Google Search Appliance

User Experience Guide

August 2014

User Experience Guide

This paper is a guide to creating the Google Search Appliance user experience. The success of your deployment depends not only on the breadth and depth of search, but also on how satisfying and effective the search experience is for users. The aim of this guide is to share the many things you can do to drive user satisfaction and increase use of the search solution.

About this document

The recommendations and information in this document were gathered through our work with a variety of clients and environments in the field. We thank our customers and partners for sharing their experiences and insights.

What's covered	This paper discusses tools for enhancing the search experience, such as selecting the appropriate presentation method and enabling various different GSA features to enhance and improve the user experience.
Primary audience	First-time and experienced GSA administrators, as well as developers and web designers tasked with creating the search user interface and experience.
IT environment	Google Search Appliance.
Deployment phases	Initial and ongoing configuration of the GSA user experience.
Other resources	● Creating the Search Experience contains information about using Google Search Appliance features to create the search experience for end users. ● Learngsa.com provides educational resources for the GSA. ● GSA product documentation provides complete information about the GSA. ● Google for Work Support Portal provides access to Google support. ● GSA help center provides a comprehensive list of resources to get help about the GSA.

Contents

Chapter 1 Presentation Methods

Overview

Sending a search query to the Google Search Appliance (GSA) returns results in raw XML. To make this XML presentable to the end user, the GSA parses the XML, along with an XSLT stylesheet, and this generates an HTML web page.

One of the first decisions to be made when deciding upon the user experience is which presentation method to use. By presentation method, we mean the platform on which the search interface is presented to the users.

There are two primary methods of delivering the search experience to your users:

- Google Search Appliance presentation layer
    - Presented directly from the GSA
    - Presented by way of an inline frame

- Application presentation layer

This section will help you choose an appropriate method for your users based on technical requirements and the outcomes you are trying to achieve.

Google Search Appliance presentation layer

The Google Search Appliance uses an XSLT stylesheet for its presentation layer.

Key considerations
Using this built-in presentation layer has several advantages and some limitations and disadvantages.

Advantages
Using the GSA presentation layer has the following advantages:

- All presentation is rendered on-box and delivered directly to the user. The search appliance does not require any additional hardware to manage presentation.

- Built-in user features, such as Query Suggestions, Document Preview, Translation, User Results, Expert Search, Dynamic Result Clusters, and so on, can be enabled and delivered to users as simply as selecting a checkbox.

- Relatively sophisticated user experiences can be delivered by means of declarative XSLT transformations through direct customization of the stylesheet.

- The search experience can be delivered very quickly, with minimal XSLT/XML/HTML knowledge.

- Using inline frames, you can present the search interface within the context of an otherwise static web page.

- Simplifies secure search authentication setup due to a reduction in the number of security hops.

Limitations and Disadvantages

However, there are some limitations to using the GSA presentation layer: most notably that highly sophisticated, interactive, or JavaScript-rich user interfaces are more challenging to deliver. This is primarily due to the declarative nature of XSLT and security restrictions that prevent uploading of content to the search appliance.

Using the GSA presentation layer has the following disadvantages:

- Not advisable for a highly sophisticated user experience.

- All JavaScript must be embedded directly into the output HTML pages, which may lead to browser inefficiencies.

- Additional resources, such as style sheets, JavaScript files, images, and so on, can be hosted on a separate server and delivered to client browsers as included resources, improving perceived performance to users.

- Additional code must be maintained if you are matching the look and feel of the XSLT to an existing website.

- Search URL is exposed to the end user and can be manipulated.

For information about using the Google Search Appliance presentation layer, see Creating the Search Experience.

Application presentation layer

The Google Search Appliance can return search results, including metadata, directly in XML. The XML can be processed and rendered on a separate application server or portal.

Key considerations

Using an application presentation layer has several advantages and disadvantages.

Advantages

Using the application presentation layer has the following advantages:

- Presentation can take full advantage of the flexibility and richness of modern programming languages, such as Java, Python, .NET, or even Flash, to provide an extremely rich and interactive user interface (UI).

- Removing the rendering of content from the search appliance also removes the processing required by the search appliance.

- Additional resources, such as style sheets, JavaScript files, images, and so on, can be hosted on a separate server and delivered to client browsers as included resources, improving perceived performance to users.

- Security can be managed at the application level by allowing the application to determine the collections and front ends a user is able to see.

- Look and feel can be maintained within host site templates.

Disadvantages
However, using the application presentation layer has the following disadvantages:

- The deployment architecture is more complicated because of the additional hardware required.

- It is more challenging to manage document-level security because credentials need to be captured at the application layer and forwarded to the search appliance.

- Incorporation of some core features may require more effort to bring into the UI. For example, enabling dynamic result clusters with the search appliance presentation layer is as simple as clicking a checkbox in the Admin Console. To enable this feature with an application presentation layer requires manual implementation in your application. The following features are not currently optimised to be deployed off box (via a separate application server):
    - Document Preview
    - User Results
    - Expert Search
    - Dynamic Result Clusters

For information about search results in XML, see XML Output in the Search Protocol Reference.

