

Google Search Appliance

Control de la implementación y modelos operativos

Marzo de 2014



© 2014 Google

Control de la implementación y modelos operativos

En este documento se explican varias estrategias relacionadas con el control de una solución 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.

Temas tratados	En esta guía se explican las prácticas recomendadas en cuanto a la publicación de contenido, estrategias de control de datos, procesos para la incorporación de contenido y modelos de propiedad de la solución
Público objetivo principal	Arquitectos de negocios, arquitectos de soluciones, arquitectos técnicos, administradores de GSA y administradores de sistemas de gestión de contenido empresarial que colaboren en el diseño de una solución de búsqueda empresarial con Google Search Appliance
Entorno de TI	GSA en un entorno de TI centralizado, descentralizado o de terceros
Fases de implementación	Planificación e implementación de GSA
Otros recursos	<ul style="list-style-type: none">• Toda la documentación Notas sobre GSA es pertinente para este manual en particular, ya que los temas tratados se explican con más detalle en otros documentos• Learngsa.com proporciona recursos formativos para GSA.• La documentación de producto de GSA proporciona información completa sobre GSA.• El Portal de asistencia de Google for Work proporciona acceso a la asistencia de Google.

Índice

[Acerca de este documento](#)

[Capítulo 1 Control de la publicación de contenido y prácticas recomendadas](#)

[Descripción general](#)

[Determinar la relevancia y el ranking](#)

[Creación y estructura del contenido](#)

[Opciones de ajuste de la relevancia](#)

[Consideraciones sobre la experiencia de usuario](#)

[Capítulo 2 Modelo de propiedad y contracargo de GSA](#)

[Descripción general](#)

[Propiedad de una organización de TI centralizada](#)

[Propiedad de una unidad empresarial/funcional](#)

[Propiedad de un tercero en un entorno alojado](#)

[Operaciones y asistencia continua](#)

[Control continuo de la búsqueda](#)

[Capítulo 3 Modelos de incorporación de contenido en GSA](#)

[Descripción general](#)

[Fuentes de contenido para la indexación](#)

[Modelo de incorporación de contenido mediante repositorio/plataforma o mediante aplicación empresarial](#)

[Proceso de incorporación y mantenimiento de contenido continuo](#)

[Gestionar datos confidenciales que ya se pueden buscar en GSA](#)

[Capítulo 4 Enfoque del entorno de GSA](#)

[Descripción general](#)

[Entorno de laboratorio/desarrollo](#)

[Entorno de pruebas/control de calidad \(opcional\)](#)

[Entorno de producción](#)

[Resumen](#)

Capítulo 1 Control de la publicación de contenido y prácticas recomendadas

Descripción general

Google Search Appliance proporciona algoritmos estándar, sin necesidad de configuración, que indexan y clasifican de forma eficaz documentos empresariales comunes. Los usuarios no tienen que crear documentos específicamente para los algoritmos de indexación y clasificación de GSA, ya que el dispositivo indexa eficazmente contenido no estructurado.

Todo es mejorable, por ejemplo, a través de indicadores adicionales de GSA que puedan mejorar el ranking. Estos indicadores podrían ser, por ejemplo, metadatos adecuados como el título, el autor, la fecha de publicación o contenido interesante en los documentos. Cuanto más diferenciadores sean los indicadores que GSA utilice para clasificar la relevancia del contenido, mejor.

Determinar la relevancia y el ranking

Google Search Appliance utiliza varios factores para clasificar los resultados de búsqueda, con lo que se ofrece una relevancia muy elevada sin necesidad de ajustar la configuración. Muchos de los algoritmos que utiliza el dispositivo de búsqueda se han optimizado para el contenido empresarial. De esta manera, a los usuarios les resultará más valioso el corpus existente con los documentos empresariales más usados, como contenido de documentos de procesador de texto, hojas de cálculo y presentaciones. La búsqueda será más eficiente.

Además, Google Search Appliance ofrece opciones de ajuste inteligentes e intuitivas, como asignación de preferencia a nodos, colecciones, fuentes, fechas y metadatos. Por otro lado, el sistema de ranking del dispositivo de búsqueda permite a los administradores usar registros de servidor y otra información específica de la empresa para ajustar el ranking de los resultados mediante programación para un número ilimitado de URLs. Gracias a los miles de millones de búsquedas que sirven para la investigación y a miles de ingenieros especializados en la tecnología de búsqueda, Google lidera el sector de las búsquedas con calidad e innovación.

Aunque no todas las Herramientas para webmasters de Google pueden usarse para la búsqueda empresarial y para el dispositivo GSA, la [ayuda sobre estas herramientas](#) incluye directrices eficaces sobre cómo un algoritmo de búsqueda de Google clasifica el contenido.

Creación y estructura del contenido

GSA indexa muchos tipos de documentos. Por lo tanto, debería tener en cuenta algunas directrices para preparar bien cada tipo de documento para su indexación.

