See Archiving Modes)

Content Server. Specify which OpenText Content Server you will be arching to.

Root Enterprise Library Folder. This is the folder location in Content Server where the document/folder will be stored. The Content Server and Root Enterprise Library Folder together form the specific target location in Content Server.

Folder Structure. (not shown). When archiving a folder with multiple documents and/or subfolders, AGA will limit the maximum number of documents per folder in Content Server, in order to preserve Content Server performance. It will distribute the documents across a number of folders. You can specify that the destination Content Server folder structure is to be configured in one of the following two ways:

Flat	Store all documents in the same folder. Maximum allowable documents per folder is AGA-configurable. Default recommended maximum is 400 documents.
Technical	Create a child folder for each X documents, where X can be specified in a configuration file.

RM Classification. Here you determine if the document/folder being archived is to be a record (declared as a record), and if so, its records classification. There are two components of this AOP:

Content Type Mappings	You can map a SharePoint content type to a Content Server RM Classification. The SharePoint Content Type of the document/folder will then determine the Content Server RM classification.
Default RM Classification	Force this specified Content Server RM classification on all documents/folders being archived, where a content type mapping as shown above is not in place.

File Versions. Here you specify how document (not folder) *versions* are to be handled. You must specify one of the following:

Archive all versions
Archive only the most recent version
Archive only last major version
Archive only last major and last minor version

Archive as New File Version. Here you can specify if the document/folder is archived as a completely new Content Server document, or if as an additional version of an existing Content Server document. You can also specify that if a document of the same name exists in Content Server, create a new version of that document.

Folder Versions. Here you specify how folder (not document) *versions* are to be handled. You must specify one of the following:

Archive all versions
Archive only the most recent version
Archive only last major version
Archive only last major and last minor version

Additional Information. A set of parameters that do not fit neatly into the above categories of settings. For example, you can specify if the user will be allowed to edit the metadata fields of the archived Content Server document. This is useful where the user may wish to change the metadata fields before storing it into Content Server.

Note that this is a highly configurable form. Most AOPs have an option called "Allow Users to Change the Default Value". If you select this option, then during archiving, the user will be presented with an onscreen prompt allowing them to edit (change) the value of that particular AOP. If you enable this for many parameters, the end user will see a lot of AOP parameters. This can add up to a lot of thinking and screen input time for the user, which may be discouraging. Try to automate as many as possible by using defaults. This way, a one-off interactive archiving operation becomes a one-click process (or at least as close to one-click as you can realistically get it).

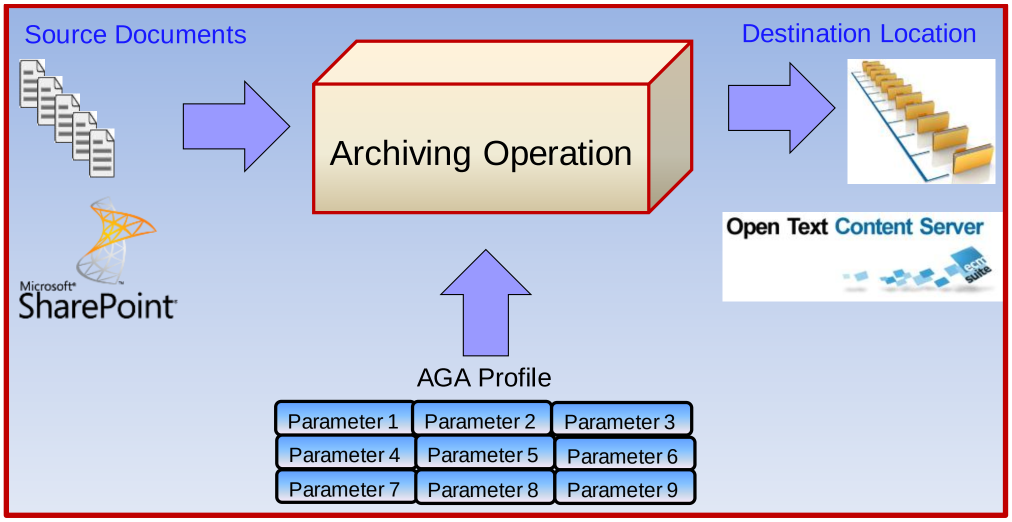
Your goal is to pre-define as much in the DFAS files as possible, which in turn will minimize the screen input of the users when they interactively archive a document or folder. Carefully plan a global DFAS that will apply to all sites, and then define a DFAS for each SharePoint site that introduces a set of circumstances to which the global DFAS parameters are not appropriate.

Profile-Based Archiving

Profile-based archiving refers to the means by which a SharePoint *library* or *site* can be automatically archived without any human intervention. You define the criteria (known as a rule) that must be met in order to initiate (trigger) the archive operation, and the remaining AOPs, and then record them in a configuration file known as a **Profile**. You assign a profile to a SharePoint site or a library, and define a schedule specifying when AGA will evaluate the profile. If the triggering conditions defined in the profile have been met, the archiving operation will proceed, based on the AOPs you recorded in the profile.

Remember that you cannot archive SharePoint **folders** using profile-based archiving – only **documents**.

To create an AGA profile, the AGA administrator answers the questions by filling out a series of mandatory inputs in the AGA configuration forms (SharePoint pages). AGA then records these answers internally in supporting XML files. When the operation is later initiated (triggered), AGA will extract the pre-determined archive operation parameters (AOPs) from the supporting XML files. The diagram below illustrates how a profile-based archiving operation requires AOPs for completion:



AGA allows you to name each unique profile, and assign it to a SharePoint site or library. You may define two types of profiles:

Site Profile Defines which library profiles (if any) will be applied to the libraries within the site. Can also specify if and how the entire site is to be archived. Must be assigned to a site.

Library Profile Defines *which* documents in a specified library are to be archived, and *how*. Must be assigned to a site profile, or directly to a library.

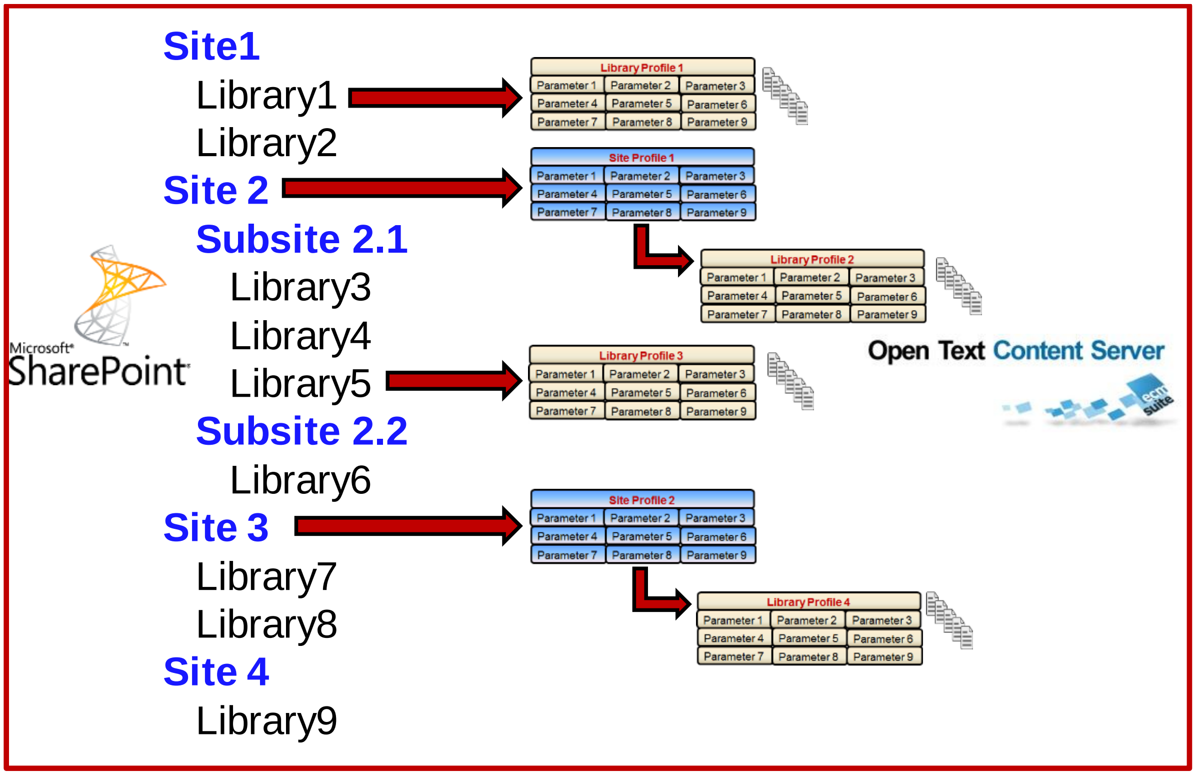
You first create site and library profiles, then **assign** them to the SharePoint structure (sites and libraries). Note that AGA site profiles do not specify *which* documents to archive, or *how* to archive them. Site profiles rely on library profiles to tell them this detail. Think of a site profile as a “container” of library profiles. Site profiles can be inherited downward within the SharePoint site hierarchy, whereas library profiles cannot. Hence you use site profiles to determine which library profiles to apply, and where to apply them.

A given SharePoint library can therefore gain a profile assignment in one of the following two ways:

Direct Assignment The library has an AGA library Profile assigned directly to it.

Site Profile Inheritance The library has no AGA library profile assigned, but is in a site that has an AGA site profile assigned above it.

An assigned library profile will over-ride any profile assignment above it in the SharePoint hierarchical structure. Refer to the diagram below for an example of profile assignments.



Note in the diagram that SharePoint has 6 sites (site 2 has 2 subsites) containing a total of 9 libraries. AGA has assigned (4) library profiles, and (2) site profiles. AGA site profiles have been assigned to *Site 2* and *Site 3*. AGA Library profiles have been assigned to *Library1, Site Profile 1, Library 5*, and *Site Profile 2*. The table below shows how each SharePoint library has received an AGA library profile:

Always refer to a profile as a **Site** Profile or a **Library** Profile, never just a profile. This will make AGA easier to understand and operate.

Library Profile Assignments	
SharePoint Library	Assigned AGA Library Profile
1	1
2	none
3	2
4	2
5	3
6	2
7	4
8	4
9	none

Library 1 has been directly assigned library profile 2. Site 2 has a site profile which has been assigned library profile 2. Assuming it inherits down (an AGA configurable option), all the libraries in Site 2 (libraries 3 through 6) will be assigned library profile 2, except for library 5 which has a direct assignment of library profile 3 (it over-rides library profile 2). Libraries 7 and 8 are each assigned library profile 4 via site profile 2. Libraries 2 and 9 do not have any library profile assignments, and will not be archived.

Similar to a DFAS, each library profile, and (indirectly via assignment) site profile contains AOPs that specify *what* documents to archive, and *how* to archive them to Content Server. However there are some major differences between a site/library profile and a DFAS:

Timing You have to specify *when* the archiving operation will occur. You do this by specifying a **schedule**[13] of operations, i.e. which days to run the operation, and when to repeat the operation. This is not required for a DFAS used for an interactive archive operation, as the user is manually initiating the operation.

Conditional Archiving may or not necessarily happen. Each automatic archiving operation specifies conditions that must be met for the operation to proceed, and/or or specifies documents that must meet certain criteria in order to be archived. It is possible that one or both conditions may not be satisfied, therefore documents may not be archived.

Parameters The AOPs are not exactly the same at those of an AOP. Some of the DFAS AOPs are the same, but different parameters are required for a site or library profile because an entire site or library introduces completely new considerations that have to be taken into account during archiving.

Programming Some custom SharePoint programming may be required to automatically trigger site archiving. Programming is not required for the actual archiving criteria (they are defined within the site/library profile). Only a very small amount of relatively simple programming is required. The programmer must be competent in CAML queries (to extend the onscreen query builder), C# programming language (if automatic triggering is required), and the SharePoint programming object model.

Establishing automatic (profile-based) archiving requires thorough planning and forethought, followed by careful implementation[14].Follow these high-level steps for profile-based archiving:

Step 1 Define and create library profiles. Specify which documents are to be archived, and how they are to archived (AOPs). Create a library profile for each library where the AOPs of its site's assigned library profile are not suitable (to create the needed exceptions to the default).

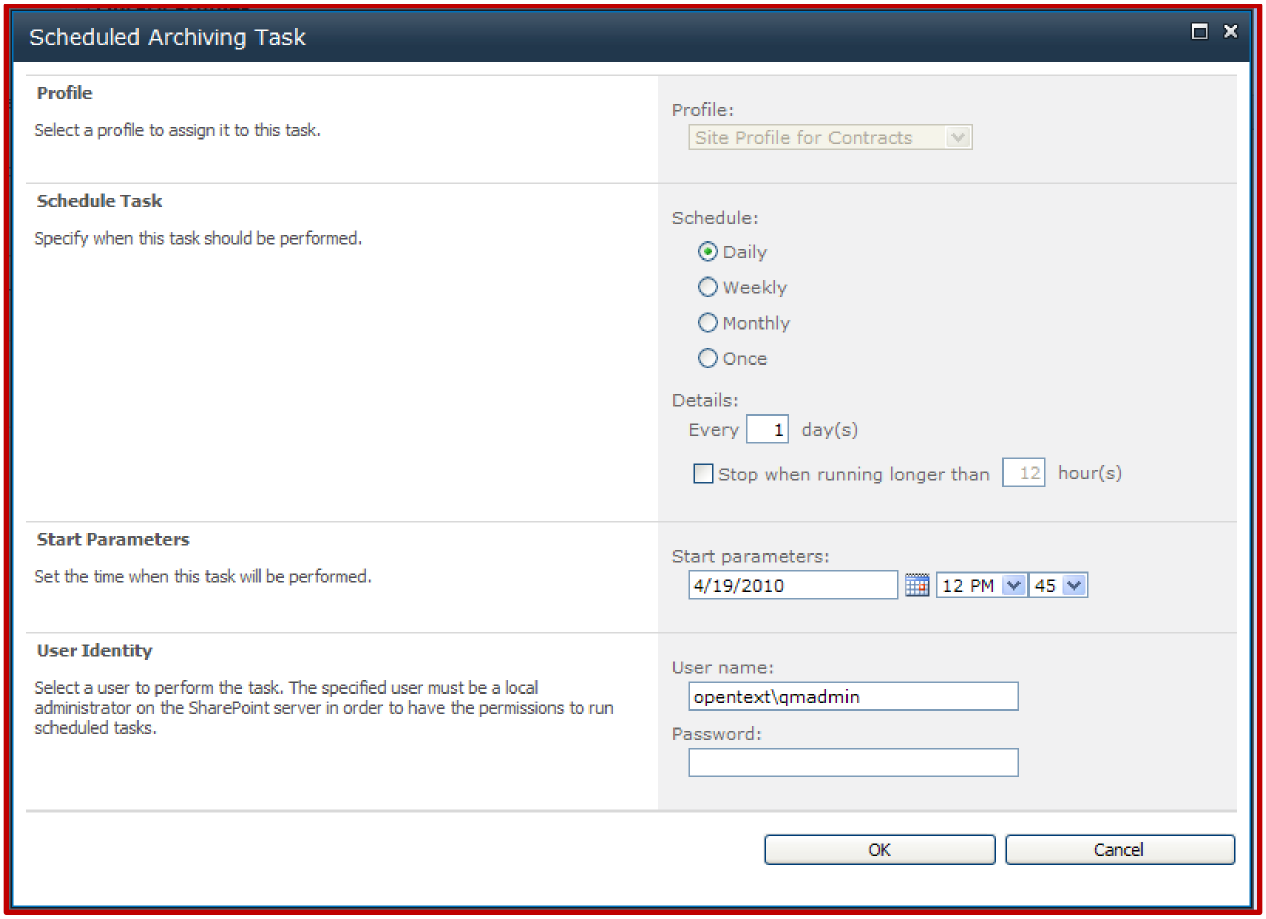
Step 2 Define and create site profiles. Specify the needed inheritance throughout the libraries of each site you wish to archive.

[13] Using **Windows Task Scheduler** facility. You can access this via the AGA Administrative pages in SharePoint.

[14] You can run AGA archiving in a test configuration before running actual archiving, in order to validate that your archiving scenario will work as expected. AGA will even generate detailed reports about the archiving operation (in both test and actual archiving modes).

Step 3	Assign the AGA site profiles to the SharePoint sites. Assign the AGA library profiles to the sites where inheritance is needed, and directly to the SharePoint libraries where exceptions are needed to the inherited defaults.
Step 4	Create a schedule of tasks, (daily, weekly, etc.), and assign the tasks to each of the site and library profiles as needed.

Automatic archiving is likely the easiest approach if your SharePoint information architecture is well defined and consistently implemented. AGA's site deployment capability can facilitate this, by using AGA to deploy sites that are structured in a consistent manner. Below is a screencap showing the options for defining a scheduled task:



Scheduled Archiving Task

Profile

Select a profile to assign it to this task.

Profile: Site Profile for Contracts

Schedule Task

Specify when this task should be performed.

Schedule:

● Daily

○ Weekly

○ Monthly

○ Once

Details:

Every 1 day(s)

□ Stop when running longer than 12 hour(s)

Start Parameters

Set the time when this task will be performed.

Start parameters: 4/19/2010 12 PM 45

User Identity

Select a user to perform the task. The specified user must be a local administrator on the SharePoint server in order to have the permissions to run scheduled tasks.

User name: opentext\qmadmin

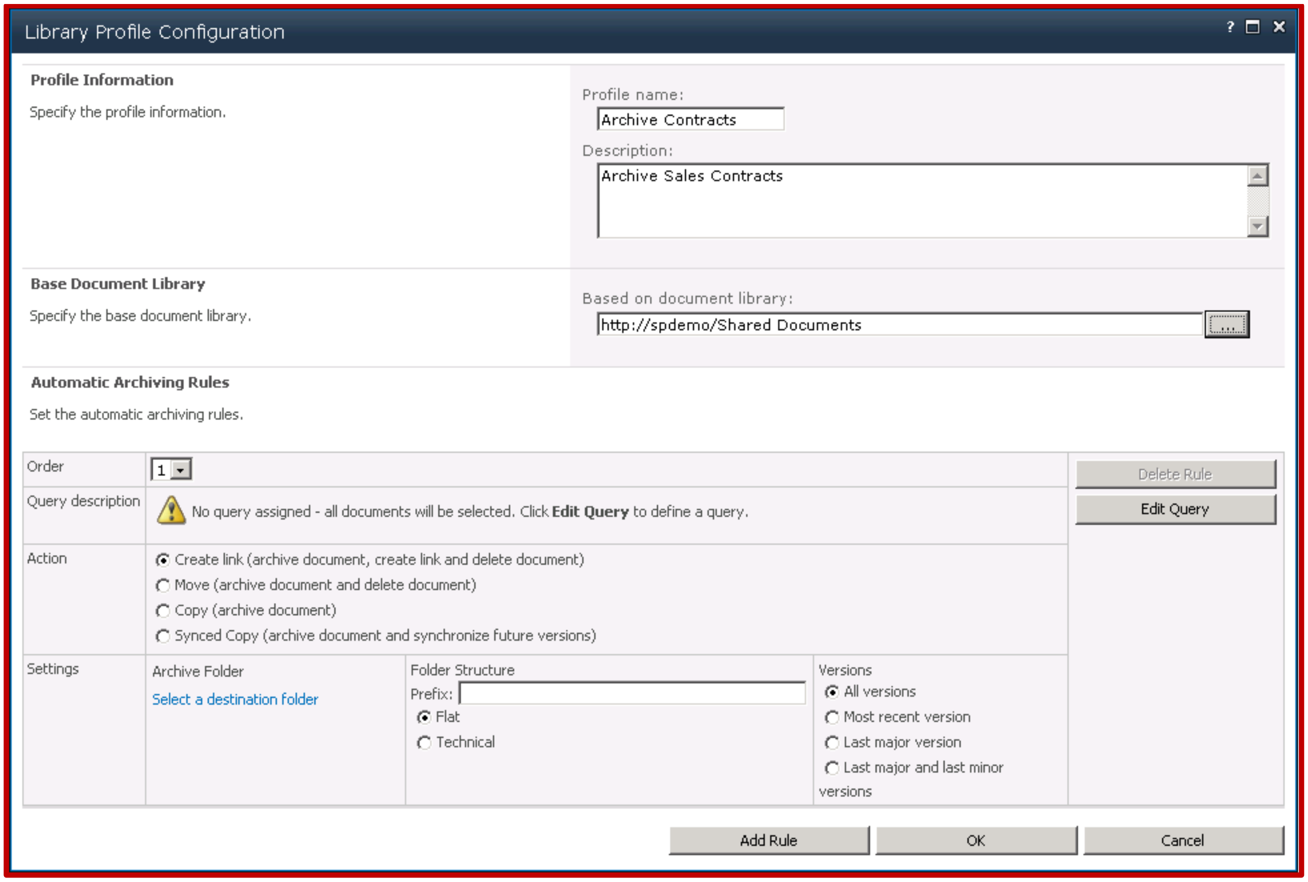
Password:

Defining a task schedule is straightforward, as you can see. You simply assign a schedule task to each of your site or library profiles.

