

# Google Search Appliance

## Diseño y planificación de la solución

Marzo de 2014



© 2014 Google

# Diseño y planificación de la solución

En este documento se trata el diseño y la planificación de la solución para una implementación de Google Search Appliance (GSA).

## Acerca de este documento

La información y recomendaciones que se exponen en este documento se recopilaron gracias al trabajo que realizamos con distintos clientes y entornos del sector. Nos gustaría agradecer a nuestros clientes y partners que hayan compartido sus experiencias y puntos de vista con nosotros.

<b>Temas tratados</b>	En este documento se incluyen actividades de planificación con principios básicos de orientación que pueden aplicarse a cualquier implementación de un sistema de búsqueda
<b>Público objetivo principal</b>	Gestores de proyectos y administradores de GSA
<b>Entorno de TI</b>	GSA, varias fuentes de contenido y mecanismos de seguridad
<b>Fases de implementación</b>	Planificar la implementación
<b>Otros recursos</b>	<ul style="list-style-type: none"><li>● <a href="#">Notas sobre GSA: arquitecturas de desarrollo</a></li><li>● <a href="#">Notas sobre GSA: seguridad</a></li><li>● <a href="#">Learnrsa.com</a> proporciona recursos formativos para GSA.</li><li>● <a href="#">La documentación de producto de GSA</a> proporciona información completa sobre GSA.</li><li>● <a href="#">El Portal de asistencia de Google for Work</a> proporciona acceso a la asistencia de Google.</li><li>● <a href="#">Google for Work Connect</a></li></ul>

## Índice

[Acerca de este documento](#)

[Capítulo 1 Descripción general](#)

[Capítulo 2 Documentar los requisitos](#)

[Requisitos de usuario](#)

[Requisitos de contenido y de seguridad](#)

[Requisitos de asignación de recursos](#)

[Requisitos de rendimiento y de escalabilidad](#)

[Sistema de conmutación por error/alta disponibilidad](#)

[Requisitos de administración y de informes](#)

[Capítulo 3 Identificar las fases](#)

[Fases de implementación](#)

[Punto de partida](#)

[Fases de las actividades](#)

[Duración de las fases](#)

[Capítulo 4 Definir los criterios de una implementación correcta](#)

## Capítulo 1 Descripción general

Desde el punto de vista conceptual, una solución de búsqueda eficaz es muy simple: ayudar a los usuarios a encontrar la información que buscan. La búsqueda debe ser rápida, fácil y relevante.

Google Search Appliance se encarga de proporcionar la rapidez, la facilidad y la relevancia. Pero hay que planificar y ejecutar el proyecto para aprovechar todo el potencial del dispositivo de búsqueda. Para ello es clave que las fases sean cortas y estructurar el trabajo en función de estos ciclos.

Cada implementación de una solución de búsqueda de Google es única. En un caso, se podría proporcionar el sistema de búsqueda para contenido de SharePoint y ampliar la búsqueda básica con órdenes de compra de SAP. En otro, se podría proporcionar un sistema de búsqueda de los millones de documentos que las empresas acumulan con el tiempo, agrupándolos con documentos de políticas y los datos de contacto de quienes los crearon.

Aunque cada implementación tiene diferentes fuentes de contenido, requisitos de seguridad y necesidades de usuario, hay actividades de planificación básicas con principios fundamentales de orientación que pueden aplicarse a cualquier implementación de una solución de búsqueda. Este documento se centra en las siguientes actividades de planificación básica:

- [Documentar los requisitos](#)
- [Identificar las fases](#)
- [Definir los criterios de una implementación correcta](#)

## Capítulo 2 Documentar los requisitos

Cuando recabe los requisitos, agrúpelos según la relación para que pueda priorizarlos y asociarlos con las fases de trabajo. En general, céntrese en estas áreas:

- [Requisitos de usuario](#)
- [Requisitos de contenido y de seguridad](#)
- [Requisitos de rendimiento y de escalabilidad](#)
- [Requisitos de administración y de informes](#)

### Requisitos de usuario

Averigüe, desde la perspectiva del usuario, qué elementos hacen que una implementación sea eficaz. En general, los requisitos del usuario se basan en lo siguiente:

- [Funcionalidad](#)
- [Amplitud y profundidad](#)
- [Comunicación y comentarios](#)

### Funcionalidad

Para los usuarios, la búsqueda no debería ser una tarea. Definir los requisitos de funcionalidad puede ayudar a que los usuarios piensen que la solución de búsqueda es intuitiva y eficaz.

