



Robert L. Bogue
MS MVP, MCSE, MCSA:Security
317-844-5310
Rob.Bogue@ThorProjects.com

How SharePoint is a Different Kind of ECM

The allure of SharePoint is pretty compelling. It is radically cheaper to acquire, radically cheaper to implement, and boasts a radically better user interface -- after all, it has great integration with the key set of productivity tools in Microsoft Office. So it's easy to see why organizations of every size would look to SharePoint to handle their enterprise content management needs. Certainly, there are reasons why SharePoint is a great tool; however, there are also reasons why you can't approach SharePoint from the perspective of a traditional content management system.

What Do You Think Of When I Say "ECM"?

Industry insiders, consultants, and implementers have developed an understanding of ECM systems as process enablers. Whether they are recording and filing invoices, contracts, or packing slips, ECM systems typically have been implemented in places where the number of documents — and the findability problem associated with those documents — is great.

It's easy for most folks to remember rows and rows of file cabinets overflowing with paper. Locating a document in these rows and rows of file cabinets was done very rarely, because the cost of manually retrieving the paper was so high. However, this meant that important checks to determine whether or not the invoice had been paid, or the work was being double billed, were skipped, causing organizations to lose millions of dollars each year. ECM systems, some of which quite literally cost millions of dollars to implement, were implemented to improve the speed of filing and retrieval, and to reduce the amount of physical space required to maintain the records.

I once worked on an ECM project where the paper in the building was quite literally crushing the limestone foundation on which it sat. Simply managing the space that record keeping required had a tangible impact to the bottom line. Due to the cost required to implement an ECM system, the technology has historically picked up small efficiencies in processes that have very high document volume. The sweet spot for ECM as a market has historically been these high-volume operations.

Similarly, because these processes often meant real money, the documents and their metadata were tightly controlled. Accounting decided what metadata would be assigned to invoices, and because of the relatively constrained set of people performing the ingest process, everything was relatively well controlled.

Expensive scanners and ingest tools meant that providing a centralized scanning facility was the most cost-effective solution to getting documents into the system. With a relatively small number of people ingesting documents, training was easy. The centralized facility generally meant a group of people who were dedicated to the process and who could be taught the importance of good metadata entry, proper scanning, and what happens when mistakes are made.

SharePoint's Roots

SharePoint's roots — where it came from — is on the opposite side of the universe from the high-volume, highly structured, narrow participation of a traditional ECM system. SharePoint cut its teeth as a collaboration tool. It was the benefactor of the empirical process. Take, for instance, responding to a request for proposal (RFP). The exact process that is followed, and the number and types of artifacts in responding, vary by the RFP itself. One RFP may require an engineering drawing, the next an artist's rendering, and so on. Where an invoice is received, processed, and approved in a defined way — the things that SharePoint grew up with — was much less defined.

SharePoint collaboration has tended to revolve around documents created in Microsoft's Office applications. Instead of having to ingest a document through a scanning process, one would simply select File-Save As from the menus in Word, and the document would be automatically added to SharePoint for others to use. Gone were the expensive tools for entering metadata about the document. Similarly, retrieval was either handled by selecting File-Open or through navigating a relatively simple Web site. SharePoint grew up as a replacement for some of the things that the file server was doing. Along the way, the difficulty in specifically defining metadata and developing comprehensive taxonomies meant that there was a very real importance to full text indexing and searching. Since most of the documents started their lives as electronic files, the problem of full text searching in a corpus of documents was very possible. So where the structure for metadata in an invoice process might be well known and well defined, SharePoint has historically had a very ad-hoc organization applied to it.

This outlines the key differences between a traditional ECM, which is high volume and uses tightly controlled metadata, and SharePoint, which is of more moderate volume and leverages a more distributed and flexible metadata strategy.

Control And Metadata

One of the most palpable differences between a traditional ECM system and SharePoint is in the level of control exerted on the users and on the information that they store. ECM systems typically have a tightly controlled set of document profiles that detail out what metadata is required for a document and what metadata is optional. These document profiles are typically centrally controlled and, from a process perspective, are difficult to change.