Chapter 2 Understanding and Using Front Ends

Overview

The Google Search Appliance feature that enables you to create different search experiences for users is the front end. A front end is a framework that manages most of the elements of a single search experience.

Front ends can be broadly classed into two categories of functionality:

- [Presentation/look and feel](#)
- [Content filtering and enrichment](#)

Presentation/look and feel

Presentation/look and feel contains many of the features described in [Chapter 3, Using Enrichment Features](#). If you are using a custom presentation layer, you may choose to skip presentation/look and feel completely.

Changing the look and feel of the UI can be as simple as updating global settings, such as font type or size, and activating the presentation features by using the search appliance's Page Layout Helper. Or it can be as sophisticated as extensive customization of the stylesheets.

The `proxystylesheet` search parameter controls how search results are rendered. The following table provides details about `proxystylesheet`.

Parameter	Possible Value	Description
`proxystylesheet`	omitted	Results are in XML format.
	front end name	Results are in custom HTML format. The XSL stylesheet associated with the specified front end is used to transform the output.

Example

```
http://GSA_HOSTNAME/search?site=default_collection&client=default_frontend&output=xml_no_dtd&proxystylesheet=default_frontend&proxycustom=%3CHOME/%3E
```

Content filtering and enrichment

This category also contains many of the features described in [Chapter 3, Using Enrichment Features](#). These features enhance the users' search effectiveness significantly. Because many of these make changes to the results XML, these features should be used regardless of whether presentation is delivered on-box or by means of a separate presentation layer, as they can generally be leveraged at the application tier.

The `client` query parameter controls which front end is used in serving search results. The following table provides details about `client`.

Parameter	Possible Value	Description
`client`	omitted	Required parameter. If this parameter does not have a valid value, other parameters in the query string do not work as expected.
	front end name	A string that indicates a valid front end and the policies defined for it, including: ● KeyMatches ● Related Queries ● Filters—Enabling Query Expansion & Results Biasing Policies. Filtering content by metadata, language, document type or domain. ● Remove URLs ● OneBox Modules ● Dynamic Navigation Notice that the rendering of the front end is determined by the `proxystylesheet` parameter.

Example

```
http://GSA_HOSTNAME/search?site=default_collection&client=default_frontend&ou
tput=xml_no_dtd&proxystylesheet=default_frontend&proxycustom=%3CHOME/%3E
```

Review: Which presentation method should you use?

Use the following table to choose the method that best suits your needs.

Method	When to use it
XSLT stylesheet	Use this option when you want to serve formatted results directly from the search appliance. You can use the Page Layout Helper to make quick and easy changes such as adding your own logo, changing the header and footer, as well as basic results options. You can also edit the XSLT stylesheet directly; this enables you to customize every aspect of the results pages. This is also useful if you want to return your own XML schema, RSS, or JSON.
XML output	If you exclude the `proxystylesheet` parameter in the request URL, the search appliance returns raw XML. This is useful if you have a multi-tier architecture, and want to handle the presentation yourself.

proxystylesheet versus client parameters

The following diagram shows the factors that each parameter controls.
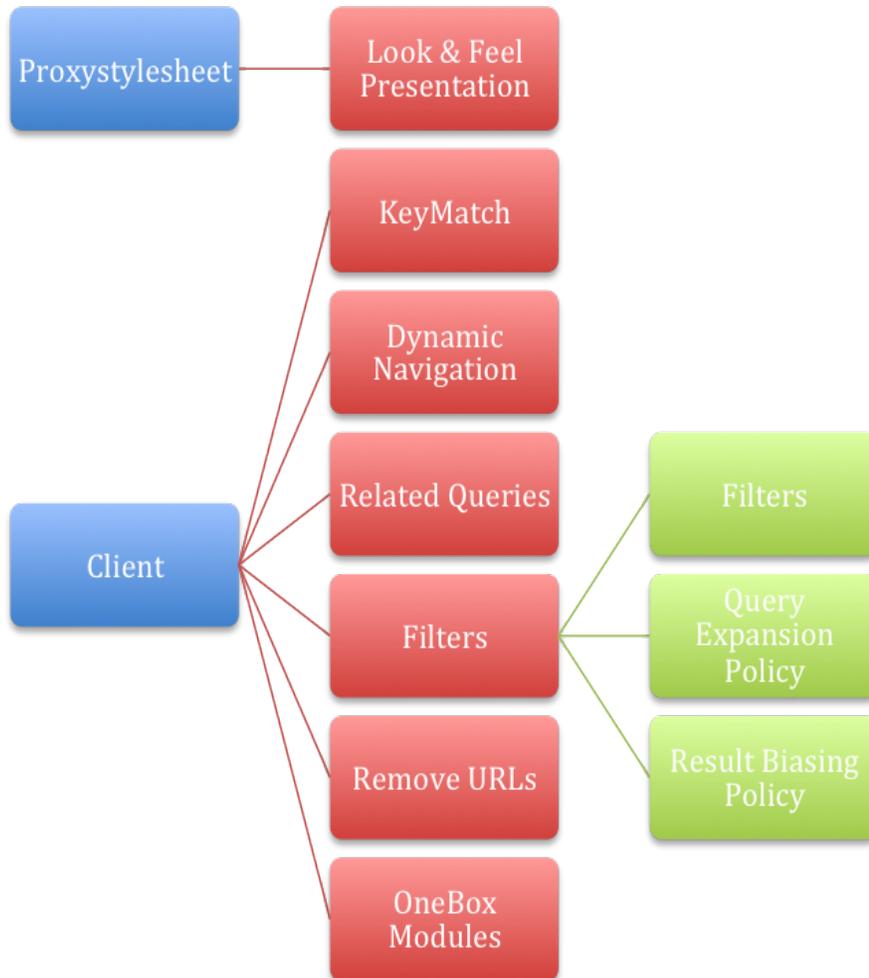
For more information about front ends, see Creating the Search Experience.

