

Secure Enterprise Search

*One Search Across Your Enterprise Repositories:
Comprehensive, Secure, And Easy To Use
January 2007*

Secure Enterprise Search

EXECUTIVE OVERVIEW

With distributed enterprise content, storage grids, and virtualization on the horizon, enterprise search takes on critical importance as the unifying engine for content and process exploration. Enterprise search is a superset of searches on specific repositories and applications – a single search tool that encompasses structured data, unstructured content, enterprise content management, messaging, media, and external repositories. Enterprise-class search must be fast, comprehensive, and secure; it needs an index for speed, the ability to search each repository of corporate content, security in crawling to create a comprehensive index and security on the index to display results to which the user is authorized.

Secure Enterprise Search has been designed as a stand-alone enterprise search solution. It incorporates best-in-class indexing, crawling, and security capabilities to create a reliable, comprehensive search solution for any organization. Its single search interface accesses any repository, with either native or federated search, to deliver results in a timely fashion to meet the increasing demand for enterprise information.

INTRODUCTION: THE CHALLENGE OF FINDING INFORMATION

Every day, knowledge workers create, collate, share, and use information. Every day, these same knowledge workers seek more information from a range of internal and external sources. The Internet, with its variety of sites and search services, has replaced many clipping and service bureaus for finding publicly available information. But internal information often lays undiscovered.

Why does internal information get lost? Here are four obvious reasons:

- Workers don't know where to look
- Workers must manually search individual repositories to which they have access
- Workers don't have permission to access the information
- Multiple versions of information exist, without easy ways to distinguish among them

With the increasing number of repositories, content can be in many places. Finding that content today requires searching each repository individually. This is both time-consuming and cumbersome, but is necessary to provide a comprehensive search for enterprise content. Enterprise search addresses this problem by providing a single, comprehensive search across all enterprise repositories.

Content Proliferation Hinders Content Discovery

Consider how many repositories are used inside a typical enterprise: databases, file servers, portals, email, content management systems and many more. Add to that a number of external repositories to which an organization subscribes. Each has its own security model. Each has its own organization for storing content. Some incorporate metadata to help track and index individual elements, while others have no metadata whatsoever.

If users cannot conduct a comprehensive search, they lose content. But, searching individual repositories is time-consuming and problematic – since critical content may exist in a repository to which a user has no access.

Now, consider the problem of a worker trying to find a specific document. Three documents with the same name but different modification dates exist on a file server, in an email repository, and in a content management system. The document in email is an attachment sent by the author. A team member saved the file into a group folder on the file server. The document in the content management system should be current, but it has an earlier modification date than the other two documents. To find the current document, the worker must search each repository, using different search criteria, find the document, and then compare the three against each other to find the proper version.

What happens when a worker lacks access to one of these repositories? For example, let's say the worker has no access to the content management system. The document might be publicly available – if it could be discovered. But, the user has no way to discover that the content management system contains a relevant document. This document is lost to this user.

Regulatory Compliance Requires Comprehensive Discovery

Regulatory compliance depends on the ability to find sensitive, inappropriate, or licensed materials stored outside compliance with organization policy. Organizations need comprehensive search capabilities to ensure all content is in compliance

While individual workers may have limited access to content – or lack the time or patience to perform a comprehensive search of all relevant content – government regulations require that organizations track and manage all their content. For instance, organizations must ensure that sensitive information is not publicly available. Next, they must ensure that no inappropriate material is stored within corporate systems. Also, they must ensure that copyrighted and licensed materials are not being used inappropriately. Executives in public companies that fail to take such steps are at risk for criminal prosecution or shareholder litigation.

These regulations require comprehensive search of all repositories to find inappropriate content or sensitive content stored incorrectly. With the number of distributed systems storing content, the effort to search each individual repository is tremendous. Organizations need a comprehensive enterprise search to address this regulatory requirement.

In addition to basic regulatory requirements, organizations may be engaged in legal action that requires comprehensive disclosure of content related to a single topic. Without a comprehensive search solution, a

company in a time jam often may produce too much unrelated content during disclosure, putting the company at risk.