Para identificar los requisitos de la funcionalidad, tenga en cuenta las cuestiones siguientes:

- ¿Qué características de la funcionalidad son importantes en una solución de búsqueda para los usuarios?
  - Google Search Appliance ofrece muchas características de funcionalidad fáciles de implementar y que no requieren configuración, como las sugerencias de consultas, KeyMatch, Navegación dinámica, Reconocimiento de entidades o Búsqueda experta, entre otras.
- ¿Cómo acceden los usuarios a la búsqueda?
  - Elija el punto de acceso a la búsqueda que sea más cómodo para los usuarios. Se podría acceder a la página de búsqueda directamente desde GSA o de forma indirecta a través de una aplicación web que analizara los resultados de XML de GSA. En el caso de los portales indexados, quizá sea mejor reemplazar la interfaz de búsqueda existente con la Búsqueda de Google (p. ej. el cuadro de búsqueda para SharePoint).
  - La búsqueda puede tener varios puntos de acceso: una intranet, una aplicación web, una aplicación para móviles o incluso un ordenador. Si hay varios puntos de acceso, la complejidad podría ser mayor al ser necesarios otros mecanismos de autenticación.
- ¿Cuáles son los requisitos de velocidad de los usuarios?
  - Las tecnologías de estilo AJAX pueden mejorar considerablemente la capacidad de respuesta y el rendimiento percibidos, al tiempo que proporcionan una experiencia de búsqueda más completa.

En general, cumpla con los requisitos de funcionalidad al principio de la fase porque no suelen estar vinculados a fuentes de contenido y pueden crear buenas expectativas sobre la solución de búsqueda entre los usuarios.

## **Amplitud y profundidad**

Una solución de búsqueda debe satisfacer las demandas de los usuarios. Definir los requisitos de amplitud y profundidad le permite asegurarse de que abarca un grupo de usuarios suficientemente amplio y, al mismo tiempo, les proporciona el contenido adecuado.

A la hora de identificar los requisitos de amplitud y profundidad, tenga en cuenta estas cuestiones:

- ¿Cómo son los grupos de usuarios?
  - Siempre que sea posible, céntrese primero en los grupos y usuarios que estén teniendo más problemas para encontrar la información que quieren.
  - Con las interfaces del dispositivo de búsqueda podrá presentar un diseño y contenido distintos a varios usuarios, en función de sus necesidades. Si desea obtener información sobre las interfaces, consulte [Crear la experiencia de búsqueda](#).
- ¿Qué información quieren buscar pero no encuentran?
  - Este contenido deberá incluirse en las primeras fases.
  - Los usuarios y el contenido que necesitan deberán incorporarse al mismo tiempo. Es decir, no proporcione a los usuarios una herramienta de búsqueda si todavía no ha incluido el contenido que buscan.
- ¿Algunos usuarios tienen necesidades de búsqueda más sofisticadas?
  - La búsqueda avanzada estándar se puede ampliar fácilmente a través del uso de metadatos y otras funciones básicas.
  - Si no es su caso, proporcione una búsqueda simple pero con funciones completas.

## **Comunicación y comentarios**

Los comentarios de los usuarios son una herramienta eficaz para identificar los problemas de funcionalidad de la solución. Cuando pide a los usuarios que le hagan llegar sus comentarios, les está diciendo que su opinión sobre la implementación de búsqueda es importante. Al definir los requisitos de comunicación y comentarios, los usuarios podrán enviar información al equipo acerca de la implementación de búsqueda, es decir, los aspectos que funcionan o no según ellos.

A la hora de identificar los requisitos de comunicación y comentarios, tenga en cuenta estas cuestiones:

- ¿Qué debe comunicar a los usuarios?
  - Además de añadir contenido nuevo y funciones interesantes, es importante que se lo diga a los usuarios para crear buenas expectativas sobre el producto y recibir cumplidos tras cada logro.
  - Dado que la mayoría de los usuarios ya saben cómo funciona la tecnología de búsqueda de Google, la necesidad de formación suele ser mínima. Sin embargo, dídeles que ahora pueden hacer búsquedas en el contenido de la empresa con la misma facilidad que buscan en Internet desde sus casas.

- ¿Cómo recibirá los comentarios de los usuarios?
  - Designe a un administrador de búsqueda para que revise periódicamente los informes de GSA y recabe estadísticas sobre las principales consultas o las consultas que no muestran resultados, entre otros elementos, para obtener información sobre cómo ajustar la experiencia de usuario. Es una forma de obtener información sobre la solución de búsqueda. Esto se podría automatizar a través del API de administración.
  - Los comentarios de los usuarios son una de las mejores formas de medir el éxito de una solución. Considere la posibilidad de realizar encuestas periódicas a los grupos de usuarios.
  - Podría proporcionar a los usuarios un enlace para comentarios (p. ej., dentro del sistema de incidencias).
  - Inserte en la página de resultados de búsqueda un informe que pregunte a los usuarios si han encontrado lo que buscaban para que puedan proporcionar sus comentarios directamente.
  - Para captar la opinión de los usuarios, también puede usar el [kit de herramientas Search Quality Feedback](#) gratuito.

## Requisitos de contenido y de seguridad

Para la mayoría de las organizaciones, los siguientes dos aspectos de una implementación de búsqueda suelen ir juntos:

- [Contenido](#)
- [Seguridad](#)

Los casos que abarcan contenido y seguridad pueden presentar varios grados de complejidad, desde páginas de sitios web públicos sin protección alguna hasta una integración compleja con un sistema de planificación de recursos empresariales, como SAP o PeopleSoft, y cualquier posibilidad intermedia.

Planifique la arquitectura de estado final en las primeras fases de la implementación, pero introduzca los elementos de contenido y seguridad por fases. Es decir, no retrase la solución de búsqueda ante los usuarios porque quiera indexar todo el contenido o implementar un proceso de seguridad que el usuario no vaya a necesitar hasta más adelante.

