



Anexo Configuración de ElasticSearch

Configuración ElasticSearch

Dentro de las preferencias de administración en Koha, busque la opción 'engine'¹ y cambie a ElasticSearch, una vez que haya cambiado a Elastic, verá un nuevo enlace en la sección Catálogo de Administración: **Configuración del motor de búsqueda**.

Aquí es donde Koha le muestra qué partes de sus registros MARC se indexan para la búsqueda y luego cómo esos índices de búsqueda se ponderan para su uso en la clasificación de relevancia de los resultados de búsqueda.

El primero reemplaza la documentación del índice de búsqueda del manual de Koha. Esto último es algo a lo que nunca hemos tenido acceso directo en Zebra.

A partir de la versión 19.11 de Koha, la mayoría de los elementos de esta página requieren alguna acción en su servidor cuando se modifican, por lo que le recomendamos que se comunique con nosotros antes de realizar cualquier cambio aquí. Dicho esto, sigue siendo un gran recurso para comprobar qué partes de sus registros se están indexando y cómo puede buscarlas directamente.

Mappings

Las pestañas etiquetadas como "Registros bibliográficos" y "Autoridades" contienen sus asignaciones entre el índice y el campo marc.

En cada pestaña, la columna "Campo de búsqueda" contiene nombres de índice de búsqueda y la columna "Mapeo" contiene campos y subcampos MARC.

¹ Requiere un servidor ElasticSearch instalado



Search field	Sortable	Facetable	Suggestible	Mapping	
abstract	Under ↓	No	No ↓	520	Delete
expnense	Under ↓	No	No ↓	852e	Delete
art	Under ↓	No	No ↓	526c	Delete
wp	Under ↓	No	No ↓	526d	Delete
author	Under ↓	Yes ↓	Yes ↓	100a	Delete
author	Under ↓	Yes ↓	Yes ↓	110a	Delete
author	Under ↓	Yes ↓	Yes ↓	111a	Delete
author	Under ↓	No ↓	No ↓	245c	Delete
author	Under ↓	Yes ↓	Yes ↓	700a	Delete
author-in-order	Under ↓	No	No ↓	245c	Delete

En la captura de pantalla anterior, podemos ver que el índice de autor contiene 100 \$ a, 110 \$ a, 111 \$ a, 245 \$ cy 700 \$ a. Entonces, una búsqueda de autor devolverá registros con nuestros términos de búsqueda en cualquiera de esos campos.

Muchos campos MARC se incluyen en varios índices diferentes. Si no está seguro de cómo buscar valores en un campo MARC en particular, puede ir a esta página y usar ctrl-F para buscar el campo MARC.

No todos los campos MARC están indexados de forma predeterminada. Si desea indexar un nuevo campo MARC o cambiar qué índices incluyen qué campos MARC, podemos cambiar estos valores por usted, por ejemplo, uno de nuestros técnicos de bibliotecas separó los 245 \$ n y los 245 \$ p en un nuevo índice llamado "title-part" para poder buscar números de temporada y volumen por separado del resto del título.

Un administrador de biblioteca especial solicitó un nuevo índice llamado "tipo de producto" por 513 \$ a, donde habían estado guardando datos internos para registros locales.

Las tres columnas restantes de estas pestañas tienen una funcionalidad limitada. Describen si estos campos se pueden utilizar para ordenar resultados, generar facetas o sugerir otras búsquedas, respectivamente.

A partir de Koha 19.11, cambiar estos valores no habilitará ningún cambio y puede romper la funcionalidad existente, por lo tanto no se recomienda cambiarlos.



Ponderación de campos

Name	Label	Type	Weight
abstract	abstract		
acqsource	acqsource		
arl	arl		
arp	arp		
author	author	String	16.00
author-in-order	author-in-order		
author-name-corporate	author-name-corporate		
author-name-personal	author-name-personal		
author-personal-bibliogr	author-personal-bibliogr		
author-title	author-title		

La pestaña denominada "Campos de búsqueda" muestra las ponderaciones aplicadas para generar la relevancia en los resultados de búsqueda.