Chapter 3 Using Enrichment Features

Overview

Several Google Search Appliance enrichment features enable you to customize search results and enhance the user experience. By using these features, you can ensure that users get search results that are appropriate to their interests, roles, departments, locations, languages, or other characteristics.

It is advisable to conduct a workshop to test and investigate these features and establish their suitability for the end users. Not all features need to be rolled out at the same time, but can be rolled out in an iterative fashion to allow users to get familiar with each one.

Feature list

The following table lists several Google Search Appliance enrichment features.

Feature	Description
Query suggestions	When query suggestions are enabled, search queries autocomplete and query suggestions, with the most commonly searched terms, appear as a user types in the search box. From 6.14 onwards, there is an Admin Console page for this. In 7.2, this is at **Search > Search Features > Suggestions**. This page enables you to see which query suggestions the search appliance is using by exporting them. You can also manage the suggestions blacklist from here.
KeyMatches	Used like advertisements to display specific URLs and text for specific phrases and queries. URLs returned as KeyMatches are not actually part of the index, and are not controlled by the crawl pattern rules. ● External: Use KeyMatches to highlight new products or specific areas of your site. ● Internal: "Advertise" new employee benefits, highlight top-level landing pages.
Dynamic result clusters	Dynamic result clusters show different topics for a specific search term. These topics enable users to focus on areas of interest while ignoring irrelevant information. When a user clicks on any of the topics, the search appliance returns a new, narrower set of results.
Result biasing (by source, collection, date, metatags and entities)	Result biasing enables you to influence the way that the search appliance ranks a result, based on the URL, document date, metadata or entities in or associated with the result. You can use result biasing to increase or decrease the scores of specified sources, or types of sources, in the search index and target the biasing at specific groups of users since biasing policies can be applied on a per front end basis.
User results	User results give users the capability to add search results for certain keywords and give admins the ability to moderate these added results. User

	results cause designated documents to always appear on the results pages for specific keyword searches.

From version 7.0, User Results replaces UAR (User-added results). |
| Alerts | Alerts allow users to monitor topics of interest by receiving search results for these topics in email messages. These can be setup as hourly, daily, or weekly. Please note alerts will only work for public data; secure content is not returned in alerts. |
| OneBox modules | "Federate" a query out to another service, bringing the results back and including them with the normal search results.

Google.com uses OneBoxes to present real-time data from a variety of sources (weather, stock quotes, flight status, and so on) and you can use the same technology on the search appliance.

A common OneBox is employee directory search. |
| Related queries | Similar to spelling suggestions. Use related queries to guide users to a different query. If a user searches for "cookies," a related query might prompt: "You could also try *snacks*." |
| Query expansion (Synonyms Data, Blacklists, Stopwords, Query Parameters, Snippets Generation, Link Results and Spelling) | Query expansion automatically expands search terms to include other terms. Query expansion modifies the actual query, and therefore the returned result set. For example, a search for "light" becomes a search for "light or lighter or lightest or lighting or lights." In most situations, you should enable query expansion. Although query expansion can have a positive impact on search results relevancy and quality, it is disabled by default. Google usually recommends selecting "Full" query expansion, which applies the active Google-supplied dictionaries along with your own. Creating your own query expansion dictionaries is a great way to provide synonyms for acronyms, jargon, and company-specific terms.

You can set up a query expansion policy for meta tags. Use a query expansion policy for meta tags to select the parts of the name/value pair in a meta tag that the search appliance expands with synonyms.

Blacklists: You can control query expansion by creating a blacklist. A blacklist is a set of words that are excluded from query expansion. A blacklist can be useful for eliminating unwanted search results that result from synonym matching and clarifying special words used in your environment.

Stopwords: Query expansion also has stopwords data. A stopword is a search term that is ignored by the search appliance. Examples of stopwords include "to," "a," and "the." Custom stopwords data can be uploaded.