Library Profile Configuration

The library profile specifies which documents to archive, and how to archive them. To specify which documents to archive, you define rules. Rules specify criteria that have to be met in order for archiving to occur. You can specify as many rules as you wish, and specify the order in which the rules must be met before archiving can proceed.

The all-important "how the documents are archived" relies on the same AOPs found in the DFAS, namely the *mode*, *version handling*, etc. The library profile configuration screen is shown below:



Library Profile Configuration

Profile Information

Specify the profile information.

Profile name: Archive Contracts

Description: Archive Sales Contracts

Base Document Library

Specify the base document library.

Based on document library: http://spdemo/Shared Documents

Automatic Archiving Rules

Set the automatic archiving rules.

Order: 1

Query description: No query assigned - all documents will be selected. Click **Edit Query** to define a query.

Action:
- ● Create link (archive document, create link and delete document)
- ○ Move (archive document and delete document)
- ○ Copy (archive document)
- ○ Synced Copy (archive document and synchronize future versions)

Settings:

Archive Folder: Select a destination folder

Folder Structure: Prefix:
- ● Flat
- ○ Technical

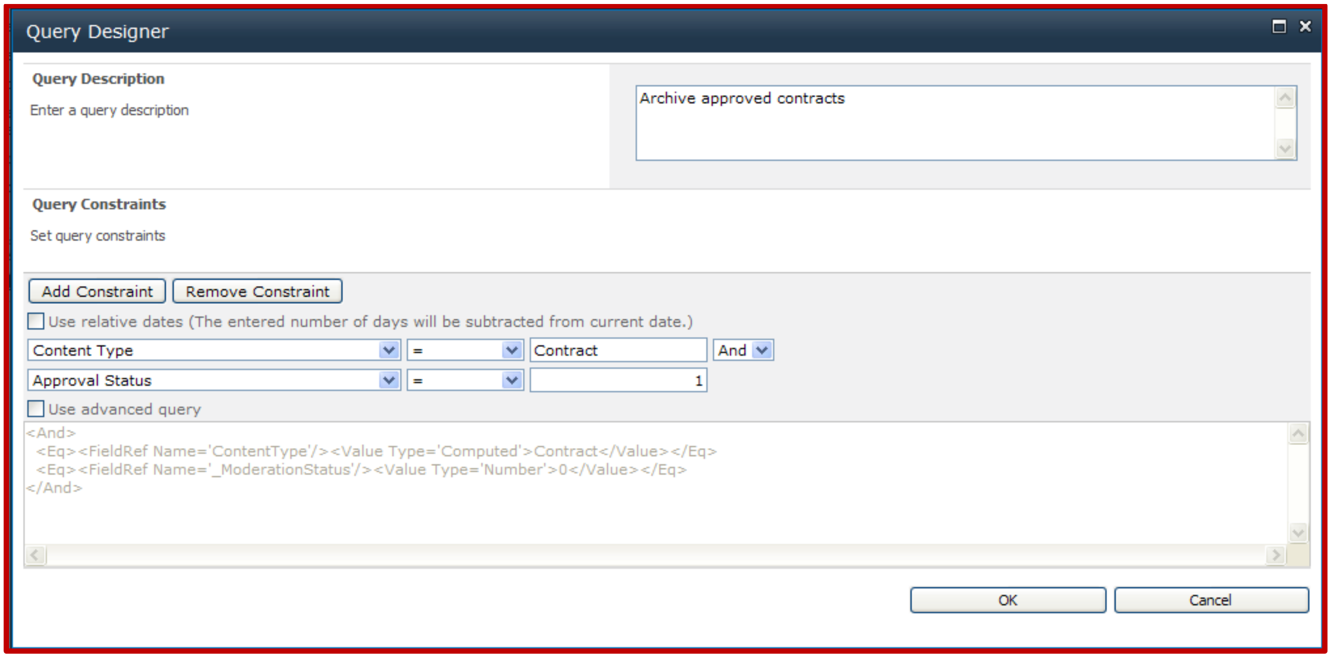
Versions:
- ● All versions
- ○ Most recent version
- ○ Last major version
- ○ Last major and last minor versions

The **Base Document Library** indicates the location (address) of the targeted library.

To build a rule, you have to define a **document query** that will identify the documents you wish to archive. You have to specify some document criteria such as:

Document Type =	**Contract**
Status =	**Finalized**
Age =	**greater than 5 years**

The query editing screen is shown below:



Query Designer

Query Description

Enter a query description

Archive approved contracts

Query Constraints

Set query constraints

□ Use relative dates (The entered number of days will be subtracted from current date.)

Content Type	=	Contract	And
Approval Status	=	1	

□ Use advanced query

```
<And>
  <Eq><FieldRef Name='ContentType'/><Value Type='Computed'>Contract</Value></Eq>
  <Eq><FieldRef Name='_ModerationStatus'/><Value Type='Number'>0</Value></Eq>
</And>
```

You can build a query by adding **constraints** as shown (document metadata fields, and Boolean relationships such as = or >). Some queries will not be possible to build this way, so you can use the CAML language to manually enter a query. By writing a query this way, you are not restricted to simple field selection. You can specify any conceivable query such as "*All documents from author "Smith" older than 5 years in library X or Y and without a legal hold and field Subject contains the phrase "Housing Crisis*"". During the archiving operation, those documents that meet the query criteria (rule) will be picked up for archiving.

You can run the profile as often as you wish (as scheduled), and each time it runs, AGA will execute the query. Each time, it may pick up some new documents that did not meet the criteria on the previous run. And it may pick up no documents at all. It's very important to note that with automatic archiving, you never really know what, if anything, is going to be archived. You therefore have to be careful to plan in advance what documents you wish have archived. Furthermore, be careful not to write a "loose" query that could pick up documents you never intended to archive. All your queries need to be explicit as possible to reduce accidental archiving.

The ACTIONS and SETTINGS parameters are similar to those in the DFAS, as illustrated earlier in this report.

Action. This is the archiving mode that will be used when archiving a *document* or *folder* (See Archiving Modes)

Folder Structure. Use this when archiving a SharePoint folder with multiple documents and/or subfolders. Select a Content Server target folder, and AGA will create subfolders according to the Content Server prefix. Content Server folders can include metadata elements such as the source SharePoint site and library names.

AGA will limit the maximum number of documents per folder in Content Server, in order to preserve Content Server performance. It will distribute the documents across a number of folders. You can specify that the destination Content Server folder structure is to be configured in one of the following two ways:

Flat Store all documents in the same folder. Maximum allowable documents per folder is AGA-configurable. Default recommended maximum is 400 documents.

Technical Create a child folder for each X documents, where X can be specified in a configuration file.

RM Classification. (not shown on the screencap). Here you determine if the document/folder being archived is to be a record (declared as a record), and if so, its records classification. There are two components of this AOP:

Content Type Mappings You can map a SharePoint content type to a Content Server RM Classification. The SharePoint Content Type of the document/folder will then determine the Content Server RM classification.

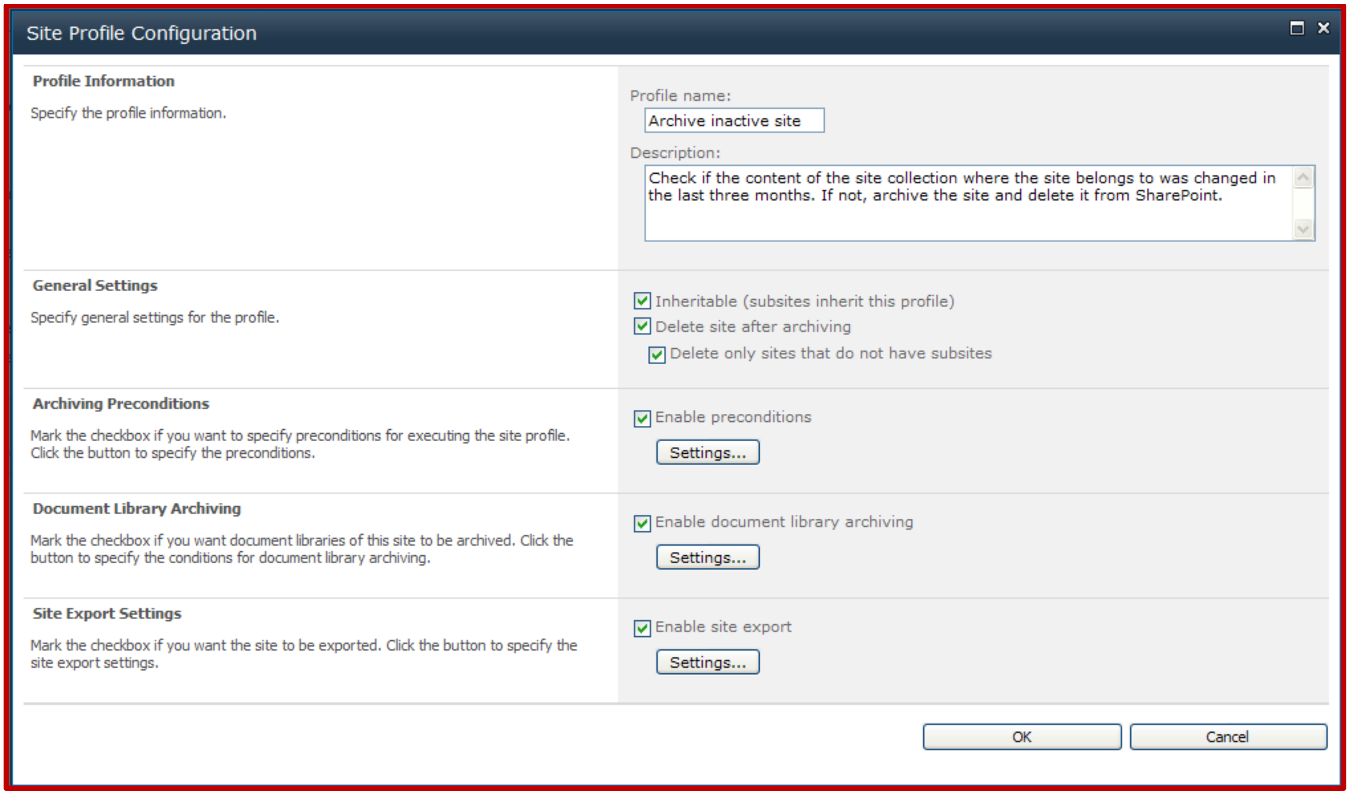
Default RM Classification Force this specified Content Server RM classification on all documents/folders being archived.

Versions. Here you specify how folder *versions* are to be handled. You must specify one of the following:

Archive all versions
Archive only the most recent version
Archive only last major version
Archive only last major and last minor version

Site Profile Configuration

Remember that a site profile is essentially a "container" for one or more library profiles. You use a site profile as a means of inheriting a library profile downward in the SharePoint library structure. In essence, you assign a library profile to a site profile, and instruct AGA to inherit it downward. If you assign an inherited site profile (with an assigned library profile), any new SharePoint libraries added to that SharePoint site in the future are automatically included in the archive operation, as they are picked up via the inheritance. The Site Profile Configuration screen is shown below:



Site Profile Configuration

Profile Information

Specify the profile information.

Profile name: Archive inactive site

Description: Check if the content of the site collection where the site belongs to was changed in the last three months. If not, archive the site and delete it from SharePoint.

General Settings

Specify general settings for the profile.

☑ Inheritable (subsites inherit this profile)
☑ Delete site after archiving
☑ Delete only sites that do not have subsites

Archiving Preconditions

Mark the checkbox if you want to specify preconditions for executing the site profile. Click the button to specify the preconditions.

☑ Enable preconditions

Document Library Archiving

Mark the checkbox if you want document libraries of this site to be archived. Click the button to specify the conditions for document library archiving.

☑ Enable document library archiving

Site Export Settings

Mark the checkbox if you want the site to be exported. Click the button to specify the site export settings.

☑ Enable site export

Here you enter all the details of a specific site profile, before you assign it to a SharePoint site. Parameters are:

Profile Information. Name and Description of this site profile.

General Settings. These settings affect the operation of the site profile as a whole:

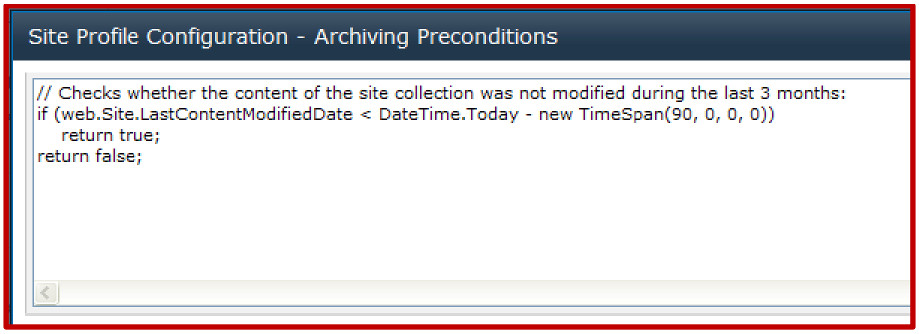
Inheritable. Forces site profile inheritance.

Delete Site after archiving. Delete entire site once it has been archived.

Delete only sites that have no subsites. Avoids loss of non-archived subsites.

Archiving Preconditions. Preconditions are the conditions that must be met before the site will be archived, such as "*No documents in this site have been modified for 2 years*". It is important to understand the difference between *preconditions* and a library profile's *query rules*. A library profile's query rules determines which specific documents in the target SharePoint site are to be archived. A site profile's preconditions however specify if the archive operation on a target SharePoint site will proceed or not. Site preconditions have nothing to do with the site's documents at all (this will be determined by the library profile). Put another way, the site profile preconditions determine if an archive operation will begin or not. Once the preconditions determine it is to proceed, the library profile's rules determine which documents in the specific target site will be archived.

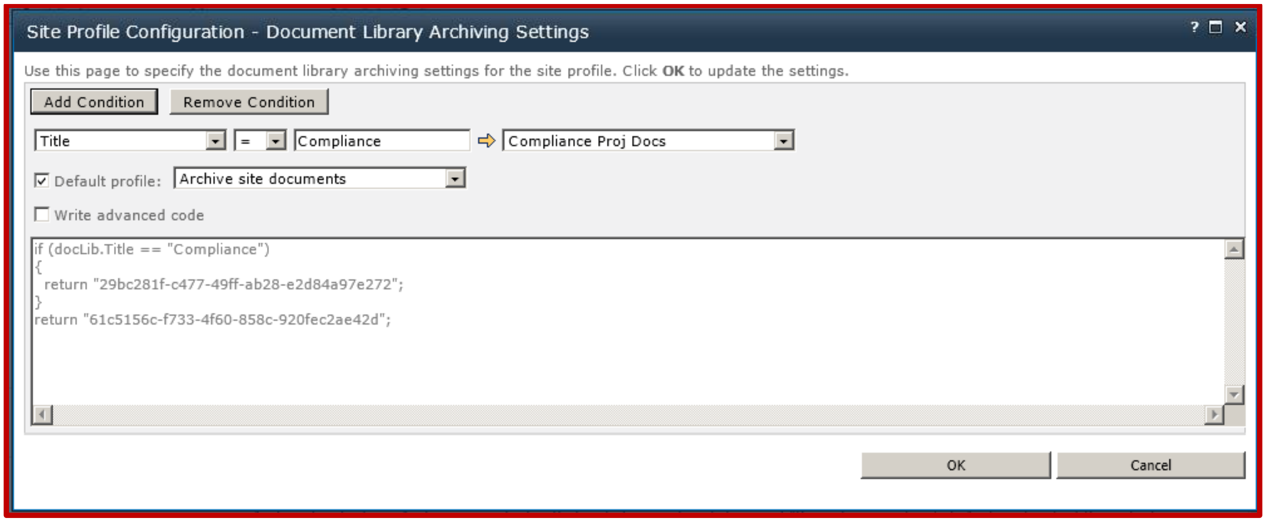
Programming is required to specify the preconditions. Using the C# programming language, enter your preconditions in the editing window as shown in the example below:



Site Profile Configuration - Archiving Preconditions

```
// Checks whether the content of the site collection was not modified during the last 3 months:
if (web.Site.LastContentModifiedDate < DateTime.Today - new TimeSpan(90, 0, 0, 0))
   return true;
return false;
```

Document Library Archiving. This tells AGA to use one or more assigned document libraries. If you click **Settings**, the following screen is presented:



Site Profile Configuration - Document Library Archiving Settings

Use this page to specify the document library archiving settings for the site profile. Click **OK** to update the settings.

Title = Compliance ⇨ Compliance Proj Docs

☑ Default profile: Archive site documents

□ Write advanced code

```
if (docLib.Title == "Compliance")
{
 return "29bc281f-c477-49ff-ab28-e2d84a97e272";
}
return "61c5156c-f733-4f60-858c-920fec2ae42d";
```

Here you can add multiple preconditions, in the same fashion as the query rules in a library profile. You also specify the ACTIONS and SETTINGS applicable to this profile. You basically select an existing document library profile as a default, then add conditions where other document library profiles should be used.

Site Export Settings. This allows you to **export** a complete site (everything within a site). An export is not an archive operation, but a complete system "dump" of a site. The site, everything it contains, and everything about that site is written out to a so-called CMP file. Sometimes referred to as an "image file", the CMP file is a more or less a "picture" of the site as it existed as of the moment it was exported. Once exported to Content Server, an export file is not searchable. If in the future you needed a document from the exported CMP file, you would have to **import** it (a SharePoint function) in its entirety back to SharePoint, then find the document in SharePoint. Click **Settings** to enter the following parameters:

Root Enterprise Library Folder. This is the folder location in Content Server where the document/folder will be stored. The Content Server and Root Enterprise Library Folder together form the specific target location in Content Server.