Sin valores de ponderación asignados, todos los campos se consideran igualmente importantes en sus resultados de búsqueda.

Esencialmente, un espacio en blanco aquí es lo mismo que ingresar uno. Al aumentar el valor de ponderación, le damos más importancia a ese campo para que los registros con nuestros términos de búsqueda en los campos ponderados se coloquen más arriba en nuestros resultados que los registros con nuestros términos de búsqueda en campos no ponderados.

En nuestras pruebas, establecimos los siguientes pesos predeterminados:

- title: 32
- author: 16
- subject: 8
- title-series: 4
- contents: 2

Estas ponderaciones se basan en la suposición de que es más probable que sus usuarios proporcionen un título o autor al realizar la búsqueda.



Si no es un título o autor, pueden dar un título de tema o serie o algunas palabras contenidas en los campos 5XX.

Recuerde que siempre puede verificar las asignaciones para ver exactamente qué campos se incluyen en cada uno de estos índices ponderados.

Los pesos se aplican en el momento de la búsqueda, por lo que se pueden ajustar sin necesidad de volver a indexar.

Siéntase libre de modificar estos ajustes. pero avísenos si se decide por una ponderación diferente para que podamos asegurarnos de registrarlo.

Facetas

Facet order

Search field	Label	Display
author	Authors	<input checked="" type="checkbox"/>
homebranch	Home libraries	<input checked="" type="checkbox"/>
holdingbranch	Holding libraries	<input checked="" type="checkbox"/>
ccode	Collections	<input checked="" type="checkbox"/>
su-geo	Places	<input checked="" type="checkbox"/>
itype	Item Types	<input checked="" type="checkbox"/>
title-series	Series	<input checked="" type="checkbox"/>
subject	Topics	<input checked="" type="checkbox"/>
location	Locations	<input checked="" type="checkbox"/>
ln	Language	<input checked="" type="checkbox"/>

Volviendo a la pestaña "Registros bibliográficos", encontrará una pequeña tabla en la parte inferior de la página que controla la visualización de sus facetas, podemos desmarcar la casilla de una determinada faceta para deshabilitarla por completo o arrastrar las facetas hacia arriba y hacia abajo en la lista para cambiar el orden.



Díganos si desea cambiar la configuración para que podamos asegurarnos de que esos cambios se conserven en caso de que necesitemos reconstruir sus índices elásticos.

Elástico cambia facetas de una forma más, construye sus facetas a partir de todos los resultados de búsqueda que tiene. Eso significa que no usa la preferencia del sistema `MaxRecordsForFacets` en absoluto.



Anexo Búsquedas en ElasticSearch

Búsqueda de palabras clave generales

Al igual que con Zebra, la búsqueda predeterminada en Elastic es una búsqueda por palabra clave general, si no especificamos un índice de búsqueda, Koha lo interpreta como una búsqueda en el índice de palabras clave, en Zebra esto significa buscar en todo el registro MARC, en Elastic sobre Koha 19.05, el índice de palabras clave contiene solo los campos MARC que están incluidos en al menos un índice.

Generalmente, esto es útil ya que nos permite ignorar por completo partes de nuestros registros que no nos importa, en 19.11, tenemos la nueva opción para especificar cuáles de nuestros índices de búsqueda se incluyen en una búsqueda por palabra clave y hacer que difieran entre el interfaz administrativo y el catálogo público.

Eso nos permitiría, por ejemplo, incluir el campo 952 \$ x (nota de elemento no público) en las búsquedas del personal, pero no en las búsquedas públicas, además en 19.11 puede agregar una opción en la búsqueda avanzada para buscar todo el registro MARC como lo hizo Zebra, si lo prefiere

Especificando un índice

Si desea realizar una búsqueda más específica, puede decirle a Koha que use algo diferente al índice de palabras clave, el proceso básico para esto no ha cambiado en Elastic, para buscar en el índice de títulos, puede configurar el menú desplegable de búsqueda en "Título".

