|  |  |
| --- | --- |
|  | 🞂Search Module  Manual de uso y configuración |
|  |  |
|  | **Sergio Raposo Vargas** 🞂SAGA Soluciones Tecnológicas 🞂 28/05/2013 |
|  |  |

**CONTROL DEL DOCUMENTOS**

|  |  |  |  |
| --- | --- | --- | --- |
| **Realizado** |  | **Revisado** |  |
| (28/05/13) |  |  |  |
|  | Sergio |  |  |
|  | Raposo |  |  |
|  | Vargas |  |  |
|  | Director Técnico |  |  |

**Control de Modificaciones**

|  |  |  |  |
| --- | --- | --- | --- |
| *Rev.* | *Fecha* | *Autor/es* | *Descripción* |
| 001 | 28/05/13 | Sergio Raposo Vargas | Realización inicial del documento. |
|  |  |  |  |

Índice del documento

[1 Objetivo del documento 4](#_Toc357615371)

[2 Descripción Funcional 4](#_Toc357615372)

[3 Instalación del Módulo 5](#_Toc357615373)

[3.1 Dependencias 5](#_Toc357615374)

[3.2 Paso a paso 5](#_Toc357615375)

[3.3 Configuración inicial 5](#_Toc357615376)

[3.3.1 Configuración de Solr 5](#_Toc357615377)

[3.3.2 Configuración de contenidos XML 7](#_Toc357615378)

[4 Recurso de configuración 9](#_Toc357615379)

[4.1 Buscador Simple 9](#_Toc357615380)

[4.2 Buscador Avanzado 10](#_Toc357615381)

# Objetivo del documento

La finalidad de este documento es realizar una guía de uso y configuración del módulo: **SEARCH MODULE** donde veremos cómo se instala y se configura para añadir un buscador simple o avanzado en nuestros sites en OpenCms 8.5.0 o superior.

# Descripción Funcional

Con este módulo podremos crear un buscador de contenido dentro de nuestro sitio web utilizando **Apache Solr** como motor de búsqueda. El módulo permitirá crear dos tipos de buscadores:

* **Buscador simple** con el que podemos crear el clásico buscador por palabra clave que buscará en todo el contenido (tanto en el XML, como en el título de los recursos que no son XML). Le podremos indicar algunos parámetros de configuración como la carpeta raíz de búsqueda, el tipo de recurso que queremos buscar o el número de resultados a mostrar por página.
* **Buscador avanzado** donde tendremos la posibilidad de especificar el filtro que deseamos añadir a nuestro formulario. Los filtros disponibles son:
  + **Filtro de texto:** filtro que buscará un texto introducido por el usuario en el campo que le indiquemos en la configuración
  + **Filtro de fechas:** permite realizar un filtro en base a un rango de fechas. El campo fecha donde se busca puede ser configurable.
  + **Filtro por categoría:** muestra una lista de categorías disponibles en base a la categoría base que le indiquemos y el usuario podrá realizar un filtro de los contenidos que pertenecen a dicha categoría.
  + **Filtro por tipo de recurso:** en este caso, podremos querer realizar un tipo de filtro por tipo de contenido, donde el usuario podrá elegir entre una lista de posibles valores que tendremos que indicar en el recurso de configuración.

Una de las características del módulo es que podremos elegir el tipo de visualización que deseamos. Podremos elegir entre los siguientes tipos de campos a mostrar en nuestra visualización:

* **Título**: Propiedad Title del contenido.
* **Descripción**: Propiedad Description del contenido
* **Fecha**: Inicialmente cogerá el campo de Solr: xmldate. En caso de no existir cogerá la fecha de última edición del contenido (contentdate).
* **Imagen**: El campo imagen se leerá del campo de Solr: xmlimage. En el caso que el recurso sea de tipo Image, se mostrará directamente la imagen correspondiente.

La visualización nos permitirá elegir que elementos de estos 4 queremos mostrar.

A parte del filtro que podamos especificar y que dependerá de la información que añade el usuario, podremos añadir campos de filtrado de Solr por defecto, criterios de ordenación de los resultados a medida y otro tipo de configuración de la pantalla como son textos de cabecera y pie, hacer la caja de filtro desplegable, etc.

# Instalación del Módulo

## Dependencias

Este módulo solo es válido para versiones de **OpenCms 8.5 o superiores** y tienes una dependencia del módulo **Template Saga Base 2.0.X**.

