

Portal del Estado Uruguayo

Guía de uso buscador.gub.uy en sitios del Estado

Introducción

buscador.gub.uy es un buscador que provee búsquedas especializadas en los contenidos del Estado uruguayo disponibles en Internet y además tiene la capacidad de generar resultados sobre un único sitio o una familia de sitios. Está implementado con el motor de búsquedas GSA (Google Search Appliance).

El presente documento intenta brindar la información necesaria para llevar adelante el proceso de incorporar buscador.gub.uy como motor de búsquedas personalizadas en un sitio. Está dirigido a los equipos de Tecnología y encargados de los sitios Web de los Organismos del Estado Uruguayo.

Conceptos Básicos

buscador.gub.uy brinda servicios a través de 3 procesos:

- **Rastreador:** Es el proceso que visita los sitios periódicamente para leer el contenido y luego procesarlo e ingresarlo en un índice.
- **Buscador:** Es el proceso que a partir de una clave de búsqueda (una palabra, número, oración), devuelve un conjunto de resultados relevantes en formato XML. Para elegir y priorizar los resultados el Buscador utiliza el índice que confecciona el Rastreador
- **Interfaz:** Es el proceso que a partir de los resultados de una búsqueda, que el Buscador entrega en XML, genera el código HTML que se desplegará en la pantalla. buscador.gub.uy puede implementar múltiples interfaces.

Funcionamiento

El rastreador de buscador.gub.uy funciona en forma constante, recorriendo todos los sitios .gub.uy y los sitios de otras terminaciones de dominio identificados como estatales. Como todos los rastreadores, descubre páginas nuevas a partir de los vínculos que encuentra en las páginas ya rastreadas.

Todas las páginas rastreadas son evaluadas y si tienen contenido válido se incorporan al índice, que actualmente reúne aproximadamente 500.000 páginas distintas del Estado uruguayo en Internet.

Para saber si un sitio fue rastreado e incorporado al índice existen 2 opciones

- **Revisar los logs del servidor:** el rastreador se identifica (user agent) como “agesic-crawler”
- **Buscarlo en buscador.gub.uy** utilizando el término de búsqueda especial *site*. Por ejemplo, para saber si en el índice se incluyeron páginas del dominio *miorganismo.gub.uy* se debe buscar

site:miorganismo.gub.uy

sin espacios. Las páginas de resultados que aparezcan son las que fueron rastreadas e indexadas. Si no arroja resultados, es porque no está siendo rastreado.

Si usted conoce algún dominio que debería ser rastreado e indexado y aún no lo está, puede solicitar que se incluya simplemente enviando un email a contenidosPEU@agesic.gub.uy

Utilizando buscador.gub.uy en mi sitio

Una de las capacidades destacadas de buscador.gub.uy es poder brindar servicios de búsqueda para cualquier sitio del Estado uruguayo. La clave para poder hacerlo es la posibilidad de crear colecciones.

Colecciones

Una **colección** es un subconjunto del índice, definido en base a una serie de reglas que indican qué URLs incluir o excluir de la colección.

Por ejemplo, se podría crear una colección que cumpla las siguientes reglas:

- Incluir todas las URLs del tipo *miorganismo.gub.uy/...*
- Excluir las URLs del tipo *miorganismo.gub.uy/testing...*

De esta forma, quedaría creada una colección que incluiría todas las páginas del organismo, pero con la prevención de no indexar aquellas destinadas al testing.

Una colección puede incluir una cantidad ilimitada de reglas y éstas pueden llegar a ser muy complejas, ya que se implementan con expresiones regulares¹. Las reglas no tienen que restringirse a un único dominio, sino que pueden abarcar cualquier dominio o URL que esté incluida en el índice.

No hay un límite en la cantidad de colecciones que se puedan crear y el proceso de búsqueda está optimizado de tal forma que la restricción de una búsqueda a una colección prácticamente no tiene costo de procesamiento.

Toda búsqueda está obligatoriamente asociada a una colección. Las búsquedas en la interfaz de buscador.gub.uy están asociadas a la colección default, que abarca todo el índice.

