

BUSCADOR SEMÁNTICO SAWA
MANUAL DE REFERENCIA

OMAR ERNESTO CABRERA ROSERO
JIMMY MATEO GUERRERO RESTREPO
MAURICIO FERNANDO BENAVIDES BENAVIDES
SILVIO RICARDO TIMARÁN PEREIRA





UNIVERSIDAD DE NARIÑO
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA DE SISTEMAS
SAN JUAN DE PASTO
2014




Sobre este trabajo



You are free:

-  to **Share** – to copy, distribute and transmit this work
-  to **Remix** – to adapt this work

Under the following conditions:

-  **Attribution** – You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work)
-  **Noncommercial** – You may not use this work for commercial purposes.
-  **Share Alike** – If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one.

Subject to conditions outlined in the license.

This work is licensed under the *Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported* License. To view a copy of this license, visit

<http://creativecommons.org/licenses/by-nc-sa/3.0/>

or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

CONTENIDO

1. Introducción	4
2. Contrucción de la aplicación	5
3. Proceso de Instalación	7

1 Introducción

La falta de significado que se maneja en la Web actual, dificulta la búsqueda eficiente de información. La web semántica ha comenzado a adquirir una gran importancia debido a que se quiere encontrar información de manera precisa y poder convertirla en información del conocimiento y representarla en recursos web que puedan estar disponibles en otras aplicaciones. Para lograr encontrar un significado claro en la búsqueda, la Web semántica hace uso de ontologías, que organizan la información de manera que se podrá interpretar lo que se quiere buscar y, por tanto, permitirá buscar e integrar datos mucho mejor que ahora.

Esta aplicación web esta basada en el uso de una ontología la cual fue construida para realizar busquedas inteligentes en los trabajos de grado de los estudiantes de pregrado de la Universidad de Nariño y servir como base para mejorar la eficiencia en la búsqueda de información sobre éste dominio.

2 Contrucción de la aplicación

El buscador semántico SAWA fue desarrollado en JavaEE¹ y liberado bajo licencia libre GPL3² el cual puede ser descargado del repositorio del portal github³ y se puede mirar el funcionamiento en la página⁴.

En la aplicación se usaron algoritmos como lematizadores y similitud de palabras para hacer corrección ortográfica, en caso de que el usuario tenga error de digitación, además de la construcción de un tesoro para que pueda hacer la búsqueda por sinónimos de palabras.

Las siguientes son las características específicas del software construido:

- Búsqueda General.
- Búsqueda por título.
- Búsqueda por autor.
- Auto completar palabras.
- Corrección de digitación.
- Búsqueda por sinónimos.
- Ordenamiento de resultados por mayor coincidencia.

En el diagrama de actividades de la Figura 2.1 se muestra como se realiza una búsqueda.

Para la construcción del software se tubo en cuenta la manera de realizar las consultas usando el lenguaje SPARQL y la extensión de postgresql pg_similarity⁵.

¹Java Platform, Enterprise Edition o Java EE (anteriormente conocido como Java 2 Platform, Enterprise Edition o J2EE hasta la versión 1.4