Folder Structure. Here you can define a Content Server **Prefix** which can represent (name) the individual folders with the SharePoint site you are archiving. AGA will automatically use this prefix to name new Content Server folders that represent the original source folders in SharePoint. For instance, an original SharePoint folder path *Safety/Incidents/Cases/Smith* may be prefixed in Content Server with the name of the SharePoint site *Corporate Safety*, to become:

Corporate Safety/Safety/Incidents/Cases/Smith

RM Classification. Here you assign a single RM classification to the entire archived (exported) site. This however can be over-ridden by any individual document library profiles that may be in place (which are assigning a RM classification). Assuming no individual library profiles are in play within this site, a single retention rule (e.g. destroy after 10 years) is applied to all the documents/folders within the archived site. Good recordkeeping practice would allow this if in fact the same rule applied to all the documents within the site, i.e. if all documents were of the same subject or case. If however the site contains documents of different subjects/cases, then a single rule would not be applicable. In this case you should use individual library profiles within the site to assign RM classifications to appropriate groups of documents of the same case/subject.

File Versions. Here you specify how document/folder *versions* are to be handled. You must specify one of the following:

Archive all versions
Archive only the most recent version
Archive only last major version
Archive only last major and last minor version

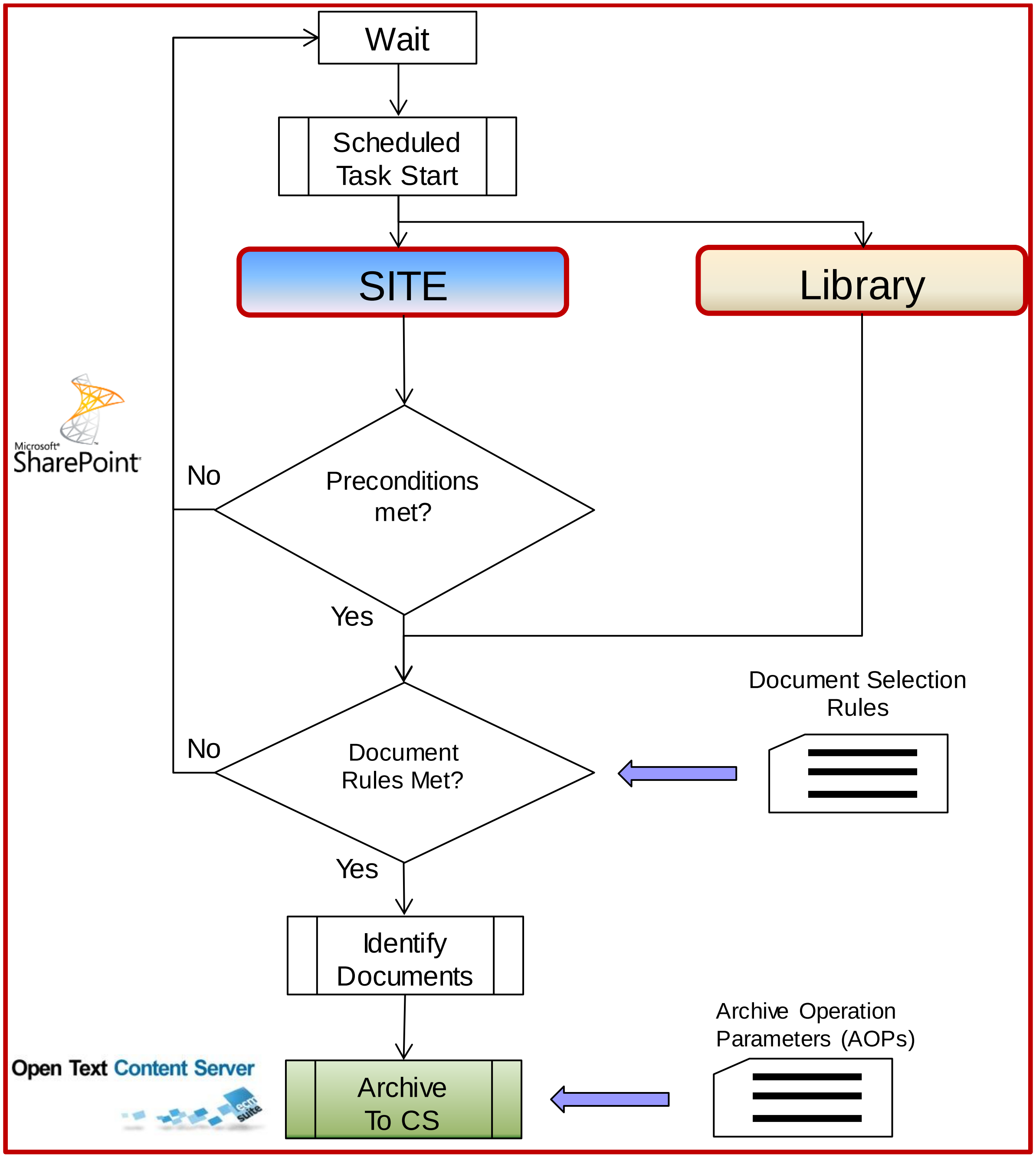
Security. Here you can specify the user and group information to be included with the exported site:

None. No user or group security settings are exported.
WSS Only. Roles only, not individual user information
ALL. Roles and user and group details.

There is a notable issue about a site export that involves SharePoint **lists**. Lists are archived as HTML files. Any attachments of documents within the list are automatically archived in subfolders. AGA creates, names, and organizes the required folders and subfolders in the same manner as document library profiles (as shown earlier). If you need to archive lists in a different manner, you can custom-build a **Provider** to do the job, using the AGA software development kit.

Process Flow

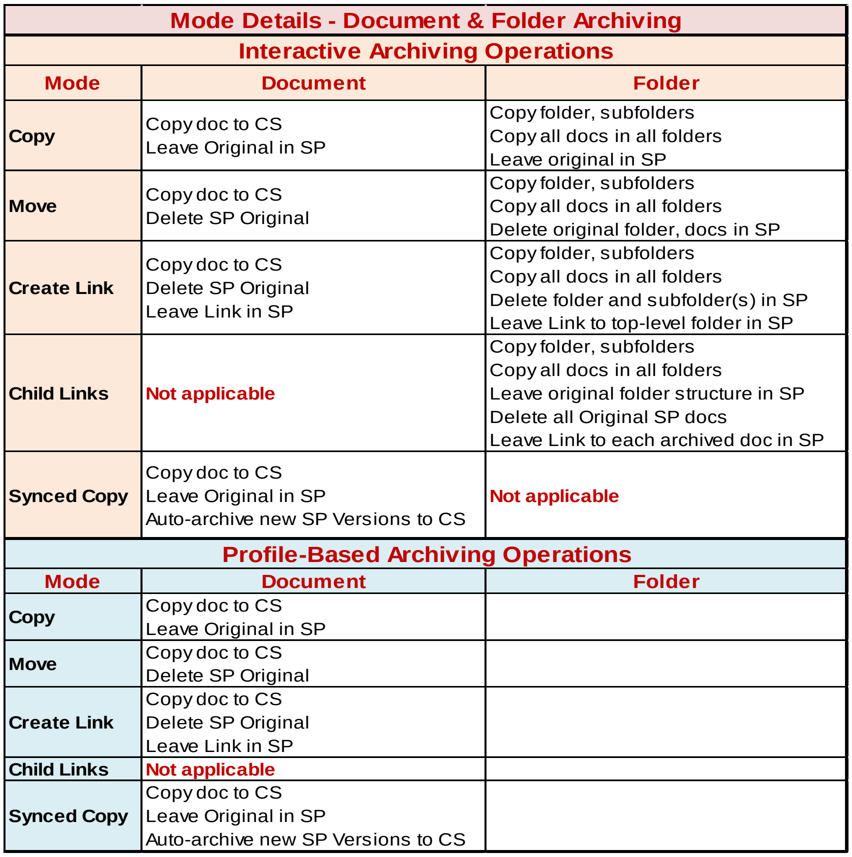
The diagram that follows illustrates the process of profile-based archiving, at a high level:



Schedule tasks are assigned to site profiles or library profiles. For a site profile, if the preconditions are met, AGA proceeds to process the assigned library profile parameters. Library profile rules are evaluated, and if one or more documents/folders are qualified for archiving, then AGA proceeds to identify the target documents/folders. It then starts the actual archive operation, based on the AOPs supplied by the library profile.

Archiving Modes

As stated earlier, AGA can move a document from SharePoint to Content Server, however there is more to it than a simple move. There are actually five different methods by which AGA can archive a given document. RIMtech defines each of these methods as a **Mode**. The mode is one of the AOPs that must be defined for each archive operation. The mode specifies how the document is to be archived. AGA refers to these modes as **Actions**. The details of each mode can differ, depending on if a *document* or a *folder* is being archived. The five modes are shown in the table below:



Mode Details - Document & Folder Archiving

Interactive Archiving Operations

Mode	Document	Folder
Copy	Copy doc to CS Leave Original in SP	Copy folder, subfolders Copy all docs in all folders Leave original in SP
Move	Copy doc to CS Delete SP Original	Copy folder, subfolders Copy all docs in all folders Delete original folder, docs in SP
Create Link	Copy doc to CS Delete SP Original Leave Link in SP	Copy folder, subfolders Copy all docs in all folders Delete folder and subfolder(s) in SP Leave Link to top-level folder in SP
Child Links	**Not applicable**	Copy folder, subfolders Copy all docs in all folders Leave original folder structure in SP Delete all Original SP docs Leave Link to each archived doc in SP
Synced Copy	Copy doc to CS Leave Original in SP Auto-archive new SP Versions to CS	**Not applicable**

Profile-Based Archiving Operations

Mode	Document	Folder
Copy	Copy doc to CS Leave Original in SP	
Move	Copy doc to CS Delete SP Original	
Create Link	Copy doc to CS Delete SP Original Leave Link in SP	
Child Links	**Not applicable**	
Synced Copy	Copy doc to CS Leave Original in SP Auto-archive new SP Versions to CS	

You need to understand, in detail, exactly what each mode does, and any possible considerations or implications of each mode. There are advantages and disadvantages to each mode. And you have to consider the differences between archiving a *Document* or a *Folder* in that particular mode.

Copy

For documents, the original SharePoint document is copied to Content Server, leaving the original behind in SharePoint. For folders, the original SharePoint folder, all documents in that folder, any subfolder(s), and all documents in all subfolders, are copied to Content Server. The Copy mode should not be used for profile-based archiving operations because the automatic profiling would re-archive the original documents/folders left behind, which have already been archived. Some might consider this mode to be essentially a means of backing up important documents in SharePoint.

Advantages

1) SharePoint users of the originals retain 100% native functionality within SharePoint, without any compromises whatsoever.
2) Ease of Retrieval. The document can be searched in either Content Server or SharePoint.

Disadvantages

1) You are consuming double the storage space, as you are storing the documents in 2 systems.
2) This mode is poor information management practice, as the original in SharePoint is an unmanaged copy.

Search Considerations

SharePoint User	No impact, all native search capabilities preserved.
Content Server User	No impact.

Recordkeeping Considerations

Avoid this mode where possible, as it results in an unmanaged duplicate of every archived record. Even if the original was a fully managed record, you have to run disposition twice, in two difference systems. Sooner or later, you'll need to dispose of the originals, which forces you to run a disposition in SharePoint, which does not have a usable native disposition process.

For folders, it is important that each document is classified correctly. If the subject of the documents changes within a folder, or among folders, you will have to ensure that all documents of the same subject are classified with the same RM Classification (**records series**). Either the user will have to manually choose the records series, or you may be able to achieve RM classification through a SharePoint Content Type Mapping.

Move

When archiving a document, the original copy of the archived document is removed entirely from SharePoint. With a folder, the original folder tree (the folder and any subfolders, and any documents they contain) are removed entirely from SharePoint. This is ideal when the work on a SharePoint site or project completely ends, and you need to relocate the documents to a permanent repository, governed by recordkeeping. If SharePoint users seek out the documents after they have been archived however, they are no longer in SharePoint. They will have to use the EL Web Part to find the documents that have been archived. Note that the EL web part can be on the same page the user normally uses to search for documents in SharePoint, making it easier to search for both SharePoint and Content Server documents.

Advantages

1) No duplication of storage space.
2) No unmanaged duplicate copy in SharePoint.

Disadvantages

1) No native search capability in SharePoint – you will have to merge the search of SharePoint and Content Server.
2) The SharePoint user must run the EL web part to work on the document in SharePoint. Will need to retrieve from Content Server back to SharePoint workspace (which then creates an unmanaged duplicate). However, if the SharePoint user has sufficient permissions to revise the document in Content Server, they can work on it directly inside Content server, without needing to retrieve a copy to SharePoint.

Search Considerations

SharePoint User Options for searching for the archived document:

1) Use the EL Web Part to search Content Server for the document.
2) Log into Content Server and use it independently of SharePoint.
3) Merge (cross-index) Content Server and SharePoint searching, so a SharePoint search yields both Content Server and SharePoint results.

Content Server User No impact, as documents are in Content Server.

Recordkeeping Considerations

From a recordkeeping perspective, this is the ideal mode. There is no unmanaged coy, and there is no doubling of storage space. There are no residual links or anything to “clean up” later. Full official disposition can carried out in Content Server.

Create Link

When archiving a document, the original copy of the archived document is removed entirely from SharePoint. A link is left behind in SharePoint that the user can click on to retrieve the document (from Content Server). With a folder, the original folder tree (the folder and any subfolders, and any documents they contain) are removed entirely from SharePoint. A link is left behind in SharePoint that the user can click on to display the archived folder (in Content Server). The SharePoint context menu of the archived folder is intact, in that you can view folder information and audit details, etc. This mode is suitable when the SharePoint users of a document or folder need to continue working on a document (in SharePoint) once it has been archived to Content Server. By leaving a link to the document/folder behind in SharePoint, the SharePoint user will be able to view the document and do most of the things on it that would have been possible if it was still physically in SharePoint. This mode is somewhat of a compromise between a COPY and a MOVE.

Advantages

1) A reference (link) to the archived document in its final location for easy retrieval.
2) SharePoint users of the original can still view, search for, and work with the original in SharePoint, even though it is physically stored in Content Server.
3) No duplication of storage space.
4) No unmanaged duplicate copy in SharePoint.

Disadvantages

1) A link has limitations on what the SharePoint users can do with the document in SharePoint (detailed later in this report).

Search Considerations

SharePoint User Options for searching for the archived document:

1) Use the EL Web Part to search Content Server for the document.
2) Log into Content Server and use it independently of SharePoint.
3) Merge (cross-index) Content Server and SharePoint searching, so a SharePoint search yields both Content Server and SharePoint results.

Content Server User No impact, as documents are in Content Server.

Recordkeeping Considerations

The considerations are the same as for a document/folder MOVE. A link is not really an unmanaged duplicate of an archived document. If however the SharePoint user retrieves a copy of the document to the SharePoint workspace, an unmanaged copy is created. Once the archived documents are deleted in Content Server (with the records disposition module), the remaining links left behind in SharePoint are now invalid, and will need to be cleaned up via an AGA link cleanup utility (detailed later in this report).

Child Links

This is a variation of the CREATE LINK mode. This mode is only available when archiving folders, not for individual documents. The original SharePoint folder tree (the folder and any subfolders, and any documents they contain) is left behind in SharePoint. Each document that was in the original folder tree is archived to Content Server and replaced with a link to the archived document in Content Server. The difference between this mode and the Create Link mode is summarized below:

Create Link	**Child Links**
Deletes original SharePoint folders	Preserves original SharePoint folder tree
One link for top folder	One link for each individual document

Advantages

1) SharePoint users of the original can still view, search for, and work with the original in SharePoint, even though it is physically stored in Content Server.
2) No duplication of storage space.
3) No unmanaged duplicate copy in SharePoint
4) User's familiar SharePoint folder structure remains, to facilitate finding documents.

Disadvantages

1) A link has limitations on what the SharePoint users can do with the document in SharePoint (detailed later in this report).

Search Considerations

SharePoint User Options for searching for the archived document:

1) Use regular SharePoint search.
2) Use the EL Web Part to search Content Server for the document.
3) Log into Content Server and use it independently of SharePoint.
4) Merge (cross-index) Content Server and SharePoint searching, so a SharePoint search yields both Content Server and SharePoint results.

Content Server User No impact, as documents are in Content Server.

Recordkeeping Considerations

The considerations are the same as for the document/folder CREATE LINK mode.

Synced Copy

This is a variation of the COPY mode. This mode is only available when archiving documents, not for folders, as folders do not have versions. The original SharePoint document is copied to Content Server, leaving the original behind in SharePoint. Like Copy, you should not use this mode for profile-based archiving operations because the automatic profiling would re-archive the original documents/folders left behind, which have already been archived.

Once archived, if the original archived document in SharePoint is edited and saved (incrementing its version number), the new (revised) version can be "automatically" archived to Content Server. The new version will be copied to Content Server, leaving the original new version behind in SharePoint. This way, the document's progress over time is automatically "followed" by Content Server. The versioning feature in SharePoint must be enabled for the target document. so that all saved changes to the original document are saved as a new versions (the version number is incremented.

Suppose you have a SharePoint document archived to Content Server in Synced Copy mode. Once you revise it in SharePoint, the newly-created revision does not immediately get archived to Content Server. You have to create a Windows Scheduler Task that runs periodically on SharePoint to examine Content Server for any documents that need to have their versions copied to Content Server. Therefore the time lag between the new revision in SharePoint and the matching synced copy in Content Server will depend on the frequency of the scheduled update task.

Suppose SharePoint document 707V1 (version 1) is archived to Content Server. Three subsequent revisions will result in the following (4) documents in SharePoint, and (4) in Content Server, for a total of (8) documents:

SharePoint (Original)	**Content Server (Copy)**
707V1	707V1
707V2	707V2
707V3	707V3
707V4	707V4

Advantages

1) SharePoint users of the originals retain 100% native functionality within SharePoint, without any compromises whatsoever.
2) Ease of Retrieval. The document can be natively searched for in either Content Server or SharePoint.
3) Revisions are automatically archived to Content Server without any user intervention.

Disadvantages

1) You are consuming double the storage space, as you are storing the documents in 2 systems.
2) This mode is poor information management practice, as the originals in SharePoint are unmanaged copies.

Search Considerations

SharePoint User	No impact, all native search capabilities preserved.
Content Server User	No impact, as documents are in Content Server.

Recordkeeping Considerations

This mode has an obvious advantage of automatically capturing each revision to the document, without having to depend on any users. On the other hand, the original (and all versions of the original) are unmanaged duplicates. If you use this mode, you will have to run disposition twice, in two difference systems. Sooner or later, you'll need to dispose of the originals, which forces you to run a disposition in SharePoint, which does not have a usable native disposition process. Consider tracking the original documents and going into SharePoint at some point in the future to manually delete them.

Link Cleanup

After running a disposition in Content Server, documents are deleted from Content Server. Any links left behind in SharePoint that reference the deleted documents are now "orphaned" – they will no longer work, as they point to documents that no longer exist in Content Server.

Obviously you would not want orphaned links to exist for very long in SharePoint. AGA lets you run a special process to clean up orphaned links. In the SharePoint Administration console, you can schedule an AGA **Link Cleanup** utility to run as often as you wish (daily, weekly, monthly, etc.).

Records Management

For recordkeeping compliance, you have to achieve two things:

1) Get the SharePoint documents that are records, declared as records in Content Server
2) Classification accuracy of declared records in Content Server must exceed minimum acceptable threshold (assume 90%).

If the SharePoint document is already a declared record, and you have confidence in the accuracy of its classification, then you have no further recordkeeping to do – you simply need to move (archive) the document from SharePoint to Content Server. You do however have to ensure you are properly transferring all the SharePoint metadata fields to Content Server (including the vital classification field), and that the destination document in Content Server is properly declared as a record.