## Paso a paso

Para proceder a la instalación del módulo debemos seguir los siguientes pasos:

1. Subir el fichero .zip del módulo a la carpeta: $OPENCMS\_HOME/WEB-INF/package/modules/
2. Entrar en OpenCms y situarnos en la vista de administración
3. Entrar en la gestión de módulos y en la opción subir módulo desde el servidor
4. Seleccionar de la lista nuestro módulo: com.saga.opencms.search

No es necesario reiniciar el servidor tras instalar este módulo.

## Configuración inicial

Uno de los apartados más importante del uso del módulo es la correcta configuración de nuestro OpenCms. La configuración la tendremos que hacer en dos sitios: La configuración de Solr y nuestros tipos de contenido.

### Configuración de Solr

Solr dispone de dos ficheros de configuración principales:

* **opencms-search.xml**: $OPENCMS\_HOME/WEB-INF/config/
* **schema.xml**: $OPENCMS\_HOME/WEB-INF/solr/conf/

Por defecto Solr dispone de unos parámetros de configuración que funcionan perfectamente, pero es posible que según nuestras necesidades tengamos que adaptar algo. A continuación mostraremos algunos ejemplos de configuración en base a necesidades reales:

**Creación de un índice Solr propio**

OpenCms trae dos índices por defecto de Solr, uno Offline y otro Online. Este índice indexa todos los sites de nuestros OpenCms por lo que podemos utilizar sin problema. Pero este índice está configurado para no indexar algunos tipos de recursos como por ejemplo las Imágenes. En muchos casos es necesario que las imágenes estén indexada para poder implementar un buscador de imágenes o incluso para que aparezcan en nuestros buscador general.

Para esto, podemos dar de alta un nuevo índice Solr con nuestra source por personalizado. Para ello, en nuestros fichero opencms-search.xml añadiremos el siguiente código XML:

<index class="org.opencms.search.solr.CmsSolrIndex">

<name>indexSolrNew</name>

<rebuild>auto</rebuild>

<project>Online</project>

<locale>all</locale>

<configuration>solr\_fields</configuration>

<sources>

<source>solr-source-new</source>

</sources>

<param name="search.solr.postProcessor">org.opencms.search.solr.CmsSolrLinkProcessor</param>

</index>

La configuración del source se puede realizar desde la vista de administración. En esta configuración es donde le diremos que carpetas vamos a indexar y qué tipo de contenidos.

**No distinguir entre palabras con tildes y sin tildes**

Cuando usamos solr para realizar búsqueda de texto nos encontramos con los clásicos errores de encoding de tildes y otros caracteres que tenemos en nuestro idioma y que pueden traernos problemas a la hora de configurar nuestras búsquedas.

El principal problema que tenemos con la configuración por defecto de Solr en OpenCms son las tildes. Por defecto, Solr es capaz de indexar las palabras con tildes pero para localizarlas posteriormente es necesario indicar la palabra exacta con las tildes correctas. Es decir, si tenemos un contenido que contiene la palabra 'vídeo', y escribimos en nuestro buscador 'video' no encontraremos ningún resultado.

El funcionamiento ideal, tal y cómo funcionan los buscadores normalmente es que sea capaz de no distinguir entre la búsqueda de vídeo y video. Para conseguir esta funcionalidad en Solr existe un filtro que es capaz de ignorar este tipo de caracteres:

<filter class="solr.ASCIIFoldingFilterFactory"/>

Este filtro convierte los caracteres unicode alfabéticos, numéricos y simbólicos que no están en los primeros 127 caracteres ASCII (el bloque de Unicode "Latín básico") en sus equivalentes ASCII, si existe.

