Search Solutions for the Enterprise

Implementing search solutions for your enterprise

BearingPoint®
Management & Technology Consultants

Search solutions for the enterprise can help users leverage information to improve productivity, decision making and innovation.

INTRODUCTION: UBIQUITOUS SEARCH REQUIRES SOLID GROUNDWORK

Improving the process of searching for enterprise information offers many real benefits, including higher employee productivity, better leveraging of intellectual capital and greater innovation through improved information discovery. A good search solution helps deliver these benefits by making all data repositories, from corporate systems to the global Internet, easily available to users. These capabilities will help drive enterprise software toward the same functionality, transportability and ease-of-use as users have come to expect from the Internet.

To achieve this state, search solutions for the enterprise must be properly planned and designed for each organization. BearingPoint has identified key considerations in this effort: focusing on effective deployment, reliability and performance, and security and privacy.

STREAMLINING DEPLOYMENT

Installing search solutions can be as simple as implementing appliances on the enterprise network, pointing them at an intranet repository and turning them on. But while making data repositories search-enabled has some value, it is only a first step toward realizing the benefit of finding information throughout the enterprise. A full deployment takes comprehensive planning and a rollout plan. The rollout plan should reduce risks and manage schedules while minimizing complexity, extending the reach of search and continually improving the value of the search solution to a broad range of users.

Minimizing Complexity

Search solution deployments should not typically begin with a "big bang." To reduce complexity, deployments should start with well-defined goals and move toward an enterprise solution in controlled stages. This gradual deployment approach relies on the expandability of the search engine. Full deployment throughout the enterprise will take time, but benefits can begin immediately as users gain access to useful information and begin to trust the solution and its interface in the same way they do their favorite Internet search engine.

Also, as the enterprise deploys effective search solutions, users will quickly begin to expect that they can find more information. The technology rollout must be accompanied by a good communications plan to set proper expectations.

IN THIS POINT OF VIEW:

As the deployment adds repositories, it must manage the presentation of results to ensure that the user experience is effective and positive. Users will not adopt the tools if search results become cluttered or not relevant.

The deployment team should resist the temptation to build information taxonomies and inter-system ontologies in order to "organize" search results. Maintaining such relational data structures is not scalable and requires relating the importance and taxonomy of data between different repositories. Because the number of relations between repositories grows exponentially with the number of repositories, this maintenance activity can rapidly become unmanageable. The search solution must be able to rank results between systems automatically.

Also, the law of unintended consequences applies to the customization of organizationwide search ranking. Deployment teams should avoid applying special-purpose relevance and ranking algorithms. Search algorithm changes should be left to search specialists.

Extending the Reach of Search

Users often do not have a good idea of exactly what information they need or where that information is stored. Therefore, if search solutions have broad reach, the search results page can provide a good starting point for finding what users need to do their job.

Internet search shows how information is often found in unexpected places, starting from one location. However, a good search solution for the enterprise is valuable to users for more than discovery. Often the searcher knows what is needed, or even where it is, but that information is inconvenient to get to because of the need to launch and use multiple applications. A properly designed search solution can make the retrieval process easier. The goal of the deployment should be to extend search to become as universal as is practical.

Much of the information stored within the enterprise changes slowly, or not at all. Search engines compile and index this "static" information through automatic crawling or by receiving feeds from repositories that cannot be crawled.

Deployment of search solutions for the enterprise typically starts with Web-based repositories. During the process, the deployment team will discover additional non-Web-based repositories that contain valuable static data. The team will need to obtain, test and install adapters to extract that information.

However, much valuable information inside the enterprise changes rapidly. A financial services organization uses a great deal of analysis and real-time data. It is important that users be able to retrieve this "dynamic" data along with the more static information through a single interface.

Indexing of dynamic data, such as financial information, is not practical. When a user makes a specific request for such data, the best course is for the search engine to query the actual data repository and return the dynamic data at the time of the search.

Organizations should determine which data is important to users. The deployment team can then select appropriate software adapters and configure them to conform to the enterprise infrastructure and access methods. Some adapters provide more complex data access, enabling "deep search" into enterprise databases. Deploying such adapters extends the reach of search beyond intranet pages.

Continually Improving Search Solutions

Search is an inherently "fuzzy" activity. Much information is stored in unstructured documents that, no matter how well written, only approximate the thoughts of the author. Searchers then attempt to locate these documents by typing short requests of a few words. Search queries usually only approximate the intent of the user.

An effective search solution tries to match the author's thoughts with the searcher's intent. This is still an inexact science. Only in rare cases today can the best search engine provide an exact match between result and intent. Search engine specialists continually improve the capabilities of their products through the use of analytics—data showing what searchers have looked for and what they have found—to improve the usefulness of their algorithms. These analytics also describe where the most often accessed information resides.

Search solutions should include a well-defined feedback loop for process improvement based on analytics generated by searchers and recorded by the search system. Search specialists should analyze these statistics to improve results. As the organization learns more about the information environment as shown by the search engine, it can continue to improve the value delivered by the search solution.

CONFIGURING FOR RELIABILITY AND PERFORMANCE

Users have done without organizationwide search in the past, so it might appear that the reliability and performance of search tools is not crucial. However, once knowledge workers become accustomed to a new and better way of finding information—once their growing expectations for ease of use are whetted—they will be impacted if the tool is not reliable and effective.