For the purpose of this report, we will assume the SharePoint document is in fact a record, but has not been declared as a record in SharePoint. We base this on the assumption we are not declaring records in SharePoint, as it does not have a compliant means of doing so. This is why we would normally wish to archive the document in the first place – to transfer it from a non-compliant system (SharePoint) to a compliant system (Content Server). So your challenge is to get the document declared as a record, and properly classified, during the archiving operation.

> Never lose sight of the fact that you have to ensure classifications are correct **9 times out of 10**.

Declaration and classification are theoretically two separate activities, however AGA combines them both into something they refer to as **RM Classification**. RM classification in AGA is one of the AOP parameters that must be specified (if desired – if not, the document will be archived in Content Server as a non-record). Combining the two makes sense, as a document is not a record without being assigned a classification. To classify a document, an appropriate retention rule from the file plan must be assigned to it. In AGA, RM Classification means that it will be declared in Content Server as a record, and that one of the valid file plan retention rules from the file plan, stored in Content Server, has been assigned to it. Content Server refers to a retention rule as a **Records Series**. All of the following four terms each refer to the same thing – a retention rule from the file plan:

Records Series	OpenText terminology for a retention rule
Category	US DoD 5015.2 definition of a retention rule in the file plan
RM Classification	Records Series assigned to a SharePoint document archived to Content Server as a record
Retention Rule	General descriptive language

We will endeavour to consistently use the OpenText terminology Records Series. Remember that the file plan is stored in Content Server, not in SharePoint. Getting the RM Classification parameter assigned to a document during an archiving operation is simple. Your real challenge is to ensure that the classification is accurate 9 times out of 10 (or whatever your classification accuracy threshold dictates).

AGA provides a different means of archiving a document as a record, depending on whether you are archiving interactively, or using the profile-based method.

Interactive Records Archiving

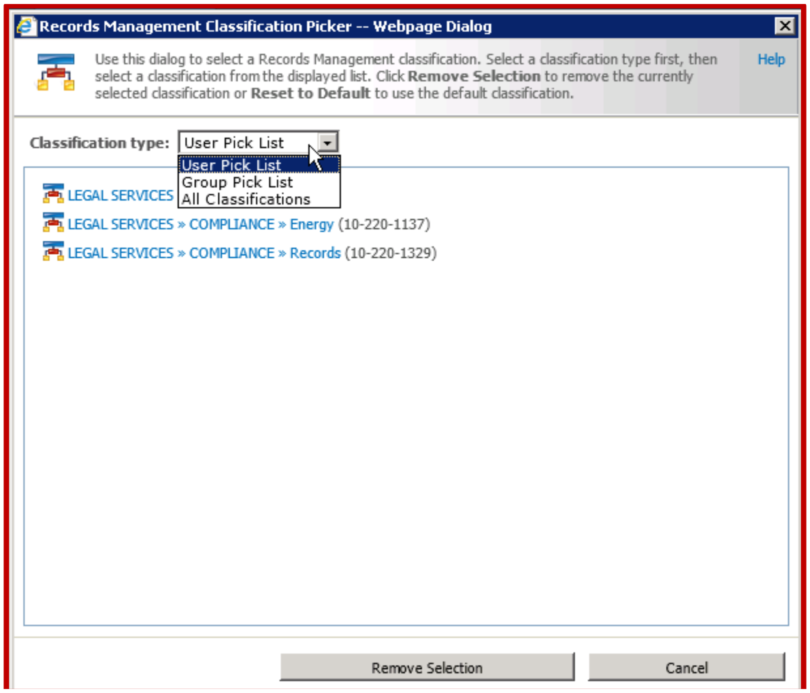
Recall that you control interactive archiving AOP parameters by recording them in a DFAS. AGA provides three different approaches to assigning an RM classification:

Default
Manual User entry
Content Type Mapping

You can specify a default RM classification in the Global or Site DFAS. This value will then apply to all documents interactively archived. The user will not have to specify a RM classification – all documents archived will be a record, classified to the RM classification that you specify as a default. Obviously, this is only applicable if all documents are of the same subject or case. If for example the default RM classification is (to use an example from the sample file plan show in Appendix 4) *Contracts – Safe-Tee Security* (04-50-1), then any document archived with this RM classification that is not about Safe-Tee's Security's contract will be incorrectly classified.

It would make more sense to specify a default value for each library, where the library is known to contain documents of like subject, that can safely be classified against the same records series (to use the Content Server language to describe the rule). This would work well for a SharePoint library for instance that stored exclusively documents about *Safe-Tee Security*'s contract.

There are times when the default will not be correct, and you need to allow the user to over-ride the default setting you pre-defined in the DFAS. In the DFAS you can specify **Allow User to Change Default Value**, and they will be able to change the default value by clicking on the field. The user is then presented with the **Classification Picker** as shown:



Use this dialog to select a Records Management classification. Select a classification type first, then select a classification from the displayed list. Click **Remove Selection** to remove the currently selected classification or **Reset to Default** to use the default classification.

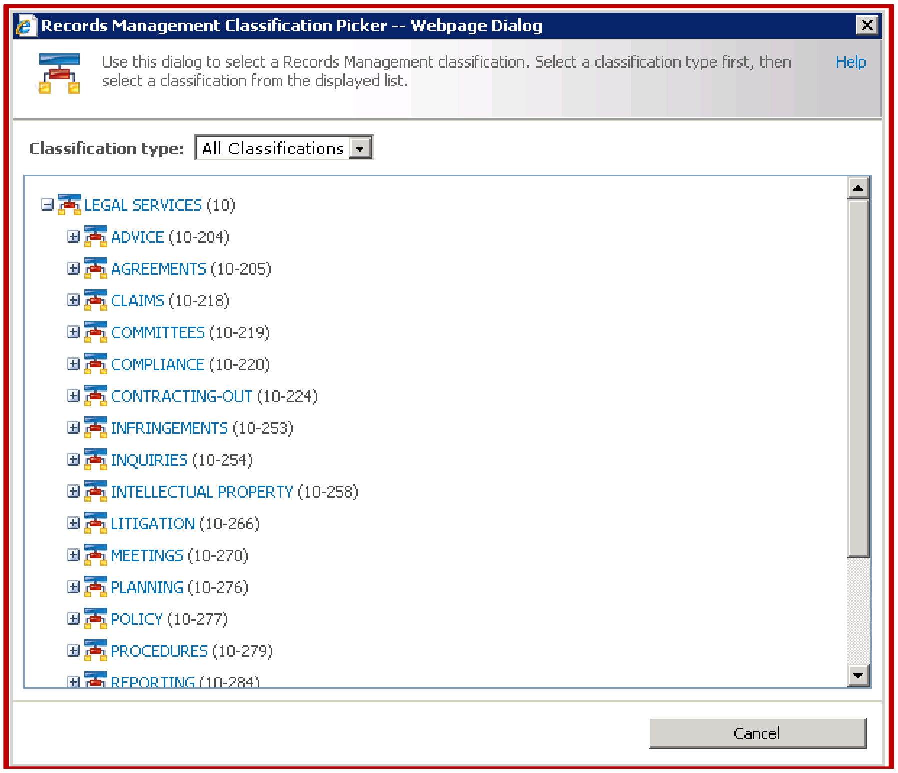
Classification type: User Pick List

LEGAL SERVICES

LEGAL SERVICES » COMPLIANCE » Energy (10-220-1137)

LEGAL SERVICES » COMPLIANCE » Records (10-220-1329)

Depending on how the user's account is set up in Content Server, they may be presented with a number of optional **classification types**[15], as shown above. The **User Pick List** is a personalized list of records series tailored to that user. The **Group Pick List** is a list of records series for that user, and others in their group. **All Classifications** presents the entire file plan for browsing. Below is a screencap of all classifications, where the user can browse the full file plan:



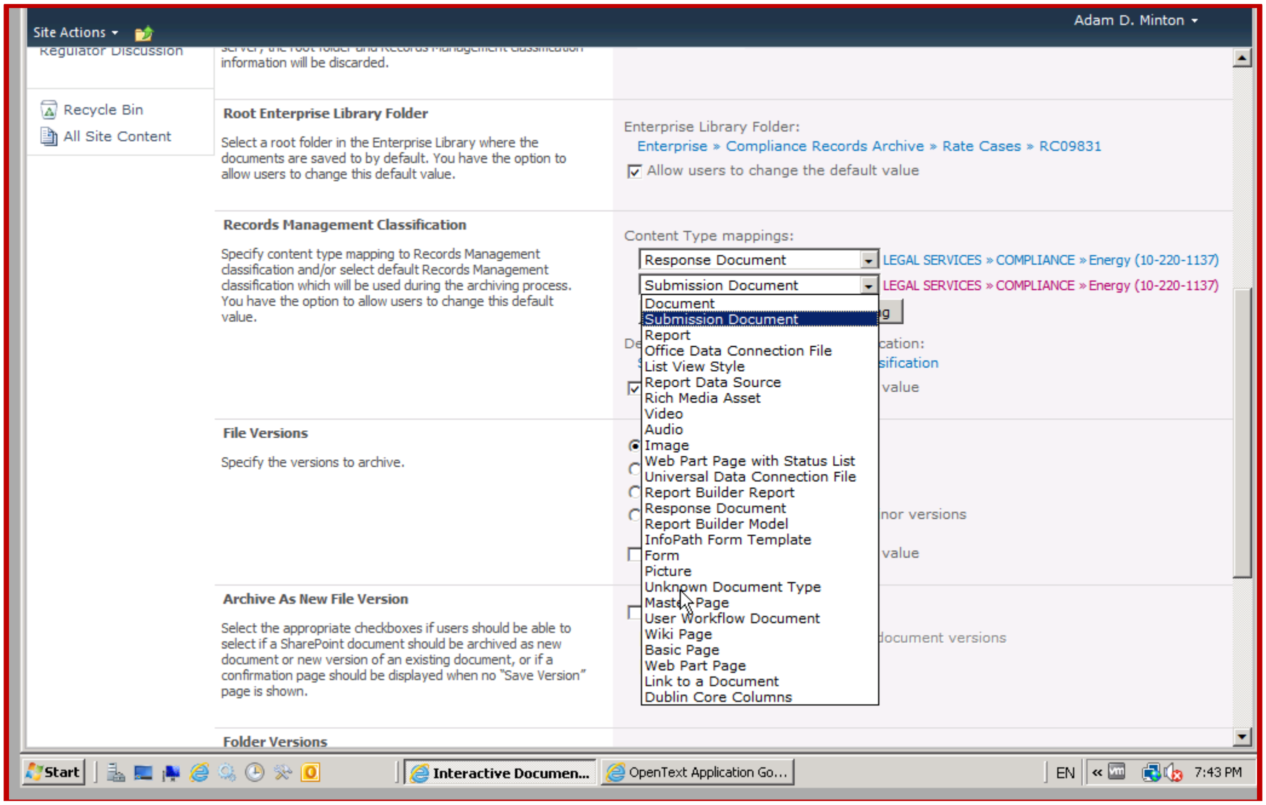
Use this dialog to select a Records Management classification. Select a classification type first, then select a classification from the displayed list.

Classification type: All Classifications

- LEGAL SERVICES (10)
  - ADVICE (10-204)
  - AGREEMENTS (10-205)
  - CLAIMS (10-218)
  - COMMITTEES (10-219)
  - COMPLIANCE (10-220)
  - CONTRACTING-OUT (10-224)
  - INFRINGEMENTS (10-253)
  - INQUIRIES (10-254)
  - INTELLECTUAL PROPERTY (10-258)
  - LITIGATION (10-266)
  - MEETINGS (10-270)
  - PLANNING (10-276)
  - POLICY (10-277)
  - PROCEDURES (10-279)
  - REPORTING (10-284)

[15] A method in Content Server whereby a subset of all possible classifications is presented to the user, as a convenience.

Note that few users typically have sufficient motivation to invest the time and effort to select a RM classification. Even if they did, they will not always choose the correct records series. Therefore, you should try to accomplish RM classification automatically without user intervention, where possible. AGA can help with a technique called **Content Type Mapping** (not be confused with AGA's other Content Type Mapping for SharePoint vs. Content Server metadata field matching). With this approach, you assign (map) a RM classification for each Content Type in the target SharePoint library, as shown:



information will be discarded.

Root Enterprise Library Folder

Select a root folder in the Enterprise Library where the documents are saved to by default. You have the option to allow users to change this default value.

Enterprise Library Folder:
Enterprise » Compliance Records Archive » Rate Cases » RC09831
☑ Allow users to change the default value

Records Management Classification

Specify content type mapping to Records Management classification and/or select default Records Management classification which will be used during the archiving process. You have the option to allow users to change this default value.

Content Type mappings:
Response Document — LEGAL SERVICES » COMPLIANCE » Energy (10-220-1137)
Submission Document — LEGAL SERVICES » COMPLIANCE » Energy (10-220-1137)

File Versions

Specify the versions to archive.

Archive As New File Version

Select the appropriate checkboxes if users should be able to select if a SharePoint document should be archived as new document or new version of an existing document, or if a confirmation page should be displayed when no "Save Version" page is shown.

Folder Versions

Recall that the SharePoint Content Type determines a document's metadata fields. Every document in SharePoint must be assigned a content type – this is how SharePoint knows which metadata fields will be assigned to the document. This AGA classification method is based on the assumption that you can use metadata values to achieve classification. For instance, suppose the SharePoint document had a content type called *Budgeting*. All documents of content type *Budgeting* will be about budgeting. We can safely assign RM classification *Finance/Budgeting* (04-30).

This technique will obviously work ***only*** where you have a one-to-one relationship between content type and document subject. You will have to confirm that all documents of that content type are of exactly the same subject in the file plan, i.e. all are about *budgeting*.

This technique will ***not*** work where there is a possibility that the content type leads to documents of more than one subject. Take case files for example. If (from our sample file plan) we had a content type *Contracts*, it would not work. This is because *Contracts* is too general. On our file plan it only takes us to *Finance/Contracts* (04-50). However we need to know the specific contractor in order to achieve proper classification. Is it *Safe-Tee Security* or *Tasty Catering*? If there were documents about each contractor with content type *Contracts*, we would be classifying all of the documents against records series 04-50, which is obviously not correct. In order for this content mapping method to work with case files, there would have to be a SharePoint content type for each case file. While technically feasible, this would amount to entering your entire file plan into the Content Type structure of SharePoint, which is very poor architecture design, on a number of levels. If for instance you had 10,000 employees, you would most certainly not wish to enter 10,000 SharePoint Content Types!

With case files, a dedicated metadata field is most likely needed to specify the RM classification. The SharePoint content type may well be *Contracts*, but a metadata field called *Contractor* will specify the particular contractor (*Safe-Tee Security* or *Tasty Catering*). If you have 10,000 employees, each employee would likely have a system-generated unique employee number. You would therefore likely have a metadata field *Employee Number* for all SharePoint content types (memos, complaints, performance reviews, etc.) that could possibly be related to (about) an employee. Any process you define to automatically assign a records classification to one of these documents will need to deal with two parts of the metadata in order to classify it correctly:

Document Type (e.g. *Contracts* takes us to 04-50). Incomplete classification.

Contractor Name (e.g. takes us to 04-50-1). Complete classification.

Remember that this technique of assigning RM classification for interactive archiving operations via SharePoint content type mapping is powerful, but can ***only*** be used for content types of identical subject/topic (against the file plan). Put another way, this approach only works where there is a one-to-one mapping in SharePoint between content type and file plan category (record series). For other situations we will have to use a different approach to achieving classification.

Profile-Base Records Archiving

Recall that an AGA site profile essentially “calls” library profiles to determine how individual libraries within the SharePoint site will be archived. In a recordkeeping setting, your goal is to ensure that all the documents within this site are properly classified against the appropriate record series (category) in the file plan (as defined in Content Server).

You could define a single RM classification for the entire site, which would be inherited downward to all subsites. However this would only be appropriate in the case where all documents in the site (and subsites) were of the same subject (administrative category, e.g. *Travel*) or case (case category, i.e. *Safe-Tee Security*). For instance, if you had a team site exclusively for documents about contractor *Safe-Tee Security*, then you could assign *Finance/Contracts/Safe-Tee Security* (04-50-1) for that entire site. Again, you would have to be 100% certain the site does not contain documents of any subject/case other than the one that fits the assigned RM classification.

In the case where the document subject or case varies within the site, you can still apply a site-wide default, but account for the differences within the site as you encounter them in the site hierarchy. You will need to traverse the site from the top down, and assign a suitable RM classification wherever the subject/case changes. Essentially, you need to “follow” the changes downward in the hierarchy, assigning a different RM classification wherever the document subject/case changes. To do this, you define and assign an AGA library profile to each library where a document subject/case changes from the default value assigned by the site profile.

With an AGA library profile, you can define rules that determine which documents in the library are to be archived. To assign RM classification, you include RM classification as part of your rule. Rules can be easily defined within the onscreen Query Builder, using simple logic statements such as:

If field <Contractor Name> = “Safe-Tee Security” then RM classification = **04-50-1**

If for any reason you are unable to construct a query builder rule to match the document with a suitable RM classification, you can use the CAML programming language to create a sophisticated rule that can go beyond a simple rule and include programming code. Developing CAML-based rules in the AGA library profile is technically more complicated, however it is unrestricted – you do not have the limitations of the Content Type mapping technique, or the query builder rules, to assign RM classifications. You can create any conceivable type of rule to accomplish RM classification, as long as SharePoint document metadata is sufficiently rich to support it.

Once again, suppose a SharePoint document has a content type called *Contracts*. All *Contracts* content types in turn have a mandatory field *Contractor Name*. Each document of this type has the two things we need to reliably assign a RM classification:

Records Series Finance/Contracts (04-50)
Name of Contractor. One of:
Safe-Tee Security
Tasty Catering
Lovely Landscaping

Because we can access all document metadata fields inside a library profile rule, we can now access the critical *Contractor Name* field, which allows us to properly and fully classify the documents about *Safe-Tee Security*'s contract. Our rule would contain the following logic:

If SharePoint **Content Type** = *Contracts* and field **Contract Name** = *Safe-Tee Security*
then
Content Server Field **RM Classification** = *04-50-1*

It is important to note that you <u>must</u> ensure that the document's metadata, which you are relying on for a match, is accurate and appropriate. As a general rule, investigate the source documents before you apply such a rule, and ensure the metadata accuracy exceeds your acceptable classification accuracy threshold of 90% or so. In our example, the metadata field *Contractor Name* must have the proper value (the right contractor), and the spelling of the contractor name must be correct. If both are correct 90% of the time or greater, then the rule will assign the proper classification 90% of the time or greater.

Your ability to utilize SharePoint document metadata for RM classification depends entirely on the document metadata. Put another way, the records classification will only be as good as your document metadata. Suppose your rule depended on the SharePoint field *Contractor Name*. Suppose further that SharePoint users were manually typing the contractor name into that field. Over time, with normal human error, you might realize the following entries in the *Contractor Name* field:

Safety Security	
Safe-tee	
Safe-Tee Security	(correct, right case)
Safetee security	
Safety	
Safe Security	
Safe-tee Security	(correct, wrong case)

Because this field is operator-entered, it is not a reliable field to use in a rule. Only the third and last entries would work, assuming your rule was insensitive to case. If it was sensitive to case, only the third entry would work.

As a best practice, only use metadata field matching on list-entered fields. This way, the user is never typing the contractor name, just selecting it from a pre-defined drop-down list, which removes any data entry errors. Similarly, it is a good idea to enumerate large lists of items that are difficult to consistently enter. Instead of 10,000 employee names, assign each employee a number. With these approaches, you eliminate data entry errors, and reduce the error rate down to selection errors only (did the user make the right selection?). Hopefully, the selection error rate for well-meaning users will be below your allowable classification error rate.

