Intelligent Metadata Enabled Solutions

By Martin Garland, President, Concept Searching

Martin Garland

One of the founders of Concept Searching, Martin Garland has more than 21 years' experience in ECM. His understanding of the information management landscape and his business acumen provide a foundation for guiding organizations to achieve their business objectives using best practices, industry experience and technology.

For the most part, gone are the days when SharePoint was a good choice to enable a department to share its documents and have a single point of access to information. As SharePoint has matured, so has the marketplace, and enterprise expectations of SharePoint have also changed. New challenges such as records management, security, migration and the cloud have changed the landscape. With regulatory requirements surrounding privacy, security and compliance increasing, the stakes are much higher. With unstructured data increasing at an average annual growth rate of 62%, search is inadequate and has made volumes of content unusable for applications, such as text analytics and collaboration.

As a metadata, classification, search and taxonomy provider, we have seen a dramatic shift in the market. One client with us for eight years has deployed the technologies to solve search, records management, data privacy and migration in SharePoint. A recent survey of all our clients and the broader marketplace indicated that organizations are seeking solutions that leverage their SharePoint investment, and provide the ability to reuse and repurpose content with an enterprise infrastructure platform for unstructured data that is flexible and easily adapted to address changing organizational requirements. In the age of limited funds and resources, organizations are also more focused on achieving a measurable ROI.

The crux of the problem has always been the inability to generate contextual metadata, either manually or automatically that results in *intelligent* metadata. Intelligent metadata is the output of automatically capturing conceptual metadata, and identifying relationships and context, enabling the proactive management of content through goal- and mission-aligned structures and workflows. This approach results in the deployment of intelligent metadata solutions that are integrated with an organization's unstructured information governance policies. The inability to identify intelligent metadata from within content affects the entire lifecycle management of the asset.

Search still ranks high on the list of market requirements. Organizations must recognize that search is not a standalone technology or application, but must be integrated with business processes and corporate objectives as a key infrastructure component. With the merging of the former FAST products into SharePoint 2013, search functionality is improved. Regardless of the enterprise search solution, search results will be irrelevant without intelligent metadata applied to the content and the content classified to achieve an organization's objectives. SharePoint's introduction of the Term Store provided rudimentary taxonomy tools, yet is still a manual process prone to subjectivity. SharePoint provides no functionality to automatically identify the concepts and relationships of the content and auto-classify it to one or more organizational taxonomies or the Term Store.

Migration is a key priority for executives in 2013. Organizations are now requiring more sophisticated techniques to ensure compliance objectives are met. It is estimated that 69% of existing unstructured content can and should be deleted. Information cannot be managed from inception to deletion without comprehensive metadata associated with the content. Documents can exist in multiple places at the same time, different revisions of the same document exist, some documents should be deleted, and others should be archived. There may be records that were never declared, as well as confidential or privacy information that was never identified. As an intelligent metadata solution, migration tasks are executed at the source and the accuracy of the content migration is greatly improved. As content is migrated it is first analyzed for organizationally defined descriptors and vocabularies residing within the content, then tagged and correct content types automatically applied. Finally, the application or changes to information rights management and/or retention policies are applied, driving document movement based on policy and workflow rules.

Records management using this approach includes the ability to identify records and fully automate the process transparently, without user involvement, and handle the appropriate disposition of the document. This includes discovering where the content resides, cleansing the content through organizationally defined concepts and descriptors, identifying the relationships within the content, applying the policies for automatic enforcement, changing the SharePoint content type where required, and routing to the records management repository. Proof of compliance and data protection can only be accomplished through the automatic enforcement of organizational policy to ensure consistency and can be implemented transparently, resulting in improved record keeping, monitoring and auditing processes.

The Risk Factor

Corporate risk is increasingly becoming a priority for senior management. The demarcation of who is responsible for the protection of data is becoming blurred. It can no longer be the responsibility of one person, and must encompass all stakeholders and end users to identify and secure privacy assets to protect an organization. Regardless of whether a breach is contained within an organization or—worst-case scenario—is exposed to the world, 76% of organizations will face investigations, forensics, law enforcement and actions to protect the victim from harm. Protection of privacy and confidential data includes the proactive identification and protection of unknown privacy exposures before they occur, as well as identifying in real time organizationally defined vocabularies and descriptors in content as they are created or ingested from any source and removing it from unauthorized access.

The development of an intelligent enterprise metadata framework provides flexibility, and enables organizations to implement governance policies and operations incrementally, with each step directly correlated to business value. From this overarching infrastructure framework, intelligent metadata enabled solutions can be deployed to address the unique challenges of organizations. These challenges may include search, compliance, information lifecycle management, sensitive information protection, migration, text analytics and enterprise/Web 2.0. From a business perspective, one set of technologies used for addressing these challenges, can be implemented incrementally, and leverages an organization's investment in SharePoint. ∎