En esta sección se incluyen las directrices para preparar los siguientes tipos de documentos para su indexación:

- [Documentos de Microsoft Office](#)
- [Documentos PDF](#)
- [Documentos HTML](#)

También se especifican las directrices para [feeds de documentos](#) y [clasificación/taxonomía de datos](#).

Estas directrices tratan principalmente sobre la indexación de metadatos adicionales y el contenido de los documentos. Los metadatos pueden mejorar el contenido del índice de GSA. Si desea ver un ejemplo del tipo de atributos de metadatos que se pueden añadir para mejorar la clasificación de los documentos, consulte el [Dublin Core](#) de elementos de metadatos, donde se proporcionan las normas para un conjunto de campos de texto que pueden describir un recurso.

Los propietarios de contenido pueden seguir algunos principios de creación de documentos que permitirán distinguir de manera eficaz y única sus documentos. Esta cuestión es importante, ya que los documentos empresariales tienen una estructura de URL menos definida, un factor que resulta útil para determinar las relaciones en Internet. Por este motivo son importantes los identificadores de documento único en un corpus de documentos empresariales.

Documentos de Microsoft Office

En las propiedades de un documento de Microsoft Office, rellene estos campos:

- Autor
- Título
- Comentarios

Si se tiende a crear un gran número de documentos a partir de plantillas, es muy importante cambiar las propiedades de documento predeterminadas.

Documentos PDF

En las propiedades de un documento PDF, rellene estos campos:

- Autor
- Título
- Creador

Si el documento no se ha creado directamente en una herramienta de edición de PDF, asegúrese de que estas propiedades se rellenen en el documento original antes de convertirlo a PDF.

Documentos HTML

En un documento HTML, incluya la información siguiente:

- <title> contenido
- descripción (manual o mediante programación): escriba descripciones de calidad, aunque no se muestren en la página

Diferencie las descripciones de las distintas páginas. Para evitar la indexación de cierto contenido de la página, utilice etiquetas googleoff.

Feeds de documentos

El protocolo de feeds de GSA permite la indexación de metadatos junto con el contenido subyacente. Incluya en el feed de documentos otros metadatos que no estén presentes en el contenido y que pueden mejorar el contenido indexado.

Clasificación/taxonomía de datos

Si los documentos que debe indexar GSA están asociados a un esquema y/o taxonomía de clasificación de datos empresariales, asegúrese de que el dispositivo indexe información ontológica. No es obligatorio incluir información ontológica junto con el contenido, pero puede mejorar la solución de búsqueda.

La taxonomía se puede usar junto con filtros de metadatos para restringir las búsquedas de documentos en determinados niveles de la jerarquía.

Según el tipo de documento y su fuente, es posible que para indexar información de taxonomía haya que aplicar un proceso de integración que permita asociar la clasificación del documento con el documento del índice. Habrá que llevar a cabo una investigación exhaustiva para ver cómo asociar esta información en el momento de indexación de GSA.

Reconocimiento de entidad

Se puede usar la función Reconocimiento de entidades de GSA para etiquetar automáticamente los documentos con metadatos en el momento de la indexación según las reglas predefinidas de los diccionarios de texto o a través de una estructura XML que permita coincidencias de términos o de expresiones regulares. Lo ideal para definir las entidades serían términos empresariales comunes o estándar de su organización. Si se ha identificado una taxonomía de empresa, también sería ideal para incluirla en las reglas de Reconocimiento de entidades. De esta forma, aunque algunos documentos no se etiquetaran con esos términos, GSA puede proporcionar la asociación en el momento de la indexación a través de estas reglas.

Opciones de ajuste de la relevancia

Google Search Appliance incluye funciones que permiten a un administrador de búsqueda ajustar la personalización de resultados para toda la organización o para determinados usuarios de la búsqueda. El ranking de los resultados de búsqueda se puede personalizar por fecha, metadatos, fuente, colección y fuente del feed.

Por ejemplo, si son varias las fuentes de contenido que envían feeds a GSA, entre ellas, SharePoint, y observa que los resultados de SharePoint no aparecen en las primeras posiciones de la clasificación, tal y como debería ser, puede personalizar la fuente de contenido SharePoint para mejorar el ranking del contenido.

Para obtener información detallada sobre cómo configurar una política de personalización de resultados, consulte la página de ayuda [Publicación > Personalización de resultados](#).

En la tabla siguiente se incluye una descripción general de las funciones que influyen en cómo se muestran los resultados a los usuarios.