Para añadir este filtro en nuestro buscador, debemos irnos a la carpeta /$OPENCMS\_HOME/WEB-INF/solr/conf/ y editar el fichero **schema.xml**. En la configuración del campo text\_general, podemos añadir dicho filtro quedando de la siguiente forma:

<!-- A general text field -->

<fieldType name="text\_general" class="solr.TextField" positionIncrementGap="100">

<analyzer type="index">

<tokenizer class="solr.StandardTokenizerFactory"/>

<filter class="solr.StopFilterFactory" ignoreCase="true" words="stopwords.txt" enablePositionIncrements="true" />

<filter class="solr.LowerCaseFilterFactory"/>

<filter class="solr.ASCIIFoldingFilterFactory"/>

</analyzer>

<analyzer type="query">

<tokenizer class="solr.StandardTokenizerFactory"/>

<filter class="solr.StopFilterFactory" ignoreCase="true" words="stopwords.txt" enablePositionIncrements="true" />

<filter class="solr.SynonymFilterFactory" synonyms="synonyms.txt" ignoreCase="true" expand="true"/>

<filter class="solr.LowerCaseFilterFactory"/>

<filter class="solr.ASCIIFoldingFilterFactory"/>

</analyzer>

</fieldType>

Una vez modificado este fichero, reiniciamos nuestro servidor, reindexamos todos nuestros índices solr y tendremos resuelto el problema.

### Configuración de contenidos XML

Para el correcto funcionamiento de nuestro buscador Solr debemos indicarle a nuestros contenidos específicos que indexe cierta información en campos de solr concretos:

* **xmldate**: En este campo debemos guardar la fecha concreta del contenido, de forma que será la fecha que luego se mostrará en los resultados finales.
* **xmlimage**: Debemos indexar la ruta de la imagen que queremos asignarle al recurso y que será la que mostremos cuando se encuentre el contenido en concreto.

Para configurar estos dos campos de solr debemos añadir en nuestros XSD la siguiente configuración:

<searchsettings>

<searchsetting element="Date" searchcontent="true" >

<solrfield targetfield="xmldate" />

</searchsetting>

<searchsetting element="Parrafo/ImagePrincipal/Image" searchcontent="true" >

<solrfield targetfield="xmlimage" />

</searchsetting>

</searchsettings>

Donde en nuestro ejemplo, el campo Date es el campo donde se guarda la fecha del contenido y la imagen principal está en Parrafo/ImagePrincipal/Image

# Recurso de configuración

## Buscador Simple

**Pestaña Datos Generales**

* **Título**: Título de la caja.
* **ID**: Identificador del buscador que nos permitirá distinguir en el caso de existir más de un buscador en la misma página
* **Clase CSS**: Clase CSS que se le aplica al contenedor principal
* **Resumen**: Resumen del buscador que se mostrará en caso de listar recursos de este tipo
* **Texto del formulario**: Texto que se mostrará por encima del formulario de búsqueda
* **Texto de cabecera**: Texto que se mostrará entre el formulario y los resultados de búsqueda
* **Texto de pie:** Texto que se muestra por debajo de los resultados de búsqueda

**Pestaña Configuración**

* **Índice Solr**: Índice que utilizaremos para las búsquedas en Online
* **Índice Solr Offline**: Índice que utilizaremos para las búsquedas en Offline. Puede ser el mismo en ambos casos
* **Tipo de Visualización**: Indica el tipo de visualización que se desea a la hora de mostrar los resultados
* **Mostrar resultados**: Si marcamos la pestaña, se mostrarán los resultados en el caso que se haya producido una búsqueda. Es útil para cajas de búsqueda que queramos redirigir a otra pantalla cuando se realice la búsqueda correspondiente. Se puede utilizar como caja en una columna.
* **Mostrar formulario**: Muestra el formulario de búsqueda para que el usuario pueda introducir el criterio de búsqueda que desee. En este caso, se da por hecho que los parámetros llegarán desde otra pantalla.
* **Formulario desplegable**: Añade un botón al formulario de búsqueda con el que se puede expandir y contraer.
* **URL Destino**: Indique en este campo la URL destino que se le aplicará al formulario (campo action). Este campo es útil en el caso que deseemos mostrar la información del resultado de búsqueda en otra pantalla.
* **Tipo de recursos**: Podremos indicar un tipo de recurso en concreto por el que deseamos realizar la búsqueda. Se puede configurar tantos recursos como queramos. Para no distinguir entre recursos, deshabilitar este campo.
* **Resultados por página**: Indique el número de registros por página a mostrar. Si se encuentran más resultados se mostrará una paginación de los resultados.
* **Ruta de búsqueda**: Indique la ruta raíz a partir de la cual vamos a realizar la búsqueda.

