



SAWA: Ontología para la Gestión de Conocimiento sobre Trabajos de Grado

Omar Ernesto Cabrera Rosero¹
omarcabrera@udenar.edu.co

¹Grupo de Investigación Aplicada en Sistemas (GRIAS)
Universidad de Nariño
San Juan de Pasto, Colombia

2 de marzo de 2014

Esquema

1 Generalidades

2 Objetivo

3 Metodología

- Recolección de Información
- Implementación de la Ontología
- Creación de instancias
- Construcción de la Aplicación
- Ejecución de Pruebas

4 Productos Obtenidos

5 Bibliografía

Esquema

- 1 Generalidades
- 2 Objetivo
- 3 Metodología
 - Recolección de Información
 - Implementación de la Ontología
 - Creación de instancias
 - Construcción de la Aplicación
 - Ejecución de Pruebas
- 4 Productos Obtenidos
- 5 Bibliografía

Esquema

- 1 Generalidades
- 2 Objetivo
- 3 Metodología
 - Recolección de Información
 - Implementación de la Ontología
 - Creación de instancias
 - Construcción de la Aplicación
 - Ejecución de Pruebas
- 4 Productos Obtenidos
- 5 Bibliografía

Esquema

- 1 Generalidades
- 2 Objetivo
- 3 Metodología
 - Recolección de Información
 - Implementación de la Ontología
 - Creación de instancias
 - Construcción de la Aplicación
 - Ejecución de Pruebas
- 4 Productos Obtenidos
- 5 Bibliografía

Esquema

- 1 Generalidades
- 2 Objetivo
- 3 Metodología
 - Recolección de Información
 - Implementación de la Ontología
 - Creación de instancias
 - Construcción de la Aplicación
 - Ejecución de Pruebas
- 4 Productos Obtenidos
- 5 Bibliografía

¿Qué es Ontologías

Ontología

- Según (Gruber, 1993) “Una especificación explícita de una conceptualización”, en donde una conceptualización es una visión abstracta y simplificada del mundo que queremos representar para algún propósito
- (Studer, Benjamins, y Fensel, 1998) “Una especificación explícita y formal de una conceptualización compartida”

- Clarity
- Coherence
- Extensibility
- Minimum Coding Bias
- Minimum ontological compromise

Elementos

- Clase
- Subclases
- Clases Disjuntas
- Propiedades
- Rango
- Dominio
- Instancias

Tipos

- Cantidad y tipo de estructura
 - Terminología
 - Información
 - Modelado del conocimiento
- Conceptualización
 - Aplicación
 - Dominio
 - Genéricas
 - Representación

Objetivo

Soportar la búsqueda inteligente de las consultas sobre los trabajos de grado de la Universidad de Nariño mediante la construcción de una Ontología de aplicación en Trabajos de Grado utilizando la herramienta de software libre Protégé.

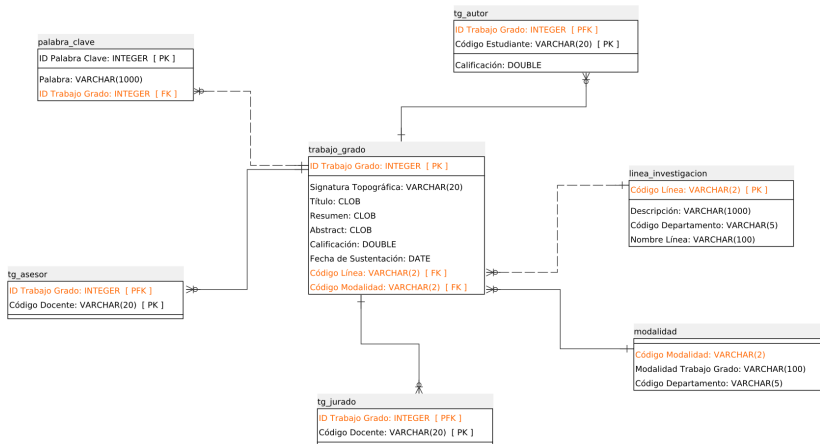
Recolección de Información

- Se recolectó la información necesaria sobre 202 trabajos de grado que se encontraban en medio digital
- Se diseñó la base de datos con el diagrama entidad relación ^a y se lo implemento usando PostgreSQL^b
- Los atributos de la base de datos contenían información como: título, autores, jurados, modalidad, nota Obtenida, línea de investigación, palabras clave, entre otros.