Función	Comentarios
Otorgar preferencia a fuentes	A través de la concordancia de patrones, dé preferencia a una fuente respecto a otra.
Otorgar preferencia a fechas	Asigne más o menos importancia a la fecha de creación del documento.
Otorgar preferencia a metadatos	Priorice documentos que incluyan determinados metadatos.
KeyMatch	Aunque KeyMatch es, técnicamente, una función de sugerencia y no forma parte de los resultados de la búsqueda orgánica, puede usarla para ofrecer documentos para ciertas consultas.
Ampliación de consultas	Utilice una política de ampliación de consultas para ampliar los términos de la consulta de búsqueda a otros términos (sinónimos).
Marcador de puntuación inteligente	Cuando se habilita la función Informes de búsqueda avanzados, GSA utiliza Marcador de puntuación inteligente para analizar los datos de seguimiento de los clics y, con el tiempo, promocionar ciertos resultados de búsqueda. Por ejemplo, en una determinada consulta de búsqueda, si los usuarios siempre hacen clic en el segundo resultado en lugar del primero, al final el resultado pasará a ocupar la primera posición en la página.
Multiplicidad/filtrado de host	Habilite GSA para que excluya cualquiera de estas combinaciones: <ul style="list-style-type: none"> ● Resultados de la misma ruta ● Resultados con títulos y fragmentos duplicados
Estructura de ranking	Especifique una personalización por URL. Tenga en cuenta que se trata de una solución muy compleja de gestionar y solo debería probarse como última opción.
Palabras irrelevantes	Utilícelas para evitar que ciertos términos de la consulta se usen en la búsqueda. Use esta función con cuidado ya que puede tener muchas implicaciones si se utiliza como solución para un problema en concreto.
Colecciones	Distribuya el contenido en varias colecciones para restringir el corpus de documentos disponible para una consulta de búsqueda.

Consideraciones sobre la experiencia de usuario

Las modificaciones en la experiencia de búsqueda del usuario también pueden afectar a la manera en que estos perciben la implementación de búsqueda.

Si desea obtener algunas prácticas recomendadas sobre opciones relacionadas con la mejora de la experiencia de búsqueda de los usuarios, consulte la sección sobre [prácticas recomendadas para la experiencia de búsqueda](#).

Para obtener información detallada sobre las opciones relacionadas con la experiencia de usuario, consulte [Notas sobre GSA: guía de la experiencia de usuario](#).

Capítulo 2 Modelo de propiedad y contracargo de GSA

Descripción general

Hay tres modelos principales de propiedad de la solución de búsqueda asociados a las implementaciones de GSA:

- [Propiedad de una organización de TI centralizada](#)
- [Propiedad de una unidad empresarial/funcional](#)
- [Propiedad de un tercero en un entorno alojado](#)

Cada modelo tiene unas características concretas y puede tener ciertas ventajas o desventajas según la situación que presente la organización. A la hora de elegir el modelo que más se ajuste a su caso, considere lo siguiente:

- Modelo de financiación y cobro
- Definición del proceso
- Responsabilidad
- Métricas que se usan para informar sobre el uso y/o el valor de la solución
- Modelo de informes
- Acuerdos existentes de proveedores de terceros para alojamiento y/o servicios

Propiedad de una organización de TI centralizada

En las organizaciones con un modelo de TI centralizado, es habitual que el servicio compartido centralizado de TI sea propietario de una solución de búsqueda GSA. En este modelo, el grupo principal de recursos conoce el funcionamiento de GSA y las consecuencias en el negocio de las funciones y la incorporación de fuentes de contenido.

Este grupo también tiene una visión del contexto de toda la empresa y no específicamente de una unidad empresarial/funcional. El grupo es responsable de estas funciones:

- Identificar y planificar un plan por fases de la solución de búsqueda en toda la organización
- Tener en cuenta las peticiones de cada unidad empresarial y propietarios de contenido para indexar el contenido en GSA y mostrarlo a los usuarios de la búsqueda de toda la empresa

El hecho de que la solución pertenezca a un solo grupo de la empresa también debería simplificar la administración y la provisión del sistema de búsqueda.

Modelos de coste y de cobro

Hay varias formas de calcular cómo financiar y respaldar una solución de búsqueda:

- Proporcionar fondos de TI a la organización con el sistema de búsqueda de servicio compartido y no cobrar el uso de GSA a las unidades de negocios

- Recuperar el coste de implementación y respaldo de GSA a través del cobro a las unidades de negocios por el uso del dispositivo:
 1. Averigüe los costes base de funcionamiento de GSA, por ejemplo, el coste de licencia, el centro de datos, los bastidores o la asistencia técnica.
 2. Calcule el porcentaje de GSA que utiliza una unidad de negocios para saber cuánto habrá que cobrar multiplicando el porcentaje por el coste total del dispositivo. El uso puede ser un factor dependiendo del porcentaje de la licencia de documentos que se use o el uso y/o volumen de las consultas de búsqueda.
 3. Cobre por las tareas de configuración/personalización de GSA según el tiempo y los materiales empleados; deben quedar cubiertas todas las horas de trabajo.

En el caso de la personalización, sobre todo para cambios XSLT de la interfaz, quizá sería más eficaz aplicar un modelo híbrido, ya que la unidad empresarial podría tener una buena idea de cómo quieren que sea su experiencia de búsqueda.