Interfaces

Tal como se indicó más arriba, la interfaz es el proceso que a partir de los resultados de búsqueda en XML genera las transformaciones necesarias para generar el HTML que se desplegará en la pantalla. Esto se puede hacer de dos formas:

- **En buscador.gub.uy:** es el servidor de buscador.gub.uy quien toma el XML y lo procesa emitiendo como salida HTML para desplegar en la pantalla.
- **En el propio sitio:** es responsabilidad del sitio que está utilizando buscador.gub.uy como servicio de búsqueda tomar el XML y formatearlo.

Interfaz en buscador.gub.uy

buscador.gub.uy implementa dos interfaces. La primera es la default, que está integrada en su diseño y en su interacción al Portal del Estado Uruguayo.

La segunda es la llamada **Interfaz mínima**. Está pensada para ser integrada a cualquier sitio utilizando alguna herramienta sencilla, como por ejemplo un iframe.

La Interfaz mínima incluye en el HTML de salida todos los elementos propios del buscador, de modo que el sitio disponga de un buscador 100% funcional, pero no tiene ningún

¹ Una expresión regular es una regla que, con comodines, mecanismos de sustitución, de repetición y otras herramientas, permite representar de una forma compacta un conjunto muy grande de cadenas de caracteres distintos.

elemento de imagen, de modo que el sitio pueda incorporar cabezal, menú y cualquier otro componente que crea conveniente.



democracia

Resultados de Búsqueda 1 a 10 de aproximadamente 1500 resultados para democracia.

1 2 3 4 5 6 7 8 9 10 [Siguiente >>](#)

[Cumbre Iberoamericana dio importante paso en la defensa de las ...](#)
... la importancia de la cláusula democrática porque defiende las **democracias** en América. ...
referidos a la iniciativa de defensa de la **democracia**, así como otros ...
www.presidencia.gub.uy/sci/noticias/2010/12/2010120407.htm - [En caché](#)

[Porque aún resuena el eco grave de su voz que exclama: "¡Viva ...](#)
... Viva la libertad! ¡Viva la republica! ¡Viva la **democracia!**". El actor Alberto
Candeau contaba con un buen prestigio a los 37 años. ...
www.presidencia.gub.uy/sci/noticias/2010/09/2010092303.htm - [En caché](#)
[[Más resultados de www.presidencia.gub.uy/sci/noticias/2010](#)]

[Conmemoración de los 25 años del retorno a la democracia](#)
Imprimir. 19 de julio, 2010. Conmemoración de los 25 años del retorno a la **democracia**. ...
www.presidencia.gub.uy/sci/fotos/2010/07/2010071903.htm - [En caché](#)

[Lanzamiento del proyecto "Alternativa Solidaria: Promoción del ...](#)

Ejemplo de resultados utilizando la interfaz mínima

Interfaz en el sitio

La forma más potente y flexible de implementar la interfaz es tomando el XML del resultado de la búsqueda en el propio sitio y generar localmente el HTML.

Además de que el XML es muy fácil de entender, está completamente documentado y provee valiosa información que habitualmente es invisible a los ojos del usuario, como por ejemplo los metadatos de la página, o los filtros que fueron aplicados a la búsqueda.

Al final del presente documento se encuentran las referencias y direcciones de la documentación disponible.

Cajas de búsqueda

Para promover una imagen uniforme de buscador.gub.uy en todos los sitios del Estado uruguayo, el equipo de trabajo del Portal del Estado creó un conjunto de cajas de búsqueda para insertar en los sitios.

La intención es que un sitio que utiliza buscador.gub.uy cumpla, dentro de sus posibilidades, con dos reglas básicas:

- **Utilice una caja de búsqueda de las disponibles**, de modo de ayudar a la homogeneidad en el uso de buscador.gub.uy en beneficio de los ciudadanos: lo que ven en un sitio del Estado les servirá para el próximo.
- **Permita buscar en su sitio o en todo el estado**. Las cajas de búsqueda proveen una cantidad de opciones con dos botones, uno para buscar en el sitio y otro en todo el estado uruguayo. El default es siempre en el propio sitio.

