



UNIVERSIDAD NACIONAL DEL ALTIPLANO DE PUNO

**FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA,
ELECTRÓNICA Y SISTEMAS**

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



**PROTOTIPO BASADO EN WEB SEMÁNTICA PARA BUSCAR
MANUALES DE PROGRAMACIÓN PARA LA ESCUELA
PROFESIONAL DE INGENIERÍA DE SISTEMAS DE LA
UNIVERSIDAD NACIONAL DEL ALTIPLANO**

TESIS

PRESENTADA POR:

Bach. JEAN MARCO CASTILLO ALEJO

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO DE SISTEMAS

PUNO – PERÚ

2021



DEDICATORIA

A mis padres Marco Castillo e Hilda Alejo, por ser el pilar fundamental en todo lo que soy, en toda mi educación tanto académica y como de la vida, por su incondicional apoyo y comprensión perfectamente mantenido a través del tiempo.

Todo este trabajo ha sido posible gracias a ellos.

Jean Marco Castillo Alejo



AGRADECIMIENTOS

A la Universidad Nacional del Altiplano y a la Escuela Profesional de Ingeniería de Sistemas, en su seno viví una de las etapas más hermosas de mi vida, la cual estará en mi recuerdo por siempre.

A mis jurados, M.Sc. Hugo Yosef Gómez Quispe, Ing. Aldo Hernan Zanabria Galvez y el M.Sc. Edwin Fredy Calderon Vilca, por sus correcciones y recomendaciones brindadas para la realización de esta investigación.

A mi director/asesor de tesis, M.Sc. Adolfo Carlos Jimenez Chura, por el tiempo dedicado y acompañamiento en el desarrollo de esta investigación.

Jean Marco Castillo Alejo



ÍNDICE GENERAL

DEDICATORIA

AGRADECIMIENTOS

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS¹⁰

ÍNDICE DE ACRÓNIMOS

RESUMEN 12

ABSTRACT..... 13

CAPITULO I

INTRODUCCIÓN

1.1. PLANTEAMIENTO DEL PROBLEMA 15

1.2. FORMULACIÓN DEL PROBLEMA 16

1.3. OBJETIVOS DE LA INVESTIGACIÓN..... 16

1.3.1. Objetivo general 16

1.3.2. Objetivos específicos 16

1.4. HIPÓTESIS DE LA INVESTIGACIÓN..... 17

1.5. JUSTIFICACIÓN DE LA INVESTIGACIÓN..... 17

CAPITULO II

REVISIÓN DE LA LITERATURA

2.1. ANTECEDENTES..... 18

2.2. BASE TEÓRICA 19

2.2.1. ¿Qué es la web?..... 19

2.2.2. Comunicación con Datos..... 20

2.2.3. La web semántica 21

2.2.4. La evolución de la Web 22



2.2.5.	Beneficios de la Web Semántica	23
2.2.6.	Arquitectura de la Web Semántica	23
2.2.6.1.	Unicode + URI	24
2.2.6.2.	XML+NS+XML SCHEMA.....	25
2.2.6.3.	RDF + rdfschema	25
2.2.6.4.	Ontology vocabulary	25
2.2.6.5.	Logic.	26
2.2.6.6.	Proof.....	26
2.2.6.7.	Trust (+ Digital Signature).....	26
2.2.6.8.	Web Semántica Confiable.....	26
2.2.6.9.	Firma y cifrado (Firma digital).	27
2.2.7.	Ontologías.....	27
2.2.8.	Tipos de Ontologías.....	28
2.2.8.1.	Ontología filosófica Histórica.....	28
2.2.8.2.	Ontología filosófica contemporánea	29
2.2.8.3.	Ontologías formales versus materiales.....	30
2.2.8.4.	Ontología de dominio.....	31
2.2.8.5.	Clasificación por el conocimiento que contienen	32
2.2.8.6.	Clasificación por el tipo de estructura.....	32
2.2.8.7.	Clasificación por motivación	33
2.2.8.8.	Ontologías genéricas:	33
2.2.8.9.	Por el grado de formalidad de la ontología	35
2.2.9.	Elementos de una ontología.....	36
2.2.9.1.	Clases.	36
2.2.9.2.	Atributos.....	36
2.2.9.3.	Relaciones.	37



