

RECUPERACIÓN DE INFORMACIÓN Y RECOMENDACIONES EN LA WEB

GRUPO 04

Fernanda Toledo

Fabian Kremer

Joaquín Cardozo

Sebastián Jara

Índice

Índice.....	2
Descripción del Problema.....	3
Solución del problema.....	3
Herramientas utilizadas	4
Ruby on Rails	4
Objective-C.....	4
Lucene - Solr.....	4
Arquitectura.....	5
Casos de prueba.....	6
Trabajo a futuro.....	8
Conclusiones	9
Referencias	10
Anexo:	11
Manual de usuario:.....	11

Descripción del Problema

Se plantea presentar una solución al problema de búsqueda sobre películas, específicamente las que se encuentran en IMDb. Consideramos que a veces es difícil encontrar una buena película para ver, que cuadre con nuestros gustos, por lo tanto es importante contar con la capacidad de realizar búsquedas avanzadas, con filtros y optimización de las mismas de manera de obtener el mejor resultado posible.

Solución del problema

Este trabajo está centrado en la realización de búsquedas de películas, las cuales pueden realizarse según: elenco, género, director, país de origen, idioma, calificación, o una combinación de ellos. A su vez también se permiten buscar con filtros “negativos”, por ejemplo buscar películas en las que no actúe cierto actor. Este tipo de filtros permitirán al usuario filtrar la búsqueda de manera flexible y amplia para así obtener una lista de resultados reducida y precisa.

Se realizó una aplicación móvil para iOS, con un backend realizado en Ruby on Rails. Los datos fueron extraídos desde una API, (<https://www.themoviedb.org/documentation/api>), la cual contiene toda la información en base a IMDb necesaria para nuestro problema. La forma de obtención de los datos es la siguiente:

De esta API nos interesan dos servicios:

- https://api.themoviedb.org/3/movie/:id?api_key=XXX
Servicio para obtener la información general de la película con id *:id*.
- https://api.themoviedb.org/3/movie/:id/casts?api_key=XXX
Para consultar la información de los participantes de una película con id *:id*

Se creó un proceso en nuestro servidor que itera para cada id de película de la base de themoviedb, llamando a estos servicios para cada una de ellas, obteniendo de esta manera nuestra propia base de datos con todos los datos necesarios para futuro manejo.

Como primera instancia de prueba se decidió obtener las 20 películas mejor calificadas desde el año 2000 hasta la fecha, obteniendo así un total de 320 películas en la base de datos. Otro detalle a destacar es que se pidieron los datos en español ya que se consideró pertinente.

Luego, nuestra base ya alimentada se integró con la librería de recuperación de información Sunspot Solr para la indexación de la misma.

Para solucionar el problema de ingresar múltiples filtros, se decidió realizar la búsqueda en una serie de pantallas en donde en cada pantalla se ingresa un filtro distinto, en la misma pantalla se puede ingresar tanto filtros positivos como negativos,

así obteniéndose una pantalla para filtrar por actores, otra para géneros, otra para años, y lo mismo para el resto de los filtros: idiomas, directores.

Adicionalmente se agregó un panel de administración brindado por la integración con la gema ActiveAdmin, con el fin de poder corregir aquellas traducciones que se consideren incorrectas. Sin embargo, debido a las facilidades que presenta la gema para la creación de un panel administración, es que además de editar, se puede crear o borrar los datos ya existentes.

Herramientas utilizadas

Ruby on Rails

Es un framework de aplicaciones web de código abierto optimizado el cual permite desarrollar aplicaciones escribiendo menos código que con otros frameworks y con un mínimo de configuración.

Se decidió realizar el backend de esta aplicación en este haciendo uso de esta plataforma debido a la experiencia y conocimiento de la misma por parte de la mayoría de los integrantes del grupo.



Objective-C

Mac OS X incluye una herramienta de desarrollo, XCode basada en Objective-C, la cual también incluye una herramienta para el diseño de la interfaz de aplicaciones iOS. Es por tal motivo que se utilizará para realizar el front end de la aplicación.



Lucene - Solr

Apache Lucene es una librería de motor de búsqueda de texto de alto rendimiento. Es una tecnología adecuada para casi cualquier aplicación que requiera la búsqueda de texto completo, o indexado, especialmente si es multiplataforma.

