

Geolocalización y filtrado de farmacias de Farmashop

*Recuperación de Información y
Recomendaciones en la Web*

Diciembre 2015

Alejandro Añón - 4.719.805-0

Carina Indarte - 4.598.456-6

Juan Ignacio Larrambebere - 4.558.505-1

Gastón Salgado - 4.801.723-3

Índice

1. Introducción	2
1.1. Definición del problema	2
1.2. Solución planteada.....	2
2. Implementación	3
2.1. Herramientas y diseño.....	3
3. Casos de prueba	4
3.1. Caso de prueba 1	4
3.2. Caso de prueba 2	5
.....	5
3.3. Caso de prueba 3	5
3.4. Caso de prueba 4	6
4. Trabajo a futuro	6
5. Conclusiones	7
6. Manual de usuario	7
7. Referencias	8

1. Introducción

1.1. Definición del problema

Esta aplicación planea brindar una solución a la búsqueda y localización de farmacias en el Uruguay que cumplan con ciertos requisitos que desee el cliente. Estos requisitos pueden ir desde características comunes como la ubicación y el horario hasta servicios más específicos como los exámenes que realizan.

Se eligió la cadena de farmacias Farmashop por ser la cadena más grande en el Uruguay y con mayor información en su página web. A pesar del hecho que dicha página web provee un buscador de farmacias, se consideró importante ampliar este buscador brindando filtros avanzados sobre las características de las farmacias.

1.2. Solución planteada

La solución que se planteó fue obtener la información de las farmacias mediante web scraping¹ sobre la página de Farmashop² para luego procesarla y poder mostrarla en la aplicación así como también localizarla en un mapa. La información de las farmacias obtenida fue: número de sucursal (identificador), horario, dirección, teléfono y departamento. El barrio, aunque está asociado en la farmacia no aparece dentro de la información de la misma, por lo tanto se tuvo que obtener realizando la búsqueda por barrio en la página. Esta información fue obtenida de la pestaña "Farmacias" dentro de la página principal. También se obtuvo lo que se denominó como opciones o características de cada farmacia, las cuales podían ser: "Farmacia", "Megafarmacia", "24 horas", "Fertilidad" y "Perfumería importada", para luego poder filtrar las farmacias según la opción deseada.

Una vez obtenida la información básica de las farmacias se procedió a obtener la información relativa a los exámenes que se realizan en las mismas. Cabe destacar que si bien la opción "Atención Farmacéutica" se podía obtener de la pestaña "Farmacias", consideramos que la sección "Área de salud" contenía más información al respecto, por lo tanto se obtuvo de esa pestaña y fue considerado como un examen, ya que tenía atributos en común con ellos. La información obtenida de los exámenes fue: nombre, precio, descripción, farmacias en las que se realizan y el profesional a cargo en cada una de ellas. Los exámenes obtenidos fueron: "INR", "Celiaquía", "Glicemia", "Hemoglobina Glicosilada", "Hemoglobina" y "Colesterol y Triglicéridos".

Una vez obtenida toda la información se procedió a procesarla, esto implicó entre otras cosas obtener la latitud y la longitud de las farmacias a partir de la dirección para luego poder localizarlas en un mapa. Si bien esto dio buenos

resultados en la mayoría de los casos, hubo ciertos casos en los que la dirección no estaba completa o estaba mal formada, lo cual no permitió la correcta localización.

Los filtros implementados fueron la búsqueda por departamento, barrio, opciones y exámenes, permitiendo cualquier combinación entre ellos.

2. Implementación

2.1. Herramientas y diseño

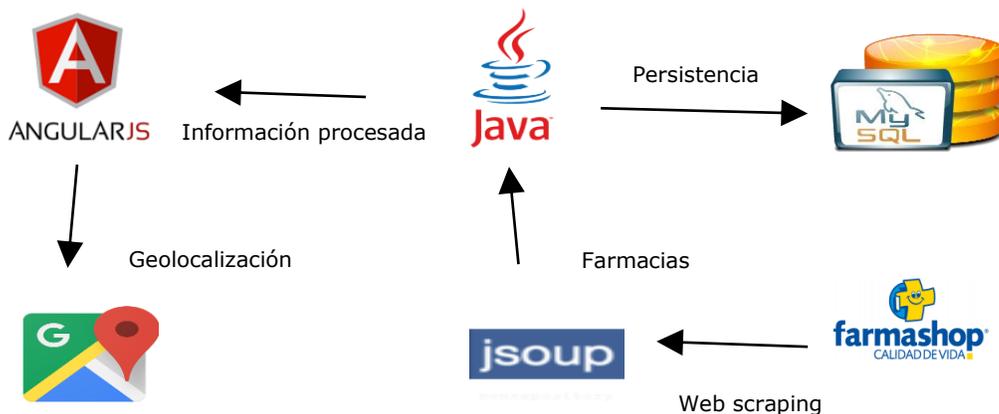
Para la aplicación se realizó un cliente fino y un servidor con la mayoría de la lógica. El back end se realizó en Java EE 7 ya que se contaba con experiencia previa con este lenguaje. Por otro lado, para el front end se decidió investigar y utilizar una tecnología de punta como Angular JS, obteniendo muy buenos resultados.