Content Server has a SAC (**S**ystem-**A**ssisted **C**lassification) capability. Content Server can scan documents and automatically assign RM classification codes to it. Content Server reads and understand the document, finds the "best-fit" match to a record series in the file plan, and if the confidence level exceeds a pre-determined threshold, assigns the RM classification automatically. In theory, you could rely on this technology to assign classification <u>after</u> archiving documents into Content Server. It is outside the scope of this report to evaluate this capability. RIMtech has not assessed or measured the efficacy of this SAC capability. RIMtech strongly recommends you achieve classification before or during AGA archiving.

Classification Rules for Case Categories

Suppose you have 1,000 employees, and 100 contractors, for a total of 1,100 case categories in your file plan. Using the same *Contracts* case example from earlier, you would have 100 different possible contractor names. This would mean you would in turn have to define 100 different "If/then" rules within the AGA profile, where each rule specifies a different contractor name. This is feasible, but laborious to enter and maintain. But what about the 1,000 employee case categories? It is impractical to enter 1,000 rules for each employee, and 100 more for each contractor, for a total of 1,100 rules.

To handle records bearing case categories (typically 50%+ of all records in most organizations), you should use AGA's CAML programming capability to assign RM classification programmatically. It can be driven by document metadata fields that specify the particular case pertaining to the document. Suppose you have SharePoint content type called *Employee Performance Reviews*. This content type has a metadata field called *Employee Number*. Each of our 1,000 employees has a 5-digt employee number (likely to be generated and maintained by some third party system). Suppose further that the *Employee Number* field is always populated by a drop-down list, ensuring that each field value is a legitimate number. You have everything needed to correctly classify the document. You would write simple CAML code that would carry out these steps:

Content Type = *Employee Performance Reviews*

Base classification = 03-30 (Human Resources/Performance Reviews)

Field *Employee Number* = 00707 (Rubble, B)

RM Classification = 03-30 + 707 or 03-30-707

All you have to do is add the employee number to the base classification in order to build the RM classification. This way, you only need a rule for each **business function**, in this case *Employee Performance Reviews*, and not for every employee. In this example we have a SharePoint content type for each business function, which is probably likely to be the case in most systems.

Advanced Records Management

You can configure AGA for “Advanced Records Management”. This is not an AGA module or function, rather it is a configuration setting that determines some of the records management behaviours during archiving. The setting is suitable for an environment such as a US DoD 5015.2 compliant environment, where more documents would be captured as records, and more document metadata fields are collected for classification purposes. The setting is called **Advanced Item Management**. When set, the RM classification throughout AGA archiving operations is more stringent, in the following ways:

1) RM Classification is mandatory for each site and library profile (for automatic archiving).
2) RM Classification is mandatory for each interactive site archiving.
3) When interactively archiving a document, the SharePoint user is automatically re-directed to Content Server to fill out additional fields mandated by 5015.2.
4) When archiving a document, or editing its categories in Content Server, the default categories for the document’s mime type are retrieved from Content Server and added to the document as an initial set of categories.
5) The SharePoint user cannot mark a document as **Official** while using the EL Web Part.

If the setting **Security Clearance** is set, the SharePoint user will be able to apply **supplemental markings** to a document as part of the archiving operation. Supplemental markings are a requirement of US DoD 5015.2. A supplemental marking is a document metadata field that further restricts access to a document, such as “***RD** – Restricted Data*”.

Search and Retrieval

With AGA, a SharePoint user can search Content Server. AGA delivers two capabilities for searching Content Server within SharePoint:

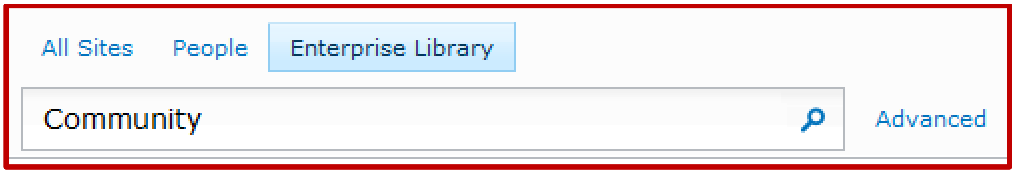
SharePoint Web Parts Content Server search components can be embedded inside SharePoint.

Merged Search The native SharePoint search capability can be merged with Content Server such that the a native SharePoint search includes Content Server as well.

SharePoint Search Web Parts

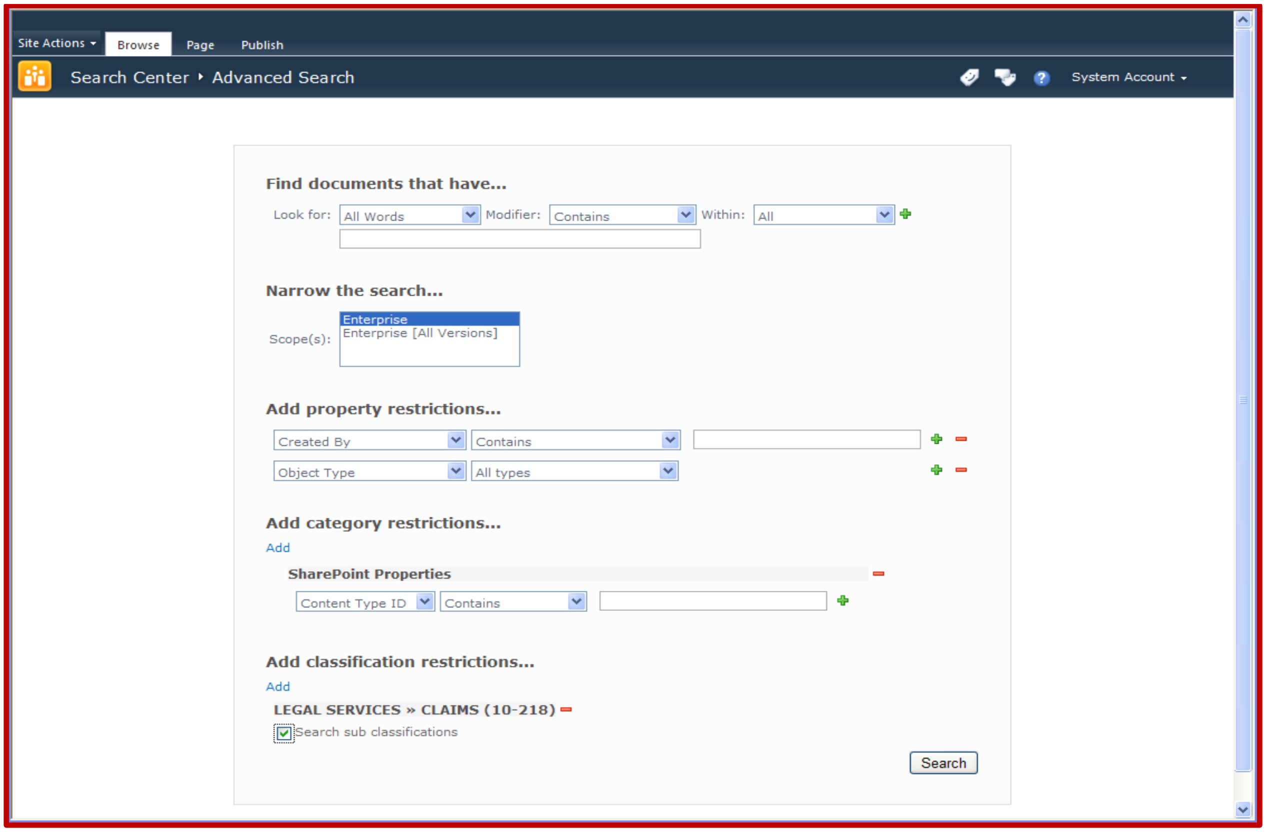
AGA provides three search-related web parts inside SharePoint. Using these web parts, a SharePoint user can search Content Server. Note that a Content Server user cannot search SharePoint. The three web parts are:

Search Box. This web part allows you to integrate a simple Content Server search capability into the SharePoint Search Center, or any custom SharePoint page, as shown:



Community

Advanced Search. This is a more sophisticated version of the simple Content Server search shown previously. The SharePoint user can narrow the search and add property restrictions as shown:



Find documents that have...

Look for: All Words Modifier: Contains Within: All

Narrow the search...

Scope(s): Enterprise
Enterprise [All Versions]

Add property restrictions...

Created By | Contains

Object Type | All types

Add category restrictions...

SharePoint Properties

Content Type ID | Contains

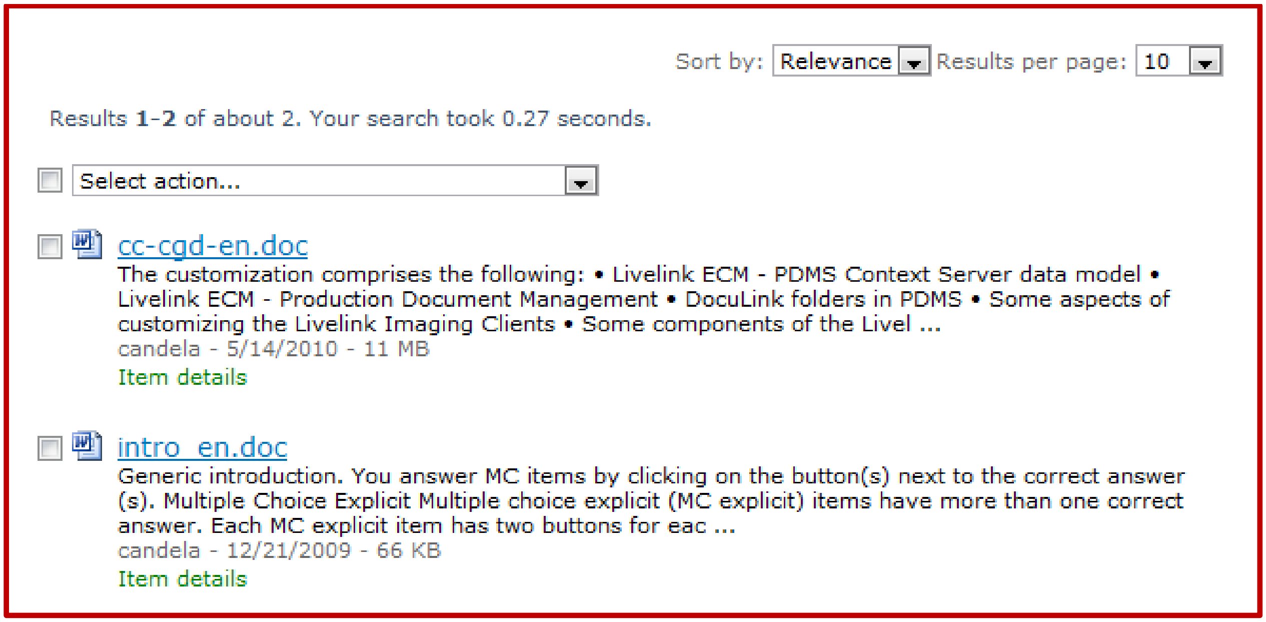
Add classification restrictions...

LEGAL SERVICES » CLAIMS (10-218)

☑ Search sub classifications

Search Results.

Following a search inside the **Search** or **Advanced Search** web parts, the results are delivered back to the SharePoint user in the **Search Results** web part, as shown:



Results **1-2** of about 2. Your search took 0.27 seconds.

□ cc-cgd-en.doc
The customization comprises the following: • Livelink ECM - PDMS Context Server data model • Livelink ECM - Production Document Management • DocuLink folders in PDMS • Some aspects of customizing the Livelink Imaging Clients • Some components of the Livel ...
candela - 5/14/2010 - 11 MB
Item details

□ intro_en.doc
Generic introduction. You answer MC items by clicking on the button(s) next to the correct answer (s). Multiple Choice Explicit Multiple choice explicit (MC explicit) items have more than one correct answer. Each MC explicit item has two buttons for eac ...
candela - 12/21/2009 - 66 KB
Item details

These SharePoint web parts appear in the SharePoint Search Center. They can be used as-is, or you can create your own custom web pages throughout SharePoint that contain any of these three web parts. These parts, and pages that contain these parts, can be highly customized using the configuration capability of AGA. For instance, you can customize the following (only a representative sample):

- Display Content Server search scope selector
- Specify which Content Server properties to search
- Allow user to specify RM classification options
- How search results are displayed (which columns and how they are sorted)
- Look and feel (style sheets)
- Maximum results per page
- Onscreen field display labels

You can greatly customize the search web parts by changing administrative settings in AGA configuration XML files, and by custom programming using the AGA **S**oftware **D**evelopment **K**it (SDK).

Merged Search

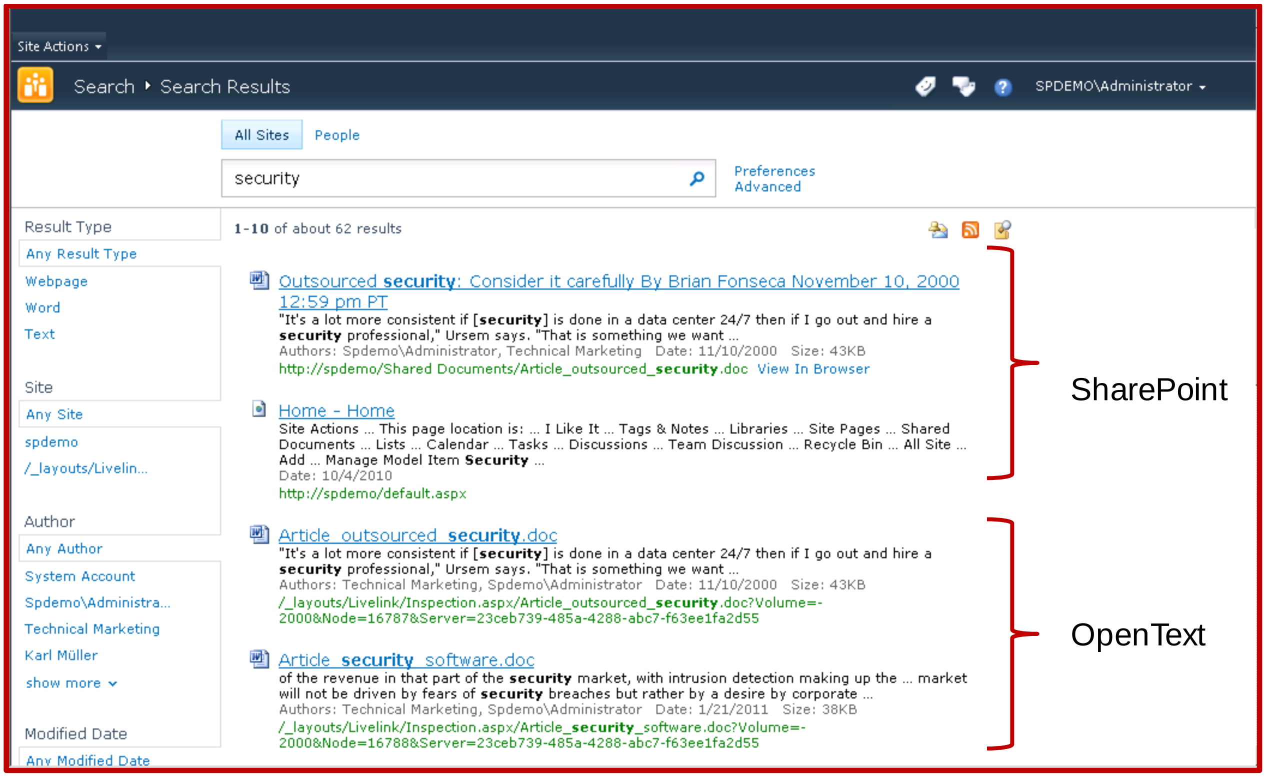
There are two forms of search capability in SharePoint 2010:

Native	Out of the box search capability included with SharePoint 2010.
FAST	An optional, extra-cost, add-on search capability for SharePoint 2010. FAST is a far more powerful search capability than SharePoint's native search.

OpenText warns that in a merged native search, it is possible for a search result to exceed SharePoint's internal processing capability. This anomaly does not occur in a merged FAST search.

With AGA, you can use SharePoint to search both SharePoint and Content Server, using SharePoint's native (built-in) search capability or optional FAST search. This is a powerful and useful feature, as it is completely transparent to the SharePoint user – nothing changes about their usual SharePoint search (native or FAST), and they do not have to do anything special to search Content Server. Content Server results are always included in the search result list.

An example of a merged search result list is shown below:



1-10 of about 62 results

Outsourced **security**: Consider it carefully By Brian Fonseca November 10, 2000 12:59 pm PT
"It's a lot more consistent if [**security**] is done in a data center 24/7 then if I go out and hire a **security** professional," Ursem says. "That is something we want ...
Authors: Spdemo\Administrator, Technical Marketing Date: 11/10/2000 Size: 43KB
http://spdemo/Shared Documents/Article_outsourced_**security**.doc View In Browser

Home - Home
Site Actions ... This page location is: ... I Like It ... Tags & Notes ... Libraries ... Site Pages ... Shared Documents ... Lists ... Calendar ... Tasks ... Discussions ... Team Discussion ... Recycle Bin ... All Site ... Add ... Manage Model Item **Security** ...
Date: 10/4/2010
http://spdemo/default.aspx

Article outsourced **security**.doc
"It's a lot more consistent if [**security**] is done in a data center 24/7 then if I go out and hire a **security** professional," Ursem says. "That is something we want ...
Authors: Technical Marketing, Spdemo\Administrator Date: 11/10/2000 Size: 43KB
/_layouts/Livelink/Inspection.aspx/Article_outsourced_**security**.doc?Volume=-2000&Node=16787&Server=23ceb739-485a-4288-abc7-f63ee1fa2d55

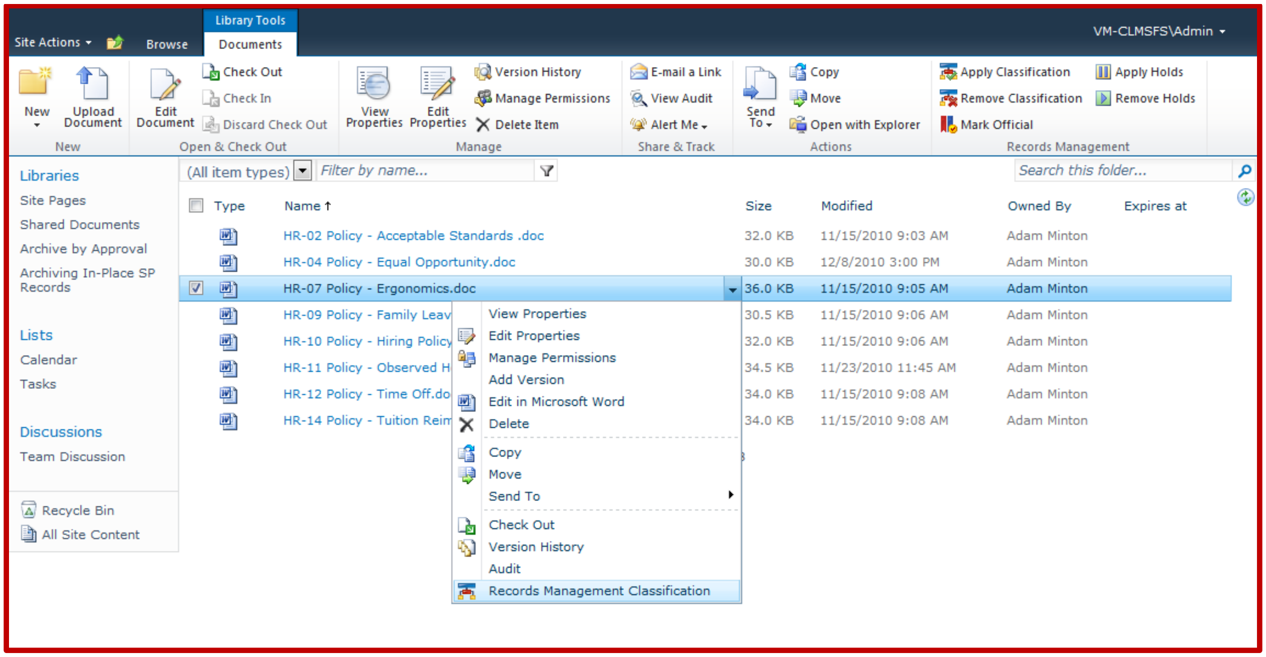
Article **security** software.doc
of the revenue in that part of the **security** market, with intrusion detection making up the ... market will not be driven by fears of **security** breaches but rather by a desire by corporate ...
Authors: Technical Marketing, Spdemo\Administrator Date: 1/21/2011 Size: 38KB
/_layouts/Livelink/Inspection.aspx/Article_**security**_software.doc?Volume=-2000&Node=16788&Server=23ceb739-485a-4288-abc7-f63ee1fa2d55

SharePoint search (native and FAST) each create a search **index**. The index is prepared each day and updated regularly, usually nightly. When you use SharePoint to search for documents, you search the index, not the system's database itself. This is what makes searching so fast. The process of updating the search index is sometimes referred to as **crawling** (the system crawls through all of the day's additions, deletions, and edits, and then updates the index).