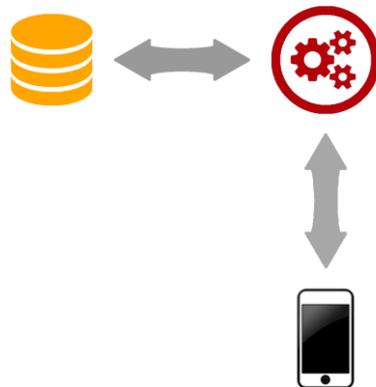


En ese caso, debido a la utilización de Rails para la implementación de la API, se hizo uso de la gema “sunspot_solr” para importar las funcionalidades de dicho motor de búsqueda.

Una vez importada, se procede a indicar sobre qué tablas y columnas de la base de datos es que se van a realizar las búsquedas, se indexan dichas columnas y finalmente se debe ejecutar el servidor Solr (independiente del backend).

Arquitectura

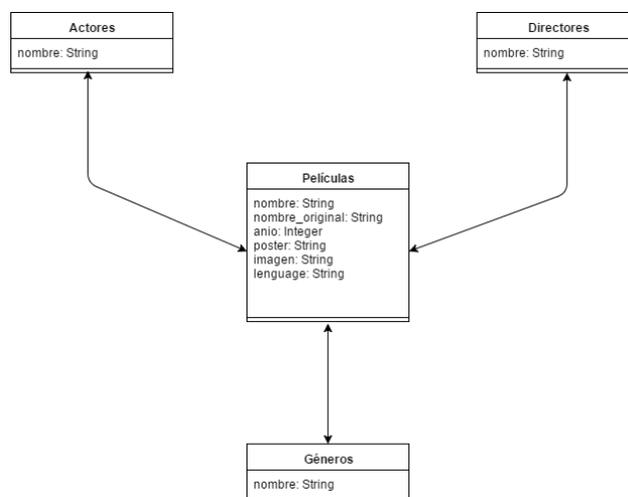
La arquitectura se compuso de 3 componentes: backend, frontend y base de datos.



El backend es el encargado de recibir los pedidos por parte de los usuarios y retornar los resultados acordes.

El frontend es una aplicación móvil la cual presenta a los usuarios los medios necesarios para realizar sus búsquedas, y luego le muestra de forma amigable los resultados obtenidos desde la API.

Finalmente, la base de datos es donde se almacena la información a consultar. Como ya se mencionó existe una tarea encargada de extraer la información de un servicio otorgado por themoviedb, la cual consulta, parsea y normaliza la información obtenida según el diseño implementado, el cual presentamos a continuación:

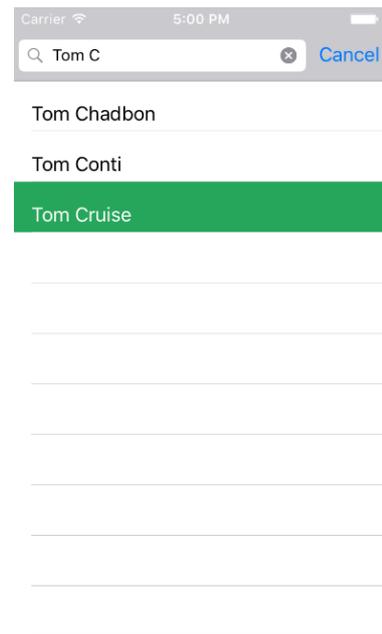


Casos de prueba

La funcionalidad más interesante que presenta la aplicación es el poder especificar qué es lo que NO se quiere dentro de los resultados esperados. Es raro encontrar una aplicación que permita reducir sus búsquedas mediante la exclusión de cierta información. Si además le sumamos que los resultados pueden ser de búsquedas mixtas entre cosas que un usuario desea incluir y excluir la aplicación se hace más atractiva.

Para mayor claridad de cómo se realizó cada paso, mirar el anexo que se encuentra al final del informe.

Un caso de ejemplo es la búsqueda de Tom Cruise. Si se quieren buscar películas donde él actuó se debe primero seleccionar sólo el filtro positivo en Actores, como muestra la siguiente captura de pantalla:



Luego, se deberá saltar todos los filtros, hasta llegar al botón "Buscar". La salida contendrá todas las películas en las cuales Tom Cruise fue actor. La página principal cambiará de modo que sólo aparecerán las películas con el filtro previamente dicho.