^a<https://sites.google.com/site/repositoriosawa/>

^b<http://http://www.postgresql.org/>

Entidad Trabajo de grado



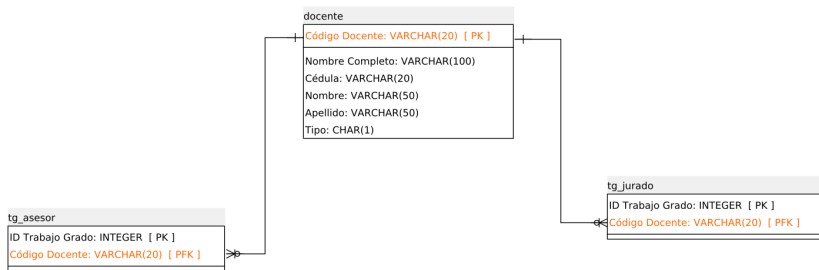
Entidad facultades, departamentos y programas



Entidad autores



Entidad asesores y jurados



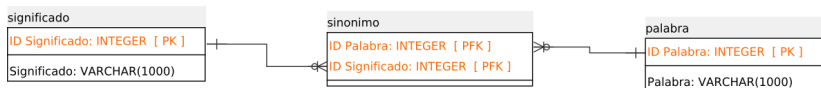
Entidad modalidad y línea de investigación



Entidad palabra clave



Entidad tesauro

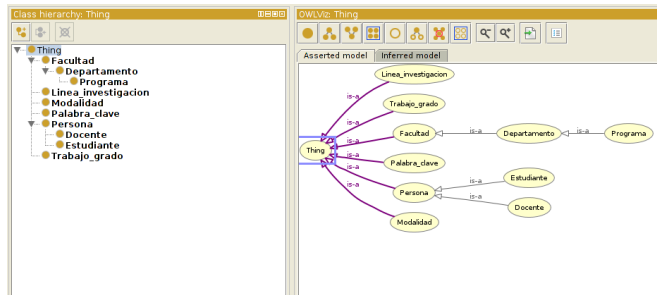


Implementación de la ontología

- Se uso la herramienta libre Protégé^a
- Se generó la ontología en formato OWL
- Se utilizó una metodología descrita en el paper “Una propuesta de metaontología para la educción de requisitos” (Zapata,Giraldo, y Mesa, 2010).

^a<http://protege.stanford.edu/>

Estructura General



Superclass Facultad

The screenshot displays the Protege ontology editor interface. The top tabs include Metadata(TesisGrado.owl), OWLClasses, Properties, Individuals, and Forms. The left pane, titled 'SUBCLASS EXPLORER', shows the 'Asserted Hierarchy' for the project 'OntologiaIntancia'. The hierarchy starts with 'owl:Thing' and branches into 'Facultad', which further branches into 'Departamento' and 'Programa'. Below these are 'Linea_investigacion', 'Modalidad', 'Palabra_clave', 'Persona', and 'Trabajo_grado'. The right pane, titled 'CLASS EDITOR for Facultad (instance of owl:Class)', shows the 'For Class' field with the URI 'http://www.owl-ontologies.com/TesisGrado.owl#Fa'. Below this, there is a table for 'Property' with one entry: 'rdfs:comment'. At the bottom of the right pane, the 'Superclasses' section shows 'owl:Thing' as the superclass.