Propiedad de una unidad empresarial/funcional

En las organizaciones donde el modelo de TI está descentralizado, es habitual que las unidades empresariales/funcionales administren la solución de búsqueda GSA según sus necesidades individuales. Dado que las unidades son las que mejor conocen su sistema, procesos y patrones de uso de los usuarios, están capacitadas para implantar una solución de búsqueda que sea de gran valor para sus usuarios. En cierta medida, puede resultar interesante que los usuarios finales decidan cómo va a ser la solución, ya que satisfacer sus necesidades es lo más importante.

Si la solución de búsqueda es propiedad de una unidad empresarial, es posible que se usurpen los procesos de control de TI comunes y salir al mercado más rápidamente. Lo ideal sería que la implementación de GSA por parte de una unidad empresarial se convierta en un modelo de aplicación para toda la empresa.

Aunque las unidades empresariales/funcionales no sigan los procesos de TI estándar, es importante tener en cuenta las consideraciones siguientes para poder medir el éxito de la implementación de búsqueda:

- Financiación
- Definición de algunos procesos básicos
- Modelo de asistencia continua
- Métricas que se usan para informar sobre el uso y/o el valor de la solución
- Modelo de informes

Aunque es posible que el sistema de búsqueda no se implante según las prácticas estándar de TI, es muy importante disponer de datos y de un punto de referencia para aplicar mejoras en el sistema de búsqueda más adelante.

Según la estructura de TI de su organización, este modelo podría implicar que la implementación de GSA no cumpla con todas las políticas de control de TI en áreas tales como la auditoría, el branding, la plataforma o la arquitectura. Un sistema de búsqueda implantado en un modelo de estas características podría suponer una restricción implícita para abarcar repositorios controlados fuera de la unidad empresarial. Este modelo es ideal para aplicaciones de búsqueda con un objetivo especial que se proporcionan a un conjunto específico de usuarios que puede no ser representativo de toda la empresa.

Modelo de costes

La implementación de GSA por parte de una unidad de negocios se puede financiar calculando el coste de proporcionar GSA y el esfuerzo que llevaría crear la solución. También habría que desarrollar un modelo de asistencia.

Aunque la solución de búsqueda pertenezca a una unidad empresarial y esta la implemente, quizá sería más eficaz documentar ciertos procesos y externalizar la asistencia continua a un tercero de acuerdo con los sistemas de externalización que ya utilice la empresa (ya sea de forma interna o a un tercero).

Propiedad de un tercero en un entorno alojado

A una empresa quizá le interese externalizar por completo la solución de búsqueda GSA a un proveedor de alojamiento de terceros. Este modelo puede definirse a partir de estos dos enfoques:

- Definir y desarrollar la solución de forma interna y transferir las operaciones a un tercero
- Recibir los servicios de un tercero para definir y desarrollar la solución y alojarla, o bien elegir entre las ofertas de alojamiento para el sistema de búsqueda GSA de un tercero

Un acuerdo de este tipo con un tercero debería regirse por las directrices y precedentes de suministro del proveedor.

Este modelo presenta ventajas porque un proveedor de terceros que esté especializado en la solución de búsqueda GSA ya contará con los recursos y la infraestructura necesarios para desarrollar y abastecer el servicio. En estos casos, además, los costes siempre son menores. Una desventaja podría ser la falta de flexibilidad y de facilidad para integrar completamente el servicio con los sistemas existentes. Este modelo podría no encajar en entornos de alta seguridad para los que el almacenamiento de datos confidenciales en las propias instalaciones exige ciertos requisitos.

Modelo de costes

Analice los modelos de costes que supone la propiedad por parte de una unidad empresarial y de una organización de TI centralizada y evalúelos frente al coste que supone alojar la solución con un tercero. Dado que este modelo se rige por los términos y condiciones del acuerdo del proveedor, habrá que negociar y definir de antemano los aspectos correspondientes.

Una cuestión importante en este acuerdo será el modelo de asistencia y los acuerdos de nivel de servicio correspondientes. En función del acuerdo, este modelo podría resultar más interesante en cuanto a la inversión inicial, pero luego ofrecería menos flexibilidad para la solución.

Operaciones y asistencia continua

Tal y como se ha mencionado en las secciones anteriores, la solución de búsqueda GSA necesita cierto grado de mantenimiento.

En primer lugar y más importante, se proporciona ayuda para cuestiones sobre asistencia empresarial de Google for Work durante la vigencia de la licencia de GSA. La empresa debe contar con una persona o con un grupo de personas que conozca y se responsabilice del sistema de búsqueda. Esta persona o grupo debería:

- Conocer la arquitectura de búsqueda de GSA, según se haya implantado en la organización
- Conocer las funciones de GSA y cómo están configuradas para los administradores de GSA (a través de la Consola del administrador o de las API de administración)
- Observar los procesos/estándares que se establezcan para incorporar contenido nuevo en GSA
- Coordinar y planificar las interfaces o las mejoras para la experiencia de usuario que se propongan
- Crear informes de la búsqueda GSA o proporcionar asistencia a quien necesite un informe de este tipo
- Estar al día de las novedades de GSA y planificar las actualizaciones de versión del dispositivo (anualmente, por ejemplo)
- Ser el contacto para cualquier problema relacionado con GSA y abrir incidencias con el equipo de asistencia de Google cuando sea necesario