²<http://www.gnu.org/licenses/gpl.html>

³<https://github.com/poldrosky/Sawa>

⁴<http://ingenieria.udenar.edu.co:8080/Sawa/>

⁵<http://pgsimilarity.projects.pgfoundry.org>

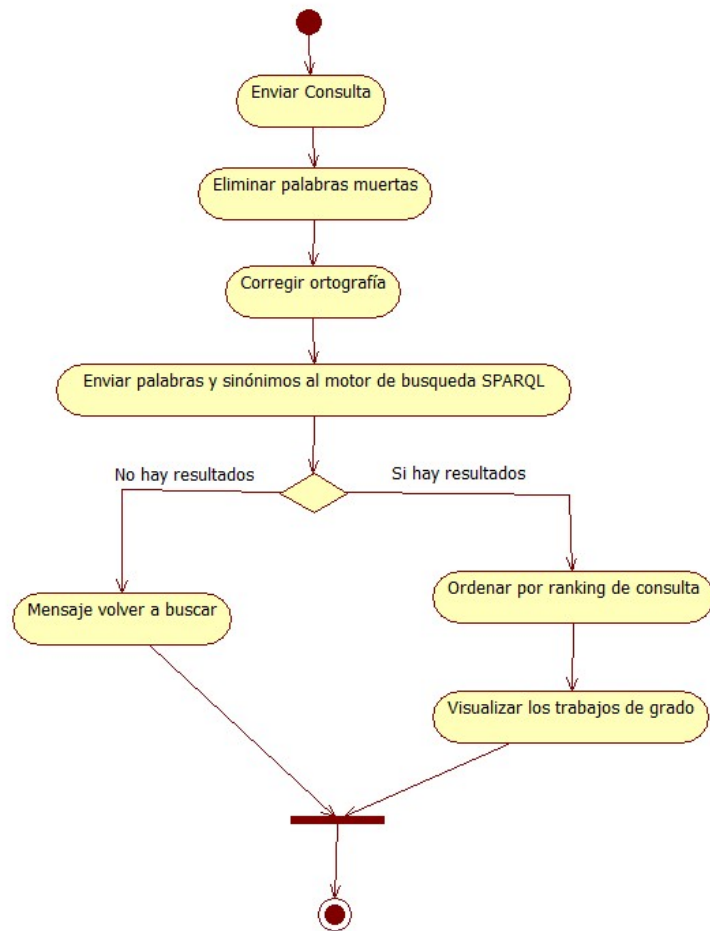


Figura 2.1: Diagrama de actividades de la búsqueda

3 Proceso de Instalación

Para poder instalar la aplicación debemos instalar PostgreSQL 9.1, Glassfish 3.1.2.2 y la extensión para postgresql llamada pg_similarity, la aplicación solo puede ser instalada bajo sistemas operativos GNU/Linux

Para la instalación de PostgreSQL 9.1 (Para Debian/Linux), en una terminal se ejecuta lo siguiente:

```
$ su
# apt-get install postgresql postgresql-client pgadmin3
postgresql-server-dev-all postgresql-contrib
```

Ahora se borra la contraseña para la cuenta de administrador de “postgres” para ello se ejecuta lo siguiente en la línea de comandos.

```
# su postgres -c psql template1
```

```
template1=# ALTER USER postgres WITH PASSWORD 'postgres1';
template1=#\q
```

Eso altera la contraseña dentro de la base de datos, ahora se tiene que hacer lo mismo para el usuario 'postgres' y colocar la misma contraseña que utilizó anteriormente.

```
# passwd -d postgres
#su postgres -c passwd
```

Ahora para instalar pg_similarity primero se descarga desde github de la siguiente manera

```
$ git clone https://github.com/eulerto/pg_similarity.git
```

Luego para que pueda ser usada la extensión se la compila como superusuario de la siguiente manera.

```
# su
# cd pg_similarity
# USE_PGXS=1 make
# USE_PGXS=1 make install
# su postgres
$ createdb sawa
```

Luego se crea las funciones y se restaura la copia de la base de datos sawa.sql

```
$ psql sawa
sawa=# create extension pg_similarity;
sawa=# \i sawa.sql
```

Instalación de Glassfish

Primero se descarga la versión 3.1.2.2 para GNU/Linux desde la página de Oracle¹ y se la ejecuta.

```
$ sh ogs-3.1.2.2-unix-ml.sh
```

Hay que descargar el driver JDBC de postgresql desde la página de PostgreSQL² y copiarlo en el directorio glassfish3/glassfish/domains/domain1/lib. Para iniciar el servidor se ejecuta lo siguiente

```
$ ./glassfish3/glassfish/bin/startserv
```

Con esto ya en el navegador se ingresa con la dirección <http://localhost:8080>, y entramos a la consola de administración, ingresamos usuario y contraseña de haber escrito una en la instalación de glassfish.

Ir a “Resources/JDBC/Connection Pools” y crear una nueva conexión con los datos que muestra la Figura 3.1 y luego clic en siguiente.

Seleccione el origen de datos de nombre de clase `org.postgresql.ds.PGConnectionPoolDataSource` y escribir a las siguientes propiedades adicionales como muestra la Figura 3.2.

Ya con esto guardamos las conexiones y damos clic en finalizar para guardar la conexión.

¹<http://www.oracle.com/technetwork/middleware/glassfish/downloads/ogs-3-1-1-downloads-439803.html>

²<http://jdbc.postgresql.org>

New JDBC Connection Pool (Step 1 of 2)
Identify the general settings for the connection pool.

General Settings

Pool Name:

Resource Type:
Must be specified if the datasource class implements more than 1 of the interface.

Database Driver Vendor:
Select or enter a database driver vendor

Inspect: ☐ **Enabled**
If enabled, data source or driver implementation class names will enable introspection.

Figura 3.1: Crear Nueva conexión

Name	Value
portNumber	5432
databaseName	sawa
dataSourceName	jdbc/sawa
roleName	
networkProtocol	
serverName	localhost
user	postgres
password	postgres1

Figura 3.2: Propiedades adicionales de conexión

Luego en “Resources/JDBC/JDBC Resources” escribimos en nombre JNDI y escogemos el Pool Name creado anteriormente como lo muestra la Figura 3.3.

Ya por último en “Applications” escogemos en donde tenemos almacenada la aplicación web que tiene extensión “.war” como muestra la Figura 3.4.

Si todo sale bien nos aparecerá una ventana con la url de la aplicación como muestra la Figura 3.5. y al ingresar a la aplicación nos desplegara la ventana de inicio de la aplicación como lo muestra la Figura 3.6.

