**Migración Dynamic**

**CONSIDERACIONES IMPORTANTES**

* En la sección *Explore* de *Jobs content & Indexation* en el menú de las herramientas de Talent, se podrán encontrar dos opciones para los ***scanids.*** El que tiene la palabra “**Legacy”** entre paréntesis enlista los *spiders* que aún se encuentran en neuvoo, mientras que la opción que solo tiene el nombre *scanids* enlista los *spiders* que se encuentran en Talent.

Imagen que contiene Interfaz de usuario gráfica

Descripción generada automáticamente

* A pesar de que la opción de migrar *scanids* con status *inactive*, *new* o *Working on it* es posible, está **restringido** realizar esta acción a menos de recibir órdenes especiales por parte de su jefe directo.
* El último ítem de la tabla presente en *Scanids (Legacy)* contiene 3 botones con funcionalidades distintas:
* ***Open Spider:*** este botón redirecciona al *spider* en la herramienta de neuvoo.

Interfaz de usuario gráfica, Sitio web

Descripción generada automáticamente

Interfaz de usuario gráfica

Descripción generada automáticamente

**Dentro de le herramienta de neuvoo** se encuentra el botón *scanid info* que despliega la información del *scanid*.

Interfaz de usuario gráfica

Descripción generada automáticamente

SCANID INFO

* ***Scanid Info:*** este botón despliega un menú al costado derecho de la página con información general del *scanid* que se seleccione.

Imagen que contiene Interfaz de usuario gráfica

Descripción generada automáticamente

Interfaz de usuario gráfica

Descripción generada automáticamente

* ***Migrate Spider***: este botón se utiliza para comenzar el *proceso de migración* y el que abre la herramienta *Spider Dynamic Migration* de Talent.

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

* La tabla de *scanids* en Talent también cuenta con el último ítem y tres botones, al igual que los *scanids* de neuvoo. Comparten dos de estos tres que son el *Open Spider* y *Scanid Info*. El primero abre el *spider* en la herramienta ***Indexing Tool*** (Talent) y el segundo cumple la misma función que en la tabla de *scanids* en Neuvoo.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

El tercer botón es *Open Crawl History* y redirecciona a la herramienta ***Crawl History***.

Interfaz de usuario gráfica

Descripción generada automáticamente con confianza media

* En la herramienta *Indexing Tool* se pueden encontrar 2 botones en la parte superior: *Scanid Info* y *Scanid Config.* El primero despliega una pestaña al costado derecho de la venta con información del *scanid y* el segundo despliega también al costado derecho de la ventana una pestaña, pero con la información de la configuración del *spider.* Adicionalmente en la parte superior derecha se puede encontrar el incono de configuración para cambiar la apariencia del editor de la herramienta *Indenxing* *Tool* y del editor de texto del código.

**NOTA:** Los campos “LoadPageTextTimeout”, “LoadPageResourcesTimeout” y “LoadPageTimeout” en la pestaña *Scanid Congif* son campos de visualización, estos no pueden ser modificados.

Imagen que contiene Patrón de fondo

Descripción generada automáticamente

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

A diferencia de la herramienta de indexación de neuvoo, que contaba con un menú desplegable para seleccionar el paso de extracción que se quisiera trabajar e incluso para realizar un *test* al *spider*. Con la herramienta de indexación de Talent.com, todos los pasos de la extracción aparecen como opciones seleccionables debajo de los tres botones mencionados en el párrafo anterior y el *test/run* del código se puede realizar desde cualquier *step*.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

Debajo de los pasos de extracción se encuentra el recuadro que muestra el código como tal y este recuadro cuenta con 5 botones en la parte superior- derecha:

* **Save:** permite guardar el código editado de cada *step*.
* **Delete:** permite eliminar todo el código escrito de cada *step*.
* **Test Step:** permite realizar un *test* en el *step* donde se seleccione el botón.
* **Test:** permite realizar un *test* a toda la información del *spider*.
* **Kill Test:** permite detener la ejecución de un step o la ejecución del spider en su totalidad. Este botón se habilita si y solo si ejecutamos un *Test* *Step* o un *Test*.
* **Run:** permiteejecuta el *spider* para la inserción de los *jobs* en la base de datos.

Texto

Descripción generada automáticamente

**Imagen de la pantalla de un celular con letras

Descripción generada automáticamente con confianza media**

* Los *spiders* migrados **DEBEN** quedar funcionando correctamente.
* Realizar análisis - tiempo indefinido. Tener en cuenta que deben agregarse todas las variables cuyos datos estén presentes en la fuente de acuerdo con los lineamientos establecidos actualmente en indexación orgánica. (Source\_location, reqid, experience required.)
* Dependiendo de los pasos realizados para la migración del *spider* de neuvoo a Talent, el status del *spider* en ambas plataformas puede variar de la siguiente manera:

|  |  |  |
| --- | --- | --- |
| **El *spider* está…** | **Status neuvoo** | **Status Talent** |
| Completamente en neuvoo | Active | N/A (No aplica) |
| Migrado (únicamente código) | Active | Migrated to Talent |
| Pendiente por terminar proceso migración código neuvoo->Talent | Active | Migration in porgress |
| Completamente en Talent | Migrated to Talent | Active |

* Validar en el *Explore Scanid* que el *spider* aparezca con status “Active”
* Validar en el *Scanid (legacy)* que el *spider* aparezca con status “Migrated to Talent”.
* Validar en el *Crawl History* que en las próximas corridas automáticas se reflejen en la columna *Excecute By* “Robot” (esto significa que el *spider* está corriendo del lado de Talent).