Metadata(TesisGrado.owl) | OWLClasses | Properties | Individuals | Forms

SUBCLASS EXPLORER

For Project: ● OntologiaIntancia

Asserted Hierarchy

- owl:Thing
 - Facultad
 - Departamento
 - Programa
 - Linea_investigacion
 - Modalidad
 - Palabra_clave
 - Persona
 - Trabajo_grado

CLASS EDITOR for Facultad (instance of owl:Class)

For Class: <http://www.owl-ontologies.com/TesisGrado.owl#Fa>

Property
rdfs:comment

Superclasses

- owl:Thing

Class Departamento

SUBCLASS EXPLORER
For Project: **OntologiaIntancia**

Asserted Hierarchy

- owl:Thing
 - Facultad
 - Departamento**
 - Programa
 - Linea_Investigacion
 - Modalidad
 - Palabra_clave
 - Persona
 - Docente
 - Estudiante
 - Trabajo_grado

CLASS EDITOR for Departamento (instance of owl:Class)
For Class: <http://www.owl-ontologies.com/TesisGrado.owl#Departamento>

Property	Value
rdfs:comment	

Properties:

- codigo_departamento (multiple string)
- Especifica (multiple Modalidad)
- nombre_departamento (multiple string)
- Pertence (multiple Facultad)
- registra (multiple Linea_Investigacion)
- codigo_facultad (multiple string)
- nombre_facultad (multiple string)
- Tiene (multiple Departamento)

URI: http://www.owl-ontologies.com/TesisGrado.owl#codigo_facultad
ontology: <http://www.owl-ontologies.com/TesisGrado.owl>
location: main ontology [OntologiaIntancia]

Superclasses

- Facultad

Subclasses

- Programa
- Facultad
- Docente

Clase Programa

The screenshot displays the Protege ontology editor interface. On the left, the 'SUBCLASS EXPLORER' pane shows the 'Asserted Hierarchy' for the project 'OntologiaIntancia'. The hierarchy is as follows:

- owl:Thing
 - Facultad
 - Departamento
 - Programa**
 - Linea_investigacion
 - Modalidad
 - Palabra_clave
 - Persona
 - Docente
 - Estudiante
 - Trabajo_grado

The 'CLASS EDITOR for Programa (instance of owl:Class)' pane on the right shows the class definition for 'Programa'. It includes a table for 'Property' with the entry 'rdfs:comment'. Below this, a list of properties is shown, each with its domain and cardinality:

- codigo_programa (multiple string)
- Es_estudiado (multiple Estudiante)
- nombre_programa (multiple string)
- codigo_departamento (multiple string)
- codigo_facultad (multiple string)
- Especifica (multiple Modalidad)
- nombre_departamento (multiple string)
- nombre_facultad (multiple string)
- Pertence (multiple Facultad)
- registra (multiple Linea_investigacion)
- Tiene

A tooltip is visible over the 'registra' property, displaying the following information:

- URI: <http://www.owl-ontologies.com/TesisGrado.owl#registra>
- ontology: <http://www.owl-ontologies.com/TesisGrado.owl>
- location: main ontology: [OntologiaIntancia]

At the bottom of the class editor, the 'superclasses' section shows 'Departamento' and 'Trabajo_grado' as superclasses of 'Programa'.

Clase línea de investigación

The screenshot displays the Protégé OWL editor interface. On the left, the 'SUBCLASS EXPLORER' pane shows the 'Asserted Hierarchy' for the project 'OntologiaIntancia'. The hierarchy is as follows:

- owl:Thing
 - Facultad
 - Departamento
 - Programa
 - Linea_investigacion**
 - Modalidad
 - Palabra_clave
 - Persona
 - Docente
 - Estudiante
 - Trabajo_grado

The 'Linea_investigacion' class is highlighted. The main 'CLASS EDITOR for Linea_investigacion (instance of owl:Class)' pane shows the class URI: `http://www.owl-ontologies.com/TesisGrado.owl#Linea_investigacion`. Below the URI, there is a table for properties:

Property
<code>rdfs:comment</code>

Below the table, the class properties are listed:

- `codigo_linea` (multiple string)
- `descripcion` (multiple string)
- `es_inscrita` (multiple Trabajo_grado)
- `nombre_linea` (multiple string)
- `Pertenece_a` (multiple Departamento)

At the bottom, the 'Superclasses' pane shows the hierarchy of superclasses for 'Linea_investigacion':

- owl:Thing
 - Trabajo_grado
 - Programa
 - Modalidad

Superclase Modalidad

The screenshot displays a web ontology editor interface with two main panels: the SUBCLASS EXPLORER on the left and the CLASS EDITOR for 'Modalidad' on the right.