Query Parameters: Allows you to edit the maximum number of query terms. The number must be an integer between 0 and 150, inclusive. |

	Snippet generation: A snippet is a small section of text in a search result. In some cases, you might see search results where snippets have not been generated for some documents. Use snippet generation to restore missing snippets for these documents. You can also edit the length of the search result snippets here.
	Link Results: Allows administrators to configure the number of returned results when using a "link" query. The query prefix link: lists web pages that have links to the specified web page.
	Spelling: By default the search appliance will not provide spelling suggestions for words that are either marked, internally, as bad words or that are diacritical equivalent. You can use the spelling settings under query expansion to allow diacritical equivalence in spelling suggestions and to allow bad words to have spelling suggestions.
Dynamic navigation	This extremely powerful feature helps users explore search results by using specific metadata attributes. When a user clicks on an attribute value, the search results are filtered to contain results from the original search query which also have that specific attribute value. You can also configure ranges to allow users to search over date or price ranges. The attribute name-bars are clickable. When clicked, this opens up an autocomplete search box through which the user can search over all the attribute values. This search is a 'contains' search. Dynamic navigation also supports adding attributes based on entities discovered by Entity Recognition.
Document Previews	This productivity improving feature enables users to view a preview image of a document's pages without leaving the search results. Users on any device (mobile, tablet, desktop) can open and review documents right in the browser without having to have the native application installed. The search appliance supports document previews for Microsoft Word (doc, docx), Microsoft PowerPoint (ppt, pptx), and Adobe Portable Document Format (pdf) documents.
Translation	Harnessing the power of Google Translate, this feature gives users the ability to translate titles and snippets in search results, as well as cached documents, into the user's language in real time. The user's language is determined by the default language set in the user's browser. Please note that internet connectivity to Google Translate is required as the translation does not occur on the GSA.
Expert Search (sidebar element)	Expert search is a feature which enables your users to find experts in your organization by searching on keywords. When the user searches on a term, a list of experts appears in a sidebar next to the search results. The list might include photos, names, and phone numbers. There might also be a more detailed list of experts on a separate page that is linked to the search results page. Experts in your organization are identified by metadata that resides in a

	collection that contains only profile information. Data sources for this collection can include Microsoft SharePoint, LDAP, or any other profile content with metadata that can be crawled or fed into the search index. From version 7.0, Expert Search replaces People Search.
Google Apps results (sidebar element)	Google Apps provide your organization with tools for collaborating on documents, spreadsheets, presentations, sites, and more. You can configure the search appliance to federate search results from Google Apps. The integrating personal content feature shows private and public content that a user has permission to view directly from your Google Apps domain.
Google Site Search results (sidebar element)	Google Site Search is a hosted search solution for publicly available content. You can configure search across your external website, your partner sites, or other sites that are relevant to your organization. You can enable the search appliance to serve results from Google Site Search in the sidebar element on the search results page by specifying your custom search engine identifier.

Key considerations

When planning your user experience, it is important to note which features will work in combination together and which features will not. In addition, certain features are only available while searching public content and certain features require your users to provide credentials. The following table lists the caveats of some of the enrichment features described in this chapter.

Feature	Caveat
Dynamic navigation	- Works for secure searches as of release 6.12—does not work for secure searches in 6.10 and previous releases. - Does not work when dynamic result clusters is enabled in GSA default front ends. Customization is required to work together. Advise against using these features together particularly for secure search due to increased load on GSA. - There are certain limitations when GSA unification is enabled. The customer needs to have the same configuration on all boxes (same attribute and front ends). If there are multiple front ends, you need to create the same front end on the secondary boxes (that is, no remote front end resolution). The configuration needs to be created on all boxes by the administrator (no auto-propagation). - Does not work when sidebar elements are enabled. - Currency ranges only support the exact format $NN, for example $23. Values such as 23$ or $ 23 are not supported for range queries. - Dynamic navigation does not support aggregating counts for similar/normalized metadata attributes. For example "pub" and "publication" might represent the same metadata attribute, but their counts are not aggregated. However, because query expansion for meta data is supported, you might add a synonym file with the entry {pub, publication} to get this behavior. - Can impact performance for secure search if there is a high number of concurrent users. - Recommended only for secure search when using ACLs for Authorization.

Dynamic results clusters	● Does not work when dynamic navigation is enabled in GSA default front ends. Customization is required to work together. Advise against using these features together particularly for secure search due to increased load on GSA. ● Must be displayed on top to work with sidebar elements via Page Layout Helper. ● Can impact performance particularly for secure search.
Alerts	● Does not work with secure content. ● Cannot edit the alerts management page. ● Requires LDAP integration.
User results	● Works for web URLs only, that is, not smb:// paths.
Sidebar elements ● People Search results ● Google Apps results ● Google Site Search results	● Does not work when dynamic navigation is enabled in GSA default front ends. Customization is required to work together. ● Dynamic results clusters must be displayed on top to work with sidebar elements.
Query suggest	● Use in public sites with caution, as suggestions are based on previous user searches and their frequency. ● From 6.12 and earlier, the blacklist is available by way of the Admin API, not in admin console. From 6.14 onwards the blacklist can be exported, edited, and imported to the search appliance from the admin console (in 7.2, **Search > Search Features > Suggestions**).