### Contenido

En general, analice todos los posibles repositorios de información empresarial en las primeras fases del proyecto. Aunque Google Search Appliance es muy eficaz a la hora de proporcionar búsquedas eficientes, rápidas y relevantes en contenido no estructurado, no excluya el contenido estructurado, como los datos de almacén o sistemas transaccionales, entre otros.

Es importante entender la relación que existe entre las fuentes de contenido, ya que le ayudará a establecer las fases de la implementación del contenido. Por ejemplo, el contenido de un sistema de gestión de casos puede complementarse eficazmente con contenido de un catálogo de productos, lo que permitirá a los usuarios ver no solo información sobre productos, sino también los tipos de problemas y cuestiones con las que se encuentra un usuario al utilizar esos productos.

En la tabla siguiente se incluyen varios tipos de fuentes de contenido estructuradas y no estructuradas, así como consideraciones para definir cómo establecer las fases de la implementación.

Fuente de contenido	Estructurada/ no estructurada	Complejidad (baja/media/alta)	Consideración
Sitios web públicos	N	B	<ul style="list-style-type: none"> <li>● HTTP</li> </ul>
Sistemas de archivos	N	M-A	<ul style="list-style-type: none"> <li>● SMB (conector) o HTTP</li> </ul>
Bases de datos	E	M-A	<ul style="list-style-type: none"> <li>● Utilice un conector para bases de datos o un feed, o habilítelas para la Web</li> <li>● Volumen de contenido</li> </ul>
Sitios web de intranet	N	B	<ul style="list-style-type: none"> <li>● Quizá haya que considerar la seguridad</li> <li>● Gestión de páginas dinámicas</li> </ul>
Portal para el personal	N	M-A	<ul style="list-style-type: none"> <li>● Quizá haya que considerar las complejidades de seguridad</li> <li>● Quizá haya que disponer de URLs que no sean únicas (que una misma URL incluya contenido distinto, según la función del usuario)</li> </ul>
CMS	N	B-M	<ul style="list-style-type: none"> <li>● Quizá se puedan aprovechar metadatos</li> <li>● Hay que determinar si se puede rastrear de forma nativa</li> <li>● Quizá haya que considerar la seguridad</li> </ul>
Aplicaciones LOB (p. ej., Lotus Notes)	E	M-A	<ul style="list-style-type: none"> <li>● Si están habilitadas para la Web, es posible que estas fuentes de contenido se puedan rastrear</li> <li>● Quizá se necesite un conector</li> <li>● Seguridad</li> </ul>
Aplicaciones empresariales (p. ej., ERPs)	E	M-A	<ul style="list-style-type: none"> <li>● Hay que identificar los datos principales</li> <li>● Quizá se necesiten feeds o un conector</li> <li>● Seguridad</li> </ul>
Otros sistemas transaccionales	E	M-A	<ul style="list-style-type: none"> <li>● Se puede acceder mediante un conector personalizado o un módulo OneBox</li> <li>● Seguridad</li> </ul>

## Security

La seguridad puede ser el área de mayor complejidad en una implementación de búsqueda. Cuando analice el contenido, averigüe si está protegido y, en ese caso, de qué manera (protegido mediante formularios, cookies, un sistema de seguridad de nivel de aplicación, etc.).

Sobre todo en el caso de proyectos con varias fuentes de contenido, es importante analizar en profundidad las opciones de seguridad antes de empezar con la implementación. Es fundamental, ya que cierta información sobre la seguridad puede cambiar por completo la perspectiva de un proyecto. Por lo tanto, asegúrese de realizar el análisis de la seguridad y de recabar los requisitos previamente.

En una solución de búsqueda, la seguridad afecta principalmente a dos áreas:

- [Rastreo y adquisición de contenido](#)
- [Publicación de resultados: autenticación y autorización de usuarios](#)

Si desea obtener amplia información sobre el dispositivo de búsqueda y la seguridad, consulte [Gestionar la búsqueda para contenido con acceso controlado](#).

### Rastreo y adquisición de contenido

Google Search Appliance puede usar protocolos de seguridad estándar, como LDAP, HTTP básico, NTLM, Kerberos, certificados, SAML o seguridad basada en formularios.

Si analiza todas las modificaciones de seguridad, podrá planificar la adquisición de contenido. Por ejemplo, la seguridad podría afectar el rastreo web y de sistemas de archivos que quizá deba planificar, como configurar un servidor proxy o asegurarse de que sus sistemas de archivos Windows tienen habilitado CIFS para la indexación de SMB.

Si la seguridad es más compleja, quizá haya que buscar formas alternativas de adquisición de contenido, como feeds o conectores.

En la mayoría de los casos, GSA estará preparado para indexar configuraciones estándar. Sin embargo, en algunos casos, quizá haya que hacer pequeños ajustes al entorno para que el dispositivo de búsqueda rastree y adquiera contenido sin necesidad de feeds ni de conectores. Por ejemplo, es posible que para ampliar el rastreo a un subdominio nuevo deba modificar un dominio de cookie, ya que el dispositivo de búsqueda rastrea contenido para permitir que las cookies se ajusten a la solicitud de especificaciones de comentarios (RFC). Estos tipos de cambios suelen ser mínimos y se pueden aplicar por medio de distintos métodos.