SUBCLASS EXPLORER
For Project: **OntologiaIntancia**

Asserted Hierarchy

- owl:Thing
 - Facultad
 - Departamento
 - Programa
 - Linea_investigacion
 - Modalidad**
 - Palabra_clave
 - Persona
 - Docente
 - Estudiante
 - Trabajo_grado

CLASS EDITOR for Modalidad (instance of owl:Class)
For Class: <http://www.owl-ontologies.com/TesisCrado.owl#Modalidad>

Property

Property
<code>rdfs:comment</code>

Properties of Modalidad:

- `codigo_modalidad` (multiple string)
- `Es_del_departamento` (multiple Departamento)
- `nombre_modalidad` (multiple string)
- `Pertenece_al_trabajo` (multiple Trabajo_grado)

Superclasses

- owl:Thing
- Trabajo_grado
- Programa

Superclass Palabra Clave

The screenshot displays the Protege ontology editor interface. On the left is the 'SUBCLASS EXPLORER' pane, showing a hierarchy for the project 'OntologiaIntancia'. The hierarchy starts with 'owl:Thing', followed by 'Facultad', 'Departamento', 'Programa', 'Linea_investigacion', 'Modalidad', and 'Palabra_clave' (which is highlighted). Below 'Palabra_clave' are 'Persona', 'Docente', 'Estudiante', and 'Trabajo_grado'. On the right is the 'CLASS EDITOR for Palabra_clave (instance of owl:Class)'. It shows the URI 'http://www.owl-ontologies.com/TesisCrado.owl#Palabra_clave'. Below this is a table for properties, with one row for 'rdfs:comment'. The bottom section of the class editor lists three properties: 'describen' (multiple Trabajo_grado), 'significado' (multiple string), and 'sinonimo' (multiple string), with the last one highlighted. At the bottom of the class editor is a 'Superclasses' section showing 'owl:Thing' as the superclass.

SUBCLASS EXPLORER
For Project: **OntologiaIntancia**

Asserted Hierarchy

- owl:Thing
 - Facultad
 - Departamento
 - Programa
 - Linea_investigacion
 - Modalidad
 - Palabra_clave**
 - Persona
 - Docente
 - Estudiante
 - Trabajo_grado

CLASS EDITOR for Palabra_clave (instance of owl:Class)
For Class: http://www.owl-ontologies.com/TesisCrado.owl#Palabra_clave

Property
rdfs:comment

Properties:

- describen (multiple Trabajo_grado)
- significado (multiple string)
- sinonimo (multiple string)**

Superclasses

- owl:Thing

Superclass persona

The screenshot displays the Protege ontology editor interface. On the left, the 'SUBCLASS EXPLORER' pane shows the 'Asserted Hierarchy' for the project 'OntologiaIntancia'. The hierarchy starts with 'owl:Thing' as the root, followed by 'Facultad', 'Departamento', 'Programa', 'Linea_investigacion', 'Modalidad', 'Palabra_clave', and 'Persona'. 'Persona' is currently selected and highlighted. Below 'Persona' are its subclasses: 'Docente', 'Estudiante', and 'Trabajo_grado'.

The main 'CLASS EDITOR for Persona' pane shows the class definition for 'http://www.owl-ontologies.com/TesisGrado.owl#Persona'. It includes a table for properties, with 'rdfs:comment' listed. Below the table, the class is defined with three properties: 'apellido_persona' (multiple string), 'identificacion_persona' (multiple string), and 'nombre_persona' (multiple string). At the bottom, the 'Superclasses' section shows 'owl:Thing' as the superclass.