A screenshot of a search bar interface. On the left, the word "Search" is displayed. To its right is a dropdown menu currently showing "Title". Further right is a text input field containing the search query "batman and robin". On the far right of the bar is a green button with the text "Go".

O si lo prefiere, puede dejar el menú desplegable solo y especificar su índice usando CCL (Common Command Language).

Search Library catalog ▾ title:batman and robin Go

Buscar "título: batman y robin" le dice a Elastic que busque "batman y robin" en el índice de títulos, la página de configuración de su motor de búsqueda le dará una lista de todos sus índices de búsqueda y qué campos MARC incluyen.

Si bien el menú desplegable de búsqueda tiene un conjunto predeterminado de opciones, podemos agregar las que desee, si hay algo que busca con frecuencia y prefiere no usar CCL.

Una vez que haya especificado un índice, se utilizará para todos los términos siguientes en su consulta hasta que especifique algo más.

Search Library catalog ▾ title:batman harley Go

Esto significa que "title: batman harley" busca tanto "batman" como "harley" en el título. Mientras que "title: batman kw: harley" busca "batman" en el título y "harley" en el índice de palabras clave.

Search Library catalog ▾ title:batman kw:harley Go

Por supuesto, podría realizar la misma búsqueda con "title de harley: batman". Dado que "harley" viene antes de que especifiquemos el campo de título, por defecto es la palabra clave.

Search Library catalog ▾ harley title:batman Go

Tenga en cuenta que puede mezclar y hacer coincidir cualquier índice de esta manera. La búsqueda de "título: batman autor: dini" nos da registros con "batman" en el título y "dini" en el autor.



Operadores booleanos

En Koha, Elastic está configurado para asumir que todos nuestros términos de búsqueda están conectados con operadores AND, podemos especificar diferentes operadores booleanos tal como lo hicimos en Zebra usando la página de búsqueda avanzada o CCL.

Keyword batman
or Keyword superman [+]

En la búsqueda avanzada configure su menú desplegable booleano con el operador que desee.

La captura de pantalla de arriba muestra una búsqueda que traerá registros con "batman" o "superman" (o ambos).

Search Library catalog batman OR superman Go

Se puede realizar la misma búsqueda a través de CCL escribiendo "batman OR superman". Su operador booleano debe estar en mayúsculas. Una búsqueda de "batman o superman" buscará la palabra "o".

Sus operadores se aplicarán en el orden Y, O, NO. Este orden de operaciones puede resultar en algún procesamiento inesperado. Puede utilizar paréntesis para forzar la agrupación explícita, lo que puede aclarar las cosas.

Search Library catalog (batman OR superman) AND wonder woman Go

Esto devuelve registros que contienen "mujer maravilla" y "batman" o "superman".

Un operador booleano hace que Elastic olvide en qué índice estaba buscando, por lo que la consulta "título: batman OR superman" es la misma que "título: batman OR kw: superman". Puede corregir esto repitiendo su índice - "título: batman O título: superman" - o utilizando paréntesis - "título: (batman O superman)".



Comodines y truncamiento

Elastic admite dos caracteres comodín diferentes.

A search bar with the text "Search" on the left, a dropdown menu showing "Library catalog", a text input field containing "batm?n", and a green "Go" button on the right.

Un signo de interrogación representa un carácter, "batm?n" coincidirá con "batman" o "batmen" o cualquier otra palabra que puedas hacer introduciendo una letra o número entre "batm" y "n". Pero recuerde que buscará exactamente un carácter para reemplazar ese signo de interrogación. Ese "batm?n" encontrará "batman" o "batmen", pero no encontrará "batwoman".

A search bar with the text "Search" on the left, a dropdown menu showing "Library catalog", a text input field containing "bat*n", and a green "Go" button on the right.