En ocasiones, las consideraciones sobre seguridad obligan a ajustar el procedimiento de indexación o a usar otro método para adquirir el contenido. Estas circunstancias pueden afectar a aspectos del diseño de la solución que no están directamente relacionados con la arquitectura de seguridad, como feeds o flujos de publicación de contenido.

Es necesario analizar la adquisición de contenido junto con la autorización de usuario, ya que los métodos de autorización con enlace en tiempo de compilación y con enlace en tiempo de ejecución pueden variar. En general, el enlace en tiempo de compilación (con ACLs) se considera una forma más rápida de mostrar los resultados de búsqueda que el enlace en tiempo de ejecución, porque en el primer caso los permisos de seguridad se almacenan en GSA. Por lo tanto, el enlace en tiempo de compilación debe almacenar en caché en GSA los permisos, lo que significa que la información tiene que estar disponible para el dispositivo. Esta información de ACL puede incluirse en feeds, en la información de cabecera HTML o en las cabeceras HTTP del servidor. El enlace en tiempo de ejecución es la comprobación que se realiza cuando los resultados de búsqueda se van a verificar, como normalmente sucede, con el sistema de fuentes para garantizar que el usuario tiene acceso inmediato, por lo que no se necesita personalizar el contenido.

Si desea obtener más información al respecto, consulte [Notas sobre GSA: seguridad](#).

### **Publicación de resultados: autenticación y autorización de usuarios**

Al mostrar contenido protegido, Google Search Appliance primero comprueba que el usuario pueda ver los resultados relevantes. Si el usuario no tiene permiso para ver un documento, este no aparece en el conjunto de resultados.

Tenga en cuenta que la publicación de los resultados protegidos se gestiona a través de dos procesos, la autenticación y la autorización:

- La autenticación verifica si el usuario es quien dice ser. Durante el proceso, GSA también reúne toda la información sobre grupos del usuario.
- La autorización utiliza el ID del usuario para asegurarse de que tiene acceso al contenido en cuestión.

Es fundamental entender este concepto si hay que indexar varias fuentes de contenido protegido. Por ejemplo, quizá solo se necesite un mecanismo de autenticación para dos fuentes de contenido protegido que utilicen el mismo ID de usuario. Sin embargo, si ambas fuentes de contenido requieren métodos de autorización diferentes, habrá que configurar GSA para que use dos métodos de autorización (p. ej., ACLs para una fuente de contenido y autorización mediante conector para la otra fuente).

Evidentemente, siempre tiene la opción de mostrar todos los resultados de forma pública y no aplicar ningún tipo de protección. En muchos casos, la búsqueda de puede implementar sin protección al inicio y luego añadir seguridad a medida que se adquiere contenido. La búsqueda pública (como un sitio de Internet externo) se suele implementar de esta manera.

Sin ajustar la configuración, Google Search Appliance se integra con muchos mecanismos de seguridad distintos y puede satisfacer requisitos de seguridad diversos, como la autenticación silenciosa (p. ej., a través de cookies o de Kerberos) y la integración con un sistema de inicio de sesión único. Para casos de integración personalizada, se proporciona la interfaz de proveedor de servicios SAML (SPI), tal y como se describe en [Notas sobre GSA: seguridad](#). La SPI SAML gestiona las comprobaciones de autenticación y de autorización en varios sistemas y protocolos. Esta capacidad ofrece una gran flexibilidad en cómo se implementará la seguridad. Dispone de estas opciones:

- Comprar proveedores ya creados (consulte ejemplos en el mercado de soluciones de Google for Work)
- Crear una solución de seguridad personalizada
- Integrar en un sistema de inicio de sesión único a través de SAML

Google Search Appliance también permite definir listas de control de acceso (ACL) para realizar comprobaciones de autorización con documentos mediante enlaces en tiempo de compilación. Las ACL no solo mejoran el rendimiento, sino que también ofrecen más opciones para gestionar la seguridad. Esta capacidad también ofrece opciones para implementar la búsqueda segura por fases. Si desea obtener información sobre la publicación de contenido protegido y las ACL de políticas, consulte [Notas sobre GSA: seguridad](#).

Una vez que haya recabado los ID de usuario y los grupos, quizá deba ajustar el orden del mecanismo de autorización, por ejemplo, quizá en un CMS no deba permitirse la comprobación de los resultados en caché, pero en otra plataforma, sí. A reglas como estas, incluido el orden de los métodos de autorización para un patrón, se les puede asignar una regla de [autorización flexible](#) y debería considerarse detenidamente en entornos complejos.

## Requisitos de asignación de recursos

Disponer de los expertos adecuados en el momento que se necesitan es fundamental para la implementación de un proyecto o, de lo contrario, podrían producirse retrasos. Por lo tanto, es importante tener las tareas asignadas a las personas adecuadas cuando se necesitan.

Por ejemplo, el administrador de SharePoint quizá deba ayudar con la configuración del conector de GSA para SharePoint o con las pruebas del sistema.

Planificar los recursos con antelación también permite saber a los expertos cuándo deberán llevar a cabo una determinada tarea, por lo que podrán encajarla en sus respectivos horarios.

