

## On the Record with SharePoint

Mike Alsup, Gimmel Group

This column focuses on the Enterprise Content Management (ECM) and Records Management (RM) dimensions of SharePoint, including both Microsoft Office SharePoint Server (MOSS) 2007 and the upcoming release of SharePoint 2010. We explore the broad solution ecosystem that exists for SharePoint content and records management. This week, we focus on Enterprise Information Lifecycles and Content Governance.

### Solving Business Problems with an Enterprise Information Lifecycle

First, what is a lifecycle? A lifecycle is a succession of conditions through which information is processed from creation or receipt to its final disposition. A lifecycle is comprised of states, such as “draft” or “work in process” or “final/record”. Each lifecycle state is a point in the life of information during which processes, ordered activities which initiate the application of a set of business rules, are carried out. The enterprise information lifecycle establishes states for all information that is created or received in an enterprise through to its final disposition.

The reason that a standard enterprise information lifecycle is so important is that it enables the management of content according to an enterprise taxonomy or file plan (providing classification and naming standards) and retention schedule (providing retention management rules) regardless of whether the content is electronic or physical or whether the decisions are made by humans or automated processes. The information lifecycle defines processes, rules, and repositories that are globally applied across the enterprise to all information and are separate from most business processes and applications. The lifecycle can be defined to support the full existence of information across multiple types of repositories and can support the infrastructure requirements of the largest organizations. The result is consistent content governance combined with consistent user expectations for how they should participate in the management and retention of enterprise information.

Below is a graphical example of an Enterprise Information Lifecycle:

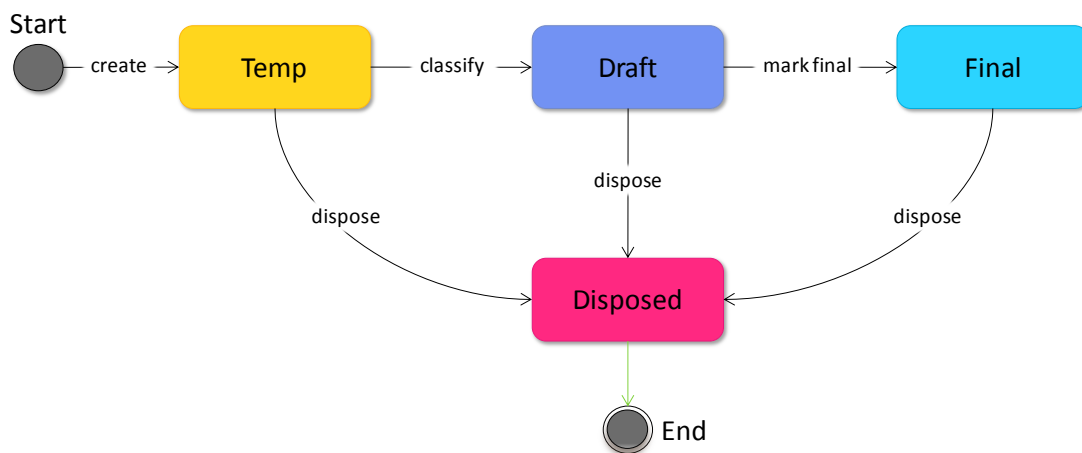


Figure 1 - Information Lifecycle Diagram

These states and process components (represented by movement between states) comprise an enterprise information lifecycle model. How these states relate to one another might be different depending on the

business unit, repository, type of content and associated business rules; but information always needs to exist in one of these states.

Business rules identify the constraints to information which prevent or enable it to move between lifecycle states. Business rules are defined for a variety of information elements such as a business event (e.g. document approved for use) or system trigger, audit and retention requirements, file format requirements, de-duplication rules, security rules, metadata rules, etc. Each of these “rules” represent a specific set of enterprise guidelines that apply based on each unique representation of the lifecycle model for an application or process. Defining business rules for each lifecycle state provides the basis for consistency in the enforcement of these rules. For example, transition to the ‘final’ (record) lifecycle state will be the trigger that ensures the appropriate retention policy is applied.