Control continuo de la búsqueda

Independientemente del tipo de modelo de propiedad del sistema de búsqueda, es importante que un equipo especializado responsable del área de búsqueda se reúna periódicamente para evaluar la eficacia de la búsqueda y revisar las prioridades más destacadas. Por ejemplo, un grupo formado por patrocinadores empresariales de GSA, propietarios del sistema, administradores y usuarios finales podrían reunirse trimestralmente (o más a menudo si la búsqueda es una cuestión prioritaria para el negocio) para revisar los comentarios de los usuarios y las estadísticas de búsqueda, y priorizar actividades de mejora

o nuevas áreas que haya que incorporar en la solución de búsqueda.

Los análisis y las métricas y cómo contribuyen a perfeccionar la solución son otros aspectos importantes del control continuo de la solución. La búsqueda puede ser un indicador empresarial importante para una empresa, donde la información se obtiene a partir de los análisis de búsqueda y de las métricas.

La integración con un servicio de análisis, como Google Analytics, o una solución personalizada puede ofrecer información de este tipo: opinión del usuario, puntos débiles, temas que interesan, orientación sobre el desarrollo de productos, etc. Si cuenta con una estrategia para recabar y analizar métricas de búsqueda, la solución de búsqueda será más valiosa en áreas donde aparentemente no lo es.

Capítulo 3 Modelos de incorporación de contenido en GSA

Descripción general

En este capítulo se tratan las estrategias y enfoques para incorporar fuentes de contenido de su empresa en GSA.

Aunque en muchos casos GSA debería descubrir, rastrear e indexar contenido sin mucho esfuerzo, es recomendable planificar una estrategia para iniciar la integración de las distintas fuentes de contenido y/o aplicaciones. Para que la integración se lleve a cabo con éxito, hay que tener en cuenta elementos tales como la seguridad, el tipo de integración, los metadatos disponibles, la definición de entidades para la función Reconocimiento de entidades y el impacto en la interfaz de usuario.

Fuentes de contenido para la indexación

Cuando deba decidir qué plataformas incorporar en el dispositivo GSA, considere los elementos de la tabla siguiente y priorice las plataformas en función de estos.

Consideración	Comentarios
Usuarios y tipos de usuarios a quienes afecta la integración del contenido	Dedíquese primero a las plataformas que tengan mayor impacto.
Facilidad de integración	Por orden de mayor dificultad para la integración: rastreo directo, conector disponible, rastreo a través del servidor proxy, necesidades que haya que desarrollar para los feed y para los conectores.
Mecanismos de autenticación y autorización de seguridad necesarios	<p>La integración más fácil es la de los sistemas que se integran con mecanismos de seguridad que se proporcionan directamente en GSA.</p> <p>Los sistemas que requieren la integración de la autenticación personalizada son más complicados de integrar.</p>
Personalización de la interfaz de búsqueda para proporcionar el máximo valor a la experiencia de búsqueda de la plataforma	<p>Hay veces que para proporcionar el máximo valor de la integración de contenido con GSA hay que modificar la interfaz para mostrar elementos tales como:</p> <ul style="list-style-type: none"> ● filtros ● categorías ● metadatos en los resultados de búsqueda ● procesos de navegación personalizada <p>Por orden de mayor dificultad de la integración: XSLT predeterminado, modificaciones menores de XSLT, modificaciones importantes de XSLT, análisis de aplicaciones personalizadas y presentación de XML que proporciona GSA.</p>
Vencimiento de los metadatos disponibles en la fuente de contenido	<p>Aunque los metadatos no son necesarios para la indexación, pueden mejorar el contenido del índice y ofrecer más opciones para crear una experiencia de búsqueda eficaz.</p> <p>Por su naturaleza, algunas fuentes de contenido incluyen metadatos sin necesidad de configuraciones, mientras que otras fuentes deben seguir ciertos procesos en el momento de mostrar contenido.</p> <p>Si lo desea, considere la posibilidad también de ampliar los documentos con metadatos durante la indexación mediante un programa y un proceso personalizado.</p>
Ampliación de metadatos en los documentos a través de la función Reconocimiento de entidades	<p>Puede usar la función Reconocimiento de entidades de GSA para mejorar el contenido con entidades extraídas mediante la definición de reglas de entidades basadas en diccionarios o en expresiones regulares.</p> <p>Una vez definidas estas reglas los documentos se etiquetarán con metadatos durante la indexación. Se puede usar esta función para asignar metadatos a los documentos que, de no ser por esta función, no se etiquetarían con metadatos.</p>

Modelo de incorporación de contenido mediante repositorio/plataforma o mediante aplicación empresarial

A la hora de seleccionar contenido para incorporarlo a GSA, hay que tener en cuenta dos sistemas de primer nivel:

- [Incorporación de contenido a través de una plataforma](#)
- [Incorporación de contenido a través de una aplicación empresarial](#)

Según estos dos modelos, hay que tener en cuenta varias consideraciones. Algunas de ellas se tratan en los apartados siguientes.