Esta salida contiene 8 películas.



A continuación agregaremos la negación como filtro adicional. En este caso negaremos los años de lanzamiento de películas de Tom Cruise. En el ejemplo, excluirémos las películas que hayan sido estrenadas en los años: 2000, 2001, 2002, 2003 o 2006.



La salida esperada es la que se muestra en la captura de pantalla. Como se puede observar, se excluyeron algunas películas en comparación a la búsqueda anterior. Efectivamente, las películas que se excluyeron pertenecían a los años que seleccionamos anteriormente. Por ejemplo, se puede observar en la captura de pantalla de los resultados anteriores las películas Misión Imposible 2 y Minory Report. Estas películas se estrenaron en el 2001 y 2002 respectivamente.



Trabajo a futuro

La aplicación final construida fue simplemente un prototipo de lo que podría ser una aplicación a utilizar en un futuro, por lo que consideramos que se podrían realizar muchas mejoras, o agregarle funcionalidades.

Como trabajo a futuro, una buena modificación sería que las opciones de los filtros dependan de los filtros previamente seleccionados, es decir, si en el filtro por actores se selecciona que Pepe Rodríguez debe pertenecer al elenco, y Pepe Rodríguez no actuó en películas con fecha de estreno en el 2000 y 2001, esos años no deberían aparecer entre las opciones posibles. De esta manera la lista de resultados se reduciría con cada filtro seleccionado, optimizando el funcionamiento de la API ya que no se estaría perdiendo tiempo en realizar consultas que no afectan el resultado final.

Agregar más filtros a futuro puede ser una alternativa interesante. Para esto habría que modificar la tarea que obtiene la información para poder extender el rango y así traer información no solo de la película, sino que también de los actores, directores, etc.

Otra mejora sería que si bien la mayoría de las películas tienen un título en español faltarían traducir algunas que no lo tienen, para realizar esto se tendría que chequear y traducir una a una las películas que no lo están.

Actualmente la base de datos no contiene todas las películas de IMDb, solo una parte de ellas. Esto es algo clave, ya que si no se tienen todas las películas entonces no es de mucha utilidad. Si esta aplicación realmente fuera a ser utilizada por el público entonces esta sería sin dudas una mejora fundamental y necesaria para el producto. Para realizar esto, es necesario actualizar el proceso que consulta todas las películas de la API externa, lo cual no es una tarea difícil.

El diseño de la aplicación podría mejorarse, por ejemplo, no hay control del tamaño del texto, por lo tanto si la descripción de la película es muy extensa, no se podrá leer en su totalidad. Debería haber un manejo dinámico del tamaño.

Por parte de los filtros, en el estado en que se encuentra la aplicación solamente realiza consultas conjuntivas (basadas en la lógica OR) sobre el mismo filtro, ya que la gema utilizada tiene ese comportamiento. Se debería realizar un trabajo más minucioso con el fin de poder implementar la lógica AND en las consultas. Para ejemplo, si se selecciona filtrar películas con el usuario A y B, los resultados también contendrán películas donde actúe solo A y donde actúe solo B.

Sin duda alguna hacer llegar esta aplicación a otros usuarios sería fundamental. En este momento solo usuarios que posean teléfonos iPhone serían capaces de hacer uso de los servicios implementados. Extenderla a usuarios con otros tipos de dispositivos, así como también computadoras de escritorios tendría una prioridad alta.

Conclusiones

Comprobamos que Solr es una herramienta muy útil para la indexación en bases de datos con gran cantidad de información, haciendo el filtrado de películas de manera rápida, esta herramienta es de vital importancia para la aplicación realizada, ya que si la obtención de resultados demorara el usuario no la usaría.

Se logró el objetivo de recuperar a través de la web datos de películas y ofrecerlos a los usuarios de una manera más amigable y accesible. Recordemos que hoy en día, en estos tiempos modernos, son pocos los que no tienen un teléfono inteligente en su bolsillo. Brindar la información a través de una aplicación móvil es brindarle practicidad al usuario, ya que puede buscar la película deseada sin la necesidad de moverse.

La aplicación móvil construida funciona tal cual lo esperado, filtrando los resultados de manera correcta según lo ingresado, por lo que la hace de gran utilidad para los fanáticos del cine.