Without a standard enterprise information lifecycle, organizations will continue down the same path they have been experiencing:

- Business units find themselves with too much information
- Productivity and business effectiveness decreases as users struggle to find the information they need to do their jobs
- Business units implement processes to control their information that are inconsistent and incompatible
- Information is often stored in a random and inconsistent locations
- Minimal information is captured about the document or it is not standardized
- Information may become orphaned, and no one knows what it is
- Users have too much or too little permission to access the information
- There is no clear determination how accurate or stale information might be
- There is no clear determination what should be kept – too much information will be retained for too long or not long enough

These issues are even more common in SharePoint environments that have been implemented without consistent standards or content governance.

### **Defining an Enterprise Information Lifecycle in SharePoint**

A SharePoint information lifecycle typically begins with the definition of a baseline enterprise information lifecycle model that defines the states and the transitions through which content flows from creation through to destruction. The enterprise information lifecycle is further defined into a specific set of components, including an enterprise retention schedule, business rules, requirements for viewing and navigating to content, tagging content for search and retrieval, and rules and procedures for metadata management for each type of content in the SharePoint repository.

Processes and procedures for information lifecycle management identify the formalized steps or actions which must be undertaken or performed to meet these rule conditions and address the application and integration of retention schedules. These processes and procedures include but are not limited to:

- Information creation
- Information storage
- Information organization, including creating taxonomies and assigning metadata
- Information searching and retrieval

- Records declaration
- Information disposition

A lifecycle model provides business units with a default set of best practices for information management within their SharePoint sites. Use cases depict how the model and its associated rules, processes and procedures would be defined using specific examples, such as the creation of Word documents, InfoPath forms, email, file shares, SharePoint lists and libraries, and paper. These should indicate default events and common exceptions, such as the Hold state, that information may enter. Use cases are also often used as tools for promoting user awareness and change management.

The information lifecycle model can be used to support many different types of applications, processes and content types. Nearly all recorded information in all sizes of organizations can be managed using fairly simple enterprise information lifecycle states. While the use of lifecycles in SharePoint typically results in a single lifecycle model, multiple models may be necessary to support the wide variety of information and record types created by the organization.

### Mapping Lifecycles and Content Governance to a SharePoint Site Hierarchy

As SharePoint sites multiply across the organization, it is imperative that the information and documents contained in these sites be managed in a consistent way. Figure 2 below illustrates how a large organization might implement a hierarchy of SharePoint sites to include My Sites and Team Sites that support the requirements of a wide variety of types of teams as well as enterprise sites that provides portal capabilities and application connectivity to an organization.

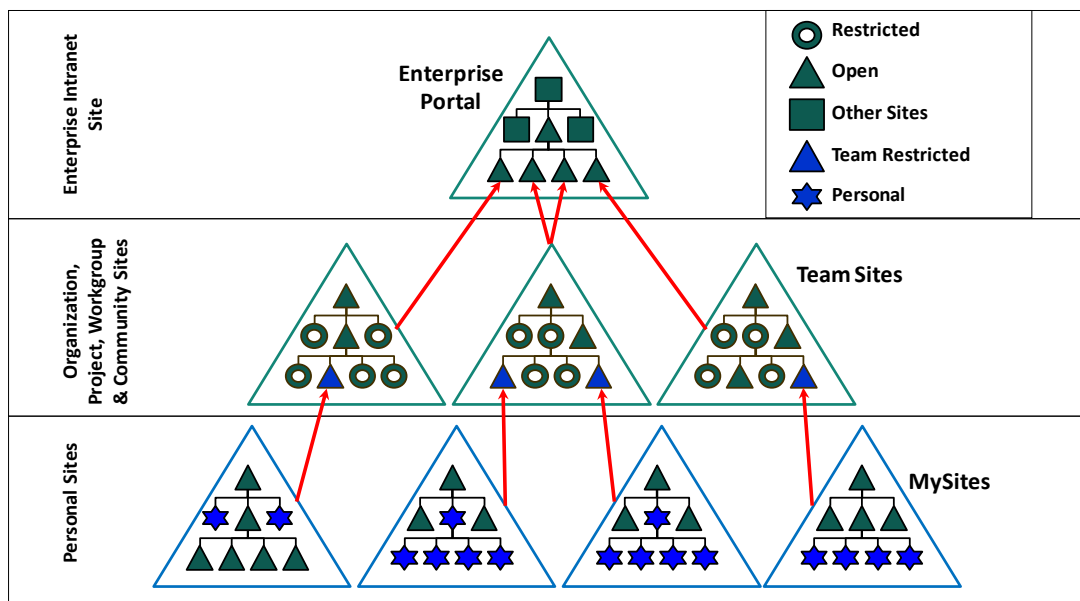


Figure 2 - Example SharePoint Site Hierarchy

SharePoint 2007 introduced new capabilities for records management. This included the SharePoint Records Center which contained the Records Routing Table and was the repository for content that had been declared as a 'record'. The information lifecycle defines standards for the flow of information into the Records Center from other SharePoint sites. The role of the Records Center is illustrated in Figure 3:

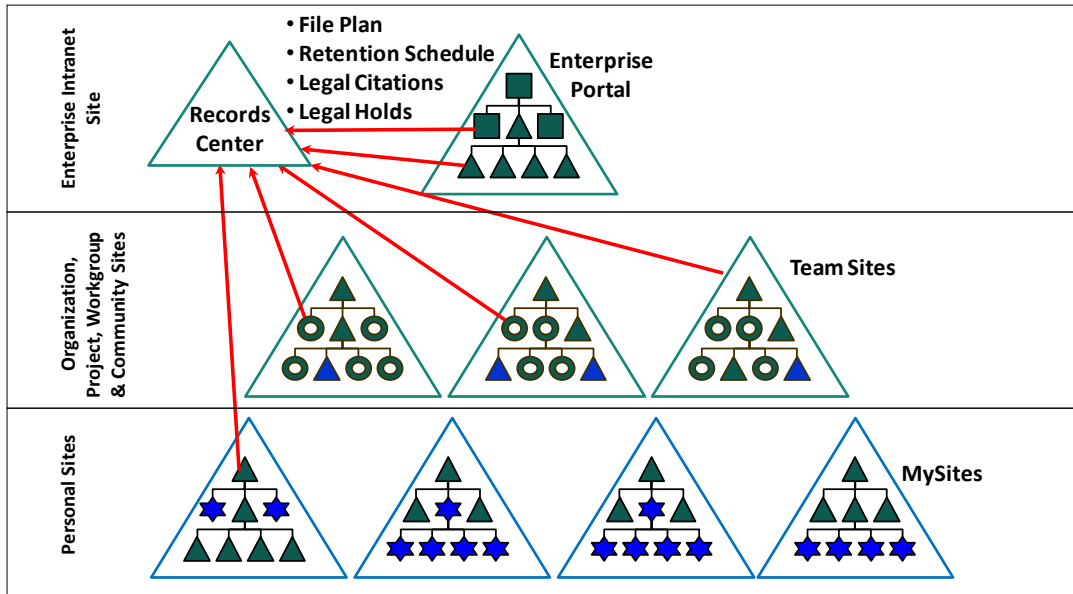


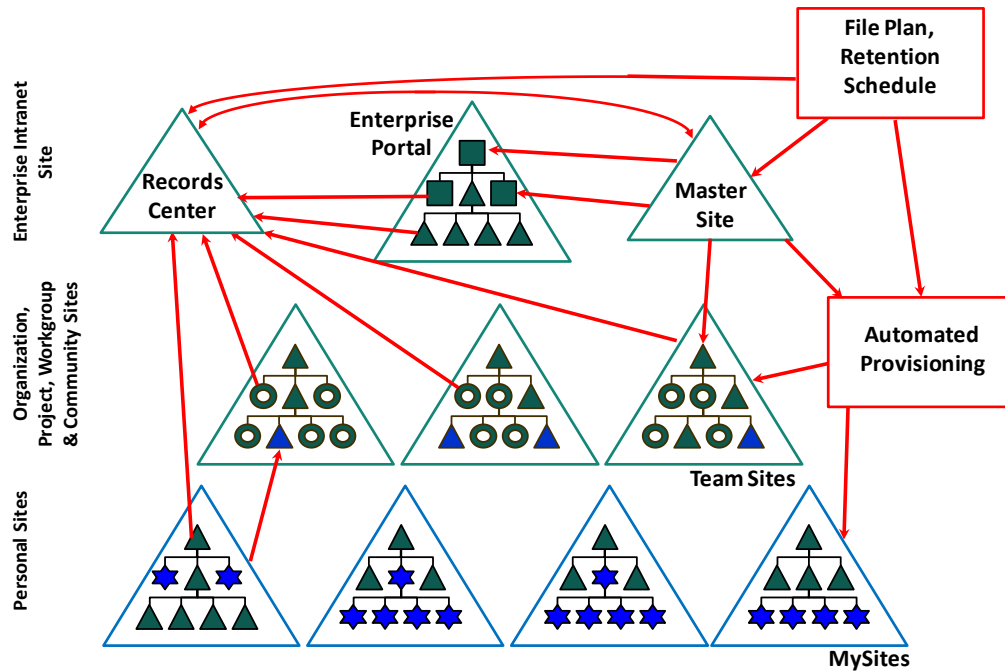
Figure 3 - SharePoint Record Center

All information is either manually or programmatically submitted to the Record Center defined for the SharePoint farm. Depending on the content types, logical content definitions within SharePoint, the information is stored in the appropriate document libraries within the Record Center based on rules defined in the Record Routing table.

The SharePoint 2007 Record Center has a unique set of default behaviors that results in the requirement for careful analysis, planning and, potentially, customization to fully support the typical lifecycle requirements for content and records management in large organizations.

### Information Lifecycles for SharePoint Sites

A primary focus of SharePoint Content Governance is to provide an information lifecycle to manage the retention and disposition of content that is maintained in SharePoint sites. The figure below illustrates an approach for SharePoint sites that incorporates the SharePoint Records Center, a Master Site for standard Content Types and Information Policy, an automated provisioning tool for SharePoint sites, and a File Plan and Retention Schedule tool to manage the synchronization and coordination of retention and disposition policies for the SharePoint Records Center.



**Figure 4. SharePoint Master Sites and Auto-Provisioning**

In SharePoint 2007, a Master Site is often used to maintain standard set of content business rules that are managed in a single location. This set of business rules includes the standard SharePoint content type structure, a base set of metadata properties for the enterprise and retention management policies. The master site contains these rules and structures and is used to publish this classification information to sites within the farm. This provides the ability to classify and manage content in a standard manner across and within numerous sites.

The master site includes the definition of metadata properties, or attributes. These attributes are inherited in the content type hierarchy, and this is integrated with the SharePoint Records Center and Records Routing Tables to enforce an enterprise lifecycle. A single Master Content Type, which is a parent that manages and controls the federation of a common definition of enterprise standard metadata, defines the subset of metadata that is required for all content to be classified. We found this to be a best practice in SharePoint 2007.