In the AGA search configuration screens, you simply have to specify which portion (content subdirectories) of Content Server to include in its daily crawl schedule. From that point forward, SharePoint will include documents from Content Server in its search.

Enterprise Library Web Part

Of the (7) web parts from AGA, this is by far the most important, and the most widely deployed. The EL Web Part essentially puts the SharePoint user into Content Server, while inside SharePoint. The SharePoint user can access and work with content living inside Content Server. The full Content Server user interface, and most functionality (but not all) is available inside this web part. The web part is shown below:



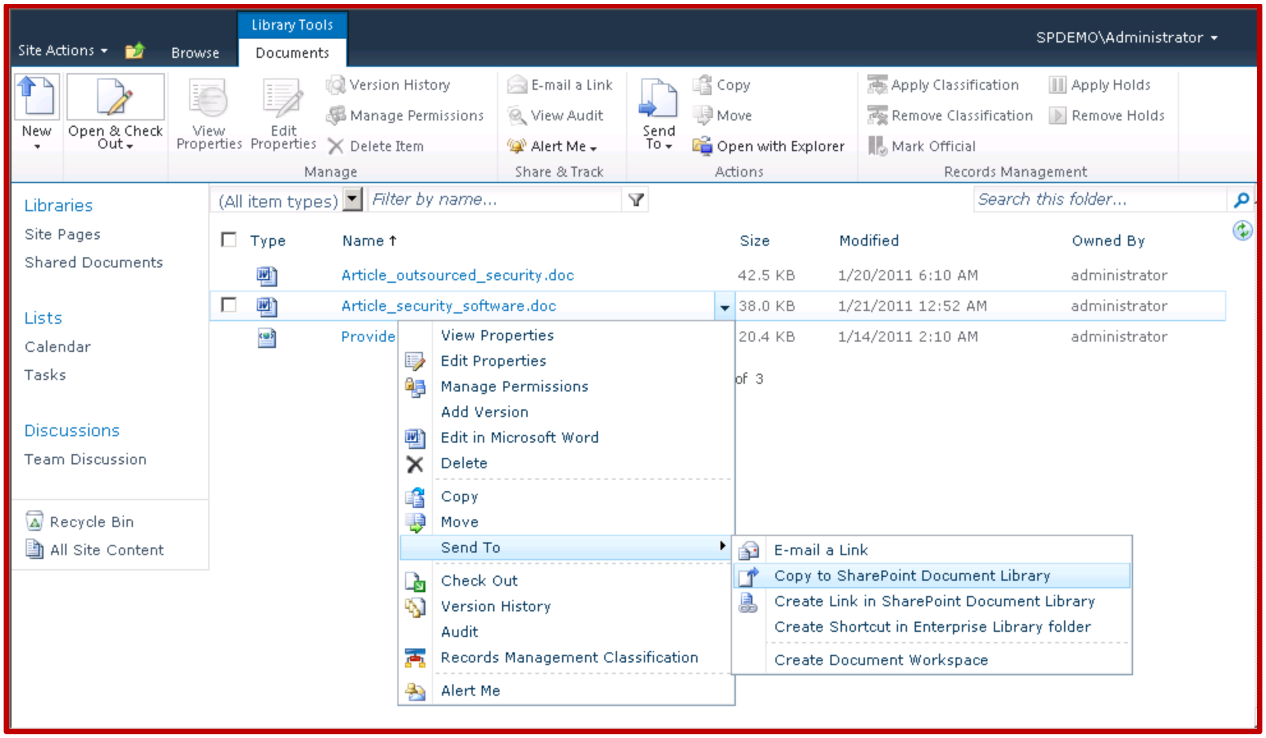
Type	Name ↑	Size	Modified	Owned By	Expires at
	HR-02 Policy - Acceptable Standards .doc	32.0 KB	11/15/2010 9:03 AM	Adam Minton	
	HR-04 Policy - Equal Opportunity.doc	30.0 KB	12/8/2010 3:00 PM	Adam Minton	
☑	HR-07 Policy - Ergonomics.doc	36.0 KB	11/15/2010 9:05 AM	Adam Minton	
	HR-09 Policy - Family Leav	30.5 KB	11/15/2010 9:06 AM	Adam Minton	
	HR-10 Policy - Hiring Policy	32.0 KB	11/15/2010 9:06 AM	Adam Minton	
	HR-11 Policy - Observed H	34.5 KB	11/23/2010 11:45 AM	Adam Minton	
	HR-12 Policy - Time Off.do	34.0 KB	11/15/2010 9:08 AM	Adam Minton	
	HR-14 Policy - Tuition Reim	34.0 KB	11/15/2010 9:08 AM	Adam Minton	

Notice the familiar SharePoint user experience including the ribbon and context menu (right-click over a document). Even though it looks like any ordinary SharePoint Document Library, the documents are physically residing in Content Server. Note the Records Management ribbon bar grouping on the right of the ribbon bar. It contains the following Content Server functions:

Apply Classification
Remove Classification
Mark Official
Apply Holds
Remove Holds

Note also the Content Server Records Management Classification option on the context menu.

Some Content Server capabilities (functions) are available exclusively in the EL web part, and not available in any other web part, i.e. they cannot be accessed anywhere in SharePoint except inside the EL web part. With this web part, you can also copy a document back to the SharePoint workspace, or to a temporary SharePoint workspace, as shown:



The user's assigned document permissions in Content Server determine which Content Server buttons are available to the user on the ribbon bar and context menu. If for example the user only has read permissions to the document in Content Server, they will not be able to edit the document or change any of its metadata fields.

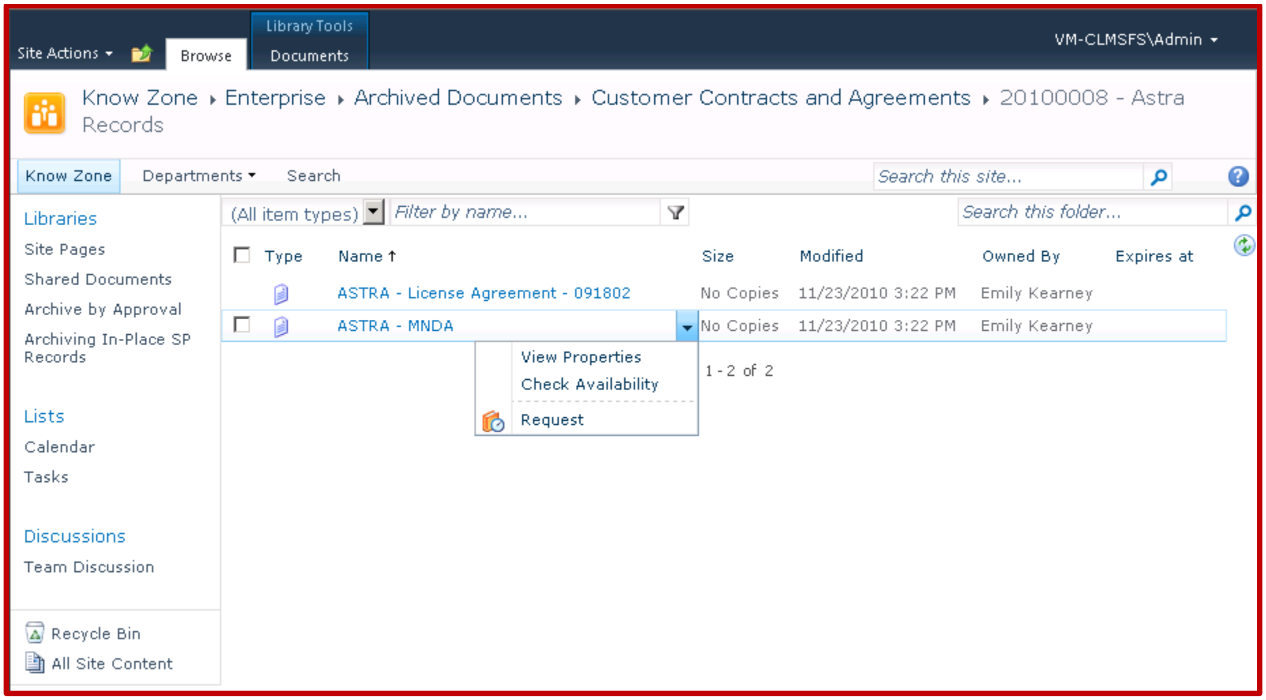
At an administrative level, AGA allows you to add/remove functionality from the users, irrespective of their assigned Content Server permissions. You can also add custom commands, such as a ribbon bar button to run a pre-defined Content Server workflow on the selected document. These administrative changes are applied via the SharePoint AGA Administrative console.

Physical Records Management

Physical Records Management (**PRM**) refers to the Content Server's native ability to store and manage physical records, such as boxes, folders, and other physical artifacts such as maps, etc. AGA refers to physical records as **Items**. Content Server has a rich set of advanced PRM capabilities, including the ability to create new records (items), apply bar code labelling to the boxes/folders, manage their location in the warehouse, and manage the space requirements.

AGA presents to the SharePoint user, those PRM capabilities that an end user would normally be interested in. The PRM administrative tasks such as creating new items, printing barcode labels, etc. are still accomplished exclusively inside Content Server. The SharePoint user can do the following inside SharePoint:

> **Browse Physical Items**. Using the Advanced Search web part, items can be included in a search.
>
> **Search for Items**. Following a search (Advanced search) that generates a result containing physical items, the user can browse and review the properties of the items, as shown:

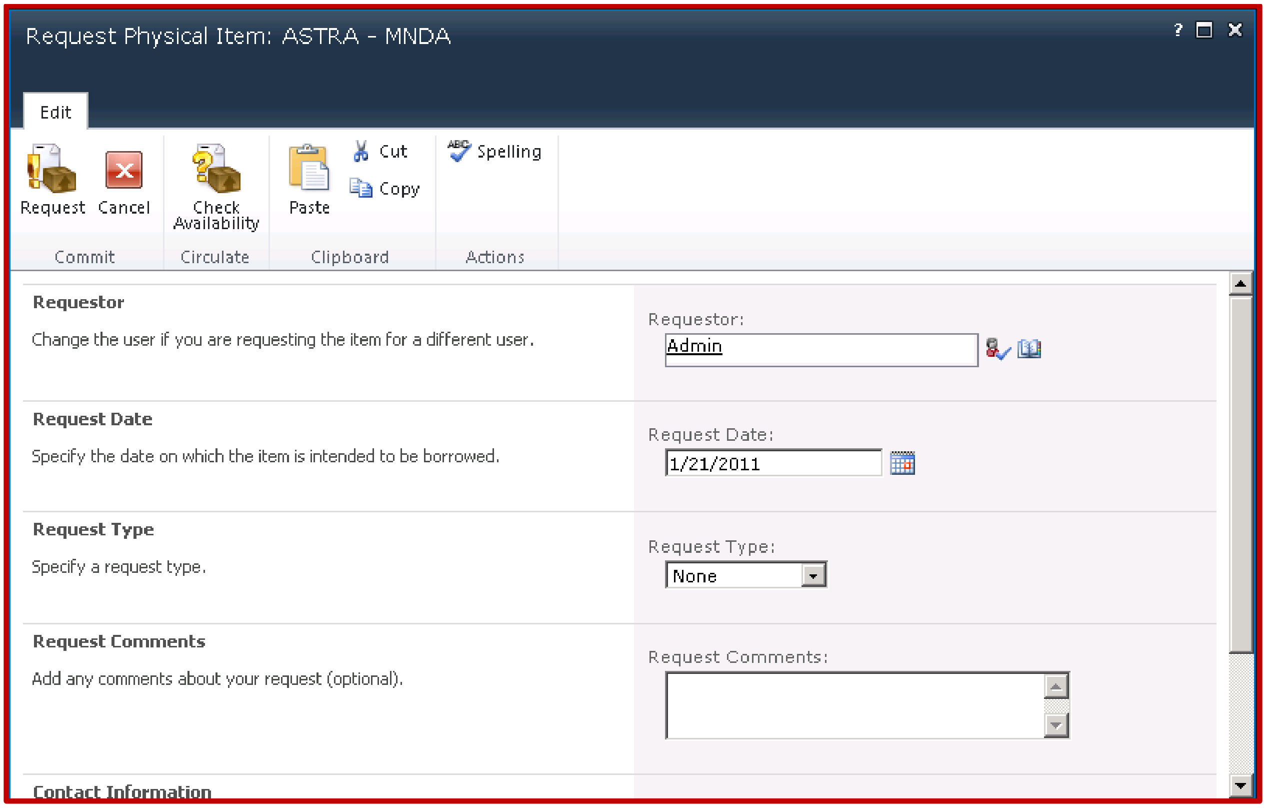


Know Zone ▸ Enterprise ▸ Archived Documents ▸ Customer Contracts and Agreements ▸ 20100008 - Astra Records

Type	Name ↑	Size	Modified	Owned By	Expires at
	ASTRA - License Agreement - 091802	No Copies	11/23/2010 3:22 PM	Emily Kearney	
	ASTRA - MNDA	No Copies	11/23/2010 3:22 PM	Emily Kearney	

1 - 2 of 2

Request (Borrow) Items. The user can click on the context menu for the item and request (borrow) the item, as shown:



Request Physical Item: ASTRA - MNDA

Requestor

Change the user if you are requesting the item for a different user.

Requestor: Admin

Request Date

Specify the date on which the item is intended to be borrowed.

Request Date: 1/21/2011

Request Type

Specify a request type.

Request Type: None

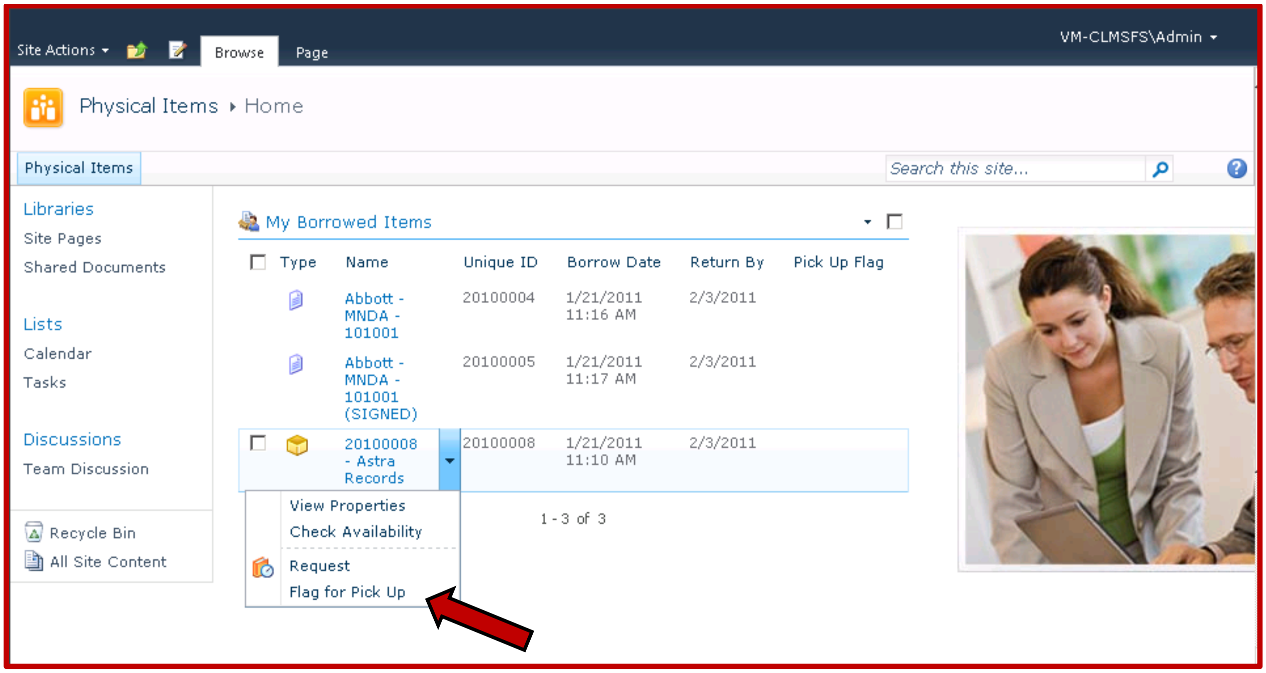
Request Comments

Add any comments about your request (optional).