Estos son los expertos que suele necesitar en un proyecto el propietario de la solución GSA:

- Director de proyecto
- Director técnico de proyecto
- Propietario de la empresa
- Analista de negocios
- Administrador de repositorios
- Administrador de red
- Experto en seguridad
- Administrador de GSA
- Desarrollador

Estos son los expertos que suele necesitar en un proyecto un partner de Google for Work:

- Director de proyecto
- Expertos en GSA
- Expertos en repositorios
- Desarrolladores

## Requisitos de rendimiento y de escalabilidad

Los requisitos no operativos suelen ser requisitos puramente técnicos. En una solución de búsqueda, estos son los requisitos no operativos más comunes:

- [Rendimiento](#)
- [Escalabilidad](#)

### Rendimiento

Los requisitos de rendimiento suelen girar en torno a la rapidez con la que la solución muestra los resultados, aunque también puede haber requisitos sobre la velocidad de adquisición de contenido.

El rendimiento suele depender de varios factores, por ejemplo, los siguientes:

- Requisitos de seguridad
- Tipo de contenido
- Tamaño del corpus
- Tipo de consultas que se realizan
- Arquitectura de red y rendimiento
- Funciones de búsqueda adicionales que se usan (p. ej., Ampliación de consultas, Navegación dinámica, Agrupación de resultados dinámica, etc.)

Por norma, si existen requisitos de rendimiento específicos, debería realizar una prueba de rendimiento al principio de la implementación para determinar los cambios que quizá deban hacerse en la arquitectura de la solución.

Aunque Google Search Appliance no se pueda modificar, estos son algunos de los cambios que puede incorporar en la implementación planificada:

- Configurar ACLs de políticas o ACLs por URL para mejorar la comprobación de la seguridad al mostrar los resultados
- Implantar un servidor proxy para almacenar en caché las búsquedas comunes siempre que sea posible; este cambio beneficia solo a búsquedas de contenido público (sin protección)
- Minimizar el tráfico de red entre Google Search Appliance y las fuentes de contenido; aunque este cambio afecta básicamente al rastreo, gracias a la latencia reducida se mejora el rendimiento de la autorización con enlace en tiempo de ejecución
- Mejorar el rendimiento percibido a través de las optimizaciones de la capacidad de respuesta (p. ej., AJAX) y mostrar un indicador de progreso para crear percepción de capacidad de respuesta
- Implantar otros dispositivos de búsqueda para distribuir la carga; este cambio reduce la demanda sobre cada dispositivo y garantiza que la capacidad no sea un factor restrictivo

Consulte [Notas sobre GSA: arquitecturas de desarrollo](#) para obtener más información sobre la arquitectura de búsqueda basada en el rendimiento.

Para los requisitos de rendimiento también habría que tener en cuenta el rastreo y la indexación. La indexación del dispositivo de búsqueda añade carga a los sistemas de contenido. Si hay ciertos momentos del día en que los sistemas de contenido no deben verse afectados, es importante saberlo para configurar los horarios de carga del host del dispositivo de búsqueda en consecuencia. Además, si el sistema de contenido tiene bastante carga o funciona con especial lentitud, considere la posibilidad de usar feeds de contenido o un conector como alternativa.

Aparte del método de indexación que elija (rastreo o feeds) y el tamaño del documento, las funciones como Reconocimiento de entidades o Vista previa del documento pueden afectar al rendimiento, ya que suelen procesar otros datos durante la indexación.

## Escalabilidad

Los requisitos de escalabilidad suelen girar en torno al número de consultas por segundo o al número de consultas por minuto. Al igual que sucede con el rendimiento, el número de consultas por segundo que la solución admite depende de los requisitos de seguridad, del tipo de contenido y de consulta, del rendimiento de la red y de otros factores.

Google recomienda que, en caso de que haya requisitos de escalabilidad, vuelva a comprobarlos; lo ideal sería con métricas derivadas de las búsquedas. En muchos casos, el requisito de escalabilidad no es tan elevado como se determina en un principio.

Si bien las soluciones de búsqueda se pueden diseñar para admitir cientos de consultas por segundo, en la práctica, no suele necesitarse. El tipo de requisitos de escalabilidad que se necesitan de una solución de búsqueda son sustancialmente distintos de los de un sistema transaccional.

Si desea consultar más información sobre el diseño de una solución de búsqueda para obtener mayor escalabilidad, consulte [Notas sobre GSA: arquitecturas de desarrollo](#). Si desea obtener información sobre el número de conexiones simultáneas que admite Google Search Appliance, consulte [Diseñar una solución de búsqueda](#).

## Sistema de conmutación por error/alta disponibilidad

La alta disponibilidad en el entorno de búsqueda es un requisito estándar para muchas empresas. GSA proporciona herramientas como la duplicación de GSA a GSA que permiten configurar y mantener fácilmente una solución con una gran disponibilidad. Sin embargo, para entornos más complejos, debe asegurarse de planificar adecuadamente la solución de conmutación por error.