Un asterisco representa cero o más caracteres. Entonces, "bat * man" encontrará "batman", "batwoman", "bathman" e incluso más cosas. Cualquier palabra que comience con "bat" y termine en "man".

A search bar with the text "Search" on the left, a dropdown menu showing "Library catalog", a text input field containing "bat*", and a green "Go" button on the right.

A search bar with the text "Search" on the left, a dropdown menu showing "Library catalog", a text input field containing "*man", and a green "Go" button on the right.

Se puede utilizar un asterisco al principio o al final de una palabra para truncar. La búsqueda de "bat*" devuelve todas las palabras que comienzan con "bat" y la búsqueda de "*man" devuelve todas las palabras que terminan en "man".

Stemming



Zebra tiene una característica llamada derivación que está relacionada con truncar, está controlado por la preferencia del sistema QueryStemming. Hace cosas como devolver "habilitados" cuando busca "habilitado". Elastic no tiene una característica como esa integrada, y la preferencia del sistema QueryStemming no hace nada mientras usa Elastic. Sin embargo, Elastic tiene sus propias opciones de derivación y eso es algo que esperamos explorar más a fondo en el futuro.

Búsqueda de frases

En Elastic puede forzar una coincidencia exacta utilizando comillas.

A screenshot of an Elastic search interface. It features a search bar with the text "bat*" entered. To the left of the search bar is a dropdown menu labeled "Library catalog" with a downward arrow. To the right of the search bar is a green button labeled "Go".

La búsqueda de "batman superman" entre comillas solo devuelve registros con esas dos palabras una al lado de la otra en ese orden. En Zebra, tenía que seleccionar una opción de búsqueda especial "como frase" del menú desplegable de búsqueda para hacer esto (como "título como frase" o "asunto como frase"). Esas opciones todavía existen en Elastic, pero todo lo que hacen es insertar comillas alrededor de su búsqueda.

Tenga en cuenta que Elastic ignorará cualquier comodín entre comillas, ya que las comillas significan que solo desea coincidencias exactas.

Rangos

Puede hacer que Elastic busque un rango de valores de varias formas. Esto se aplica principalmente a índices de campos numéricos, como la fecha de publicación, que contiene el año de publicación del campo 008.

A screenshot of an Elastic search interface. It features a search bar with the text "date-of-publication:[2010 TO 2012]" entered. To the left of the search bar is a dropdown menu labeled "Library catalog" with a downward arrow. To the right of the search bar is a green button labeled "Go".

Los corchetes como los que he usado aquí son inclusivos, por lo que mi búsqueda es para cualquier cosa publicada en 2010, 2011 o 2012. Si usa corchetes como "{2010 TO 2012}", el rango sería exclusivo, lo que significa que solo encontrará cosas publicadas en 2011.



Incluso puedes ponerte elegante y mezclarlas, como "[2010 TO 2012]", que sería inclusivo en el extremo inferior pero exclusivo en el extremo superior. Al igual que sus operadores booleanos, "TO" debe estar en mayúsculas.

A screenshot of a search interface. On the left, there is a 'Search' label and a dropdown menu set to 'Library catalog'. The main search input field contains the text 'date-of-publication:>=2010'. To the right of the input field is a green 'Go' button.

También puede usar símbolos de mayor y menor que para búsquedas basadas en números. La búsqueda anterior devuelve todo con una fecha de publicación mayor o igual a 2010.

Estas búsquedas no funcionarán bien con campos que no sean estrictamente numéricos. Si sus etiquetas 245 contienen un subcampo para datos como "Volumen 6" o "Temporada 2", Elastic no sabrá cómo descartar la palabra y observar solo el número. Sin embargo, este es exactamente el tipo de funcionalidad en la que la gente de la comunidad está trabajando actualmente, ¡así que eso puede cambiar!

Negar y exigir términos de búsqueda

Si desea asegurarse de que sus resultados no incluyan un término específico, puede negarlo con un signo menos.