SUBCLASS EXPLORER
For Project: **OntologiaIntancia**

Asserted Hierarchy

- owl:Thing
 - Facultad
 - Departamento
 - Programa
 - Linea_investigacion
 - Modalidad
 - Palabra_clave
 - Persona**
 - Docente
 - Estudiante
 - Trabajo_grado

CLASS EDITOR for Persona (instance of owl:Class)
For Class: <http://www.owl-ontologies.com/TesisGrado.owl#Persona>

Property
rdfs:comment

apellido_persona (multiple string)
identificacion_persona (multiple string)
nombre_persona (multiple string)

Superclasses

- owl:Thing

Clase Docente

The screenshot displays the Protege software interface, specifically the CLASS EDITOR for the 'Docente' class. The interface is divided into several panels:

- SUBCLASS EXPLORER:** Located on the left, it shows the asserted hierarchy for the project 'OntologiaIntancia'. The hierarchy is as follows:
 - owl:Thing
 - Facultad
 - Departamento
 - Programa
 - Linea_investigacion
 - Modalidad
 - Palabra_clave
 - Persona
 - Docente** (selected)
 - Estudiante
 - Trabajo_grado

- CLASS EDITOR for Docente (instance of owl:Class):** The main panel on the right, titled 'For Class: http://www.owl-ontologies.com/TesisGrado.owl#Docente'. It contains a table for properties:

Property
rdfs:comment
- Superclasses:** A panel at the bottom right showing the superclasses of the 'Docente' class:
- Persona
- Trabajo_grado
- Programa
- Modalidad
- Linea_investigacion

Clase Estudiante

The screenshot displays the Protege ontology editor interface. On the left, the 'SUBCLASS EXPLORER' pane shows the 'Asserted Hierarchy' for the project 'OntologiaIntancia'. The hierarchy is as follows:

- owl:Thing
 - Facultad
 - Departamento
 - Programa
 - Linea_investigacion
 - Modalidad
 - Palabra_clave
 - Persona
 - Docente
 - Estudiante** (highlighted)
 - Trabajo_grado

The main 'CLASS EDITOR for Estudiante' pane shows the class definition for 'http://www.owl-ontologies.com/TesisGrado.owl#Estudiante'. It includes a table for properties:

Property
<code>rdfs:comment</code>

Below the table, a list of properties is shown with their domains and cardinalities:

- `calificacion` (multiple float)
- `codigo_estudiante` (multiple string)
- `fecha_grado` (multiple date)
- `Pertenece` (multiple Programa)
- `Realiza` (multiple Trabajo_grado)
- `apellido_persona` (multiple string)
- `identificacion_persona` (multiple string)
- `nombre_persona` (multiple string)

At the bottom, the 'Superclasses' pane shows that 'Estudiante' is a subclass of 'Persona'.

Superclase trabajo de grado

SUBCLASS EXPLORER

For Project: OntologiaIntancia

Asserted Hierarchy

- owl:Thing
 - Facultad
 - Departamento
 - Programa
 - Linea_investigacion
 - Modalidad
 - Palabra_clave
 - Persona
 - Docente
 - Estudiante
 - Trabajo_grado

CLASS EDITOR for Trabajo_grado (instance of owl:Class)

For Class: http://www.owl-ontologies.com/TesisGrado.owl#Trabajo_grado

Property	
rdfs:comment	

- Es_calificado (multiple Docente)
- Es_dirigido (multiple Docente)
- Es_inscrito (multiple Modalidad)
- Es_realizado (multiple Estudiante)
- fecha_sustentacion (multiple date)
- id_trabajo (multiple int)
- resumen (multiple string)
- Se_inscribe (multiple Linea_investigacion)
- signature_topografica (multiple string)
- tiene (multiple Palabra_clave)
- titulo (multiple string)

Superclasses

- owl:Thing

Creación de instancias

Para la creación de las instancias se utilizó un script, tomando la información de la base de datos que fue construida anteriormente.