Nota: Si no desea instalar la aplicación y quiere ver su funcionamiento la aplicación esta alojada en <http://ingenieria.udenar.edu.co:8080/Sawa/>

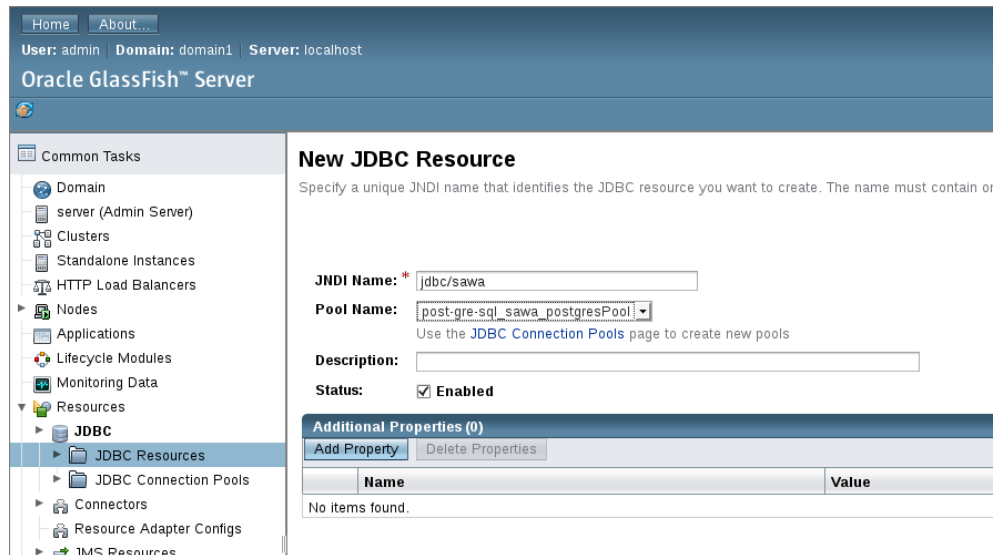


Figura 3.3: Recursos JDBC

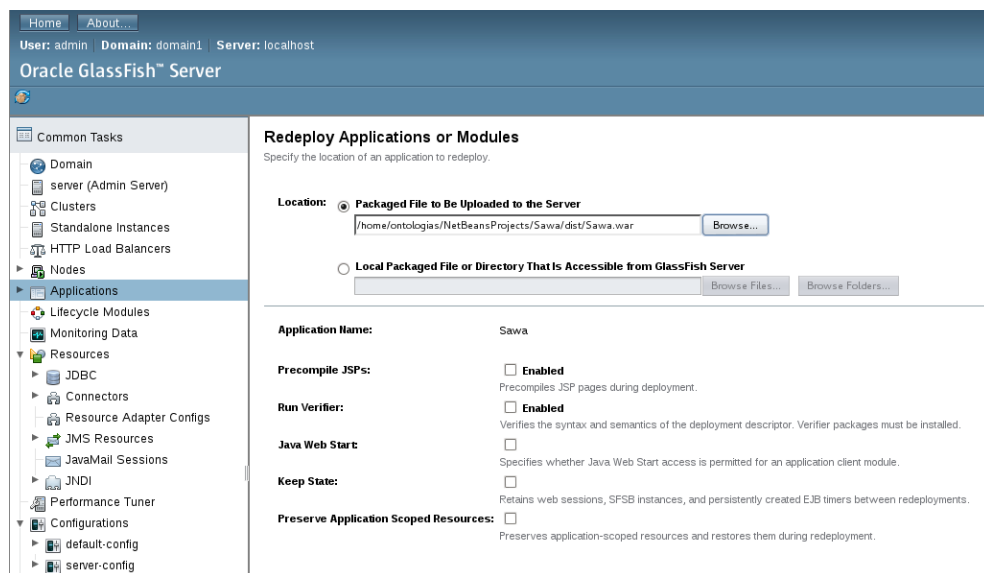


Figura 3.4: Subir Aplicación

Web Application Links

If the server or listener is not running, the link may not work. In this event, check

Application Name: Sawa

Links:
[server] <http://grias:8080/Sawa>
[server] <https://grias:8181/Sawa>

Figura 3.5: Url de la Aplicación



The screenshot displays the web interface of the 'Buscador Semántico SAVA' application. At the top, a banner features the 'Universidad de Nariño' logo and the tagline 'Entregamos Futuro con Calidad' alongside a scenic image of a campus. Below the banner, the interface is divided into two main sections. On the left, a sidebar contains the university's official seal and lists the authors (Omar Cabrera, Mateo Guerrero, Mauricio Benavides) and the director (Ricardo Timarín Pereira, PhD). The main content area on the right is titled 'Buscador Semántico SAVA' and includes three tabs: 'GENERAL' (selected), 'TITULO', and 'AUTOR'. A search input field labeled 'Buscar' is positioned below the tabs, with a 'Buscar' button to its right. At the bottom of the page, a footer identifies the 'Grupo de Investigación Aplicada en Sistemas (GRIAS)' at the 'Universidad de Nariño'.

Figura 3.6: Aplicación