2.2.9.4.	Axiomas.....	37
2.2.9.5.	Instancias.....	37
2.2.10.	Lenguajes para la representación de ontologías.....	37
2.2.10.1.	Simple HTML Ontology Extensions (SHOE).....	38
2.2.10.2.	Resource Description Framework (RDF).....	38
2.2.11.	El modelo RDF.....	39
2.2.12.	El Lenguaje de Ontologías Web (OWL).....	40
2.2.13.	Información de modelado en software.....	41
2.3.	MODELOS DE RECUPERACIÓN DE LA INFORMACIÓN.....	42
2.3.1.	Modelo booleano.....	42
2.3.2.	Modelo Vectorial.....	42
2.3.3.	Modelo probabilístico.....	43
2.3.4.	Modelos para documentos estructurados.....	44
2.4.	HERRAMIENTAS DE DESARROLLO.....	46
2.4.1.	RDF/XML.....	46
2.4.2.	OWL.....	46
2.4.3.	PROTÉGÉ.....	47
2.4.4.	JOWL.....	47
2.5.	DEFINICIÓN DE TERMINOS BÁSICOS.....	47
2.5.1.	XML.....	47
2.5.2.	XML Schema.....	48
2.5.3.	RDF.....	48
2.5.4.	RDF Schema.....	48
2.5.5.	OWL.....	48



CAPITULO III

MATERIALES Y MÉTODOS

3.1. TIPO Y DISEÑO DE INVESTIGACIÓN	49
3.1.1. Tipo del problema de Investigación	49
3.1.2. Diseño del problema de Investigación	49
3.2. POBLACIÓN Y MUESTRA DE INVESTIGACIÓN	49
3.2.1. Población de la investigación	49
3.2.2. Muestra	50
3.2.3. Método de recolección de datos	50

CAPITULO IV

RESULTADOS Y DISCUSION

4.1. ANÁLISIS DEL PROBLEMA	51
4.2. DESCRIPCIÓN DEL PROTOTIPO DE LA APLICACIÓN.	52
4.3. EL MODELO DE DATOS.....	53
4.3.1. Ontología basada en metadatos LOM	53
4.3.2. Definición del patrón de expresión para las búsquedas.....	53
4.3.3. Construcción de la ontología.	56
4.4. ELABORACIÓN DE UN MODELO DE APLICACIÓN BASADO EN WEB SEMÁNTICA.	57
4.4.1. Módulo de representación y anotación semántica.....	58
4.4.2. Módulo de indexación	59
4.4.3. Módulo de extracción de términos	59
4.4.4. Módulo de evolución de ontologías	59
4.4.5. Modulo motor de búsqueda semántico.....	59
4.5. PRUEBA DE LA INSTANCIA DE BÚSQUEDA.....	59
RESULTADOS DE LA ENCUESTA AL USUARIO	61



V. CONCLUSIONES.....	70
VI. RECOMENDACIONES.....	71
VII. REFERENCIAS.....	72

Área: Ciencias de la ingeniería

Línea: Ingeniería Computacional y Sistemas

FECHA DE SUSTENTACION: 19 de marzo 2021



ÍNDICE DE FIGURAS

Figura 1:	Arquitectura de la Web Semántica.....	24
Figura 2:	Clases y relaciones básicas de la ontología LOM2OWL.....	53
Figura 3:	Clases y relaciones básicas de la ontología LOM2OWL.....	56
Figura 4:	Ontología del alumno	57
Figura 5:	Modelo de aplicación basado en Web Semántica.	58
Figura 6:	Disminución de tiempo de búsqueda de información	61
Figura 7:	La aplicación ayuda en la búsqueda de información	62
Figura 8:	La aplicación es de utilidad.....	63
Figura 9:	La aplicación tiene un buen desempeño.....	64
Figura 10:	La aplicación es de fácil uso	65
Figura 11:	La aplicación es amigable	66