**Proceso de migración**

1. Para migrar un *spider* se debe seleccionar el botón ***migrate spider*** que será con el que comenzaremos el proceso de migración del *scanid* sobre el que se quiere trabajar. Al seleccionarlo se abrirá la herramienta ***Spider Dynamic Migration*** en el primer paso de 3 a completar.
2. Al encontrarse en el primer paso: *step1:Info* (datos del *feedcode* y del *spider*), se debe asegurar que la información presente en estos campos sea la misma que se encuentra en el *spider* de neuvoo (botón scanid info).

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

Confirmada la similitud entre la información presente en la herramienta de neuvoo y la requerida en la herramienta ***Spider Dynamic Migration***, se debe seleccionar el botón *Submit* al final de la página y pasar al *step 2.*

1. En el segundo paso *Step 2: Config* se encontrarán 3 secciones distintas y estas son: *Migrated* *Config*, *Config* *From* *neuvoo* e *Info* *about* *changes*. El último permite visualizar mensajes de advertencia que debe tener en cuenta la persona que esté migrando el *spider*. Mientras que las dos primeras son los datos de configuración del *spider* en Talent.com y en neuvoo, respectivamente. Estas dos partes comparten la misma información, pero con un nombre diferente para los campos. La información del lado de neuvoo aparece en formato Json mientras que la información del lado de Talent.com aparece como campos seleccionables en su mayoría. La información de estos campos debe coincidir para poder seguir al paso 3.

|  |  |  |
| --- | --- | --- |
| **En Talent.com** | **En neuvoo** | **Definición** |
| **scrapingEngineOptions** |  | |
| waitForPageText | ignoreLoadErrors | Permite continuar con el procesamiento aun cuando la página tiene menos de 100 caracteres. |
| waitForPageLoadEvent | waitForPageLoadEvent | Espera por el evento *load* del DOM. Es decir, espera a que el documento HTML se haya cargado y analizado por completo |
| InjectDefaultCodes | inactivateJQuery | Se asigna por defecto con la opción “Yes” para evitar errores. Se muestra la sección *inject* *Defaults* con opciones de selección simple. Hace referencia a todos los códigos que hay disponibles al momento de crear un spider o código *javascript* que va a inyectarse en las páginas.   * activateJQuery (default active) * GetEmails * msg (default active) * MD5 (default active) * ModifyParamURL * cleanHTML (default active) |
| waitForPageResources | waitForResources | Espera a que se resuelvan todos los *requests* hechos por la página. |
| **browserSettings** |  | |
| noImage | noImage | Se utiliza para evitar la descarga de imágenes. |
| skipResourcesMatch | skipResources | Se utiliza para filtra ciertos recursos mediante *blacklist* de dominios. |
| skipResourceTypes | noUnnecessaryResources | Se utiliza para bloquear ciertos recursos que se estén solicitando de la página. |

1. En el último paso de la migración *Step3: Code* se podrá únicamente visualizar la información del código del lado de neuvoo; se debe revisar que toda la información esté bien para finalmente seleccionar el botón *Submit* y terminar la migración del *spider*.

**NOTA:** El código desde el lado de neuvoo se supone que debe estar funcionando bien si se encuentra en *active.* Sin embargo, si se llega a encontrar un error, este debe ser corregido desde la herramienta *Indexing Tool* una vez el *spider* se encuentre del lado de Talent.com.

1. Culminado el proceso de migración, se debe ir a la herramienta *Explore Scanid* (*spiders* en Talent) y confirmar que esté migrado y en la columna “Status Talent” aparezca ***migrated to Talent,*** para posteriormente abrir el *spider* desde la herramienta *Indexing Tool.*
2. Con el *spider* testeado se debe inspeccionar la información desde la consola y en la misma confirmar la extracción correcta de los *jobs* (new, old, en error, total, etc). Allí se despliega información relevante de la corrida en 3 secciones principales:

* **Scraping Performance:** es un resumen del tiempo de ejecución para la extracción de los trabajos desde la herramienta de indexación *Indexing Tool*.
* **GhostProcess Summary:** es un resumen del procesamiento de los trabajos realizado por *ghostProcess.*
* **GhostProcess Performance:** es un resumen del tiempo de ejecución del procesamiento de los trabajos realizado por *ghostProcess*.

También existe un link que permite abrir la herramienta *Spider Results* para corroborar que la información de los *jobs* sea la correcta. Aquellos *jobs* que caigan en error serán mostrados con letra roja y también dirá el porqué del error.

1. Una vez verificada la información, desde la herramienta *Indexing Tool* en el *scanid info*, se debe cambiar el status de *Migrated to Talent* a ***Active*** para que las corridas se procesen de forma automática y darle RUN para comenzar con la primera corrida.

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

1. Finalmente se debe dirigir a la herramienta de neuvoo (Boo 3.1) para seleccionar el *status* ***Migrated to talent*** en el *spider* recién migrado.

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente