

Search Solutions for the Enterprise

Security considerations in deploying search

A search solution for the enterprise can help **boost organizational productivity** by making useful information readily available to people who need it.

INTRODUCTION: THE SECURITY REQUIREMENT FOR SEARCH SOLUTIONS

A well-deployed search solution for the enterprise can put a trove of information at the fingertips of knowledge workers and others within an organization—meeting the growing expectation of workers that enterprise systems provide the same level of functionality, transportability and ease of use as Internet services. The trick is to make sure knowledge workers can get to the information they need, with the proper security, to access only the information they should.

A guiding principle of organizationwide search security is compliance with underlying enterprise security policy and regulatory requirements. It is also important that search security be compatible with the security schemes of various content sources within the enterprise. In creating a security framework, BearingPoint believes significant attention should be focused on authentication, authorization, auditing, and identity and access management.

AUTHENTICATION

Authentication can take many forms or methods. Whether it is directory- or application-based, the most common standard for authentication is a username and password. The inherent insecurity in this method may or may not be a concern depending on the enterprise environment. The challenge with organizationwide search is associating the proper method or form of authentication with the requested search results. Not all information is created equal, nor does every company have the same security requirements for data access controls.

Strong authentication can play an important role as more sensitive data becomes available to users. A simple username and password may no longer be an acceptable method of authentication. At the same time, user vetting, or truly knowing who the users are, is limited by cost or because it is not part of a corporate philosophy. Public key infrastructure (PKI) certificates, biometrics and multifactor authentication are tools that may be required to overcome this.

There are three basic authentication-level concepts that can be scaled to meet a given company's needs.

- **Authentication level 1**—Accepting that the user was able to log in to his or her workstation with a correct username and password is considered enough.

IN THIS POINT OF VIEW:

INTRODUCTION: THE SECURITY REQUIREMENT FOR SEARCH SOLUTIONS	1
AUTHENTICATION	1
Authentication and the Search Engine	2
AUTHORIZATION	2
Authorization and the Search Engine	2
AUDIT	3
Audit and the Search Engine	3
IDENTITY AND ACCESS MANAGEMENT	3
Identity and Access Management and the Search Engine	4
CREATING A SECURE SEARCH ENVIRONMENT	4

- **Authentication level 2**— Requiring login, whether manually or via single sign-on (SSO), to each link that requires more sensitive data.
- **Authentication level 3**— Using multi- or two-factor authentication that utilizes an x509 multipurpose certificate or biometrics to display certificate information, convert certificates to various forms, sign certificate requests such as a mini certification authority, or edit certificate trust settings.

Authentication and the Search Engine

Out of the box, a search solution may support two methods of authentication for crawling content: basic/NTLM (NT LAN Manager) authentication and form-based authentication. Basic/NTLM authentication works in nearly all Web server implementations that support at least HTTP/1.0. It is also supported by servers based on the Microsoft® operating system (OS). Generally speaking, if a user credential set resides within a Microsoft OS and uses NTLM, the search engine can leverage the user set contained within.

Form-based authentication is typically deployed in Web-based SSO environments. A current limitation of search engines is the capacity to leverage only one form-based SSO system at a time.

After the search engine has crawled content and a user wishes to search the results, the search engine must serve the content in a secure fashion. Search engines leverage basic/NTLM and form-based authentication by using HEAD requests to a Web server for Web-based content.

The search engine typically can be configured to host both public and secure content. When the user attempts to access content that has been defined as secure, a dialog box appears within the browser session to challenge the user for the required credential set. This happens once a session.

In the case of form-based authentication, the search engine can either use cookie forwarding or full user impersonation. In both cases, the engine captures the login information in a cookie and forwards it to the systems being crawled.

For more complex implementations using outside content, custom adapters and application programming interfaces (API) will be required. Vendors offer authorization service provider interfaces (SPIs), allowing Web services to translate between the search engine authorization SPI and the server that provides access control services.

AUTHORIZATION

The challenge with authorization is balancing the rights of users so they can accomplish their work. Organizationwide search is no exception to this.

Mapping the right data or search results within the realm of the user's rights, and then presenting only that data, remains a challenge. Federated directories and SSO and PKI certificates are examples of authorization services that can be used to help identify users and validate their identities.

Below are examples of authorization-level concepts. They do not represent a complete list of the options:

- **Authorization level 1** — Internal public information that is accessible to all and requires nothing other than network access to view.
- **Authorization level 2** — Confidential information that requires a secondary login to view.
- **Authorization level 3** — Sensitive data, such as corporate intellectual property or payroll information, that only specific groups have access to.

Authorization and the Search Engine

Authorization SPIs allow search engines to leverage user credentials stored outside of typical NTLM authentication schemes or one-source-only form-based authentication. Implementation of an authorization SPI relies on the security assertion markup language (SAML) 2.0 standard as its basis and is coded with that standard.

When a user performs a search and the search engine must determine if it can serve the result, the search engine will contact the target host, or “access connector,” with the URL or target in question and the user's identity.

Each time, abiding by SAML 2.0 standards, the target host will either permit, deny or reply as indeterminate. Simple object access protocol (SOAP) over hypertext transfer protocol secure (HTTPS) enables this capability. However, delays may result from this because the search engine caches these results during the session. Cache-time is configurable.

AUDIT

Effectively implementing any security solution requires the ability to audit.

With enterprisewide search, audit focus shifts. While most organizations focus on external threats, greater threats may come from within. A search solution must do secure searches while limiting user searches to collections if necessary. Data access auditing, while still important, becomes secondary to user rights.