El back end es el responsable de realizar el web scraping sobre la página de Farmashop mediante la librería JSoup³, almacenar los resultados en una base de datos MySql relacional y realizar el filtrado de la información cuando el front end lo solicite. En contraparte el front end permite al usuario interactuar con el sistema; se encarga de listar las farmacias en pantalla y brindarle la opción de filtrar las farmacias según los criterios mencionados anteriormente. Entre ellos se comunican mediante web services REST.

El web scraping se realizó sobre el código HTML de la página de Farmashop. Para esto, se analizó el código fuente de la página y se extrajo la información deseada utilizando las etiquetas HTML del mismo. Es importante aclarar que el web scraping se realiza una vez y luego se interactúa con la información almacenada en la base.

El front end se integró con Google Maps, brindando la funcionalidad de geolocalizar las farmacias a partir de su dirección.

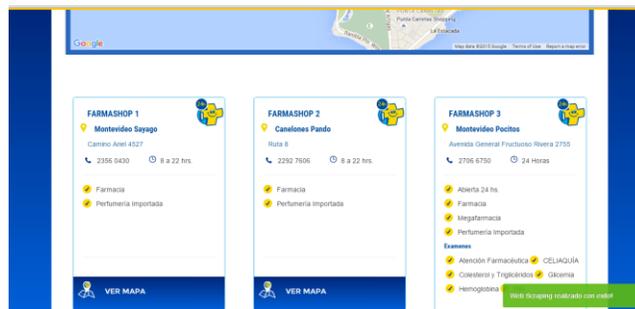
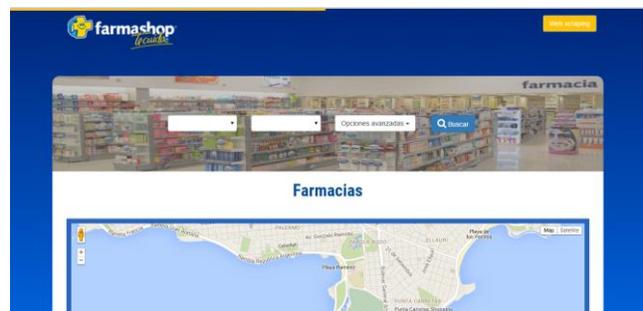
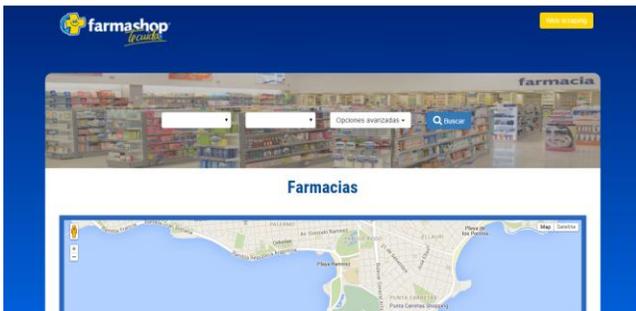
Se eligió como servidor de aplicaciones WildFly 8.