Request Comments:

Contact Information

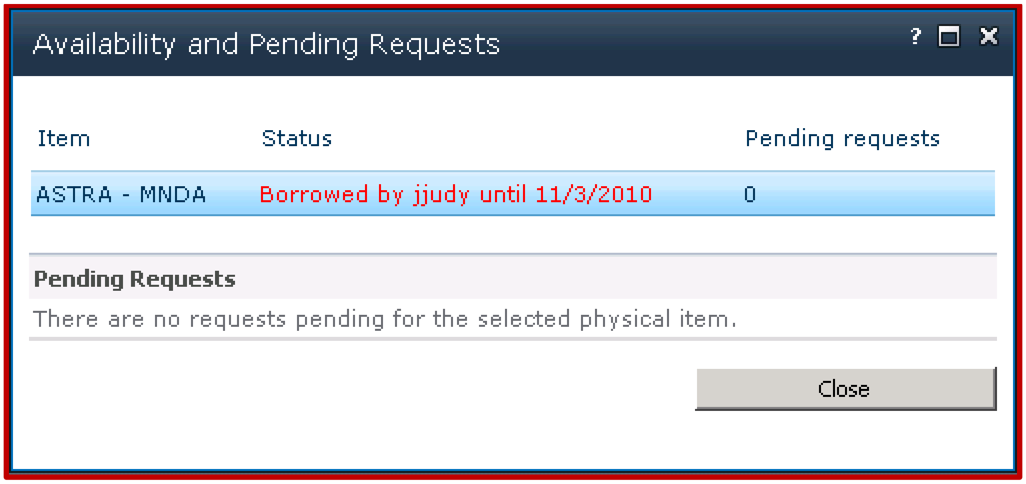
Flag Item for Pickup. Once you select the item on the result list, you can flag it for pickup as shown:



□ Type	Name	Unique ID	Borrow Date	Return By	Pick Up Flag
	Abbott - MNDA - 101001	20100004	1/21/2011 11:16 AM	2/3/2011	
	Abbott - MNDA - 101001 (SIGNED)	20100005	1/21/2011 11:17 AM	2/3/2011	
□	20100008 - Astra Records	20100008	1/21/2011 11:10 AM	2/3/2011	

1 - 3 of 3

View borrowed Items. The search result web part will display physical items along with electronic documents. You can check the availability as shown:



Item	Status	Pending requests
ASTRA - MNDA	Borrowed by jjudy until 11/3/2010	0

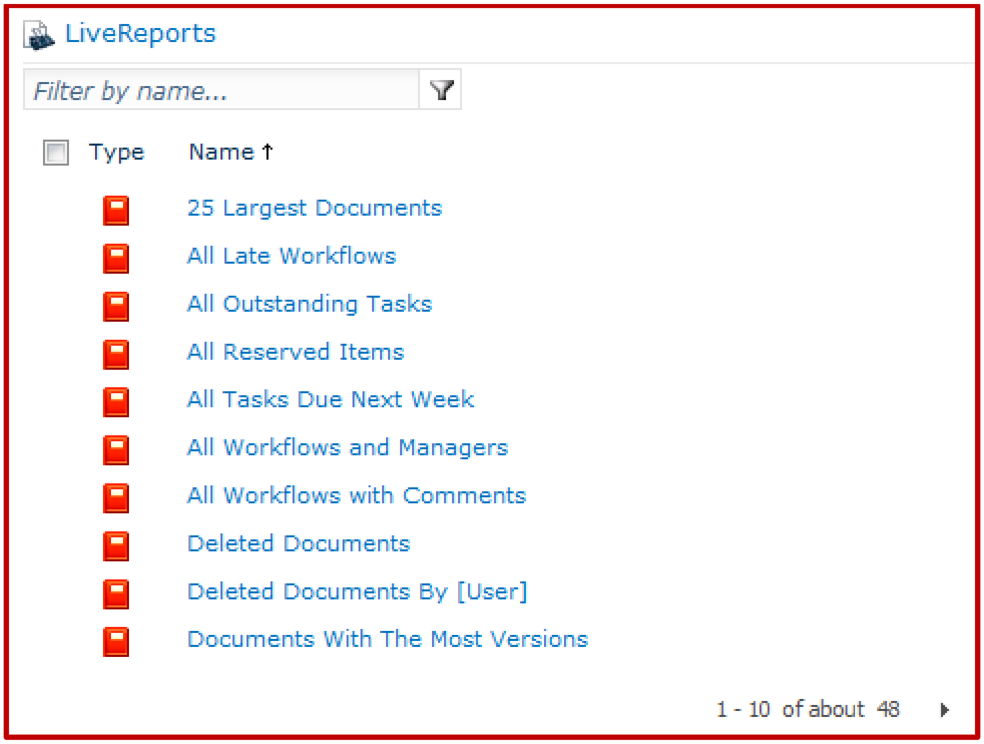
Pending Requests

There are no requests pending for the selected physical item.

Content Server Reports and WorkFlow

With AGA you can generate and view Content Server reports (called **LiveReports**) within SharePoint. You can also view and manipulate Content Server **workflows**. There are three SharePoint web parts involved; **LiveReports**, **LiveReports Results**, and **Workflow Inbox**.

LiveReports. Here you can view and run pre-defined live-reports (that you have access to in Content Server), as shown:



LiveReports

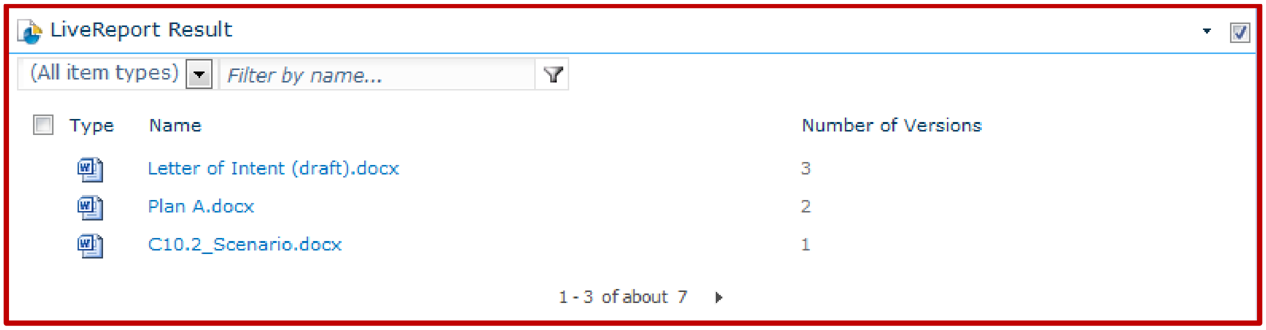
Filter by name...

□ Type Name ↑

- 25 Largest Documents
- All Late Workflows
- All Outstanding Tasks
- All Reserved Items
- All Tasks Due Next Week
- All Workflows and Managers
- All Workflows with Comments
- Deleted Documents
- Deleted Documents By [User]
- Documents With The Most Versions

1 - 10 of about 48

LiveReports Results. Once you've requested a particular LiveReport, the web part displays report results, as shown:



LiveReport Result

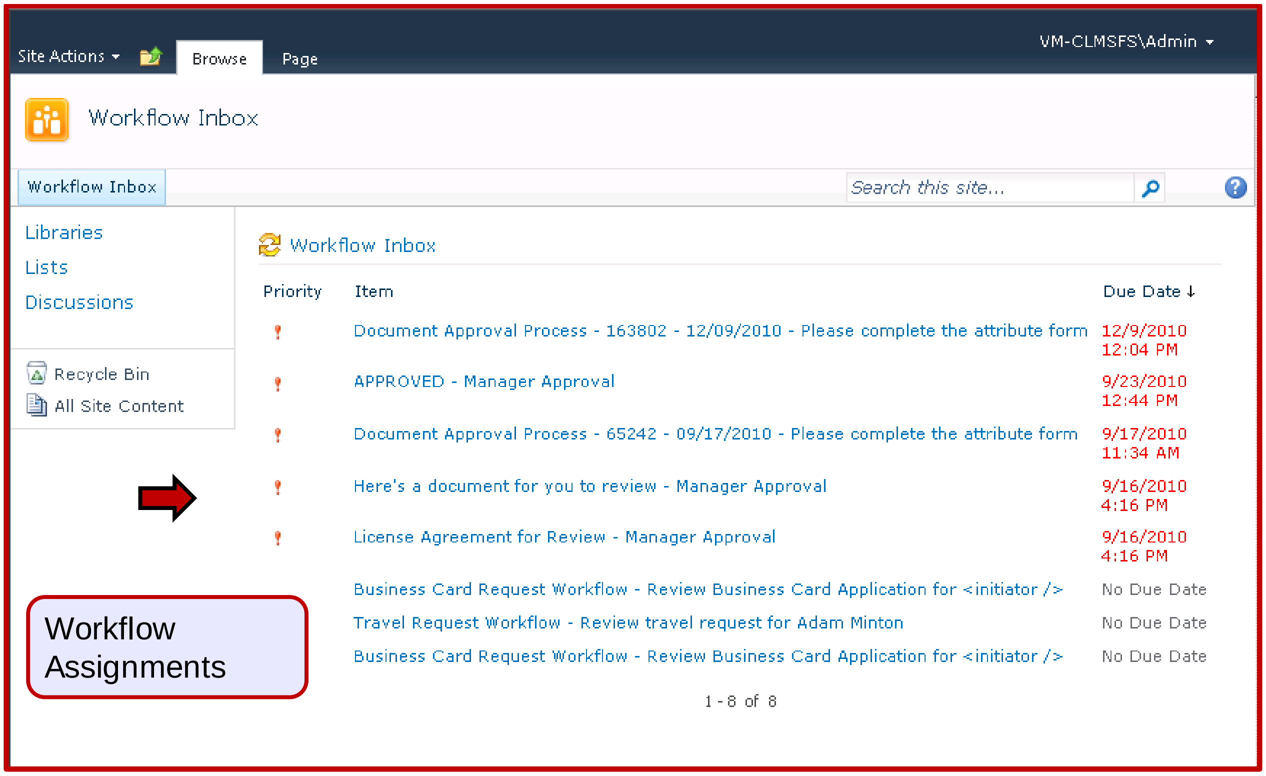
(All item types) Filter by name...

□ Type	Name	Number of Versions
	Letter of Intent (draft).docx	3
	Plan A.docx	2
	C10.2_Scenario.docx	1

1 - 3 of about 7

How the report will look will largely depend on the report definition as defined and saved in in Content Server. You can modify the report output to some extent in the AGA administrative configuration settings, by specifying such things at the columns to display, filter, etc.

Workflow Inbox. With this web part you can display all pending workflows that are assigned to you, in ascending due date order as shown:



Workflow Inbox

Priority	Item	Due Date ↓
!	Document Approval Process - 163802 - 12/09/2010 - Please complete the attribute form	12/9/2010 12:04 PM
!	APPROVED - Manager Approval	9/23/2010 12:44 PM
!	Document Approval Process - 65242 - 09/17/2010 - Please complete the attribute form	9/17/2010 11:34 AM
!	Here's a document for you to review - Manager Approval	9/16/2010 4:16 PM
!	License Agreement for Review - Manager Approval	9/16/2010 4:16 PM
	Business Card Request Workflow - Review Business Card Application for <initiator />	No Due Date
	Travel Request Workflow - Review travel request for Adam Minton	No Due Date
	Business Card Request Workflow - Review Business Card Application for <initiator />	No Due Date

1 - 8 of 8

Extensibility

Extensibility refers to the extent that a software product can be customized. The greater the degree to which the product can be customized, the greater is its so-called *extensibility*. AGA is phenomenally extensible. Part of this no doubt has to do with the nature of what AGA does. AGA is not an application in the typical sense of the word. It is a "bridge" of sorts between two very different products that work very differently. There is a nearly endless variety of circumstances and situations that can present themselves when trying to move documents from SharePoint to Content Server. Because of this, it's hard to predict what might usually work, and deliver these features in a well-defined application package.

OpenText has made AGA sufficiently "open" that you should be able to do pretty much anything you might encounter. There are eight distinct ways in which AGA is extensible:

Administrative Settings. Onscreen setting the RIM administrator can change within the SharePoint administrative console. No technical assistance is needed.

XML Configuration files. Stores settings used by various AGA modules. Over 500+ settings are easily changed by editing these files. Technical skills and some basic programming skills are required.

CAML and C# Queries and Preconditions. Through basic programming, you can customize the core elements of automatic archiving – what documents to archive. Basic SharePoint programming skills are needed.

XSLT Style Sheets. You can alter the look and feel of any AGA screen by modifying the style sheets, using an modern XSLT editor. Web page design skills are needed.

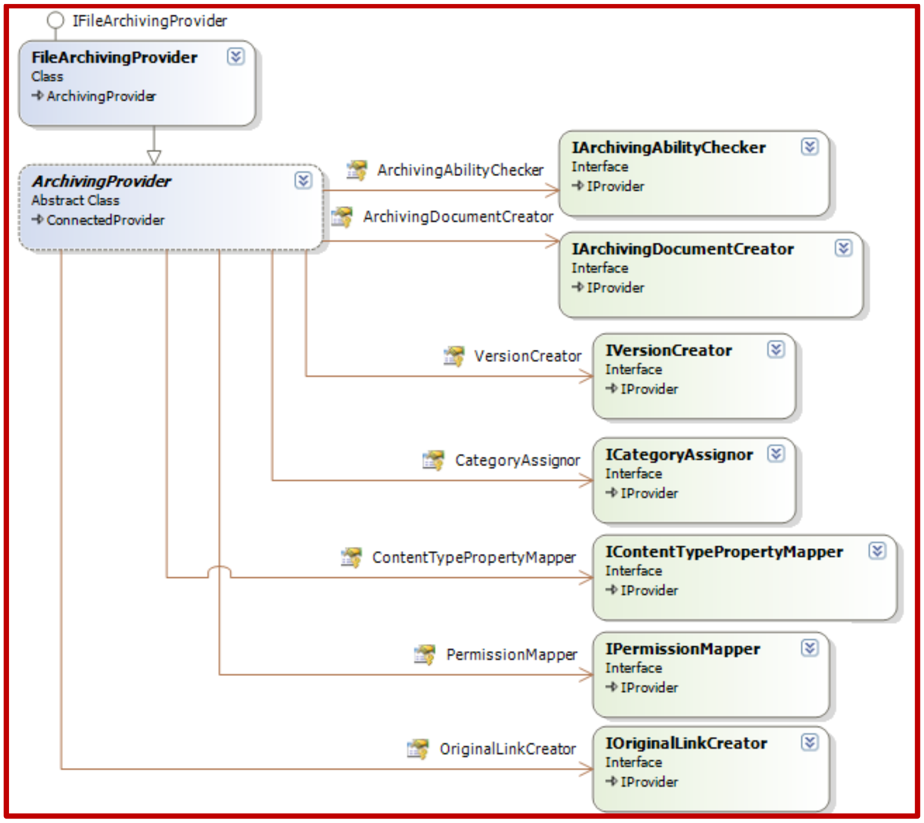
GUI Strings. GUI (Graphical User Interface) strings refer to the onscreen words and labels the user sees on each AGA screen. These words are all stored in 17 files that you can edit to change the language or wording on the AGA screens. Basic programming skills required.

Web Parts. Using standard web page design techniques, you can custom-create any SharePoint page you need, and drop AGA’s web parts into the pages. Basic Web page design skills are needed. Programming is not required.

ASPX Pages. You can design and custom-build SharePoint templates and pages using the Microsoft .NET framework programming language.

SDK. This is AGA’s comprehensive SDK, which makes any of AGA’s internal functions available for custom programming. Advanced programming skills are required to develop to this level.

The diagram below shows an example of the so-called **Provider Concept** inside the AGA SDK:



A provider is a well-defined package of functionality that you can replace or extend in a prescribed manner, while being assured it will work well inside AGA. You can easily create an entirely new provider for archiving documents, or you can modify some aspect of an existing AGA archiving method, such as altering the master permissions mapping. The AGA SDK gives you the framework and structure you need to build your own providers, and easily incorporate the new capability into AGA. Extending AGA this way means that your custom provider will work with future releases of AGA, without having to modify your custom code with each release.

Site Deployment

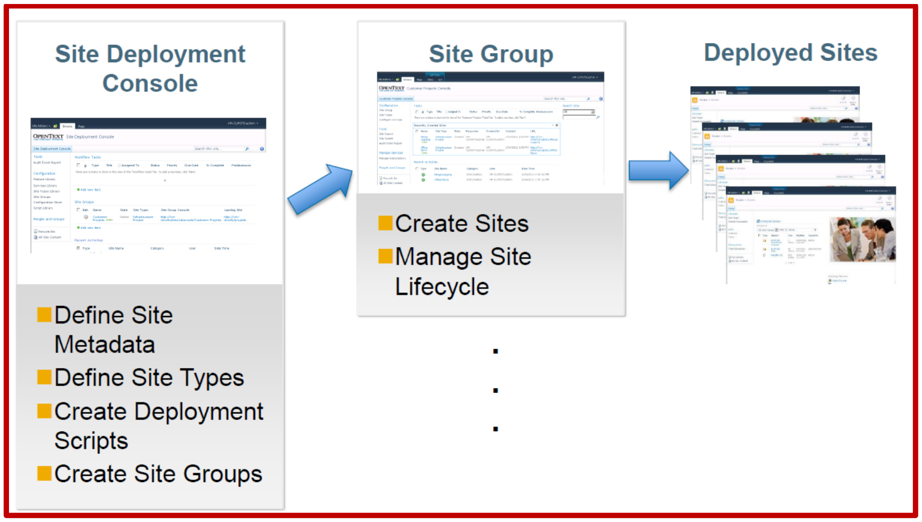
A large (more than 10,000 users) SharePoint installation can have thousands and thousands of sites. Even in a smaller SharePoint deployment with fewer sites, a great deal of careful site design and construction may have taken place. Many of these sites are highly customized, and used for all manner of work. There could be millions of documents across thousands of libraries. Some of these sites will be hierarchically structured so their properties are passed along to dozens of subsites. Some sites will be carefully planned and well organized, some will be more chaotic. In the case of a mature AGA deployment, there could be hundreds or thousands of DFAS settings, site profiles, and library profiles associated with many of these sites and/or libraries.

When somebody wants to create a new project, site, or workspace inside SharePoint, you would not often want to create a whole new site from scratch. Most often you can replicate a structure already deployed or configured somewhere in SharePoint. With AGA's **Site Deployment** capability, you can take an existing site and easily re-deploy it to another location under a different name, while preserving the structure and configuration that you've already established. This way, site deployment will save you a great deal of time, and lead to greater consistency in site structure, as more of the sites are enforcing the same structure and rules in exactly the same manner. It's an especially useful capability when enforcing corporate-wide governance such as recordkeeping across all sites in the organization.

AGA has a powerful site deployment capability. A comprehensive review is beyond the scope of this report, however it lets you:

Define minimum metadata for new sites
Define custom workflows to control site lifecycle
Include all AGA configuration in a site to be deployed
Group sites by department, business process, or other criteria
Deploy sites

The diagram below gives a high-level overview of AGA's site deployment capability:



Project Implementation

Too many EDRMS (**E**lectronic **D**ocument and **R**ecords **M**anagement **S**ystem) projects fail to deliver on the recordkeeping objectives. Most projects start out with the goal of delivering recordkeeping accountability, but invariably end up abandoning recordkeeping.

There are certain inviolable principals that are essential to successful electronic recordkeeping success. They are:

Qualification How users decide which documents to manage as business records.

Declaration Users have to decide that a particular document is a business record. Which ones? Similarly, when creating a document, at which point will they turn it over to the organization as a business record?

Classification Each record must be assigned an appropriate retention rule from the approved list of rules, or retention schedule.

If users consistently **qualify** appropriately, **declare** the documents as records, and properly **classify** them, we can proceed with ***Disposition***, where we will destroy records in accordance with the official approved retention schedule. Disposition is the "end game", where everything comes together in a way that allows us to accurately destroy massive quantities of recorded business records with full legal confidence. If declared records are not classified correctly, too many would be destroyed too early or too late, which could prevent any disposition from taking place at all.

Each of these three factors (*qualify*, *declare*, *classification*) is in fact a **critical project measurement metric**[16]. ***Everything*** in any electronic recordkeeping endeavour contributes in some way, however indirectly, to one of these three factors. Each factor can, and must, be measured against an acceptable threshold. Any organization can readily agree upon a minimum acceptable threshold for each of these three factors. The following acceptability thresholds are very common, based on RIMtech's experience across 50+ EDRMS projects:

Qualification	**5%** of all incoming email, **95%** of all documents created and downloaded.
Declaration	**95%** of all business records should be declared into SharePoint.
Classification Accuracy	A minimum **85%** of all declared records must be properly classified

Each of these three factors is easily measured, and **must** be regularly measured, by department, on a regular basis. Organizations must "hit" all three to succeed. A failure of any of the three, results in project failure. Put another way, if we cannot run Disposition with complete confidence, which is based on the accuracy of the classification, the project simply is not successful.