For comprehensive information about all Google Search Appliance enhancement features, see Creating the Search Experience.

For information on the dynamic result clustering service /cluster protocol, see Search Protocol Reference.

For information on the query suggestion service /suggest protocol, see Search Protocol Reference.

Chapter 4 Using Collections and Front Ends to Manage the User Experience

Overview

Collections are used to create logical groupings of content within the index. A document can appear in as many collections as needed, and collections can include all documents, or be as narrow as a single document. You can enable users to search all collections or restrict them to a specific one.

Deciding on your collection content

Understanding how your users are searching and what they are looking for can help you to use collections effectively. For example, users in Tech Support may need to see all your product documentation. For them it might be useful to separate results into the following collections:

- Product documentation
- Support requests
- Bugs and issues

These collections could be presented as separate tabs on the same page.

Bringing collections and front ends together

Different user groups may have dramatically different search needs with different presentations, results filtering rules, and content.

For example, a marketing or public relations department might want a visually rich, interactive UI that enables them to search for previous communications, video, audio, and images. On the other hand, IT support might want a fast, light, UI that enables them to search for technical content quickly.

To meet the different user interface needs of each department, a search appliance could have two different front ends. To meet the different content needs of each department, a search appliance could have multiple collections. Collections could be used to segment the index in ways that serve the different departments.

If both departments need to search the same content, then filtering, enrichment, and biasing profiles can be used to provide a different set of results for each. While public-facing marketing content is of primary interest to the marketing department, this content may be of secondary interest to Tech Support, which should be able to find it, but as a secondary priority to current support tickets.

Using front ends and collections together effectively can substantially improve the search experience for all users through a powerful and flexible range of deployment options.

For more information, see [Using Collections with Front Ends](#).

Chapter 5 Public versus Secure Search User Experience

Overview

The search experience can vary greatly depending on how the search results are being served. Understanding differences between viewing public search results versus viewing secure search results is essential for planning your search experience. There are certain limitations to the features that can be used when performing secure search, and the user experience is different in several key ways which are described in this chapter.

In addition to public content that is available to everyone, the search appliance can crawl and index documents that require a login and password or another form of authentication. To protect confidentiality at serving time, the search appliance determines whether the user performing the search is authorized to view each document before it displays results.

The `access` query parameter specifies whether to search public content, secure content, or both.

Public search

What is a public search? A public search, in the context of the GSA, is a search performed by a user without the need for authentication or authorization over content indexed by the GSA and marked as public in the GSA's index. This can occur on a publicly accessible website or behind the firewall over internal content that has been marked as public on the GSA.

Parameter	Value	Description
`access`	`p`	Search only public content

The following example shows a search request with `access=p`.

```
http://GSA_HOSTNAME/search?q=google&site=default_collection&btnG=Google+Searc
h&access=p&client=default_frontend&output=xml_no_dtd&proxystylesheet=default_
frontend&sort=date%3AD%3AL%3Ad1&entsp=a__gsa_doco_policy&oe=UTF-8&ie=UTF-
8&ud=1&exclude_apps=1
```

Secure search

What is a secure search? A secure search is a search in which a user is authenticated and the GSA only returns results that the user is authorized to see. This process can be silent or the user may be prompted for credentials. The search appliance also needs credentials to crawl the secure content and marks the data as secure in the index.

Parameter	Value	Description
`access`	`s`	Search only secure content
	`a`	Search all content, both public and secure

The following example shows a search request with `access=s`.

```
http://GSA_HOSTNAME/search?q=google&site=default_collection&btnG=Google+Search&access=s&client=default_frontend&output=xml_no_dtd&proxystylesheet=default_frontend&sort=date%3AD%3AL%3Ad1&entsp=a__gsa_doco_policy&oe=UTF-8&ie=UTF-8&ud=1&exclude_apps=1
```

Differences between public and secure search experience

The table below outlines specific features and data present on the search results page and how it works with public search and contrastingly how it behaves with secure search results.

Feature	Public Experience	Secure Experience
Result count Example: results 1–10 of about 4420 for query	The estimated total result count for the search request.	By default, the search appliance doesn't provide corpus estimates to secure search queries since this number doesn't take into consideration if the user can access the documents or not. However, since version 7.0, it's possible to configure the search appliance to return corpus estimates for secure queries either globally (meaning, for all queries), or on a per-query basis. Use with caution as it can indicate to users that documents exist which contain their query but to which they do not have access.
Pagination	Normal pagination based on estimated number of results. User can be offered result page 1,2,3,4, and so on.	Pagination is in the form of "Previous" and "Next."