Construcción de la Aplicación

Para realizar las pruebas a la ontología se construyó un buscador^{1 2} desarrollado en JavaEE³ y liberado bajo licencia libre GPL3⁴ . Para realizar las consultas se utilizó el lenguaje SPARQL y la extensión de postgresql pg_similarity⁵

¹<https://github.com/poldrosky/Sawa>

²<http://ingenieria.udenar.edu.co:8080/Sawa/>

³Java Platform, Enterprise Edition o Java EE (anteriormente conocido como Java 2 Platform, Enterprise Edition o J2EE hasta la versión 1.4

⁴<http://www.gnu.org/licenses/gpl.html>

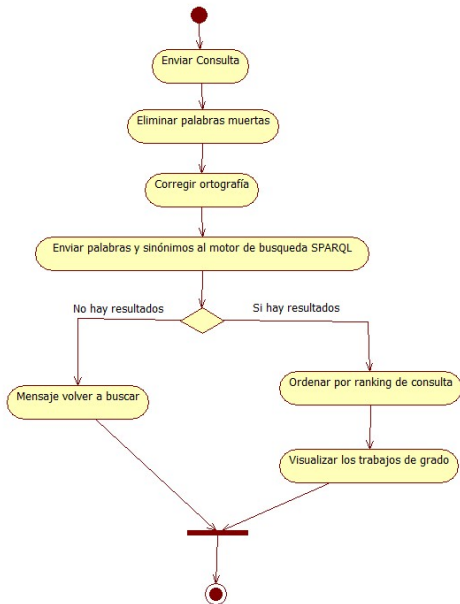
⁵<http://pgsimilarity.projects.pgfoundry.org>.

Construcción de la aplicación

Se utilizaron algoritmos como lematizadores y similitud de palabras para hacer corrección ortográfica, construcción de un tesoro para que pueda hacer la búsqueda por sinónimos de palabras. El software permite:

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

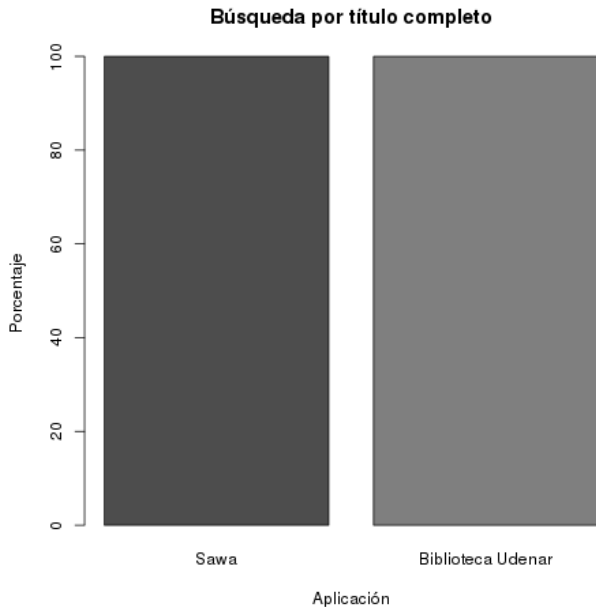
Diagrama de Actividades



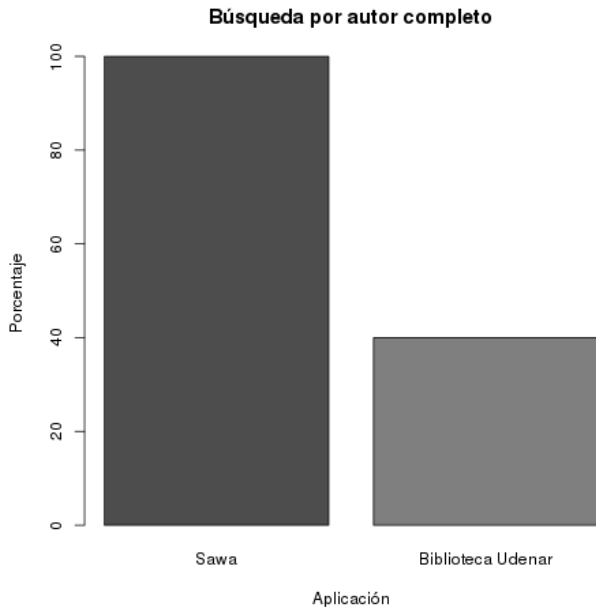
Ejecución de Pruebas

- Se realizaron siete casos de prueba con diez iteraciones cada una, para los cuales se elaboró una tabla comparativa para ver la eficiencia del buscador desarrollado en esta investigación y el buscador de la biblioteca de la Universidad de Nariño.
- Las pruebas se hicieron llevando casos de prueba y se calificó como éxito o fracaso, teniendo en cuenta que el éxito se lo califica si la búsqueda a realizar está en los quince primeros resultados.

