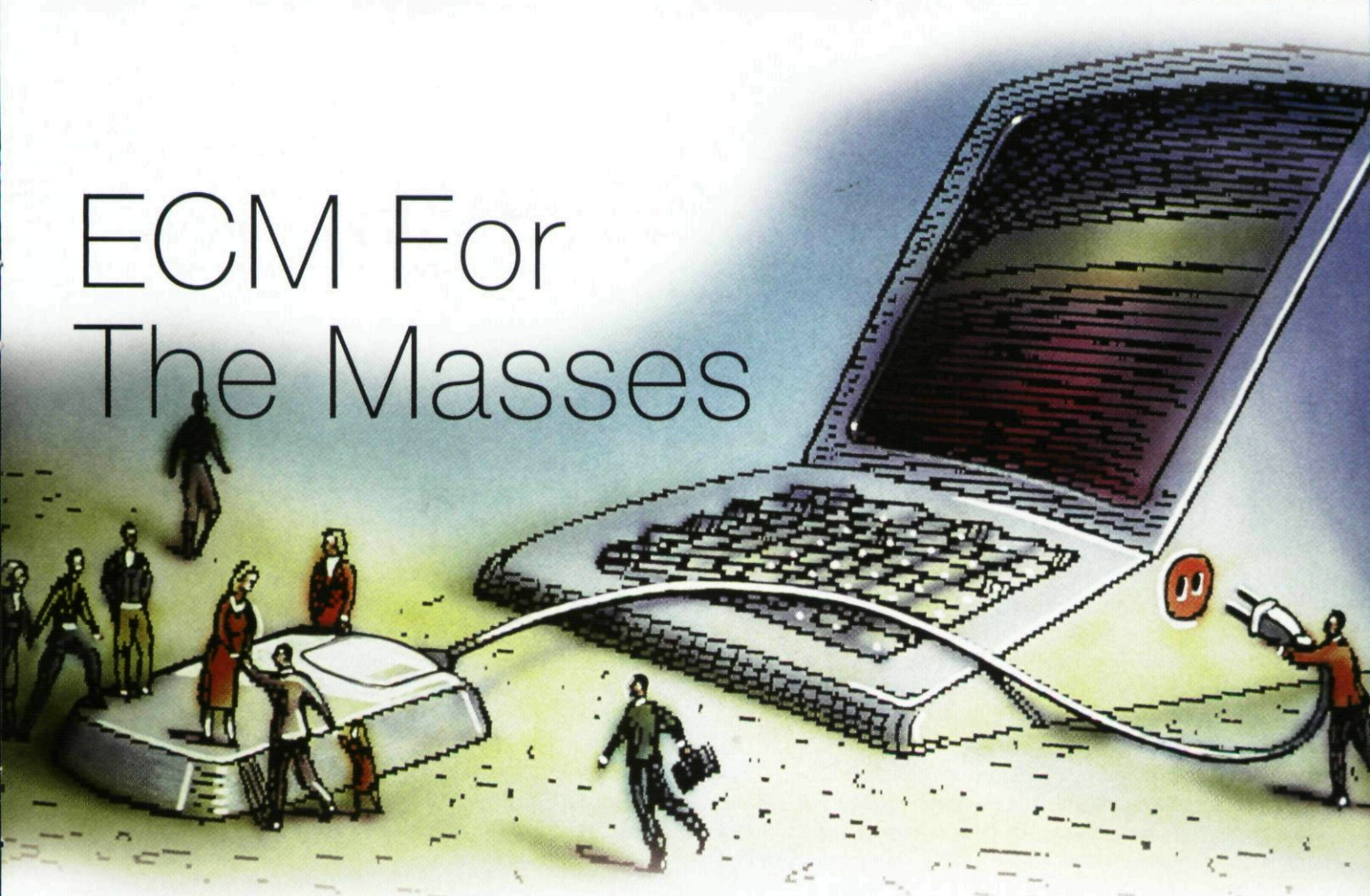
ECM For The Masses



Becomes A Reality

SOA, XML, and Web services are three technologies tipping the scales toward ECM for everyone.

By Russell Stalters and Alex Holcombe

So, what do we mean by enterprise content management (ECM) for the masses? ECM technologies have matured over the last several years significantly, but for the most part most ECM solutions still approach managing enterprise content as point applications. So what's wrong with that?

Well, with the exponentially growing volumes of electronic content and the heightened risk of not effectively managing this content, organizations cannot afford having this content "slip through the cracks." We need ECM services (capture, manage, store, maintain, etc.) to be available pervasively throughout the organization. Not just available physically, but also available conceptually throughout the organization.

ECM services should be available to every application that creates unstructured content and should be accessible at any point in the lifecycle of that piece of content. Creating a document, reviewing it, making it final (a record), and destroying it when it comes to the end of its life should happen where I create it, when I create it, and as a result of my normal business processes. This is what is meant by "ECM For The Masses."

Why Is This Important?

Much of the content created and maintained by workers lives outside of any corporate ECM system because of the difficulty of using these solutions or because of lack of access to these systems. They store things on their PCs and decide when to save and when to delete content. This poses significant risk when this content is subject to regulatory requirements and it lives outside of the corporate ECM system. This information is not available to the organization and can represent locked-away corporate intellectual property. To mitigate these risks and make all information useful to the organization a "leak-proof" ECM system that is available throughout the organization is needed.

How Does It Become Reality?

IT departments struggle to meet the demands to provide a robust ECM capability that is easy to use and available pervasively throughout the organization. The tools for creating content have long existed, but effectively capturing and managing this content has largely remained a significant problem.