Dynamic navigation	We verify 30K documents for public searches (for the purpose of creating the facets)	Works for secure searches as of release 6.12; does not work for secure searches in 6.10 and previous releases.
		Recommended only for secure search when using ACLs for Authorization.
		We verify 10K documents for secure searches (for the purpose of creating the facets). If there is a high number of concurrent requests this might have a negative impact on system performance.
Dynamic results clusters	Behaves as described.	Dynamic result clusters work with secure queries. Before displaying dynamic result clusters for secure queries, the search appliance ensures that the user has permission to view secure results by triggering authorization checks. These authorization checks might have a negative impact on system performance.

Chapter 6 Using Reports to Enhance the User Experience

Overview

You can significantly improve the ROI (Return On Investment) of your search deployment by spending a little time examining what your users are doing with the search appliance and what kind of search experience they are having.

By examining this information, you can:

- Understand the business value and criticality of your search application. It is much easier to assign a business value and priority to search if you know how it is benefitting users.

- Understand what your users are searching for and whether they are finding it effectively. Insights that you gain will help you understand which features to use, and how to use them. Giving your users a great search experience increases user satisfaction and therefore the overall success of the solution.

- Understand what content is important to your users.

Types of search reports

The following table lists the different types of search reports you can generate on the GSA.

Report type	Description
Search Report	The search report contains the following information: ● The number of searches per day ● The average number of searches per hour ● The top 100 (or whatever specified number) keywords and queries and the number of occurrences of each If you have enabled ASR, then it will also contain the rank of selected results, page of selected results, IP address of most frequent users and most popular URLs based on user clicks.
Advanced Search Reports	Each entry in an advanced search report represents a single user click in the search appliance user interface, whether on a search or results page. An entry is composed of values for the following items: ● Time of the click to the nearest 100th of a second ● The user name of the search user. This can be included only if search is secure and the user has authenticated. ● IP address of the search user responsible for the click that was recorded ● Holding place for the session ID, which is always blank ● Click type ● Click start, which is the results page where the user clicked ● Click rank, which is the rank of the result the user clicked ● Click data, which is usually blank ● Query, which is the user query that returned results ● URL of the user click
Search Logs	The logs provide the following information about user search requests: ● Type of user queries ● Parts of the user interface users click ● Speed with which results are served ● Whether users are receiving the result they want ● Whether users would find the Related Queries, KeyMatch, Query Expansion, and OneBox features useful

For more information, see Enabling the Search Appliance to Record User Identities for Queries.

Search report retention time

The search appliance retains information for varying periods in:

- Search reports
- Search logs

Search reports

Search reports remain available for one year from the creation date. You can create up to 500 reports across all collections.

Search logs

You cannot generate search log reports for time periods more than 90 days in the past. If you need to analyze log data for longer time periods, generate reports at least once every 90 days, then export and store the reports. You can then analyze the aggregated reports, which were collected over longer time periods. Alternatively, you can set up a syslog server to collect and store raw search log data as it is generated.

Advanced Search Reporting (ASR)

The Google Search Appliance provides an analytics feature, advanced search reporting (ASR), that captures detailed information about user search and navigation activity. ASR is not enabled by default but can be activated easily through the search appliance's web-based Admin Console. Analytical information can then be extracted from the search appliance and consumed into your existing analytical tool.

You can use these reports to analyze data such as:

- Searches being executed
- Number of results returned
- How many pages a user clicked through
- Which result on the page they clicked on and where they went

Whatever your solution, Google highly recommends that you provide a rich analytics capability, regularly examine the data to refine your search deployment, and identify ways to add additional value.

For information about advanced search reporting, see Gathering Information about the Search Experience.

Scenarios using Advanced Search Reports

The following scenarios describe situations where using advanced search reporting can enhance the user experience:

- Using core capabilities to help users find content more efficiently
- Demonstrating ROI and quantifying business value
- Understanding what content is important

Using core capabilities to help users find content more efficiently

In this scenario, user feedback is consistently positive and users are able to find the content they are looking for. They report that they spend substantially less time searching for results.

However, after examining your users' search behavior by using an ASR report, you notice that 90% of searches for "widget" are immediately followed by a second search for "gadget." Similarly 50% of users searching for "vacation" click on the fifth link—to your policy database.

Based on these observations, there are two immediate actions that you might take to increase user effectiveness:

- Activate query expansion (Off by default) and upload your own synonyms list, including an expansion that equates widget and gadget, so that a search for "widget" automatically becomes a search for "gadget."

  As a result of this enhancement, 90% of users running a search for widget or gadget will find search twice as effective. Using query expansion, and adding your lexicon to the Google Search Appliance is a quick way to increase search effectiveness immediately.

- Create a KeyMatch for "vacation" to return the vacation policy.
  This enhancement presents the vacation policy at the top of the search results. For 50% of users, the result at the top of the page is the one they want to see.

For information about using query expansion and KeyMatches, see Creating the Search Experience: Best Practices.