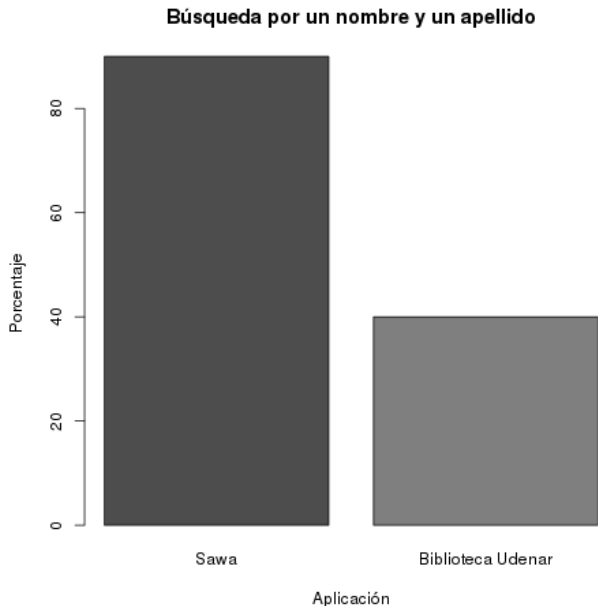
Búsqueda por título completo



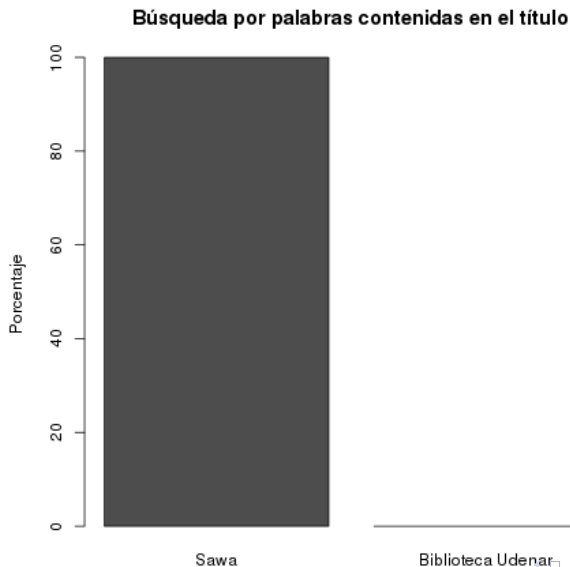
Búsqueda por autor completo



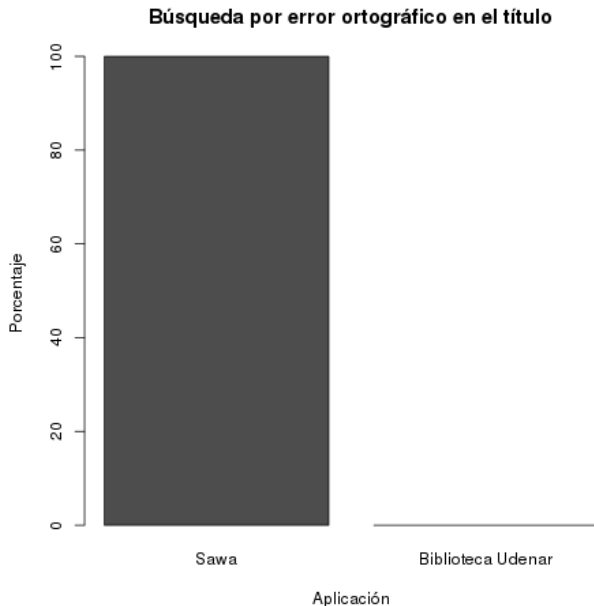
Búsqueda por un nombre y un apellido



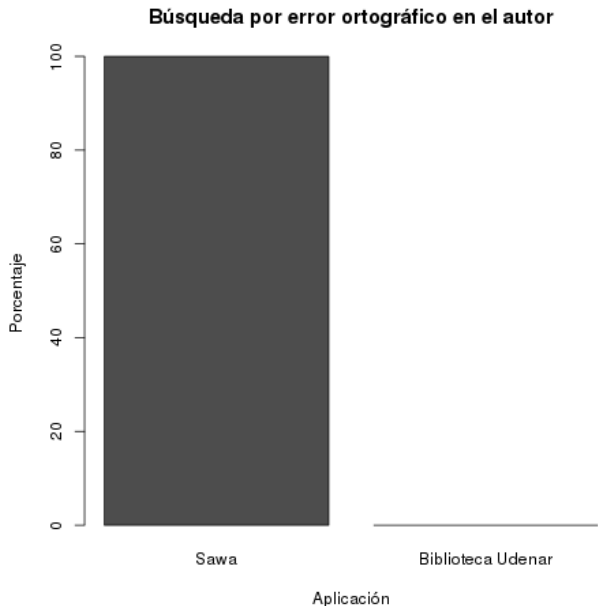
Búsqueda por palabras contenidas en el título



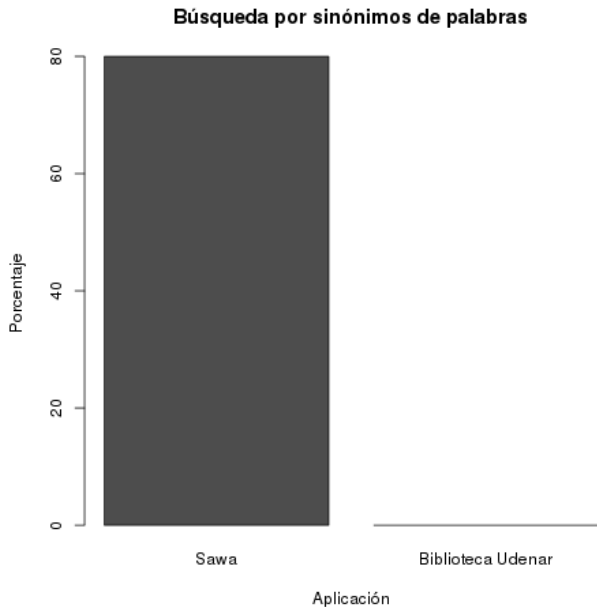
Búsqueda por error ortográfico en el título



Búsqueda por error ortográfico en el autor



Búsqueda por sinónimos de palabra



Productos Obtenidos

- Se apropió el conocimiento sobre la funcionalidad y desempeño de la herramienta libre Protégé en el desarrollo de ontologías.
- Se creó una base de datos con toda la información de los trabajos de grado del programa de Ingeniería de Sistemas.
- Se construyó e implementó la ontología para realizar las búsquedas de los trabajos de grado.
- Se realizó un buscador semántico que muestra el mejoramiento en el sistema de consultas de trabajos de grado.

Bibliografía



A. Alcina, E. Valero, and E. Rambla.
Terminología y sociedad del conocimiento.
Peter Lang Pub Incorporated, 2009.



Thomas R. Gruber.
Toward principles for the design of ontologies used for
knowledge sharing.
1993.



Rudi Studer, V.Richard Benjamins, and Dieter Fensel.
Knowledge engineering: Principles and methods.
Data and Knowledge Engineering, 25(1-2):161 – 197,
1998.



Carlos M Zapata, Gloria L Giraldo, and Jhon E Mesa.
Una propuesta de metaontología para la educación de
requisitos.
Ingeniare. Revista chilena de ingeniería, 2010.