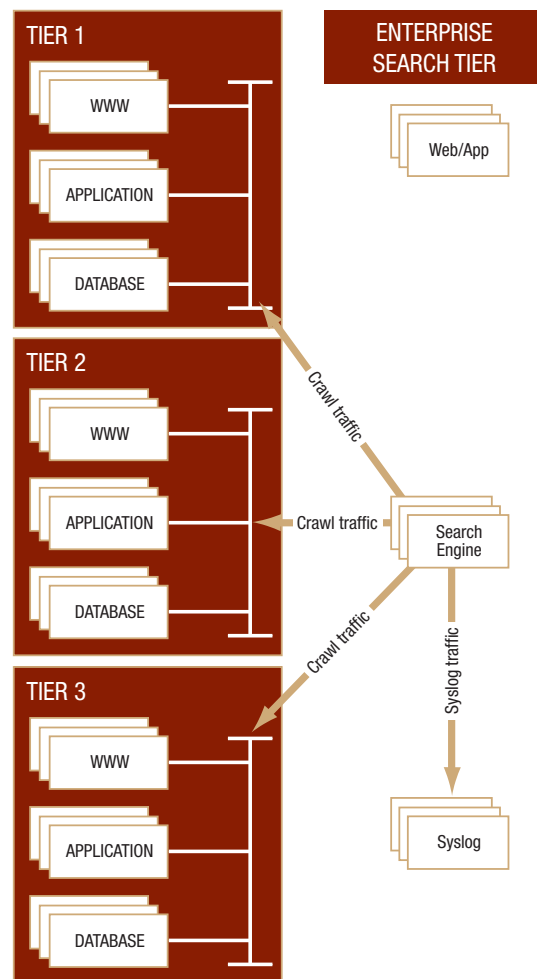
Focusing on user rights provides greater assurance that people who need information are conducting secure searches. Key areas to consider include unauthorized rights elevation; user moves, adds and changes; users accessing sensitive data with false credentials; and users holding on to rights from previous roles and gaining access to sensitive data.

Audit and the Search Engine

To implement audit functions organizationwide, proper auditing services must be deployed on the existing infrastructure. Systems to be accessed in the search deployment must be reviewed for proper audit implementation.

Available search engines have functionality built in to provide audit-logging capabilities via an external syslog server, which provides output into a text file that must be sent to a separate syslog server. Deploying a dedicated syslog server to enable proper auditing functionality within the search engine is recommended. Additionally, if audit logging servers and messages are not being captured on the accessible systems, this functionality needs to be implemented. Figure 1 depicts a typical deployment in which the search engine will use a separate syslog server to gather log files created by end-user search activity.

Figure 1. Search Engine Audit Functionality



IDENTITY AND ACCESS MANAGEMENT

A search solution can actually create security issues if proper controls have not been put in place prior to implementation. But there can be a positive outcome of this. As data is crawled and presented, long-standing security holes or unprotected data stores can be exposed, and then remedied.

Authentication, authorization and auditing are key pieces of identity and access management. Other important aspects of a well-rounded approach include:

- **Password configuration and aging policies.** With username and password being used for access to secure data, strong password policies are critical. Passwords should be alphanumeric and should be aged based on data being accessed—the more sensitive the data, the more frequent the password changes.
- **Single sign-on.** SSO can be used to combat the cost of resetting user passwords for infrequently used but secure applications because they can auto-generate expiring passwords. This feature can eliminate soft or simple passwords and prevent passwords from being passed around. It can eliminate the need for the user to log in multiple times when accessing secure data, providing a more secure environment and encouraging the user community to use search.
- **Separation of duties (SoD).** Separating roles without creating additional overhead through added staff can be a delicate balance. SoD is designed to prevent users from performing potentially hazardous actions. Except in small companies, one person usually does not have access to both accounts payable and accounts receivable. In an organizationwide search environment, the person granting user rights should not be the same one who decides what collections users have access to.
- **Role-based access control.** Roles engineering is no small undertaking, but it can create a more secure environment. Rights are granted in three ways—explicit, implicit and inherited. Rules can be set up to prevent applying conflicting roles to the same user, thereby enforcing the SoD policies that have been created. The ability for users to filter a search with a specific role, even if they have multiple roles, can provide cleaner but still secure results. Role-based access is directly linked to SoD. By having users in defined roles, with rules preventing conflict of duties, the search results presented will be directly related to what they need and should have access to.
- **User provisioning/de-provisioning.** A safer environment can be created through centralized and delegated user management, workflows, password management and role-based access control models. A twofold goal is to make sure new users have immediate access to the information they need and access is taken from

users who no longer have authority as quickly as possible. Though not directly related to search, proper provisioning has a direct impact on roles and SoD. This is the starting point for many security initiatives.

Identity and Access Management and the Search Engine

Search engines are available with leading identity management solutions using form-based authentication and leveraging an authorization SPI.

Still to be determined is how many search engines will integrate with non-Web-based SSO and SSO-like legacy systems. Many organizations have multiple security systems deployed in different forms—Web-based SSO, internal SSO, lightweight directory access protocol (LDAP) and other username/password repositories.

Because search solution requirements are unique to each organization, proper scoping of deployment is imperative. A mix of technologies may be used, including an authorization SPI, custom APIs and adapters, and the search engine form-based authentication mechanism.

CREATING A SECURE SEARCH ENVIRONMENT

Deploying search solutions for the enterprise raises new security considerations. By addressing these requirements at the outset with a comprehensive security approach, organizations can capitalize on the benefits of search solutions, while protecting sensitive information.

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