RIMtech Recommendations

AGA is a technical triumph in the sense that it manages to integrate two very different and very complex technologies, and it somehow manages to account for a seemingly endless number of possible permutations and circumstances of integration. There are countless combinations of *what* to archive, *how* to archive, and *when* to archive. Integrating Content Server and SharePoint is a complicated business, and OpenText gets it done with AGA. As is always the case with software, the price we pay for flexibility is complexity, as AGA presents such a large number of possible permutations of circumstances that we have to deal with. From a Records Manager's perspective, AGA may not be very easy to initially understand, due to the sheer number of mathematical permutations of circumstances, and product differences, all of which must be accounted for within the integration. Configuring and deploying AGA in a non-records setting will be fairly straightforward, as it involves only the (substantial) mechanical details of getting documents from one environment to a second (different) environment.

To maintain recordkeeping compliance however adds another layer of detail to the business of archiving via AGA. Even after the mechanical details are taken care of, we have to get the documents properly classified according to business activity. This is a potentially substantial information management challenge. In a complete absence of recordkeeping infrastructure in the SharePoint system, an AGA deployment in a records-compliant setting will be challenging to implement and maintain, as you have to achieve recordkeeping compliance after the fact. The more recordkeeping infrastructure in place in SharePoint, the easier it will be to implement and maintain an AGA deployment. By infrastructure we mean:

1. Documents are organized (grouped) by like subject
2. The system distinguishes case files from administrative files
3. Records are distinguished from non-records
4. Ideally, there is some form of records classification taking place.

The degree to which a real-world records-compliant AGA deployment will be challenging to implement and maintain will be driven by the degree to which the source SharePoint environment is structured for recordkeeping. The less the SharePoint structure, the more challenging the AGA deployment will be.

AGA is a very technical product that gamely handles a seemingly impossible hookup of two very different products. AGA is left with the unenviable task of having to deal with all these differences. In so doing, it

[16] See details in RIMtech's forthcoming book "*Managing Electronic Records*"

has had no choice but to provide a rather complicated series of archiving permutations. AGA however does a good job of delivering automated processes that, if properly planned and implemented, can automate a good deal of the archiving.

AGA requires a lot of care and feeding to plan, implement, and maintain. It requires a good deal of technical IT talent for each of these three stages.

AGA is not a records management product. It likely cannot be used solely by a typical RIM professional – it requires constant IT attention to plan, implement, and maintain. A RIM professional must understand AGA, but it will take an IT professional to plan, implement, and maintain the “nuts and bolts” of the integration itself. RIMtech sees the role of the RIM professional with AGA as providing critical input, namely the *what*, *when*, and *how* to archive, particularly the crucial RM classification. AGA supports recordkeeping by providing the means of moving documents from a non-compliant (SharePoint) system to a compliant system (Content Server), but if recordkeeping is not applied, it becomes simply an IT tool for moving content (active or inactive, managed or unmanaged) from SharePoint to Content Server (active or inactive, managed or unmanaged), where at the very least it is stored in a sophisticated, optimized storage environment.

The RIM professional has to see that the recordkeeping process is applied to all the documents being archived. And that means the RIM professional is back to the basics of deciding what to archive and why, how and when to declare as a record, and how to ensure classification happens correctly.

There is no reason why AGA cannot be used in a full formal recordkeeping setting to achieve recordkeeping compliance. The organization must be diligent in qualifying SharePoint documents, and assigning RM classification correctly, 90% or more of the time. Following are some recommendations we believe would facilitate reaching these important RIM goals:

1. Capture Document Status in SharePoint

A target document of recordkeeping value will be either a *record*, a *work in progress*, or a *reference* document. Configure SharePoint to capture the status in a metadata field for all documents. This will allow you to filter out genuine records from reference documents and non-records. This in turn will help with the AGA archiving, so you can identify those documents you’ll need to declare as records (and classify them if not already classified)

2. Pre-Structure in SharePoint

The more planning and structure that exists in SharePoint in support of recordkeeping, the easier the archiving will be. There are many things you can do in SharePoint to help, such as:

1) Group documents by like subjects in folders and libraries
2) Distinguish case files from administrative files via a folder (minimum) or document (ideal) metadata field. This will make classification easier and more reliable.
3) Plan, name, and implement SharePoint document metadata that will help downstream AGA RM classification mapping rules in AG library profiles.
4) Match security permissions as closely as possible.
5) Pre-classify. Have SharePoint users explicitly or implicitly classify documents at creation time.

3. Expose Content Server file plan to the SharePoint declaration process.

If you plan to achieve declaration in SharePoint, you’ll need to use the file plan, which resides in Content Server. Use AGA web parts to expose the file plan to the SharePoint declaration process. This way, you will not have to duplicate the file plan in SharePoint.

4. Declare Early

The sooner documents are declared/classified, the better. Do not let SharePoint documents languish too long before getting them classified, either in SharePoint or via archiving to Content Server. The further you get from the creation date, the greater the challenge to get them classified. Some SharePoint users may tend to keep their documents as long as possible, until they are “no longer needed”. Remember that the longer you wait, the more difficult it will be to apply a proper RM classification (unless you are baking the classification into the SharePoint document creation process).

It would be a good idea to implement an automatic deletion process in SharePoint that incents people to finish with their documents and declare them as records, or archive them to Content Server.

5. Establish metadata consistency

With a goal of classification accuracy of 90% or higher, RM classifications achieved via AGA archiving must be correct 9 of every 10 times. Where library profile rules are used to select a RM classification, the classification will only be as reliable as the metadata upon which the rule is based. Strive to have the metadata you’ll need for rule collected in SharePoint. Also try to ensure the data is entered as consistently and accurately as possible.

6. Avoid Duplicates

Avoid where possible the use of archiving modes that leave behind an unmanaged copy of a declared record in Content Server. The existence of Orphaned unmanaged copies of business records constitutes poor information management practice.

7. Learn the systems

In order to direct a successful deployment of AGA in a recordkeeping environment, the RIM professional must ensure the archiving achieves the target classification accuracy. The RIM professional will need to understand the following, not to an IT level, but to a strong level of understanding and competence:

AGA	How it works, capabilities and limitations
SharePoint	Data structures, metadata, folders, site/library structure. Basic SharePoint operation. Searching, reporting, export.
Content Server	Data structures, metadata, folders, site/library structure. Basic Content Server operation. Searching, reporting, export. Full recordkeeping capability, to administrative level, including disposition.

This is a tall order for anyone. A RIM professional should be prepared to invest a significant amount of time into the initial learning, in order to implement an effective AGA project.

8. Carefully monitor default RM classifications

If you are using DFAS or site profiles to apply a default RM classification, there is a potentially significant risk that you can inadvertently lower the classification accuracy level, via a high level of incorrect classifications. This could happen if the document subject/case changes throughout the various libraries or folders within the site containing the default RM classification.

Monitor all defaults closely, and apply an exception via a library profile to all libraries where the document subject/case changes.

9 Make library profile rules airtight

AGA library profile rules decide which documents get archived, and how they are classified. Write rules that are as bullet-proof as possible, so AGA will not inadvertently pick up documents outside the subject of interest. A rule that picks up too many documents of the wrong subject will result in potentially thousands of misclassifications, which can drive down your overall classification accuracy rate. Write them carefully, and test them thoroughly.

10 Cleanup aggressively

Run link cleanup utilities on an aggressive schedule as soon as possible following a disposition to ensure orphaned links do not stay around long following a Content Server disposition.

11 Configure SharePoint for Case Files

Out of the box, SharePoint does not have the required infrastructure for case file handling[17]. Be sure to configure SharePoint so that all documents in all folders can be readily identified as members of a case or subject file plan category. This will greatly help you achieve proper classification during AGA archiving. If the SharePoint documents are already classified, this will not be necessary, as the classifications will implicitly contain the subject vs. case categorization.

12 Coordinate SharePoint and Content Server Administration

Administrators of each system should coordinate high-level architecture and configuration to facilitate future AGA archiving. The goal is to minimize the differences or conflicts between the two systems. Some examples:

1) Use Active Directory to create users and groups. This minimizes the differences of user accounts.
2) Implement consistent security models. Plan an approach to security that is highly similar in both systems
3) Establish common nomenclature. Try to match up content types and metadata fields where possible. Ideally, create a master data dictionary that specifies site, folder, and document metadata fields that each system will draw from.

[17] See RIMtech RIMTAR report or ARMA book "*Managing Records with SharePoint 2010*"

APPENDIX

Appendix 1 - Definitions

AOP See *Archive Operation Parameter*.

Admin Files Opposite of case files. A category containing documents whereby the business activity is ongoing without any defined activity end date, e.g. “Travel”. Documents are qualified for disposition once they reach a certain age, i.e. after 2 years. Individual documents reach disposition and are processed independently of the other documents of the same subject. Individual documents can reach the age of disposition, while those not yet aged to disposition are left behind.

Archive To move a document from SharePoint to Content Server, using AGA. The document(s) being moved may or may not necessarily be records, or they may be records that are not managed as records (undeclared records).

In traditional recordkeeping, an archive refers to a storage location for long-term records. The process of moving records to an archiving facility is called **Transfer** or **Accessioning**.

Archive Operation An Archive operation is the process of moving one or more item from SharePoint to Content Server. It can be initiated manually, or via an automated process.

Archive Operation Parameters **AOP**. A set of parameters that AGA needs to complete an archiving operation. The parameters are essentially the answers to a set of (9) questions that must be answered for AGA to know how to handle a particular archiving operation.

C# (C Sharp) is an object-oriented computer programming language, part of the Microsoft .NET framework.

CAML **C**ollaborative **A**pplication **M**arkup **L**anguage. The XML-based language used to build and customize Web sites based on SharePoint Team Services from Microsoft. Used to create custom queries inside configuration pages of AGA.

Case Files Opposite of Admin files. A category containing documents whereby the business activity has a defined end date. Typically a *Person*, *Place*, *Event*, or *Thing*. Disposition is triggered by a date, such as “*End of Useful Life*” (A machine), or “*Close of all Legal Matters*” (A workplace Accident). Example – employment files, typically qualified for disposition 3 years after termination of employment. All records within a case file reach disposition and are processed as a complete, intact group – they are never separated or processed as individual records. Many organizations have 50% or more of all business records as case files.

Classify The process whereby a user assigns a formal retention rule to a document, as part of the Declaration process. Typically achieved indirectly, by virtue of selecting a storage location (typically a folder) that matches the subject of the document, and which bears the appropriate retention rule for that subject.

CMP **C**ontent **M**igration **P**ackage. Filename extension applied automatically by SharePoint to a file used to store a complete SharePoint site during backup or export.

Declare Make a document a record. Once declared, a document is generally tracked by software, and prevented from deletion (locked), except via a formal disposition process. It also has been classified against the corporate retention schedule.

Deletion As distinguished from ***Disposition***. An arbitrary deletion by a SharePoint user/administrator, or process. Deletion of a document by any method other than Disposition. Some examples of deletion:

1) Retention schedule has expired and SharePoint automatically deletes the document in accordance with its retention policy.
2) End user with deletion privileges deletes the document
3) A SharePoint administrator deletes a document

DFAS

Document/**F**older **A**rchive Settings. A configuration file that records AOPs to be used during AGA archiving. There are two types of DFAS files – Global, and Site. The Global DFAS specifies AOP settings for all SharePoint sites. Site DFAS Files specify AOP settings for specified sites, which overrides global DFAS settings.

Disposition

As distinguished from ***Deletion****.* Formal, structured process of determining what happens to records at the end of their retention period. The process is human-initiated, and the decision as to what is destroyed/transferred is ultimately decided by a human (the Records Administrator). Disposition yields (3) possible outcomes:

1) **Destroy** (typically 95-99%)
2) **Transfer** to Archive (typically 1-5%) for permanent archival storage
3) **Unknown**. Retain until disposition is known. Some Possibilities:
   - Held for legal review
   - In Dispute
   - Disposition simply not yet know or decided

Document

A single piece of recorded information stored in SharePoint. A document may be in one of the following "states":

- Being managed as a record (has been declared)
- Meets the criteria of a record but is not being managed as a record (not declared)
- Not a record. Does not meet the qualification of a business record, according to the organization's definition of a record.

Document Status

Indicates if a document, at the time of declaration into a system, is a **Record**, a **Work In Progress** (records not yet ready to be declared), or a **Reference** document.

Document Type

An attribute of a document that determines the metadata fields assigned to the document. Allows a common set of metadata fields to be applied to related documents.

DoD 5015.2

A US Government standard specifying minimum capability of software for the purpose of managing electronic records. Can be found at http://jitc.fhu.disa.mil/recmgt/standards.html.

EL Web Part

A SharePoint Web part that allows the SharePoint user to work directly with documents in Content Server. It presents a user interface into Content Server. It presents such functionality as Search and Retrieve, View, and Edit documents in the Content Server workspace.

Enterprise Library

The component of Content Server 10 used to organize and store documents. The repository of Content Server 10. All documents stored in Content Server reside 2010 somewhere in an Enterprise Library. "*Storing* into Content Server" and "*Storing into* EL" mean the same thing, i.e. store a document into Content Server (into Content Server's library).

Event Date

The date that triggers the Disposition of a set of record(s) in a case file. Suppose for example a retention rule for Contracts of "*Destroy 2 years following contract end*". The trigger date for records about contracts is therefore the contract end date. If the contract for Safe-Tee Security for instance ends Jan 15 2009, the records are qualified for disposition on Jan 15 2011.

Expunge

Means of deleting (destroying) records from the system such that they cannot be recovered by any manner. Contrasted with Deletion, in which there is no guarantee that the deleted record may be recoverable in some manner.

F1000 RIMtech term referring to the subset of DoD 5015.2 capabilities required to satisfy the recordkeeping needs of Fortune 1000 firms.

Hyperlink or **Link**. A reference to a document that the reader can follow (click on) to access the document. Instead of having a copy of the document, a user can instead click on a link to the document, and display the document.

Interactive Archiving The process whereby a SharePoint user can, at a time of their choosing, initiate an archive operation of one or more documents from SharePoint to Content Server. The user must supply operation parameters when initiating the operation.

Library A location in a SharePoint site used to store and manage documents, and to collaborate on their creation.

Merge Merge the searching of both Content Server and SharePoint within SharePoint such that documents in both Content Server and SharePoint are searchable within the native search function of SharePoint. Internally, the search indexes of both products are merged inside SharePoint. AGA refers to this as *Protocol Handling*.

Mode A method or approach of moving a document from SharePoint to Content Server. There are five modes; *Copy, Move, Create Link, Child Link*, and *Synced Copy*. The mode is one of the AOPs that must be defined for each archive operation, whether interactive or profile-based.

OPR **O**ffice of **P**rimary **R**esponsibility. The “home” location of the (electronic or paper) records of a specific category. These records are considered the “master” records, as distinguished from secondary (transient) records classified to this specified category. Constituent records contributing to this category may well be stored outside the OPR’s location. Typically the OPR represents the business unit with the most operational need for these records, and the hence the most authoritative position regarding their storage, use, and disposition. It is a common practice to apply the official applicable retention rule to OPR records and a different, more temporary retention rule to records of the same category stored outside the OPR (such as 2 years then destroy).

Profile A set of Archiving Operation Parameters (AOPs). Used by AGA during an archiving operation to determine *what* to archive, *where* to archive in Content Server, and *how* to archive it Profiles are assigned to a SharePoint Library, or a SharePoint Site.

Profile-Based Archiving The process whereby an archive operation can be triggered (initiated) by a pre-defined set of criteria (rules) in SharePoint. Operation parameters are pre-determined. User input is not required.

Provider A well-defined module of AGA functionality. All of AGA’s features are delivered in the form of 70+ providers, and you can custom-develop your own provider to extend AGA with custom functionality.

Record A document that meets the criteria of a record. A document that is being managed as a record. Smallest atomic unit that can reach disposition. Characteristics of a record:

1) It is assigned a retention rule from the official retention schedule
2) It cannot be deleted, except by the formal records disposition process (“locked down”).

Possible forms of a record:

1) *Electronic Document*. Any electronic format (.DOCX, .PDF, TIFF, etc.)
2) *Non-Electronic Document*. Metadata of a physical document may be recorded and tracked, but the physical document is not available in electronic form. Single document, electronically profiled, tracked, recorded, but content not stored or readable. Could be paper, photograph, or a non-paper artefact (microfilm, map, core sample, etc.)

3) *Box*. Multiple documents, stored in a container (box), tracked to box level only.

Record Series A category within the Content Server file plan. A retention rule. For example *Contracts, Safe-Tee Security 04-50-1* would be called a **Record Series**. Also known within AGA as a *RM Classification*.

Retention Schedule Also known as a **File Plan**. The list of approved retention periods and disposition rules for each business activity or subject within the organization. Driven by legislative obligation (various laws and regulations that apply to the business), and operational corporate policies.

Schedule Two meanings:

1) **Noun**. The “schedule”. The retention schedule (list of categories and their retention rules)
2) **Verb**. To “schedule”. The act of applying the retention rules to qualified records.

Site A SharePoint **Site** is a collection of SharePoint pages, lists, and libraries configured for the purpose of achieving an express goal. A site is a place where documents, and items related to those documents, are created and stored for some common purpose or project.

Real-Time Disposition Process of SharePoint whereby documents with an assigned retention schedule, where the conditions for end-of life have been met, are automatically destroyed without any human intervention.

RM Records Manager/Administrator.

Transfer Move record to a physically separate storage location for permanent storage. Remove from e-records repository. Also known as “accessioning”. Not needed for most F-1000 recordkeeping

Web Part A section or “part” of a SharePoint page that is custom-built outside of SharePoint by a third party.

Appendix 2 - Sample File Plan

The sample file plan fragment below shows admin files in blue shading, and case files in green shading;

File Plan Fragment						
Title			Description	File Number	Retention	Disp.
Primary	Secondary	Tertiary				
Administration 02	Travel 10		Travel application forms, approvals, Policies	02-10	2 years	Destroy
	Insurance 20		Insurance coverage, policies, renewals	02-20	7 years	Destroy
Human Resources 03	Employees 10	Smith, J 707	Employee files including offer letters, contracts, disciplinary records, promotions, etc.	03-10-707	2 years after termination of employment	Destroy
		Rubble, B 909	Employee files including offer letters, contracts, disciplinary records, promotions, etc.	03-10-909	2 years after termination of employment	Destroy
	Policies 20		General policies on employment and workplace safety, business conduct, etc. Drafts and approved copies	03-20	2 years or until superceded	Destroy
Finance 04	Accounts Receivable 10		Accounts Receivable	04-10	7 years	Destroy
	Account Payable 20		Accounts Payable	04-20	7 years	Destroy
	Budgeting 30		Budget preparation, proposed and approved budgets for current year	04-30	5 years	Destroy
	Reporting 40		Reporting obligations such as SOX, Environment Ministry, etc.	04-40	3 years after fiscal year ends	Destroy
	Contracts 50	Safe-Tee Security	Drafts and Master contracts, Negotiations, Cancellations and terminations to all contracts	04-50-1	2 years after contract end	Destroy
		Tasty Catering	Drafts and Master contracts, Negotiations, Cancellations and terminations to all contracts	04-50-2	2 years after contract end	Destroy
		Lovely Landscaping	Drafts and Master contracts, Negotiations, Cancellations and terminations to all contracts	04-50-3	2 years after contract end	Destroy
	Tax 60	Return 2008	Tax Return 2008, inc. challenges, supplementary supporting documents, notices of assessments, etc.	04-60-2008	7 years after receipt of Assessment	Destroy
		Return 2009	Tax Return 2008, inc. challenges, supplementary supporting documents, notices of assessments, etc.	04-60-2009	7 years after receipt of Assessment	Destroy

🟦	Subject Files
🟩	Case Files