taming
MULTIPLE SEARCH ENGINES
IN YOUR ORGANIZATION

BY JEAN BEDORD

O ne-size-fits-all enterprise search is dead—if it ever existed. At the same time, search that makes an organization's assets more readily accessible has become a critical mission. Customers, partners, and employees all expect to be able to "find answers" via search technologies.

Thus, organizations increasingly deploy search engines to help satisfy these expectations. These engines are typically called site search, department search, intranet search, or application search, rather than enterprise search.

However, the fact is that organizations that have one search application usually have several. A 2007 enterprise search survey I did for the publisher of the *Enterprise Search Sourcebook*, Information Today, Inc., and my employer, Shore Communications, found that 62% of respondents had more than one search solution in place. Noted industry analyst Steve Arnold reports that the typical Fortune 500 company uses solutions by at least five search vendors.

This proliferation (and multiplicity) may not be obvious, even to those within the companies deploying the solutions, since responsibility for search strategy is frequently undefined within the organization. One reason for this is because search doesn't exist independent of the content that is being accessed. It creeps in with add-ons to content management systems (CMSs), business intelligence (BI) applications, records management, document management, or knowledge management systems. In addition, ecommerce systems with baked-in search capabilities can add to the growing number of search vendors used by an organization. As an independent application, search can be tuned to specific business-process needs, which means it comes in many guises.

Thus, many companies face more than the already challenging issue of managing a single search engine: As the number of solutions deployed within an organization increases, distinct issues emerge in reconciling these disparate solutions. The major problem is that each search product builds indexes to the existing content using a "secret sauce" that's incompatible with other vendor indexes. Thus, finding answers across different application repositories, each with its own index, is not a straightforward proposition.

Strategic Planning for Search

One enterprise search myth is that all information in the organization should be accessible via search. Organizational content, however, is more difficult than the relatively simple world of consumer search on the open web, which is primarily HTML webpages and

unstructured content. Information created and controlled by the organization is complex, from the content perspective and the technology perspective.

There are some facts to consider about content and how it relates to search functions:

The value of content within the organization varies. Indexing a server that contains years of cafeteria menus is a waste of resources, creating yet more information overload. Indexing technical reports representing the intellectual capital of years of R&D, thus enabling idea discovery, has potentially high ROI. Improved findability for ecommerce companies, particularly those with large product catalogs, can result in increased sales.

Job functions require different content. A financial analyst may need current sales by product line, even down to the individual part number. Customers and partners want data sheets with technical specifications. Defining subsets of content for job functions is the domain of management, not technology.

Security complicates search. Privacy of employee records is essential, though employee benefit information should be made available to everyone. Access to customer records may prove useful for certain job functions, but it may be regulated by federal and state laws. Access to trade secrets is typically restricted to those with a need to know. Real-time identification of fraudulent transactions may be mission critical. Thus, access and security policies need to be in place before unleashing the power of search technology.

Technology Planning for Multiple Vendors

Another myth of enterprise search is the feasibility of standardizing on a single search vendor. Search engine software creates proprietary indexes and relevancy ranking, with each having different strengths and weaknesses. Products from the same company, say IBM or Autonomy, do not necessarily provide compatible upgrade paths as search applications grow. There are some simple facts every company must face when attempting to reconcile search solutions:

Legacy systems are part of the search landscape. Many business intelligence and content management vendors integrate search into their product suites with OEM relationships with enterprise search companies. This ensures best-of-breed functionality within the suite, but not necessarily between vendor suites.

Mergers and acquisitions play havoc with standardization. Acquiring a company is based on business fundamentals, not compatibility of software systems. Provided the existing systems work well at the application level, there is little ROI in switching search software. However, the consolidation of software as search companies are acquired by major players can result in product "orphans," leaving your organization without support.

Federated search needs to be part of the technology plan. Federated search refers to the capability to search multiple indexes without creating yet another index. Also called meta-, blended, or universal search, this capability can be implemented in various ways: real-time federated searching using on-the-fly queries, portals, webpage mashups, or structured top-50 queries. Implementation is highly dependent on access and security requirements, which can vary with each of the underlying content repositories.

Care and Feeding of Indexes

Search engine indexes and the information retrieved from enterprise repositories rely on the accuracy of the underlying data. Maintaining integrity of the applications that create content is crucial for business decision making, yet organizations typically underestimate the time and effort it takes to maintain clean and relevant content repositories.

Here are some issues that must be considered:

Records management is an underappreciated aspect of search. Regulatory compliance and ordinary retention schedules should be part of managing business risk. Indexes exist separately from the content, so they have their own update schedule. Overnight or weekly—rather than real-time—updates can make search results appear out-of-date. Search technology makes misspelled, non-compliant, and dirty data glaringly obvious.

Indexes grow as the numbers of records applications grow. Index files vary in size but can be an additional 50% to 200% larger than the size of the primary data files, depending on the expansion factors. Scalability of the search solution should be a major consideration in selection and implementation.

Relevancy of search results depends on context, not popularity. The purchasing department may work with a dozen "business card" vendors worldwide. A "business card" query for most individuals in an organization, however, means a short how-to procedure for ordering new business cards. Search technology alone doesn't solve the job of interpreting the query to provide relevant answers.

Managing expectations is a major challenge for search implementations within the enterprise. Search is a tool to provide answers to problems. Finding those answers across enterprise repositories requires a higher level of relevance than open web consumer search, yet the underlying content structure is more complex and access is more constrained. Just as there are no one-size-fits-all answers to organizational questions, enterprise search is not one-size-fits-all, and most organizations will need to manage multiple solutions to supply the different answers individuals seek.

JEAN BEDORD (jean@bedord.com) is a findability and search consultant with EContent Strategies, senior analyst at Shore Communications, Inc., and faculty lecturer at San Jose State University. Her report on "Enterprise Search Deployment, Usage, and Trends" is available at www.enterprisesearchcenter.com/Reports/Details.aspx?ResearchReportsID=38.]