Incorporación de contenido a través de una plataforma

En este caso, se indexará la mayoría o todos los documentos del repositorio de contenido, independientemente de las aplicaciones que active. Este modelo puede ser válido para una organización donde la búsqueda se financie de forma centralizada y el propietario sea un equipo especializado en el área de búsqueda, también centralizado.

En este caso, la incorporación del contenido se hará para cada repositorio o plataforma cada vez, independientemente de las aplicaciones o unidades de negocio que tengan contenido en la plataforma. Este enfoque también puede facilitar la implantación del concepto "búsqueda universal" en la empresa, ya que se considerará una solución global y no específica de una aplicación.

Un ejemplo de este enfoque es la indexación de contenido de Livelink ECM. En este caso, toda la plataforma Livelink se indexaría en GSA en el transcurso de un solo proyecto de implementación, independientemente de qué unidades o aplicaciones empresariales almacenan documentos en el sistema.

Debería dar prioridad a las fuentes de contenido que esté considerando para la integración. Para ello, piense detenidamente en los factores que se enumeran en el apartado [Fuentes de contenido para la indexación](#).

Incorporación de contenido a través de una aplicación empresarial

Con este modelo, se indexa la mayoría o todos los documentos de una determinada aplicación empresarial, independientemente de qué plataformas son la fuente del contenido. Integrar contenido en GSA de determinadas aplicaciones empresariales puede ser válido para una empresa en la que la búsqueda se gestiona de forma centralizada, pero luego hay que cobrar el servicio prestado a las unidades de negocio correspondientes.

Es también un modelo válido si el propietario de la implementación de búsqueda GSA es una unidad funcional de la empresa y no un equipo de servicios compartidos. Con este sistema, el objetivo y la prioridad de indexación en GSA serían aplicaciones empresariales concretas.

Un ejemplo de ello es la indexación de contenido procedente de un sistema de gestión de la aplicación de contratación de candidatos de Recursos humanos, cuyas aplicaciones se almacenan en un sistema de gestión de contenido empresarial. Si la incorporación de contenido se realiza a partir de una aplicación y no de una plataforma, el objetivo sería indexar únicamente estas aplicaciones de contratación y no todo el sistema de gestión de contenido.

Habría que dar prioridad a la indexación de los candidatos de la aplicación, similar a las plataformas, a partir de los factores que se enumeran en el apartado [Fuentes de contenido para la indexación](#).

Uno de los principales objetivos de indexar el contenido de una aplicación en GSA debería ser aumentar la productividad del usuario con dicha aplicación. Cuando piense en casos prácticos o el proceso de trabajo del usuario final, determine en qué áreas podría simplificar el proceso GSA.

Por ejemplo, en una aplicación que permite editar documentos, además de mostrar el documento en los resultados de búsqueda, incluya un enlace al modo de edición. De esta forma, si el usuario quiere modificar el contenido, se le puede dirigir directamente y se salta un paso intermedio. Cualquier tipo de optimización o de simplificación que GSA aporte a los procesos del usuario será una excelente ventaja.

Proceso de incorporación y mantenimiento de contenido continuo

La solución de búsqueda GSA evoluciona al ritmo de su empresa y a medida que se crea más contenido. La organización debe estar preparada para hacer frente a peticiones para incorporar contenido en GSA después de las primeras fases de implementación. Aprenderá mucho con la primera implementación y le servirá de muestra para ampliar o ajustar la solución en fases posteriores.

No se requiere de un proceso elaborado, pero se recomienda disponer de una plantilla con especificaciones de búsqueda que sirva para orientar el diseño y la implementación de la integración de contenido. Este documento de especificaciones debería incluir información operativa y técnica sobre la integración. Sin un documento de orientación, la implementación puede resultar imposible de gestionar, o se pueden perder los requisitos y no implementarse.

Google recomienda un proceso con las fases siguientes:

1. [Análisis inicial del contenido](#)
2. [Desarrollo/configuración](#)
3. [Pruebas e implantación](#)

El proceso recomendado no tiene en cuenta tareas relacionadas con recursos o gestión de proyecto, consideraciones importantes al emprender una implementación de tal magnitud.

Análisis inicial del contenido

La investigación previa a la incorporación de contenido en GSA debería estar motivada por una petición o por la identificación de un área con contenido que no se encuentra en GSA y que, por tanto, no se muestra en los resultados de búsqueda.

