



TechAspect



White Paper

# Searching Protected Content using Google Search Appliance

**Kumar K.N.**  
Sr. Technical Architect

## Problem

Many companies use their website as a main driver for obtaining sales leads. One of the favorite ways to do this is to show a registration form when the user tries to download a document, a video or a free tool from their website. One of the main ways to drive users to these pages is via search. We need to feed all the document URLs to the search engine to do the indexing, but how do we tell the search engine to show a registration page instead of the document when the user tries to view the content?

The problem gets trickier if we want only a subset of users to register but show a direct link to the document for other users.

For this article, let us assume two sets of users.

### Scenario 1:

First group of users (public group) would see the search results containing links to a registration page instead of links to the actual documents.

### Scenario 2:

Second group is a privileged group (partners) that would see the results directly linked to the actual documents.

Although the set of documents (collection) is the same and the search term is the same, the search results need to behave differently.

## Background

Google Search Appliance (GSA) provides a web-based facility to search documents located in an organization's repository like a database. It does this by getting the list of documents from the database and indexing the contents of those documents. When a search is done on a collection consisting of those documents, GSA searches for the search phrase in its index for that collection and returns the set of results matching/containing the specified phrase. The results are typically in the form of links to documents containing the matched phrase.

## Solution

There are several ways to solve the problem. One is by writing your own custom pages in your web application that does all the talking with the GSA and displays the formatted search results perhaps by requesting a XML result from the GSA for the specified search term. The second and entirely non-programmatic way is by doing some configurations on the GSA. We will describe the second (and quicker) approach which works like a charm.

GSA provides a way to associate metadata with the data indexed from a document. This metadata can be picked up from different sources like the document itself, a database, etc. and returned along with the search results. This feature was used in the following manner to provide the solution. The different links that were to be included in the search results for the two contexts were populated in a database. The link to the set of documents that were indexed (primary documents) was provided in the same database. This database was used to generate a feed that was supplied to the GSA. With this, the result links were associated with the indexed data.

When the search was issued, the search was setup to include the metadata along with the data returned for the search. A front-end style sheet on GSA would parse the metadata and return the appropriate URL to the user, based on the user group that issued the search.

As a database was already being used to store information about the documents for other purposes, this provided an elegant solution to the problem.



## About TechAspect

TechAspect is a leading IT company with main focus on implementing and integrating enterprise content management systems, websites and web applications. We build client-facing websites that are fast, easy and compelling. Our content management solutions enable business users to have full control on development, publishing and management of web assets. Some of the world's renowned organizations have chosen TechAspect, including: Anritsu, Affymetrix, Avon, BCBSMA, Genentech, Hertz, Oracle, Intermec and Phillips-Van Heusen.

To learn more about TechAspect, please visit [www.techaspect.com](http://www.techaspect.com) or send us an email at [sales@techaspect.com](mailto:sales@techaspect.com)

### Americas

#### Headquarters

**TechAspect Solutions Inc.,**  
46720 Fremont Blvd,  
Fremont, CA 94538, USA  
Tel +1 510 962 3200 | Fax +1 510 405 2007

#### Midwest Region

**TechAspect Solutions Inc.,**  
2200 N. Canton Center Road, Suite 120,  
Canton, MI 48187-5065, USA  
Tel +1 734 238 0005 | +1 734 238 0011

### Europe

**Techaspect Solutions UK Ltd.**  
No.17, Crawley Business Centre  
Stephenson Way, Three Bridges  
RH10 1TN, UK  
Tel +44 20 8432 2955 | +44 20 8432 9670

### Asia

**TechAspect Solutions Pvt. Ltd.**  
The V, Mariner - B3, Madhapur,  
Hyderabad, AP 500 081, India  
Tel +91 40 6668 2828  
Fax +91 40 6663 8383

[www.techaspect.co.uk](http://www.techaspect.co.uk)