Because search is an increasingly important application, performance is a key to adoption and continued use. Search tools cannot slow with increased search volume or become unavailable in the event of system failure. Also, the process of growing and extending the solution must be straightforward.

MEETING SECURITY AND PRIVACY REQUIREMENTS

Effective enterprise search solutions should be *universal* and *scalable*. But what does this mean?

Universal does not mean that search queries display all results to all inquirers. It certainly is not desirable to return salary figures or pre-rollout product details for a non-authorized user. Instead, good universal enterprise search solutions provide the right results that are appropriate for each user.

A scalable enterprise search solution automatically indexes and returns relevant search results. To do this, a good solution must be capable of accessing and appropriately displaying data from secured repositories without manual intervention.

These requirements create two primary security and privacy considerations for search solutions for the enterprise—access to secured data repositories and presentation of results. These concerns can be addressed with a two-layer security architecture comprising a *data security layer* and a *presentation security layer* (see Figure 1).

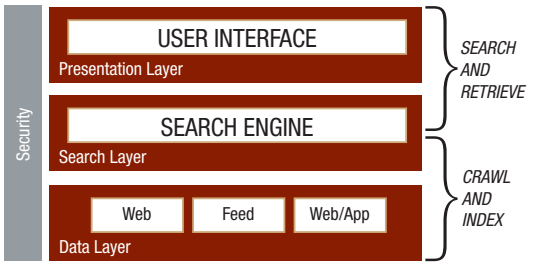
The *data security layer* governs access to data repositories. "Crawling and indexing" is the process of automatically discovering documents to build an index of search results. Search engines that use Internet technologies crawl the systems within the internal intranet just as search engines crawl the Internet. The crawl process must be automatic. That is, the solution must scale well by building search indexes without manual intervention.

Search engines based on Internet search technologies use Web access to most systems during the crawl process. Typically, a list of "seed" URLs tells the search engine where to start crawling. The use of "do not crawl" lists can control the size and extent of the resulting index.

However, many enterprise databases are not Web enabled, but instead use legacy or proprietary access methods. Sometimes, special-purpose adapters must be used to provide access to these repositories or "push" data to the search engine through data feeds created in extensible markup language (XML). These adapters can add to the complexity and expense of deployment, but two approaches can minimize the cost.

First, preprogrammed, pre-tested, ready-to-use adapters are available from certain systems integrators. Second, while large enterprises have many legacy systems, existing business or presentation layer wrappers that

Figure 1. The Layers of Organizationwide Search Security



provide Web access to legacy data often enable easier access. In this case, the search deployment team should point the search engine at the Web-enabled business layer, rather than directly at the legacy back end.

The preferred data security approach is for the enterprise security system to grant the enterprise search engine as much access as is necessary to search and index the source data. The overall security infrastructure will then protect data privacy by denying user access during return and presentation of search results.

While you may not choose to expose certain extremely sensitive data to the search engine, the security infrastructure should generally grant the search engine the broadest access to as much data as possible so that it has the opportunity to make the most complete and useful relevance calculations.

The *presentation security layer* controls access to the search results generated by the search engine. Deployment teams should plan the configuration of this layer carefully, because a search engine eliminates "security by obscurity." If a search engine is implemented without proper planning, users may find a great deal of sensitive information. Denying access to certain sensitive repositories using "do not crawl" lists reduces the amount of exposed sensitive material. However, presentation security is also needed to ensure that searchers can view only material to which they are entitled.

To aid in enforcing security and to enhance usability, the search solution should categorize results into logical sub-categories. Data available for viewing within each of these collections is appropriate to the security role of the searcher who, as a result, sees only the data indexed in that category of information.

CLEARING IMPLEMENTATION HURDLES

While implementing a basic search solution within an enterprise is straightforward, organizations need to keep these points in mind:

- Business processes may need to be modified and extended to ensure search reach spans meaningful organizational areas.
- Operational processes may need to be modified and extended to ensure that information is made available to be searched.
- Training is often needed for systems administrators to help them incorporate new content, as well as leverage new features and capabilities.
- Security assessments must be performed so that powerful search technology does not unintentionally expose sensitive data.
- Security models must be examined, understood and complied with so that the correct repositories are accessed and indexed.
- Integrating legacy systems and specialized content management systems should be included to expose all enterprise data in search results.
- Customer privacy restrictions must be considered so that sensitive data is protected.
- Disaster recovery and business continuity plans need to be extended to include search applications.

By taking these steps and addressing the enterprise search issues outlined in this paper—quality, reach, security, performance and privacy—organizations can set a course to achieving the benefits in productivity, organizational knowledge and innovation that enterprise search offers.

To learn more about how our solutions can empower your company, Let's Talk.

GLOBAL MANAGEMENT AND TECHNOLOGY CONSULTING FOR TODAY'S BUSINESS ENVIRONMENT

BearingPoint is a leading global management and technology consulting company that serves the Global 2000 and many of the world's largest public services organizations. Our experienced professionals help organizations around the world set direction to reach their goals and create enterprise value. By aligning their business processes and information systems, we help our clients gain competitive leadership advantage—delivering results in an accelerated time frame. To learn more, contact us at 1.866.661.FIND (+1.603.589.4089 from outside the United States and Canada) or visit our Web site at www.bearingpoint.com.

BearingPoint provides strategic consulting, application services, technology solutions and managed services to Global 2000 companies and government organizations.

BearingPoint
1676 International Drive
McLean, VA 22102
www.bearingpoint.com