En esta fase realice el análisis siguiente:

1. Analice los factores de incorporación de contenido que se enumeran en el apartado [Fuentes de contenido para la indexación](#).
2. Documente los requisitos e indique las especificaciones de búsqueda para el sistema en cuestión.
 - a. Piense en los requisitos que pueden contribuir a la experiencia de usuario.
 - b. Averigüe cómo se integrará el contenido adquirido y las especificaciones de la interfaz con el contenido y la interfaz que ya existen.
3. Analice qué metadatos están disponibles en el sistema identificado y cómo pueden afectar a la solución de búsqueda o mejorarla.
 - a. Piense en qué procesos se necesitan para eliminar metadatos o ampliar el contenido con metadatos, si lo desea.
 - b. Defina entidades en la función Reconocimiento de entidades para aumentar metadatos que, de lo contrario, no estarían disponibles.
4. Identifique los puntos de contacto que harán las veces de PYME en adelante.

Desarrollo/configuración

Siga estos pasos en la fase de desarrollo/configuración:

1. Integre el contenido: incluye la implementación de componentes necesarios externos a GSA, como conectores o SAML Bridge.
2. Desarrolle los componentes necesarios para la interfaz de usuario.

Pruebas e implantación

En la fase de pruebas y de implantación, siga estos pasos:

1. Realice una prueba de sistemas/integración: ¿el contenido se indexa y se muestra correctamente?
2. Realice una prueba de relevancia: ¿cómo se integra el contenido adquirido con el que ya existía en la colección predeterminada o universal, que incluye contenido de muchas fuentes?
3. Realice una prueba o análisis de rendimiento/carga: ¿cómo va a afectar al rendimiento de la solución existente la carga adicional del usuario de búsqueda?

Gestionar datos confidenciales que ya se pueden buscar en GSA

GSA es una herramienta de búsqueda eficaz que permite a los usuarios encontrar documentos a los que, hasta el momento, no podían acceder. Hay casos en los que, poco después de la implementación, los usuarios empiezan a ver documentos en los resultados de búsqueda a los que no deberían tener acceso. Son los casos en los que si bien GSA respeta los permisos de los documentos, estos son demasiado amplios en el sistema fuente. Para que no aparezcan en los resultados de búsqueda, debe restringirse el permiso de estos documentos.

Al implementar GSA se recomienda aplicar un plan para mitigar las brechas en los datos confidenciales. Esto permitirá la planificación y mitigación a través del desarrollo de procesos, comunicación y formación de usuarios en la empresa.

A menudo el contenido de Recursos humanos es el que más riesgos supone en este tipo de situaciones. Inspeccione los permisos a través de una simple auditoría de una muestra de datos de Recursos humanos. Excluya el contenido confidencial de la primera fase de implementación para asegurarse de que los usuarios y administradores se familiaricen con GSA. Esto permite a los administradores prepararse mejor ante las fugas de datos inintencionadas debido a permisos no restrictivos de los documentos.

A continuación se describe un proceso de ejemplo para hacer frente a incidencias después de que GSA haya mostrado resultados confidenciales; son documentos que deberían tener permisos restringidos:

- Configure un alias de correo electrónico que gestione los casos de los que informen los usuarios donde el permiso de los documentos debería estar restringido.
- Cree la función de analista de negocios para que gestione los casos que vayan sucediéndose.
 - Esta persona debería poder asociar el documento en cuestión a la fuente de contenido en la que reside.
 - Esta persona debería poder comunicarse con eficacia con los propietarios de la fuente de contenido para establecer el tipo de permiso adecuado para el documento en cuestión.
- Antes de la puesta en marcha definitiva habrá que explicar con claridad el proceso de gestionar aquellos documentos cuyo permiso debe restringirse.
 - Hay que hacer llegar a los usuarios el alias de correo electrónico al cual podrán enviar las incidencias.
 - Hay que proporcionar a los usuarios información de ejemplo que deberán indicar cuando informen de una incidencia.
- El analista de negocios deberá determinar la gravedad de los informes que se presenten, con la posibilidad de eliminar al instante el documento de la interfaz de GSA, ya sea de forma temporal o definitiva. El documento podría eliminarse temporalmente hasta que los propietarios del contenido confirmen qué nivel de acceso debe tener.
- El seguimiento de las incidencias puede hacerse en el sistema que se prefiera.

Capítulo 4 Enfoque del entorno de GSA

Descripción general

Los cambios deberían probarse en un entorno aparte antes de incluirlos definitivamente en la solución de búsqueda. Como sucede con cualquier tipo de servidor o aplicación, un pequeño cambio en la configuración podría traer consecuencias inintencionadas. Por lo tanto, se recomienda aplicar una estrategia de pruebas adecuada y recurrir a un entorno temporal para los cambios que se quieran introducir en un dispositivo de búsqueda.

Se recomienda usar los entornos siguientes para desarrollar, probar, representar e implantar correctamente la solución para producción:

- [Entorno de laboratorio/desarrollo](#)
- [Entorno de pruebas/control de calidad](#) (opcional)
- [Entorno de producción](#)

Aunque es lo ideal, no es necesario trabajar con entornos físicos. Por ejemplo los entornos de laboratorio/desarrollo y pruebas/control de calidad se pueden crear en el propio dispositivo GSA con una combinación de varias interfaces y colecciones.