3. Casos de prueba

3.1. Caso de prueba 1

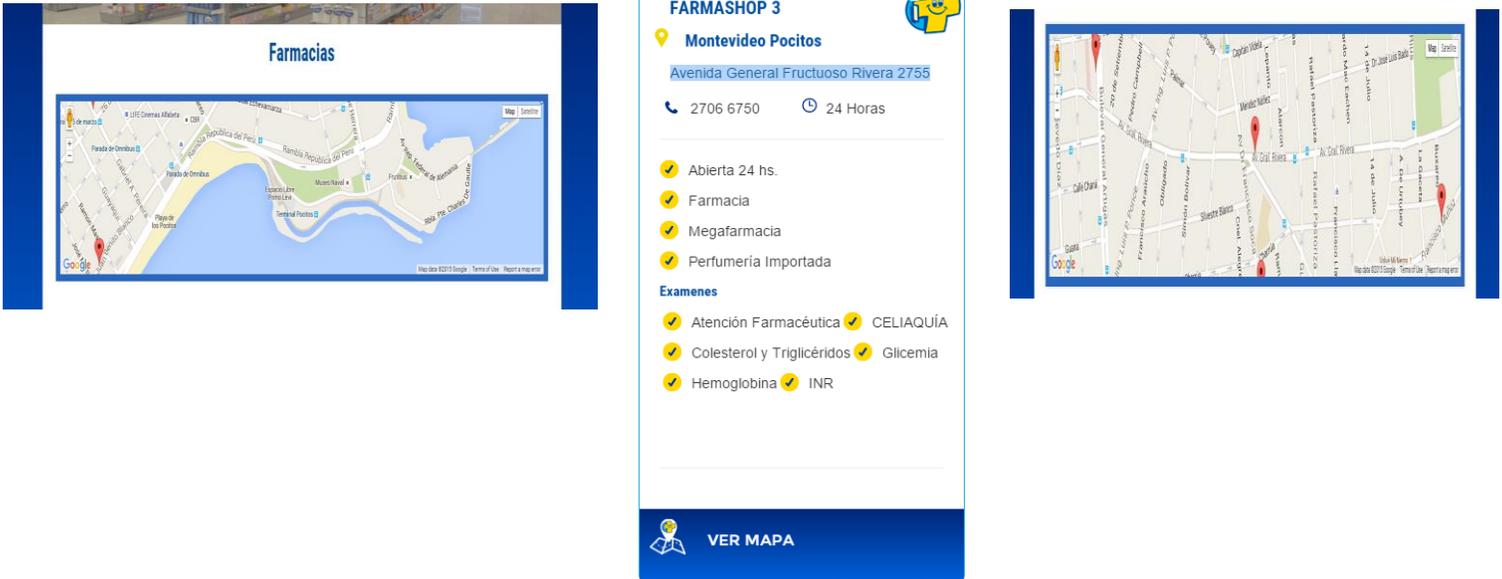
Al hacer click en el botón de “Web Scraping” se comienza a realizar las consultas necesarias al servidor de la cadena de farmacias para obtener la información. Se espera que la información de las farmacias se guarde en la base de datos del sistema y se listen las farmacias ubicando a cada una de ellas en el mapa.



Geolocalización y filtrado de farmacias de Farmashop

3.2. Caso de prueba 2

Se toma como precondition que exista la información de alguna farmacia en la base de datos del sistema. Al hacer click en una farmacia, el mapa se debe centrar sobre la misma.



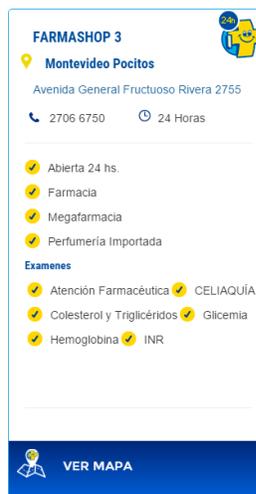
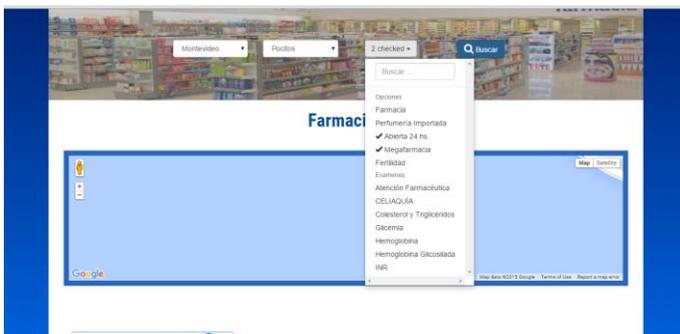
3.3. Caso de prueba 3

Se toma como precondition que exista la información de alguna farmacia en la base de datos del sistema. Al seleccionar un filtro y hacer click en el botón "Filtrar" se listan solamente la información de las farmacias que cumplen con ese filtro. Solo se deben mostrar en el mapa esas farmacias.



3.4. Caso de prueba 4

Se toma como precondition que exista la información de alguna farmacia en la base de datos del sistema. Al seleccionar varios filtros y hacer click en el botón "Filtrar", se listan solamente la información de las farmacias que cumplen con todos los filtros seleccionados a la vez. Solo se muestran en el mapa las farmacias listadas.