A screenshot of a search interface. On the left, there is a 'Search' label and a dropdown menu set to 'Library catalog'. The main search input field contains the text 'batman -joker'. To the right of the input field is a green 'Go' button.

Una búsqueda de "batman-joker" devuelve registros que contienen "batman" pero no "joker". También puede hacer que los términos sean obligatorios agregando un signo más, pero eso es redundante porque por defecto conectamos todos nuestros términos con AND, lo que también los hace obligatorios.

Búsqueda difusa

En Zebra, la activación de la preferencia del sistema QueryFuzzy hizo que todas sus búsquedas buscaran palabras escritas de manera similar. Estaba algo mal definido e impredecible y solíamos sugerirle a la gente que no lo usara. En Elastic, activar QueryFuzzy no cambia los resultados de la búsqueda por sí solo, pero le da la opción de



hacer que cualquier término individual en su búsqueda sea borroso colocando una tilde después.

Search Library catalog Go

Proximidad

Una búsqueda de proximidad le permite encontrar dos palabras dentro de una cierta distancia entre sí

Search Library catalog Go

Para realizar una búsqueda de proximidad, pero sus términos entre comillas y luego agregue una tilde y un número. Entonces, "batman robin" ~ 1 nos da registros en los que "batman" y "robin" aparecen dentro de una palabra entre sí. Eso incluiría "batman y robin" o "batman y robin". Tenga en cuenta que cuando decimos palabras aquí, usamos los términos de manera vaga, básicamente para referirnos a un grupo de caracteres separados de otros caracteres por espacios. Entonces, en este contexto, un ampersand se considera una palabra.

Ahora, técnicamente, el número aquí no es un recuento de palabras entre nuestros términos. Es un recuento de cambios necesarios para que nuestro registro coincida con nuestra búsqueda (algo así como cómo la borrosidad contaba los cambios en las letras). Se necesita una edición (eliminando el "y") para que "batman y robin" coincidan con "batman robin". Siguiendo esta idea de contar ediciones, dos ediciones nos permiten transponer nuestras palabras. Entonces, "batman robin" ~ 2 coincidiría con "robin batman". Y "batman robin" ~ 3 coincidiría con "robin y batman".



Impulsar la relevancia

En nuestra publicación de configuración de Elastic, hablamos sobre cómo definir ponderaciones para configurar cómo se ordenan los resultados de búsqueda. Elastic también le permitirá usar la función boost para darle a un término específico una importancia adicional en cualquier búsqueda.

A search bar interface with a dropdown menu set to 'Library catalog' and a search input field containing the text 'batman robin stephanie^10'. A green 'Go' button is located to the right of the input field.

Escapar de puntuación

Muchas de las funciones de búsqueda que se analizan aquí utilizan signos de puntuación para que Elastic sepa que estás haciendo algo especial. Si desea realizar una búsqueda que incluya uno de estos signos de puntuación, debe indicarle a Elastic que ignore el significado especial de la puntuación. En la codificación, esto se conoce como escapar del signo de puntuación.

A search bar interface with a dropdown menu set to 'Library catalog' and a search input field containing the text 'ti:batman\ : year one'. A green 'Go' button is located to the right of the input field.

Para escapar de un signo de puntuación, coloque una barra invertida antes. Por lo tanto, una búsqueda de "título: batman \: año uno" busca "batman: año uno" directamente sin intentar usar los dos puntos para hacer algo especial. Los siguientes signos de puntuación deben omitirse si se incluyen en su búsqueda: +, -, =, &&, |, >, <, !, (,), {, }, [,], ^, ", ~, *, ?, :, \ y /.

Por supuesto, normalmente debería poder dejar esos signos de puntuación fuera de su búsqueda por completo, en lugar de preocuparse por escapar de ellos. En mi ejemplo anterior, "título: batman año uno" sin los dos puntos habría encontrado el mismo título sin ningún problema.