Con respecto al entorno de búsqueda, una solución de conmutación por error suele abarcar dos requisitos:

- [Alta disponibilidad en la indexación](#)
- [Alta disponibilidad en la publicación](#)

## Alta disponibilidad en la indexación

La alta disponibilidad en la indexación garantiza que el índice de GSA esté actualizado aunque el entorno principal no pueda prestar servicio. Si bien GSA es un dispositivo fundamental para esta configuración, normalmente también tienen que configurarse para un sistema de conmutación por error otros elementos colaboradores, como conectores o sistemas de feeds. Aunque la alta disponibilidad en la indexación suele ser una preferencia, no es muy habitual encontrarla en las implementaciones dado que implica un cierto grado de complejidad y de recursos.

## Alta disponibilidad en la publicación

La alta disponibilidad en la publicación garantiza que los usuarios sigan recibiendo resultados aunque una de las aplicaciones necesarias para ello esté inactiva. Para mostrar contenido público, en los casos de enlace en tiempo de compilación (ACLs de políticas y ACLs por URL) y en la mayoría de los casos de enlace en tiempo de ejecución (solicitudes HEAD), suele bastar con la función de duplicación de GSA. Sin embargo, hay casos en los que no está tan claro:

- Cuando se utilicen grupos que haya devuelto un conector (p. ej., Active Directory o SharePoint)
- En algunos casos de enlace en tiempo de ejecución, como la autorización por medio de conector, SAML Bridge o acceso a cookies, también se recomienda tener un sistema de conmutación por error, ya que, de lo contrario, los usuarios quizá no reciban los resultados de búsqueda si el conector principal está inactivo.

Si desea consultar más información sobre el diseño de una solución de búsqueda para un sistema de conmutación por error, consulte [Notas sobre GSA: arquitecturas de desarrollo](#).

## Requisitos de administración y de informes

Los informes son una parte importante de cualquier aplicación empresarial, y la búsqueda no es ninguna excepción. Además de los requisitos de los informes de búsqueda, identifique otros requisitos de informes. En concreto, preste atención a los requisitos no operativos. Estos son los requisitos que debería tener en cuenta:

- Tecnología analítica que hay que usar (p. ej., Google Analytics, Informes de búsqueda avanzados u otras herramientas de terceros)
- Frecuencia y distribución de los informes
- Otros tipos de informes que puedan ser necesarios (p. ej., supervisión de sistemas o eventos de administración)

Asegúrese de comprender bien los procesos empresariales que se usarán en estos informes. Por ejemplo, debería conocer los casos para los que se solicitan los requisitos de informes y asegurarse de que la estrategia de informes los cubre adecuadamente.

## Capítulo 3 Identificar las fases

La mayoría de las implementaciones de búsqueda se clasifican en una de estas categorías, que se enumeran de menor a mayor complejidad:

- Las implementaciones especializadas se centran en proporcionar una experiencia de búsqueda eficaz y conocida a los clientes de información externa pública o protegida de una empresa
- Implementaciones de búsqueda independientes centradas en ofrecer beneficios de productividad generales a empresas y hacer un mejor uso de los activos de información
- Implementaciones de búsqueda impulsadas por un evento interesante o una implementación de mayor envergadura, como un portal nuevo, un sistema de gestión del contenido empresarial (ECM) o un nuevo proyecto de arquitectura de la información

Una implementación de búsqueda suele tener como objetivo apuestas seguras y rápidas para poder ofrecer a los usuarios una experiencia de búsqueda completa rápidamente, con un valor incremental y repetitivo durante la vigencia de la implementación.

El valor empresarial se deriva de la amplitud de contenido para la que se proporciona la capacidad de búsqueda y la funcionalidad y eficacia de la experiencia de búsqueda.

### Fases de implementación

La clave para una implementación de búsqueda eficaz es proporcionar el servicio cuanto antes y con frecuencia. No intente hacerlo todo a la vez. Es importante que los usuarios tengan acceso al contenido que necesitan lo antes posible. Proporcionar el servicio cuanto antes es una apuesta segura y rápida que le ayudará a recibir el apoyo de las personas interesadas y a generar buenas expectativas entre los usuarios.

El ámbito de las fases podría definirse en los términos siguientes:

- Fuentes de contenido
- Seguridad
- Grupos de usuarios
- Características de la funcionalidad

Cada fase debería incluir una tarea de evaluación, donde se mida explícitamente la satisfacción del usuario y peticiones de funciones de los usuarios, unidades empresariales, departamento de TI, etc. Como siempre, evalúe las peticiones de funciones, incluidos los riesgos de implementarlas o no.

En general, dado que las fases son relativamente cortas, puede usar la mayoría de las metodologías de presentación, desde Agile a Life Cycle.

Si utiliza una variedad más clásica de metodología de desarrollo, tenga en cuenta que las fases de desarrollo de un proyecto de búsqueda de Google son relativamente cortas. En ese caso, deberá aplicar ajustes para poder proporcionar con eficacia una experiencia de búsqueda de calidad de una manera flexible. Muchas de las tecnologías que usará, como transformaciones de lenguaje de hoja de estilo extensible (hojas de estilo XSLT), módulos OneBox, etc., se pueden implementar y ajustar rápidamente. La flexibilidad es clave para crear prototipos rápidamente y repetir las entregas.

En esta sección se explica cómo estructurar las entregas y los planes del proyecto para ampliar la presencia del sistema de búsqueda y aumentar el uso de la solución. Con cada entrega, la implementación es más valiosa.