4. Trabajo a futuro

Consideramos que nuestra aplicación puede mejorarse de varias maneras. La primera y más importante sería agregar más cadenas de farmacias, logrando centralizar la información sobre las mismas, para poder brindarle al usuario más opciones para poder elegir mejor. Si bien agregar dichas cadenas sería un paso importante en el desarrollo y avance de la aplicación nos encontramos que no todas tienen página web y de tenerla no suelen estar tan completas como la página web de la cadena elegida.

Siguiendo con esta idea, consideramos que sería de gran utilidad que las farmacias pudieran tener una API REST, de la cual se pueda consumir la información de las farmacias, sin necesidad de tener que obtener la misma mediante web scraping, cuyo resultado puede variar al cambiar la página web.

Otro punto que consideramos importante para mejorar la experiencia del usuario es brindarle una aplicación para los teléfonos inteligentes, ya que permite el acceso inmediato a la información de manera sencilla. A su vez, otra característica

que podría ser de interés para el usuario sería la sugerencia de farmacias que se encuentren cerca del mediante la obtención de su ubicación.

5. Conclusiones

Cabe destacar el hecho de que muchas veces la información obtenida de un sitio no es lo suficientemente completa o correcta para ser luego utilizada en otra aplicación sin antes realizar algún tipo de procesamiento. Este fue el caso con las direcciones obtenidas de la página de Farmashop y el formato de dirección que utiliza Google para poder realizar la geolocalización. Si bien en este caso era sencillo de corregir, ya que la cantidad de farmacias era reducida, este escenario se vuelve inviable en el caso de contar con una cantidad mayor de farmacias.

Esto está relacionado con el hecho que no existe soporte para el usuario o desarrollador, es decir no existen API predefinidas por parte de las cadenas importantes de farmacias, lo cual dificulta el objetivo de centralizar y normalizar la información. Dado que el resultado del web scraping depende del código HTML de la página web y este puede variar, se deberían hacer chequeos regulares sobre las páginas para comprobar que no hayan cambiado y por lo tanto el web scraping deje de funcionar. Otra desventaja con respecto a este tema es que el web scraping se debe específicamente para una página web, es decir en el caso de agregar nuevas cadenas de farmacias se debería realizar un web scraping personalizado para cada una de ellas. A su vez, es importante tener en cuenta que no todas las cadenas brindan la misma información sobre sus farmacias, por lo tanto el normalizar la información se vuelve más complicado.

A pesar de esto, consideramos que de lograr la incorporación de nuevas farmacias se lograría obtener una aplicación con una utilidad real para los usuarios.

6. Manual de usuario

Cuando se inicia la aplicación por primera vez, la base de datos se encuentra vacía. Lo primero que se debe hacer es cargar los datos en la base de datos. Para esto, basta con dirigirse a la esquina superior derecha de la web, clicar el botón amarillo que dice "Web Scraping" y esperar a que cargue la barra amarilla en la parte superior de la pantalla; este proceso puede tardar hasta un minuto.

En caso de iniciar la aplicación pero que no sea la primera vez que se ejecuta y la base de datos ya se encuentre cargada, no es necesario clicar realizar el proceso anterior aunque si lo desea realizar no hay ningún problema.

Al finalizar el proceso anterior, la base de datos se encontrará llena con las farmacias de Farmashop y las mismas se listarán en pantalla. Ya es posible proceder a filtrarlas según se desee.

Existen tres filtros actualmente en el sistema:

- Por Departamento
- Por Barrio
- Opciones avanzadas

El primero se utiliza para filtrar las farmacias por el departamento en el que se encuentran. Por defecto viene preseleccionado Montevideo. El segundo filtro depende del valor que se encuentre seleccionado en el primero y contiene una lista de los barrios (en los que existen farmacias Farmashop) del departamento en cuestión. Por último, el tercer filtro contiene una lista con las características de las farmacias de Farmashop (por ej: 24 horas, megafarmacia, etc) y una lista con exámenes que se realizan en las mismas. Este último filtro, a diferencia de los anteriores, permite seleccionar más de una opción de modo que se puede buscar aquellas farmacias que sean 24 horas, megafarmacia y que tengan exámenes de Glicemia por ejemplo.

Los filtros no son excluyentes entre sí, es decir, se pueden combinar los mismos para generar una búsqueda más específica.

Si se desea ver la ubicación de una farmacia de interés, basta con clicar la opción "VER MAPA" que se encuentra debajo de la información de la misma, y automáticamente se muestra la farmacia en el mapa.

7. Referencias

¹ http://www.webopedia.com/TERM/W/Web_Scraping.html

² <http://www.farmashop.com.uy/>

³ <http://jsoup.org/>