Enterprise Navigation And Distributed Systems

In automating systems, organizations run into an unintended consequence of automation – users don't know the structure. System designers can put tremendous thought into the hierarchy of information displayed to a user, but the hierarchy is often an impediment to the user's desire to perform a specific action. Take, for example, the issue of filing an expense report. Is there a corporate page for completing an expense report online? If so, how does a user find that page? Then, what happens when the page moves, because designers want to implement a better system?

Users are challenged to find applications and processes as well as content. The distributed nature of these applications calls for an enterprise search solution.

REQUIREMENTS FOR ENTERPRISE SEARCH

Enterprise search delivers a search solution for the enterprise that is both comprehensive and secure, delivers highly relevant results, is easy to use and maintain, and leverages existing investment in search technologies. Here is a detailed look at some of the key capabilities required to make an enterprise search solution work.

**Enterprise search must be secure,
comprehensive across all repositories,
relevant, easy to use and maintain,
scalable and fast.**

Secure

While Internet search assumes that all content is public, enterprise search must work within a world of authentication, access control, and security policies. Enterprise search must ensure the security of each repository as it builds a search index, ensure the security of the search index itself, and reveal contents of the search index based on user authentication and repository-specific access controls for each searched repository.

Enterprise search must be able to deliver information ordered by a user's preference and limited only by a user's access restrictions. Without security integration, users are either limited to public information or are provided uncontrolled access to sensitive information.

Comprehensive

Enterprise search must be able to access any enterprise repository – located inside or outside the enterprise firewall. In some cases, the search solution has built-in access to repositories through native or custom interfaces. In others, enterprise search federates search to repository-specific search solutions and integrates the results from those tools in

delivering search results. Without comprehensive search, users are required to manually integrate results from searches of multiple repositories.

Relevant

On the Internet, users are mostly anonymous. “Relevance” in Internet search means content retrieved relevant to a specific keyword or search term. In the enterprise, users have unique identities – their job, workgroup, regular activities, and information needs can be defined uniquely within the work environment. Enterprise search must be able to take this user-level information into account when retrieving results for a given query.

Easy To Use And Maintain

Search encompasses a number of critical technologies, notably content indexing, index database management, information retrieval, security and authorization management. Each of these areas has existed previously as a separate discipline requiring different technology management capabilities. Search within the enterprise must incorporate these technologies without requiring an expert in each to manage the search solution.

Scalable

As more users access more content, enterprise search must scale without impacting performance. Scaling may require hardware upgrades, but the performance must scale with the number of simultaneous users and the total content indexed in the search system. Search is an indispensable enterprise solution and cannot crash due to the weight of user demands.

Fast

Speed of search is directly related to the underlying search technologies. Fast search requires an index of previously searched content. The alternative to prior searching is real-time searching – which is both inherently slow and becomes even more so as multiple users search content repositories in real time.

Secure Enterprise Search is a standalone search solution that provides highly relevant results, that securely searches all enterprise repositories, and is easy to use and manage.

THE SOLUTION: SECURE ENTERPRISE SEARCH

Secure Enterprise Search (SES) addresses all the critical capabilities of enterprise search today, with the goal of meeting an organization's search needs as those needs grow. SES delivers a complete search solution – addressing the critical enterprise search capabilities described in the previous section.

Secure Enterprise Search is a new product based on established Oracle technologies. Oracle Text and its predecessors have delivered state-of-the-art high-speed information retrieval for well over a decade. Oracle Text provides the “meaning engine” to Secure Enterprise Search – ranking the importance of text and metadata in numerous languages and repositories.

Oracle Ultra Search, the direct predecessor to Secure Enterprise Search, incorporates a repository crawler to build an index from a number of standard repositories – including Oracle databases, portals, IMAP mail servers, HTML web servers, and file servers.

Secure Enterprise Search extends these prior products to address the growing grid of content and applications in the enterprise. Content and applications are migrating from fixed structures to a reliable, scalable grid of storage and virtual application servers. This is Oracle's vision of the Enterprise Information Architecture, which consists of storage, middleware, applications and enterprise management – integrated in a virtualized environment to deliver processing and storage as needed within the enterprise.

Here are some of the features and capabilities of Secure Enterprise Search.