## Punto de partida

Google Search Appliance está diseñado para implementarse rápidamente en las principales fuentes de contenido. Mediante el uso de estándares y protocolos abiertos se puede integrar rápidamente el contenido de una gran variedad de fuentes e implementar características de funcionalidad completas, como Sugerencias de consulta, Navegación dinámica, Resultados añadidos por el usuario y Agrupación de resultados dinámicos.

Las fases pueden durar tan solo una o dos semanas, o un mes. Google recomienda estructurar el programa de trabajo en fases cortas para proporcionar rápidamente funciones, contenido o grupos de usuarios repetitivos.

En muchos casos, solo se necesitará una única fase de entrega rápida. Sin embargo, aunque la implementación forme parte de un programa más amplio que proporcione búsqueda universal en todos los activos de la empresa, igualmente debería estructurar las fases con apuestas rápidas y seguras.

Antes de empezar la implementación de búsqueda, realice las siguientes tareas básicas para que los especialistas en implementaciones de búsqueda puedan poner en marcha la solución lo antes posible.

Antes de empezar:

- Monte el dispositivo de búsqueda en el bastidor correspondiente.
- Ajuste la configuración de red.
- Haga el inventario de las fuentes de contenido (incluido un recuento de los documentos).
- Haga el inventario de los sistemas de seguridad, principalmente de los mecanismos de autenticación y de autorización.
- Configure la red para que el dispositivo de búsqueda pueda acceder a todas las fuentes de contenido y, en caso necesario, restringir el acceso a las áreas protegidas.
- Cree los ID de usuario que el dispositivo de búsqueda necesite para rastrear el contenido.

Como partner de Google for Work, utilice las herramientas que se proporcionan en las [actualizaciones de producto](#) para asegurarse de que cuenta con toda la información necesaria para llevar a cabo un proyecto de éxito.

## Fases de las actividades

Quizá le interese aplicar las actividades por fases, tal y como se describe en las secciones siguientes:

- [Implementación inicial](#)
- [Ampliaciones](#)
- [Entregas avanzadas](#)

### Desarrollo inicial

Los elementos de la entrega que se enumeran en la lista siguiente suelen ser rápidos y fáciles de proporcionar. Considérelos posibles elementos para un desarrollo inicial. Muchos de estos podrían considerarse obligatorios. Una interfaz personalizada, por ejemplo, aunque sea muy simple, debería formar parte siempre de la entrega principal.

Tenga en cuenta que en la primera fase se suele tener que indexar una fuente de contenido clave (alta prioridad). Sin embargo, sobre todo si no está familiarizado con GSA, le recomendamos que estudie el dispositivo usando una fuente de contenido que sea fácil de indexar. Si primero gestiona una fuente de contenido simple, esto le servirá de ayuda en fases posteriores, y obtendrá conocimientos clave y básicos sobre el sistema de búsqueda y los procesos relacionados.

En la tabla siguiente se incluyen algunos elementos para el desarrollo inicial.

Elemento de entrega	Tipo	Complejidad
Rastreo de HTTP básico: <ul style="list-style-type: none"> <li>● Intranet</li> <li>● Extranet</li> <li>● Sitio web</li> <li>● Wiki</li> <li>● Bases de conocimientos habilitadas para la Web (p. ej., Lotus Notes)</li> </ul>	Fuentes de contenido	Baja
Sistema de archivos	Fuentes de contenido	Media-alta
Sitios de SharePoint	Fuentes de contenido	Media-alta
Módulos OneBox básicos (p. ej. Búsqueda de perfiles)	Fuentes de contenido	Baja-media
Búsqueda experta	Fuentes de contenido	Media
Conector Active Directory	Seguridad	Media
Autenticación mediante LDAP	Seguridad	Baja
Integración con Kerberos	Seguridad	Media
Sugerencias de consulta	Funcionalidad	Baja-media*
Resultados añadidos por los usuarios	Funcionalidad	Bajo
Interfaz personalizada	Funcionalidad	Baja-media
Informes de búsqueda avanzados	Funcionalidad	Bajo
Navegación dinámica	Funcionalidad	Baja
Reconocimiento de entidades	Funcionalidad	Baja-media
Vista previa de los documentos	Funcionalidad	Baja-media
Búsqueda con comodines	Funcionalidad	Baja
Usuarios del sistema principal	Grupos de usuarios	Bajo
Propietarios del negocio	Grupos de usuarios	Bajo

\*Dependiendo de si utiliza funciones estándar o una implementación personalizada. La complejidad puede variar en función de la estructura y de la configuración del entorno.

## Ampliaciones

En la tabla siguiente se incluyen algunos elementos para la entrega de ampliaciones. Tenga en cuenta estos elementos y programe su implementación según la prioridad (normalmente en función del volumen de contenido y de la importancia del negocio), y el nivel de esfuerzo.

En muchos casos, se puede acelerar la entrega mediante herramientas de terceros (como conectores) y la ayuda de partners cualificados de Google for Work, que están especializados en problemas con la integración de Google Search Appliance. Para algunos de estos elementos de entrega (p. ej., búsqueda avanzada personalizada) quizá sea necesario conocer los comentarios de los usuarios antes de implementarlos por completo. En la tabla siguiente se incluyen posibles elementos para las ampliaciones.