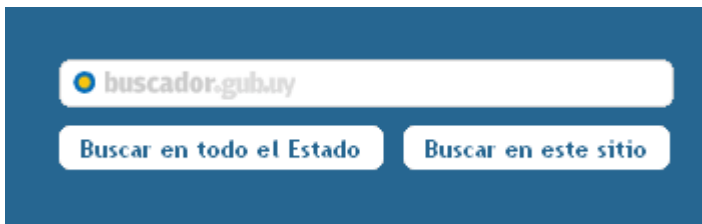
Se muestran a continuación algunos ejemplos. La lista completa de cajas está disponible en <http://peu.agesic.gub.uy/buscador>, donde además cada caja cuenta con el código HTML y la hoja de estilo para insertarla en su sitio.



Dos botones, sin cuadro



Dos botones, con cuadro



Dos botones, aplicado en negativo



Un botón, con cuadro



Un botón, sin cuadro



Un botón, aplicado en negativo

Implementando buscador.gub.uy en mi sitio

La implementación de buscador.gub.uy en el sitio de un organismo del Estado uruguayo es relativamente sencilla y cuenta con cuatro pasos:

1. Contacto

A través de un correo electrónico a contenidosPEU@agesic.gub.uy el organismo manifiesta su interés de utilizar buscador.gub.uy y describe brevemente las necesidades a cubrir.

AGESIC evalúa el requerimiento y contesta rápidamente acerca de la viabilidad y la disponibilidad de recursos para implementarla.

2. Creación de la colección en testing

Se interactúa con el organismo para definir qué dominios y patrones de URL incluir y cuáles excluir. De todos modos en el futuro estas definiciones se pueden modificar.

AGESIC crea en un servidor GSA de testing una colección como la solicitada, y proporciona una interfaz mínima de pruebas.

3. Implementación en el sitio

El organismo implementa la búsqueda en el sitio, ya sea utilizando la interfaz mínima o procesando el resultado de la búsqueda en XML. Realiza además todas las pruebas y validaciones que crea conveniente, ya que dispone de un servidor GSA de testing que tiene el 100% de la funcionalidad del de producción.

El equipo de AGESIC cuenta con una amplia experiencia en este tipo de implementaciones, por lo que el organismo contará con ayuda calificada en el caso de que sea necesario. En nuestra experiencia, las implementaciones han sido relativamente sencillas.

4. Puesta en producción

Se coordina con el organismo la puesta en producción. Desde el punto de vista de AGESIC, implica la implementación en el GSA de producción de la colección y otros elementos que se hubieran implementado en testing.

Desde el punto de vista del organismo, implica todas las tareas de puesta en producción que haya planificado, adicionando la redirección de las búsquedas desde el GSA de testing al de producción.

Información adicional

La información disponible sobre GSA es completa y de alta calidad, con el único defecto de que solo se encuentra disponible en idioma Inglés. Si la información no le resulta lo suficientemente clara o requiere información adicional, no dude en contactarlos a través de contenidosPEU@agesic.gub.uy.

La referencia completa del protocolo de búsqueda de GSA está en la dirección:

http://code.google.com/apis/searchappliance/documentation/64/xml_reference.html

La documentación general de GSA se puede encontrar en la dirección:

<http://code.google.com/apis/searchappliance/documentation/64/index.html>

El foro de discusión sobre GSA se encuentra en:

<http://groups.google.com/group/google-search-appliance?hl=en&pli=1>

Los ejemplos de cajas de búsqueda para buscador.gub.uy se encuentran en:

<http://peu.agesic.gub.uy/buscador>

La interfaz mínima en producción se encuentra en:

http://buscador.gub.uy/search?q=tramites&spell=1&client=version1&proxystylesheet=minima&output=xml_no_dtd&sort=date%3AD%3AL%3Ad1&entqr=3&oe=UTF-8&ie=UTF-8&ud=1&site=default_collection

La interfaz mínima en testing se encuentra en:

http://buscador.test.agesic.gub.uy/search?q=pepe&spell=1&client=default_frontend&proxystylesheet=minima&output=xml_no_dtd&sort=date%3AD%3AL%3Ad1&entqr=3&oe=UTF-8&ie=UTF-8&ud=1&site=default_collection

Por más información, comentarios y sugerencias: contenidosPEU@agesic.gub.uy