SharePoint offers a concept that defines a document type in a similar way. It defines the columns (metadata) associated with the document as well as the workflows that should run on the document. In this way, content types are similar to a traditional ECM's document profile. Content types differ in that anyone who is an administrator or owner of a site can create their own content type. The content type they create has its visibility constrained to the site in which it was created.

The net effect of this is that you are likely to have more content types in SharePoint than in a traditional system — the same content type name may mean different things in different site collections, and just because a document profile exists in one site doesn't mean it will exist in another.

However, this represents a change in context. SharePoint is an enabling and empowering tool designed to deliver flexibility to its users in terms of how they use the tools and how they configure them. It's possible, with some XML "coding," to ensure that content types do exist in every site. The harder task is deciding to restrict access to users adding their own content types to the system.

It's harder not because the facilities aren't there in the system to restrict access, but because it goes against the general premise of the tool. By breaking the ability to define content types, you effectively limit the users' ability to create their own solutions to problems for which a traditional ECM system would not be considered. For instance, what about a help desk system? Perhaps there's an InfoPath form that is associated with a workflow. Generally speaking, a help desk system wouldn't be big enough to be part of a traditional ECM system, but in SharePoint the same technologies brought to bear on invoices can apply to much smaller collections of documents.

Scale And Performance

SharePoint's primary design criteria weren't the management of large-scale data sets. Whether that data is row-and-column type data, as might exist in a database table or a collection of documents, or metadata like an ECM system manages. Some of the primary design criteria were ease of use and ease of change. As a result, SharePoint's performance characteristics, in general, aren't the same as other systems. SharePoint was optimized for Office documents. In fact, by default SharePoint doesn't allow documents to be uploaded that are larger than 50 MB.

In some ECM systems, a great deal of time has been spent with viewers that allow progressive viewing of a file, so the entire file doesn't have to be downloaded to the client machine. Similar optimizations are yet to be made in SharePoint, which generally assumes that the document may be edited (due to the software's collaborative roots).

So a traditional ECM system is quite likely to have an edge on SharePoint when it comes to both scalability and performance — in the large scale. SharePoint will perform well if you just need to store a few hundred thousand documents in a library, assuming that you're breaking down folders in a reasonable way. If you break the problem down in the right way — across document libraries — you can easily scale to millions of documents. So for the medium-sized organization or medium-sized ECM need, SharePoint is likely to be good or good enough.

What Does This All Mean?

In short, SharePoint is a capable ECM system for medium-sized organizations or departments that don't have tight process optimization tied performance constraints. Its sweet spot is in organizations that don't need to organize one or two processes but instead need to organize a plethora of processes, each of which doesn't generate a ton (say more than a few million pieces) of content in a single year.

Here we've been focused on the differences between SharePoint and other ECM solutions. While that's a valid comparison, it ignores the larger reality that SharePoint offers a plethora of other features that can't be compared to a traditional ECM system. Full text search and collaboration tools (including version control, checkin and checkout, Web content management, and others) may bring SharePoint into your organization. Once you have the platform, it's up to you to determine whether or not it should be a part of your ECM strategy.

About The Author

Robert Bogue, MCSE (NT4/W2K), MCSA:Security, A+, Network+, Server+, INet+, IT Project+, E-Biz+, CDIA+, is president of Thor Projects LLC, which provides SharePoint Consulting services to clients around the country. He has contributed to more than 100 book projects and numerous other publishing projects. His latest book is The SharePoint Shepherd's Guide for End Users. (You can find out more about the book at www.SharePointShepherd.com.)

Bogue has been part of the Microsoft Most Valuable Professional (MVP) program for the past 5 years. He was most recently awarded for Microsoft Office SharePoint Server. Before that, Bogue was a Microsoft Commerce Server MVP and Microsoft Windows Servers-Networking MVP.

Bogue runs the SharePoint Users Group of Indiana (SPIN, www.spindiana.com), and he is also a member of the steering committees for the Indiana Windows Users Group and Indianapolis .NET Developer Association. In addition to speaking at local and regional events, Bogue speaks at national conferences. He blogs at www.thorprojects.com/blog , and you can reach him at Rob.Bogue@thorprojects.com .