Demonstrating ROI and quantifying business value

In this scenario, your workforce consists primarily of knowledge workers. By examining your analytical information provided by ASR, you can see that users typically:

- Execute on average 1.3 searches
- Retrieve result number 3
- Spend on average 45 seconds using search

This is a substantial improvement on the estimated 25% of time spent on search by the average knowledge worker. (Source: "Hidden Costs of Information Work: A Progress Report.")

Understanding what content is important

In this scenario, you have a large number of content repositories across a variety of internal websites and file shares. By analyzing your data provided by an ASR report, which records user clicks, you see that most users click on content from your corporate wiki. This immediately tells you where your users find the content most compelling.

Because you know that the corporate wiki is where your most useful content is, you can create a result biasing profile that moves it higher in search results. By doing this, you ensure that the corporate wiki appears in results where users can most quickly find it.

For information about creating result biasing profiles, see Using Result Biasing to Influence Result Ranking.

You can also understand what content is important by using analytics to gather information about user clicks. If your organization has an existing analytics solution in place, it may be possible to use this

solution to provide analytical insight into the user search experience. In many cases, integration with a third-party analytical solution requires some effort to get search-specific reporting, but there is substantial value that can be derived from the data.

Reporting in multiple GSAs

The following sections describe considerations for gathering reporting data from deployments involving multiple GSAs.

- Reporting [with GSA Mirroring](#)
- [Reporting with GSA Unification](#)
- [Reporting with Distributed Crawl and Serve](#)

Reporting with GSA Mirroring

In an active-active serving scenario, search queries performed on replica search appliances are combined into the master's search logs and search reports. A search log or generated search report on a replica search appliance only contains data gathered from that individual appliance.

Because these search logs and search reports are not replicated between GSAs, if you need to save search logs or generated reports, export them periodically or send them to a syslog server. For additional information, see the search appliance help system topics on logs, reports, and syslog servers.

Reporting with GSA Unification

- Scenario A: If you only search from the primary GSA (the master), you only need to run reports on the primary.

- Scenario B: Each GSA is acting as a primary and a secondary. Reports would have to be run on each GSA and then be aggregated together off-box.

Reporting with Distributed Crawl and Serve

You need to run reports from any GSA that is serving results to users. For example, if you have a three GSAs set up, and only serve from GSA 1, then you only need to run reports from GSA 1. However, if you serve from all three GSAs, you need to run reports from all three GSAs.

Chapter 7 User Feedback

Overview

Users are conditioned to have a great search experience with continuous innovation on Google.com. By using the search appliance, you can deliver a similarly innovative and rich experience to your users.

If the GSA is replacing an existing search solution, a great way for gathering feedback and metrics is by conducting a review of the existing solution for comparison with the GSA solution later.

One of the best ways to innovate is by capturing user feedback on what they like and don't like about the search solution, as well as understanding how they are using it. User feedback is critical to a successful deployment. To deliver value, not only must you deliver a great search experience, but also you need to have users actively using it.

There are several methods for getting user feedback, including:

- Implicit feedback
- A feedback link
- A user survey

Implicit feedback

By activating Advanced Search Reporting, or another analytical capability, you can automatically see what your users are doing, where they are succeeding, and how you can help them be more effective. However, it's important not only to capture this data, but also to use it.

Feedback link

Make it easy for users to provide feedback by providing a link or an email address for submitting their comments.

User survey

A user survey is a great tool to analyze how satisfying users find your search solution. Surveys should be sent out regularly, and after each phase in your deployment, so that you can iterate rapidly, and continue to delight your users. See the sample survey in the following section, which you could use as a template to make your own.

Search satisfaction survey

1. **What is your role?**
   Engineering
   Finance
   Human Resources
   Sales
   Marketing
   Research
   Other <customer to fill in>

2. **What percent of your time is spent looking for information?**
   More than half my time
   A quarter to half my time
   30 minutes to 2 hours per day
   10 minutes to 30 minutes per day
   Less than 10 minutes per day

3. **How often does your result show up in the top 10 (first page)?**
   100% of the time
   80% of the time
   50% of the time
   20% of the time
   Never

4. **How often does your result show up as the first result?**
   100% of the time
   80% of the time
   50% of the time
   20% of the time
   Never

5. **How often do you click on one of the Recommended Links (the shaded key matches at the very top of the results)?**
   Whenever I see one
   Sometimes
   Never

6. **What would you like to see improved?**
   Make the results come back faster
   Make the results more relevant
   Add more content sources

7. **How is the query response time?**
   Excellent

Sufficient
Unacceptable

8. Which content sources would you like to see indexed (added to the search results)?

_____

_____

_____

_____

_____

_____

9. Have you ever had documents that you knew existed but couldn't find them with search?
Yes
No

10. How satisfied are you with your overall search experience?
I love it—it works great!
It's alright but I'm not happy—it could be improved.
It's the worst search ever.

11. What would you like to see added?

_____

_____

_____

Summary

This guide provides detailed guidance on the factors to consider when creating the user experience with the Google Search Appliance. A user's search experience and needs on a public website differ vastly from a user's search behind the firewall. As such, when designing your user experience, it is important to keep the end user in the foremost of your mind. You can use this guide to ensure that not only are the right type of features deployed but the users' pain points and needs are addressed by these features.