**Pestaña Valores por Defecto**

* **Campos por defecto**: Indicar una serie de parámetros solr por defecto. Ej: fq=Title\_prop:Test

**Pestaña Etiquetas**

* **Título caja formulario**: Título que se ve en la caja del buscador
* **Texto de relleno**: Texto de relleno que se muestra en el input de búsqueda cuando está en blanco
* **Etiqueta del botón**: Etiqueta del botón de búsqueda
* **Texto de No Resultado**: Texto que se muestra cuando no se encuentran resultados
* **Texto Error**: Texto que se mostrará cuando se produce un error al buscar
* **Fichero de idiomas**: Fichero de idiomas que utilizaremos para traducir las keys utilizadas en el formatter.

## Buscador Avanzado

**Pestaña Datos Generales**

Misma configuración del buscador simple

**Pestaña Configuración**

* Campos del buscador simple
* **Orden**: Podemos especificar los criterios para ordenar. Si no se especifica ninguno por defecto ordenará por puntuación de la búsqueda. Se pueden indicar más de un criterio de ordenación, por ejemplo, ordenar por puntuación descendientes, y en el caso que haya dos iguales ordenar por título(score desc; Title\_prop asc). Para especificar el orden hay que indicar el nombre del campo de solr y luego si es ascendente (asc) o descendente (desc).

**Pestaña Filtro**

* **Filtro de texto:** Nos permite realizar un filtro de un campo de tipo texto
  + **Etiqueta**: Etiqueta que se mostrará junto al input de entrada de datos
  + **Clase CSS**: Clase CSS que se aplicará al div contenedor del campo
  + **Campo Solr**: Campo solr sobre el que se realizará la búsqueda.
* **Filtro de categoría:** Realiza un filtro por categorías
  + **Etiqueta:** Etiqueta que se mostrará junto al input de entrada de datos
  + **Clase CSS:** Clase CSS que se aplicará al div contenedor del campo
  + **Categoría raíz:** Indica la categoría principal a partir de la cual mostrar la lista de posibles categorías a elegir
  + **Mostrar en árbol:** Si no marcamos esta casilla, se mostrará solamente las categorías hijas de la categoría raíz seleccionada. En el caso que se marque se hará un recorrido recursivo por todas las subcategorias.
  + **Mostrar padre:** Podemos elegir entre varias opciones: No mostrar la categoría raíz como opción a elegir, mostrar el título, mostrar la etiqueta ‘Todos’ o mostrar una opción en blanco
  + **Tipo de campo:**  Indica cómo queremos mostrar la lista de categorías
* **Filtro por 1 fecha:** Realizará un filtro de todos los contenidos de esa fecha
  + **Etiqueta**: Etiqueta que se mostrará junto al input de entrada de datos
  + **Clase CSS**: Clase CSS que se aplicará al div contenedor del campo
  + **Campo Solr**: Campo solr sobre el que se realizará la búsqueda.
* **Filtro por 2 fechas:** Muestra un formulario con 2 fechas con el que se establece un rango de fechas a buscar Si alguna no se rellena automáticamente utilizará la fecha actual
  + **Etiqueta fecha inicio**: Etiqueta que se mostrará junto al campo de fecha de inicio
  + **Etiqueta fecha fin**: Etiqueta que se mostrará junto al campo de fecha fin
  + **Clase CSS**: Clase CSS que se aplicará al div contenedor del campo
  + **Campo Solr**: Campo solr sobre el que se realizará la búsqueda.
* **Filtro por tipo de contenido:** Permite al usuario elegir entre una lista de tipos de contenidos por los que desea realizar el filtro
  + **Etiqueta**: Etiqueta que se mostrará junto al input de entrada de datos
  + **Clase CSS**: Clase CSS que se aplicará al div contenedor del campo
  + **Tipo de campo**: Nos permite elegir qué tipo de campo queremos usar para mostrar la lista de tipos de recursos configurada
  + **Opciones:** Lista de valores que se le mostrará al usuario
    - **Etiqueta:** Etiqueta que se va a mostrar en la lista. Ej: Noticia, Texto
    - **Valor:** Valor del campo que se enviará cuando el usuario seleccione. En este caso habrá que poner el nombre exacto del tipo de contenido. Ej: tsnoticia, v8textblock, …

**Pestaña Valores por Defecto**

* **Campos por defecto**: Indicar una serie de parámetros solr por defecto. Ej: fq=Title\_prop:Test

**Pestaña Etiquetas**

Misma configuración del buscador simple