Automated Provisioning is an approach to rapidly create and deploy structured and governed SharePoint sites. These tools insert all of the enterprise rules and SharePoint features that enable a standard information lifecycle, taxonomy elements, site navigation, search, security and other required capabilities into every site in the same way. This ensures that the information lifecycle and governance policies are consistently deployed and consistently enforced.

In addition to the consistent deployment of a common set of SharePoint structures, the Automated Provisioning includes the definition of site administration features that ensure that enterprise governed controls are enforced and not circumvented through user or site administration activities.

SharePoint Content Governance needs to let SharePoint be SharePoint. Deployment standards are required, but users need to be able to customize their SharePoint for their purposes. Organizations need to manage

the consistency of the overall Information Architecture while enabling site owners to create both customizations and documents using a variety of standard document templates in a governed and structured way. We believe it is a best practice to enforce standards through automation and by limiting the choices that are available to both users and local site administrators.

In SharePoint 2010, the Managed Metadata Service and the Content Type Hub use a similar model to incorporate content governance into overall SharePoint governance across SharePoint deployments. This is an important advancement and will make site content administration much more consistent. Additionally, records management will be extended into local sites and records will be able to be declared outside of the Record Center, so records management goals will be able to be achieved more flexibly. We believe the new capabilities of SharePoint 2010 make it even more important to think through your requirements for enterprise content and records management across the organization before deploying hundreds or thousands of sites.

A standard enterprise information lifecycle provides a powerful approach to content governance and records classification that enables users to manage their information consistently across the SharePoint sites of an organization in SharePoint 2007 today and in SharePoint 2010 in the future.

---

Mike Alsup is a Sr. Vice President with Gimmel Group ([www.gimmel.com](http://www.gimmel.com)), an ECM and RM systems integrator. He blogs at ([kj109.wordpress.com](http://kj109.wordpress.com)). He welcomes comments or scathing remarks at [malsup@gimmel.com](mailto:malsup@gimmel.com).