What is needed is a framework that allows IT to take

Brand X

advantage of existing functionality and corporate systems as well as provide the flexibility to respond as both the organization and technology continue to evolve. This framework must be extensible to all parts of the organization, independent of the underlying technology, and provide a foundation for future applications. The framework must also provide robust ECM capabilities as a service so that IT developers and software vendors can tap into the needed functionality when creating new business applications to support the organization's business. It must surface the needed capabilities and functionality in desktop applications, browsers, and email clients.

As content is created this framework must provide easy and seamless capture of content that is easy for end users. The ideal would be that every piece of content would be captured at the exact right time, have its lifecycle managed by the ECM system, and be available to anyone who needs to use the information from any application or business context. This ideal is a pretty lofty goal!

Why Now?

Over the last several years a technical revolution has been quietly brewing that is making this dream of "ECM for the masses" possible. The three major reasons, or tipping points, that are making this goal technically feasible are broad adoption and migration towards: Service Oriented Architectures (SOA), Web services, and Extensible Markup Language (XML) technologies.

The major ECM vendors will continue their move toward providing a comprehensive set of core Web services while smaller ECM and other ECM-related vendors will have to migrate their solutions towards providing specialized SOA compatible services to survive in this evolving market.

SOA What?

Service Oriented Architecture (SOA) is a framework of services that provide application functionality made available through well-defined interfaces. SOA is not new; however, as advances have been made in application development and architectural technologies, as well as advances in managing business processes, SOA is poised to become "the next big thing" which enables ECM services integration with all parts of the enterprise.

However, there were still limitations. In the past, components were often difficult to discover, integration interfaces were not standardized, and accessibility was often dependent upon the technology of the underlying systems platform. All of this changed with Web services.

The Explosion Of SOA With Web Services & XML

Web services provide the missing link between reusable components and accessibility across the organization. Their characteristics include standardized interfaces, easy accessibility, and irrelevance toward the underlying platform technology. This allows for reuse of functional components from multiple applications, transforming them from independent silos of business functionality into a horizontal collaborative platform available across the organization.

XML has become an enabler for Web services to support a SOA model. These services are broadly usable and available now to developers through a standard called WSDL. WSDL is an XML format for describing network services as a set of endpoints operating on messages containing either document-oriented or procedure-oriented information. Now the underlying infrastructure and developer tools required for creating and deploying Web services have advanced such that the average developer can now do so with ease.

Capturing Content With Web Services

In order for ECM for the masses to become a reality, content must first effectively be captured throughout the organization. The major stumbling block for doing this has been the fact that content was dispersed throughout multiple locations and within multiple systems throughout the organization. In addition to identifying and capturing typical content, such as email, documentation, and records, there is also valuable information in the system-system and human-system interactions which occur but is often difficult to obtain. Finally, information that relates to business processes as they happen within the organization must be obtained and is necessary for providing an accurate snapshot of how the organization is operating.

Content Contained Within Systems

Although the location of content needed for corporate ECM is often identifiable in back-end systems, inconsistent interfaces and technologies of these systems have resulted in the inability to easily access and capture this information. Such systems include those for email, Web content management, collaboration, and records management. Web services open the door to these systems by abstracting the underlying technology and standardizing the interfaces. Once made available, this content is then available to be captured and managed.

Content Created By Interactions

System-system and human-system interactions contain useful information that must also be captured and managed for true ECM. In the past, this information has been difficult to obtain and is often lost. In recent years, Web services have been used to manage these interactions as they occur both within the organization and with external entities. As this trend continues, the result will be the ability to capture and manage this information as it is being moved throughout the organization.

Business Process Information

As SOA starts to grow throughout the organization, the discrete pieces of functionality these services provide become strung together to form larger orchestrations that make up business processes. A new layer on top of this SOA starts to form and becomes the basis for managing and driving these business processes. Until recently there was no standardized way to manage the state of a process across these services. This all changes with the emergence of business process execution language (BPEL), an XML-based language for the formal specification of business processes. BPEL extends the Web Services interaction model and enables it to support business transactions. This opens the SOA platform and turns it into the foundation for a new application layer without regard to the technologies underneath.

What All of This Means

As these technologies and standards continue to mature, capturing content across the organization will become much more efficient and effective. The next missing link between what is available today and what is necessary for true "ECM for the masses" is the ability to capture content from the desktop when end users create it. We think this is changing.

New content management solutions from both Microsoft and Oracle make extensive use of, and interact directly with Web services. The latest generation of Office products from Microsoft has some built-in ability to directly interact with Web services and can be extended by 3rd party applications or custom-developed extensions. Microsoft's next version, Office 12, becomes a Web services ECM system and will include workflow, document and records management, and Web content management. Oracle is also fielding a Web services ECM system with the release of Oracle Collaboration Suite 10g including records management. Oracle also provides the ability to interact directly with Web services and BPEL via Oracle Forms.

We expect as these technologies and standards mature and SOA-based solutions are implemented throughout the organization, ECM for the masses will become reality. The major ECM vendors will continue their move toward providing a comprehensive set of core Web services while smaller ECM and other ECM-related vendors will have to migrate their solutions towards providing specialized SOA compatible services to survive in this evolving market. It should be an interesting time!

Russell Stalters (russ.stalters@compliancesolutionsgrp.com) is the president and founder of Compliance Solutions Group (CSG); a firm focused on delivering records management and compliance solutions for Microsoft customers. He was also president of TrueArc, a developer of digital preservation and records management software that was acquired by Documentum. Alex Holcombe is an ECM architect with CSG and has built numerous SOA-based ECM solutions for their clients.