Elemento de entrega	Tipo	Complejidad
Contenido de portal	Fuente de contenido	Media-alta
Bases de datos no habilitadas para la Web (p. ej., muchas bases de datos de Lotus Notes)	Fuente de contenido	Media
Sistemas de gestión de contenido	Fuente de contenido	Baja-alta
Módulos OneBox personalizados (pueden estar protegidos)	Fuente de contenido	Baja-alta
Contenido de aplicación personalizada	Fuente de contenido	Baja-alta
Conectores adicionales (FileNet, Livelink o Documentum)	Fuente de contenido	Baja-media
Búsqueda avanzada personalizada	Funcionalidad	Media
Características de funcionalidad avanzadas (p. ej., interfaz activada por AJAX)	Funcionalidad	Baja-media
Usuarios adicionales que dependen de contenido nuevo	Grupos de usuarios	Baja-media
Gestión de varios dispositivos GSA (unificación/GSA^n)	Arquitectura	Alta

## Entregas avanzadas

En la tabla siguiente se incluyen algunos elementos para la entrega avanzada. Es posible que estos elementos requieran más tiempo o esfuerzo de implementación o quizá no exijan ni tiempo ni esfuerzo. Si forman parte de la implementación de búsqueda, puede implantarlos en paralelo con otras tareas de implementación. De esta forma, podrá ofrecer a los usuarios contenido importante en seguida.

En algunos casos, los elementos son fuentes de datos estructuradas que requieren de un análisis para saber cuál es la mejor forma de integrarlos en la experiencia de búsqueda (p. ej., plataformas de Business Intelligence). En la tabla siguiente se incluyen posibles elementos para la entrega avanzada.

Elemento de entrega	Tipo	Complejidad
Seguridad avanzada (ACLs, acceso a cookies, autenticación silenciosa)	Seguridad	Media-alta
Proveedor de SPI SAML personalizada	Seguridad	Media-alta
Diferentes nombres de usuario (según la fuente de contenido) para un mismo usuario	Seguridad	Alta
Aplicaciones de confianza	Seguridad	Alta
Sistemas de gestión de registros	Fuentes de contenido	Media-alta
Sistemas ERP (SAP, Oracle o PeopleSoft)	Fuentes de contenido	Media-alta
Sistemas CRM (p. ej., Siebel)	Fuentes de contenido	Media-alta
Almacén de datos/plataformas de Business Intelligence	Fuentes de contenido	Media-alta
Otra línea de sistemas empresariales	Fuentes de contenido	Media-alta
Alta disponibilidad para la indexación	Arquitectura	Alta
Alta disponibilidad para la publicación de contenido	Arquitectura	Media

## Duración de las fases

En general, las fases deberían durar de varios días a varias semanas. Aunque los esfuerzos pueden variar y requieren estimaciones concretas, la duración de una tarea depende de la complejidad, tal y como se ilustra en la tabla siguiente.

Complejidad	Duración
Baja	2 a 8 horas
Media	1 a 5 días
Alta	1 a 4 semanas

Las horas de esta tabla se ofrecen a modo de orientación únicamente y variarán en función del entorno y de los requisitos. Google recomienda que lleve a cabo un análisis para determinar el esfuerzo que requiere su implementación.

Además, reserve el tiempo necesario para adquirir el contenido. Intente incluir todo el contenido posible en el índice de las fuentes de contenido objetivo. No deberá esperar a tener todas las fuentes de contenido posibles en la solución de búsqueda, pero sí debería tener en el índice todo el contenido de los sistemas que va a incorporar en la versión en cuestión.

Es difícil predecir la rapidez con la que Google Search Appliance puede adquirir contenido, ya que depende de muchos factores, por ejemplo:

- Rendimiento de la red
- Rendimiento del servidor
- Carga del host
- Tipo de contenido
- Tamaño del contenido
- Funciones (p. ej., Vista previa del documento, Reconocimiento de entidades)

Google recomienda llevar a cabo algunas pruebas al principio del proyecto para poder calcular la velocidad de adquisición del contenido. Utilice esta información para planificarse.

## Capítulo 4 Definir los criterios de una implementación correcta

Antes de empezar con el proyecto, defina con claridad los criterios de una implementación correcta que acepte por su parte. Estos son algunos de los criterios típicos de una implementación de búsqueda correcta:

- Percepción de una nueva solución basada en GSA y la solución de búsqueda anterior (p. ej., pruebas A/B)
- Evaluaciones de relevancia realizadas por usuarios (p. ej., valoraciones de los usuarios)
- Pruebas de seguridad (la autenticación y la autorización funcionan para todos los sistemas protegidos)
- Amplitud del contenido (volumen del contenido en el índice., p. ej., 95% del contenido en un sistema)
- Amplitud de la implementación (porcentaje de usuarios activados)

Defina y establezca los criterios de una implementación de éxito junto con el propietario de la empresa y, para ser más realista, tenga en cuenta los comentarios de las anteriores soluciones de búsqueda. Asegúrese también de realizar análisis con periodicidad, por ejemplo, después de cada fase, para hacer los ajustes con la debida antelación.