It is essential to revisit the user experience from time to time as the users' needs change and new features become available on the GSA. Leveraging the reporting data that the GSA provides on the box can be a great way to spot any gaps and improve the overall experience for the user.

For further details on any of the topics discussed within this guide, refer to the GSA documentation website, or contact [Google for Work Support](#).

Appendix More Information

Technical solutions for common challenges

This appendix presents some technical solutions for common challenges in the following areas of search appliance deployment.

- [Interfaces and Front End Customization](#)
- [Document Relevancy](#)

Interfaces and Front End Customization

I just changed my front end but when I view my results, they still show the old one. What is wrong?

- A front end will only reload itself into memory every 15 minutes (or even longer). Therefore, in order to force a reload of the front end, you must use the parameter `proxyreload=1` in the query URL at least once after the style sheet has been modified. This parameter should only be used for a refresh during development and not in production, as it negatively impacts the performance of the Google Search Appliance.

How can I give developers access to the front end so that they can make changes without being able to affect my KeyMatches, and so on?

- You can create two front ends, using some naming convention. For example, use the one called "my_frontend" to manage KeyMatches, related queries, filters, remove URLs, and OneBoxes (collectively known as `"client"`). Then create another one called "my_frontend_ss" to manage the user interface (or output as it is denoted in the Admin Console), which is referred to as `"proxystylesheet."`

- Give the UI developer access to "my_frontend_ss" only so they can update their style sheet there.

- Retain control over "my_frontend" where user's search experience is managed by a non-UI developer.

- Make a modification to the application so it uses two front ends: `client=my_frontend&proxystylesheet=my_frontend_ss`. By default, it would be `client=my_frontend&proxystylesheet=my_frontend`.

How can I reuse some Front End configuration?

- If you want to use most configurations in a front end for different user interfaces, while you want to have different options for query expansion policies and/or result biasing policies, do not create multiple front ends for this. Use `entqr` and `entsp` instead.

How can I see the XML that the Google Search Appliance is sending back before it gets transformed?

- For results, remove the proxystylesheet parameter and value. For example:

  - ```
    http://GSA_HOSTNAME/search?
    q=query&btnG=Google+Search&access=p&client=default_frontend&outpu
    t= xml_no_dtd&sort=date:D:L:d1&entqr=0&oe=UTF-8&ie=UTF-
    8&ud=1&site=default_collection
    ```

- For dynamic results clustering, you can directly query the Google Search Appliance for the XML output. For example:

  - ```
    http://GSA_HOSTNAME/cluster?q=query&site=default_collection&clien
    t=default_frontend&coutput=xml
    ```

How can I integrate the Google Search Appliance into a non-web application?

- The Google Search Appliance accepts HTTP requests, and can return XML (or other formats after having been transformed by means of an XSLT). The returned results can then be parsed by an application, written in the language of your choosing, and then used for whatever purpose the application requires.

Document relevancy

How can the Google Search Appliance sort the results by other criteria than relevancy and date?

- It is exactly the purpose of a search engine to sort the search result by relevancy. Everything else is the output of a data base query.

  If you need to sort the results by any other numeric value you have, you can make use of the date sort feature. To do so, convert the value to an ISO-8601 date format (YYYY-MM-DD) and insert it to a meta tag in your document. You can search any dates between 1900-01-01 and 2079-06-06. Then set up the respective name of the meta tag in the section "Document Dates" in the admin console. The Google Search Appliance considers the value of this meta tag as the document date and can sort it by this value.

  Starting in 7.2, it is also possible to sort by metadata.

I want to promote a URL to the top of the results. How can I do this?

- Use KeyMatches.
- Implement User Results.
- Create a result biasing policy, which increases the relevancy of documents based upon the URL, Collection, Date, or MetaData. Attach this policy to the appropriate front end.

How can I increase the relevancy, in the search results, of more recent documents?

- Create a result biasing policy, which increases the relevancy of documents based upon the date that they were last modified. Then, attach this policy to the appropriate front end.

- For URLs which have been added via a content feed you can add the following attribute to the feed: `pagerank`

    - This attribute specifies the PageRank of the URL or group of URLs. The default value is 96. A higher value specifies a higher PageRank. To boost the PageRank of the URL or group of URLs, increase the value to a number between 97 and 100. To lower the PageRank, decrease the value. The PageRank for a URL overrides one for a group.

- By using the [automatic self-learning scorer](#).

    - If you enable ASR, the search appliance uses its automatic self-learning scorer. This feature automatically analyzes user behavior and the specific links that users click on for specific queries to fine tune relevance and scoring. The search appliance uses advanced statistical regression to determine the statistical significance of user behavior, and adjusts for trust bias (that is, users clicking on the first result solely because it's first). Thus, over time, results become more and more precise without the need of administrator intervention.