Referencias

<https://www.themoviedb.org/documentation/api> - API utilizada para consumir los datos

<http://www.myapifilms.com/> - API considerada pero descartada para la solución

<https://github.com/sunspot/sunspot> - Gema utilizada para la integración con Solr

<http://lucene.apache.org/solr/> - Sitio oficial de Solr.

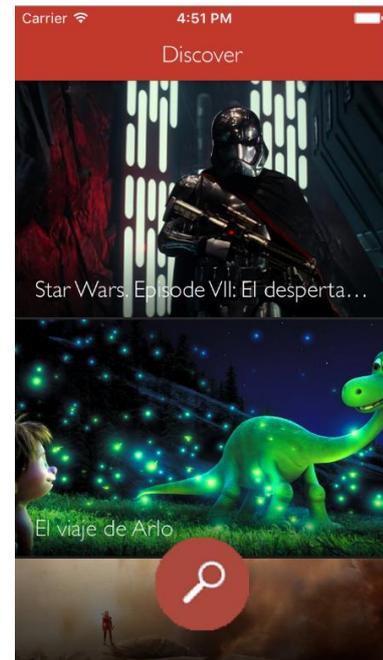
<https://github.com/activeadmin/activeadmin> - Gema ActiveAdmin

<http://activeadmin.info/> - Sitio oficial de ActiveAdmin

Anexo:

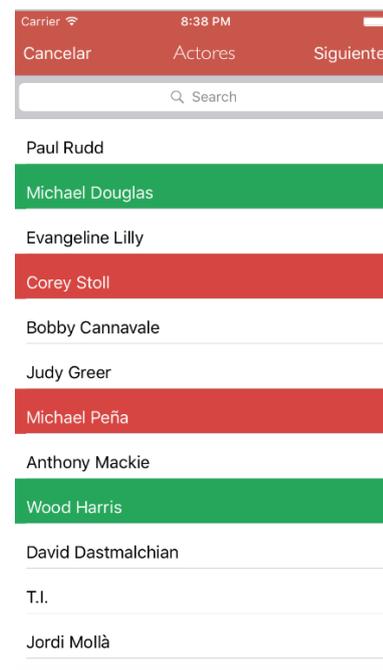
Manual de usuario:

Al acceder a la aplicación la pantalla principal lista todas las películas existentes en el sistema, ordenadas por orden de creación (o sea, al principio aparecerán las películas más nuevas). Para filtrarlas deberá tocar la lupa que se encuentra en la parte inferior de la pantalla. Luego aparecen pantallas donde se pide que se ingresen los distintos filtros, como ser, directores, actores y género.



Este es el tipo de pantallas en el que se pide que el usuario filtre según lo desea, en este caso, se quiere buscar una película que tenga a Michael Douglas o Wood Harris en el elenco, pero que no tenga a Corey Stoll ni Michael Peña. Para marcar la fila en verde se deberá tocar la fila. Para marcarla en roja (es decir, para excluirla) en este caso el actor de los resultados se debe mantener presionada la fila. Para elegir el próximo filtro se debe tocar *Siguiente* en la barra de navegación, y continuar de la misma manera hasta llegar al último filtro. Finalmente tocar *Buscar*, lo cual hará el pedido al servidor y llevará a la lista de resultados.

Además cada una de las pantallas de filtros cuenta con una barra de búsqueda interna a la pantalla para poder filtrar dentro de los actores, directores, géneros, años e idiomas. De esta manera es más fácil y amigable para el usuario elegir los filtros deseados.



Luego de realizada la búsqueda se muestra la lista de películas resultado del filtrado. En caso de querer ver mayor información sobre alguna película en particular se debe solamente tocar la fila para que aparezca una pantalla diferente donde se muestra toda la información existente en la base de datos de dicha película.



La guerra de los mu...
Aventura, Suspense, Cie...



Basada en una popular novela de HG Wells, narra la historia de una invasión del planeta tierra por los marcianos, y la extraordinaria batalla de la humanidad por la supervivenc...

Para poder acceder al administrador se deberá ingresar a [http:// URL_DEL_SERVIDOR/admin](http://URL_DEL_SERVIDOR/admin). Las credenciales para poder acceder son:

USUARIO: admin@webir.com

PASSWORD: password

Una vez que se ingrese se podrán crear, ver, editar y borrar todos los registros que se encuentren en la base de datos, para esta primera versión: películas, actores, directores y géneros.