Secure Search

Secure Enterprise Search is designed to meet the security requirements of big and small enterprises. With built in security APIs, Secure Enterprise Search can meet the custom requirements of searchable repositories today, with the goal of incorporating standard security infrastructure natively in the near future. The security infrastructure mirrors the repository security, displaying only those results that match the user's access rights from each individual repository. Individual repositories can signal Secure Enterprise Search to reset access information when access rights are changed for any individual or group of users. Secure Enterprise Search supports LDAP natively, with ActiveDirectory and NTLM support in the near future. It can also work with any user authentication scheme using APIs to link to customized authentication interfaces.

Secure Enterprise Search is a standalone search solution that provides highly relevant results, that securely searches all enterprise repositories, and is easy to use and manage.

Any Repository

Secure Enterprise Search delivers search for any repository of structured or unstructured content. It provides native search to a number of critical repositories today, including SQL databases, HTML websites, and IMAP email servers. It also provides native search for specific repositories, including Oracle Portal, Oracle Files, and Documentum repositories. Outside the natively supported repositories, Secure Enterprise Search has APIs to link with custom crawlers. It also includes federated search – allowing existing repository searches to be included in search results and integrated for the user. Native search for popular integrated messaging platforms, such as Microsoft Exchange and Lotus Notes is also supported.

Results Ranking

Secure Enterprise Search includes a ranking engine that ranks results within the enterprise context. Secure Enterprise Search can crawl and index many document formats, including web documents (.html, .asp, .php), XML archives, the Microsoft Office formats (.doc, .exe, .ppt), Adobe Acrobat .pdf, and compressed archives (.zip). By incorporating context-specific content within these formats, the Oracle Text capabilities within Secure Enterprise Search rank the documents relative to specific keywords to prioritize the retrieval of information based on a search request.

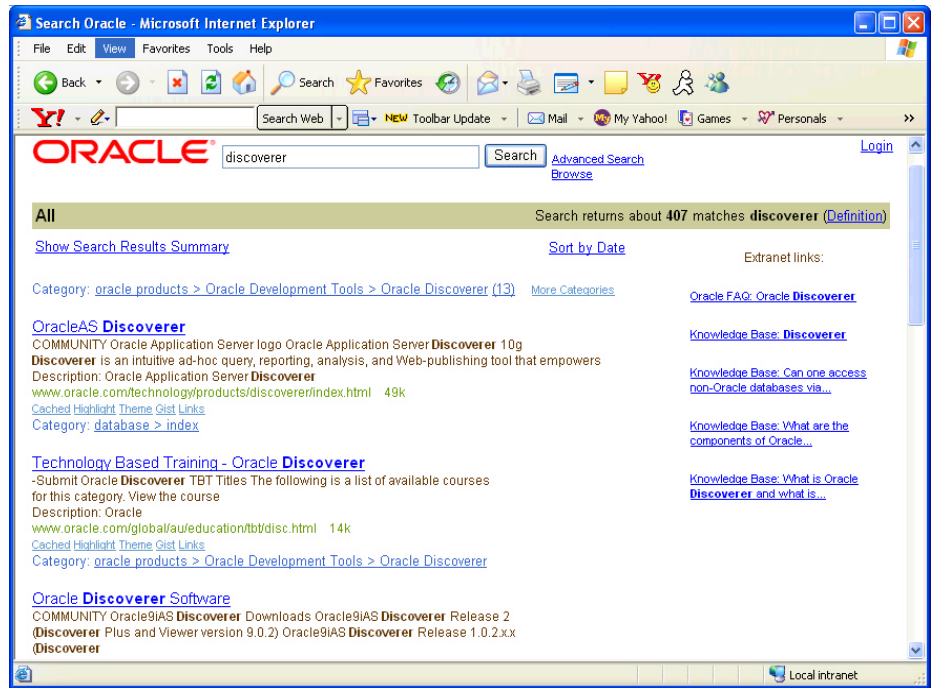
Simple

Secure Enterprise Search is designed from the ground up to be both easy to manage and easy to use.

For administrators, Secure Enterprise Search installs quickly on a standalone system. Once installed, the basic mechanisms for crawling standard repositories are easy to launch, including scheduled crawls for new content and managing repository access. The system includes logs of successful and problematic searches, and includes tools for increasing relevance, listing related links, managing user feedback, and creating standardized subgroups of repositories for searching.