Entorno de laboratorio/desarrollo

Un entorno de desarrollo para Google Search Appliance simplemente significa replicar el entorno de producción para disponer de un área aparte para desarrollar e integrar cambios de configuración y mejoras. El entorno de desarrollo debería incluir acceso a los mismos tipos y fuentes de contenido que el entorno de producción, pero con un conjunto de documentos restringido o reducido. Todos los componentes de integración que se van a implantar, como SAML Bridge o los conectores, deberían contar con las mismas instancias de desarrollo.

El tamaño del entorno de desarrollo no debe ser como el de producción. Se puede reducir a escala, ya que no es necesario que tenga en cuenta el tamaño de los documentos ni los requisitos de rendimiento del documento de un sistema de producción.

Una configuración común de desarrollo incluye un dispositivo de búsqueda que no está diseñado para producción y que no muestra resultados a la mayoría de los usuarios finales. Todos los cambios de configuración, actualizaciones y mejoras se prueban en este dispositivo de búsqueda y, cuando está todo preparado, se pasa al siguiente entorno.

Siempre que sea posible, evite usar nomenclaturas codificadas que pudieran complicar la migración de las configuraciones. Por ejemplo, colección_des o interfaz_prueba deberían renombrarse antes de pasar al entorno de producción.

Las API de administración de Google Search Appliance, que se introdujeron en la versión 6.0 del software, permiten exportar e importar configuraciones muy específicas (p. ej., concordancias concretas de KeyMatch) o configuraciones más generales (toda la configuración de un dispositivo de búsqueda). Si desea obtener información sobre las API de administración, consulte las guías para desarrolladores de APIs de administración, enlazadas a la página de documentación pública de [GSA](#).

El entorno de desarrollo también se puede usar como entorno de prueba antes de poner en marcha el entorno de producción de manera definitiva.

Entorno de pruebas/control de calidad (opcional)

Este entorno es parecido al de desarrollo, ya que se trata de un entorno replicado no diseñado para producción. Este entorno debería ser estable para los testers que accedan al sistema. La publicación de componentes de la implementación y los cambios en la configuración o en la interfaz deberían hacerse por fases según un proceso de gestión de la configuración. Cualquier cambio que se introduzca en este entorno que pudiera afectar a las pruebas debería comunicarse al equipo de pruebas.

Puesto que GSA es un producto que indexa y no modifica las fuentes de contenido, el entorno de control de calidad de GSA puede apuntar a fuentes de contenido de control de calidad o de producción. Las consideraciones son distintas según el caso.

Se pueden llevar a cabo las pruebas de rendimiento en este entorno para desarrollar un punto de referencia, teniendo en cuenta que el entorno de producción será mayor en cuanto a las especificaciones de hardware.

Entorno de producción

El entorno de producción es un entorno de usuario operativo que debería adaptarse a los requisitos de tamaño del índice y de rendimiento.

Los cambios en el entorno de producción deberían introducirse bajo un estricto proceso de gestión de la configuración. Todos los cambios que se migren a producción deberían probarse y aprobarse en el entorno de control de calidad. Para la migración a producción, hay que tener en cuenta posibles interrupciones en el funcionamiento. Los cambios deberán hacerse en el momento oportuno, con el menor impacto posible para el negocio.

Entorno de producción con copia de seguridad/recuperación tras desastres

En la mayoría de los casos, especialmente cuando la búsqueda es una aplicación clave para los usuarios, es importante disponer de un entorno de copia de seguridad/recuperación tras desastres en el entorno de producción. Este entorno es una réplica del entorno de producción y se pueda usar en caso de interrupción planificada o no planificada. Suele ser una copia exacta del entorno de producción, con una infraestructura y una capacidad similares.

Consideraciones para planificar la capacidad del entorno de producción

Los entornos de producción que se replican no siempre se utilizan solo para funciones de copia de seguridad/recuperación tras desastres. También pueden servir para planificar la capacidad. Si un dispositivo GSA no puede hacer frente a toda la carga de búsqueda necesaria para mostrar el contenido, se puede usar una réplica de GSA en una configuración de carga distribuida para planificar la capacidad. De esa forma, la carga de publicación se distribuirá entre los distintos GSA, y se mostrará el mismo contenido.

Es importante tener en cuenta que, debido a la función de duplicación de GSA, solo un dispositivo GSA suele adquirir contenido nuevo. Los cambios se publicarán y se sincronizarán en los GSA de réplica de la configuración.

Resumen

El control no es una cuestión universal y varía según la empresa. Los apartados de este documento presentan algunos conceptos que podrá adoptar para controlar la nueva solución de búsqueda GSA. Tenga en cuenta que estos conceptos pueden integrarse con los procesos que ya existen de varias formas.

Como siempre, igual que sucede con la arquitectura de GSA, los procesos de control deben ser lo más simples posible. Así, las implementaciones se podrán hacer de manera rápida y repetirse. Pruebe a usar GSA para usurpar algunos procesos de TI existentes que le permitirán implementar y repetir rápidamente el sistema de búsqueda creado en el dispositivo GSA.