A continuación se muestran algunas imágenes que muestran cómo se presentaría la información al administrador. Debido a que las películas son la estructura más completa es que las imágenes serán sobre la página de las mismas.

Poster	Title	Original Title	Year	Overview
	El viaje de Ario	The Good Dinosaur	2015	¿Qué hubiera pasado si el asteroide que cambió para siempre la vida en la Tierra no hubiera impactado en nuestro planeta? ¿Qué pasaría si los dinosaurios gigantes nunca se hubiesen extinguido? El viaje de Ario nos trasladó a un mundo donde el asteroide que terminó de forma dramática con los dinosaurios, pasa de largo. La consecuencia de este colosal cambio es que ahora los dinosaurios y los humanos tienen que vivir juntos en la Tierra.
	Marte	The Martian	2015	El astronauta Mark Watney (Matt Damon), un botánico e ingeniero mecánico de la NASA, se ve atrapado en el planeta Marte cuando la tripulación del Ares 3 debe evacuar su lugar de aterrizaje ante una tormenta de arena. Sin su tripulación, se ve obligado a encontrar la forma de regresar a la Tierra y sobrevivir confiando en sus habilidades científicas y técnicas, antes de que se le agoten los suministros. Mientras que sus compañeros astronautas piensan que podrá estar muerto, Watney comienza a escribir un diario divulgando sobre el futuro, después de su muerte. Cuando la NASA descubre a través del satélite que Mark podría estar vivo, comienzan a trabajar sobre su rescate, ocultándole al resto del grupo.
	Eternal	Selfless	2015	A un poderoso millonario, enfermo de cáncer, le ofrecen la posibilidad de conectarse a un procedimiento médico radical para transferir su conciencia al cuerpo de un hombre más joven. Al principio la operación va bien y el hombre disfruta de su renovada juventud, pero todo se complica cuando empieza a descubrir los detalles en torno a la procedencia de su nuevo cuerpo.
	Pivots	Pivots	2015	Pivots cuenta como unos entremetidos malinterpretan las imágenes de las máquinas recreativas como una declaración de guerra y deciden atacar la tierra, empleando dichos juegos como modelos para el asalto. El presidente de EEUU, Vice Cooper (Kevin James), recurre entonces a su gran amigo de la infancia y campeón de las mojarritas de los años 80, Sam Brenner (Adam Sandler), quien actualmente trabaja como instalador de sistemas de home cinema, para encabezar un equipo de expertos jugadores de su época (Dinklage) y (Said) cuya misión será derrotar a los entremetidos y salvar al planeta. Monaghan interpreta al experto en armas del equipo.
	San Andrés	San Andreas	2015	La falla de San Andrés acaba cediendo ante las terribles fuerzas telúricas y desencadena un terremoto de magnitud 9 en California. Atrial tal catástrofe, el piloto de helicóptero de búsqueda y rescate Ray (Dwayne Johnson) y su esposa Emma (Cara DiLorenzo) viajan juntos desde Los Ángeles hasta San Francisco para ir a su única hija, Blake (Isabella D'Adamo). Pero su tortuoso viaje hacia el norte solamente es el comienzo del desmoronamiento de todo lo que creían firme en su vida. (FILMAFFINITY)
	Tomorrowland	Tomorrowland	2015	Unidos por el mismo destino, una adolescente inteligente y optimista, y un antiguo rfo prodigio inventor, hastado por las discusiones, se embarcan en una peligrosa misión para desenterrar los secretos de un enigmático lugar localizado en algún lugar del tiempo y el espacio, conocido en la memoria colectiva como "Tomorrowland", y así salvar a la humanidad.
	Misión Imposible: Nación secreta	Mission: Impossible - Rogue Nation	2015	Con la FBI desuelta y Ethan Hunt (Tom Cruise) abandonado a su suerte, el equipo tiene que enfrentarse contra el Sindicato, una red de agentes especiales altamente preparados y entrenados. Estos grupos están empeñados en crear un nuevo orden mundial mediante una serie de ataques terroristas cada vez más graves. Ethan reúne a su equipo y a sus aliados con la agente británica renegada Ilsa Faust (Rebecca Ferguson), quien puede que sea o no miembro de esta nación secreta, mientras el grupo se va enfrentando a su misión más imposible hasta la fecha. ¿Cuánta entrega de la saga Misión Imposible.