For users Secure Enterprise Search matches the ease of Internet search solutions, with a simple user interface for entering requests and standard hitlist-like results.

For application developers, Secure Enterprise Search provides Web Services access that allows the user interface to be customized and the search results to be sorted and segmented by a number of criteria, including repository, context, or date.



Secure Enterprise Search – Standard User Interface

APPLICATIONS

Secure Enterprise Search delivers high value in everyday use, because it provides unified access to all corporate content. Users become more productive when they can find relevant content more quickly.

In the lost file example, a user can have the filename and use Secure Enterprise Search to find all instances of that filename across all repositories to which that user has access. By finding those files quickly, the user can rapidly determine the differences between files and resolve any discrepancies.

Productivity extends beyond resolving content versions. Users can use Secure Enterprise Search to navigate to specific business applications used within the enterprise. As different applications have been automated, they become part of a hierarchy of business activities. Users typically must navigate this hierarchy to access a specific business application. To enter a travel expense report, a user may have to log into the business system, select an activity group (such as “Employee Activities”), select “Expense Reports,” then select “Employee Travel Expense Report” – multiple clicks just to get to the travel expense report section. With Oracle Enterprise Search integrated into business applications, users type “travel expense” and, after sign-on, access the travel expense report form directly. This reduction in effort frees users from having to remember each step in a process – using Secure Enterprise Search to determine the next step. As a corollary, users can use Secure Enterprise Search in the

middle of a business process to access content or jump to critical, but disconnected business processes to speed selling or problem solving

Example: Customer Call Center

In customer support, call center agents must access a variety of online resources to provide the information needed by a customer. This includes:

1. The company's product database, which holds basic warranty information and some technical detail about all products for which service and maintenance is provided
2. The company's customer database, holding information about past sales and maintenance activities related to the caller (the caller's 'customer history').
3. Web pages, both inside and outside the company's firewall/intranet. The service department maintains the company-internal web pages on the intranet. These pages have detailed information about the most frequent product breakdowns. Outside the company's firewall are manufacturer's own web pages and documents, including manuals, latest revisions, detailed technical specifications and also reports of problems filed by other retailers.
4. Emails from service technicians on related issues, which include knowledge and experience gained from resolving related breakdowns, warranty cases, and a description of how similar issues were resolved in the past.
5. A regulations resource, comprising of PDF files circulated by government consumer protection agencies.

Within the call center, agents receive calls that require them to formulate answers within 45-60 seconds, while simultaneously listening to caller information and elaboration. Searching all the company's product-related information, from internal databases and emails to internal and external websites, requires the comprehensiveness of Secure Enterprise Search. The comprehensiveness of this search speeds getting a list of similar problems resolved, activities to perform, and processes to address the customer issue. Secure Enterprise Search also helps agents step beyond standard processes to aid in escalation by finding appropriate processes that may not be covered in existing solutions.

Example: Extranet Searches

The comprehensive solution in Secure Enterprise Search also aids in external searches of company content. The same customer service call might even be avoided if the relevant problem solving content is available in a searchable file store or database accessible to external customers.

Since Secure Enterprise Search resolves searches across multiple sources, a single search on a product type can deliver a list of known problems and their resolution, along with documentation, collateral, and links to online customer help for next steps. Secure Enterprise Search delivers not just content, but a complete process for users to address a known – or potentially unknown issue.

CONCLUSION

The need for a single, comprehensive enterprise search solution will only grow as users continue to generate and store content in various file servers, databases, workspaces, emails, and a host of other repositories. Secure Enterprise Search delivers a comprehensive, secure, and fast solution designed for the needs of organizations large and small that want to find information quickly, improve their productivity, and gain the benefit of an enterprise search solution.



Secure Enterprise Search

January 2007

Authors: Vishu Krihnamurthy, Stefan Buchta

Contributing Authors: Meeten Bhavsar

Oracle Corporation

World Headquarters

500 Oracle Parkway

Redwood Shores, CA 94065

U.S.A.

Worldwide Inquiries:

Phone: +1.650.506.7000

Fax: +1.650.506.7200

oracle.com

Copyright © 2007, Oracle. All rights reserved.

This document is provided for information purposes only and the contents hereof are subject to change without notice.

This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle, JD Edwards, and PeopleSoft are registered trademarks of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.