ÍNDICE DE TABLAS

Tabla 1: Alumnos del 1er semestre Grupo A 2019-I EPIS-UNAP	50
Tabla 2: Relación entre los recursos LOM y las dimensiones de aprendizaje.....	55
Tabla 3: Disminución de tiempo de búsqueda de información.....	61
Tabla 4: La aplicación ayuda en la búsqueda de información	62
Tabla 5: La aplicación es de utilidad	63
Tabla 6: La aplicación tiene un buen desempeño	64
Tabla 7: La aplicación es de fácil uso	65
Tabla 8: La aplicación es amigable.....	66



ÍNDICE DE ACRÓNIMOS

- XML** : Sintaxis básica para la estructura de la información
- XML Schema** : Lenguaje que provee y limita la estructura y contenido de los elementos de un documento XML.
- RDF** : Lenguaje para expresar modelos de información que se refieren a objetos y a sus relaciones.
- RDF Schema** : Vocabulario para describir propiedades y clases de los recursos.
- OWL** : Añade más vocabulario para describir propiedades y clases de los recursos RDF.



RESUMEN

Con el desarrollo de la presente investigación, se logró el objetivo de desarrollar un prototipo basado en Web Semántica para buscar recursos de Programación para la Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano, con el fin de facilitar la búsqueda de recursos de programación para los estudiantes que buscan aprender a programar, de tal forma que la información proporcionada sea la que ellos requieren de acuerdo al nivel de programación en el que se encuentran. Se realizó el Análisis del Problema donde se estableció las consultas predefinidas, en base a los metadatos, considerando la ontología de datos y la ontología de dominio, así como el modelo del repositorio RDF que luego de la búsqueda proporciona los resultados ordenados. Luego se describe el prototipo de la aplicación. Considerando el modelo de datos y la ontología basada en metadatos LOM, para luego definir el patrón de expresión para las búsquedas y finalmente construir la ontología considerando el nivel del alumno quien busca los recursos para aprender a programar, y también el estilo de aprendizaje del mismo. Los resultados de la investigación fueron obtenidos en base a las métricas de tiempo, eficiencia y usabilidad, obteniéndose un resultado satisfactorio que justifica la utilidad de la web semántica para buscar recursos de aprendizaje para aprender a programar.

Palabras Clave: Prototipo, Web, Semántica, recursos, programación, ontología



ABSTRACT

With the development of the present investigation, the objective of developing a prototype based on Semantic Web was achieved to search for Programming resources for the Professional School of Systems Engineering of the National University of the Altiplano, in order to facilitate the search for resources of programming for students looking to learn to program, so that the information provided is what they require according to the level of programming they are in. The Problem Analysis was performed where the predefined queries were established, based on metadata, considering the data ontology and domain ontology, as well as the RDF repository model that after the search provides the ordered results. Then the prototype of the application is described. Considering the data model and the ontology based on LOM metadata, to then define the expression pattern for the searches and finally build the ontology considering the level of the student who seeks the resources to learn to program, and also the learning style of the same. The results of the research were obtained based on the metrics of time, efficiency and usability, obtaining a satisfactory result that justifies the usefulness of the semantic web to find learning resources to learn to program.

Keywords: Prototype, Web, Semantics, resources, programming, ontology



CAPITULO I

INTRODUCCIÓN

En la Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano, todos los semestres ingresan nuevos estudiantes que deben hacer frente al difícil reto de aprender a programar, muchos de ellos incluso concluyen la Carrera sin haber asimilado satisfactoriamente los conceptos básicos, intermedios y avanzados de programación. Se propone desarrollar un Prototipo de aplicación basado en Web Semántica que permita apoyar a los estudiantes de los diferentes cursos básicos de programación para que puedan encontrar información apropiada al nivel de conocimientos que tienen.

Para el desarrollo del software se empleó el framework Protege, se realizó el Análisis del Problema donde se estableció las consultas predefinidas, en base a los metadatos, considerando la ontología de datos y la ontología de dominio, así como el modelo del repositorio RDF que luego de la búsqueda proporciona los resultados ordenados. Luego se describe el prototipo de la aplicación. Considerando el modelo de datos y la ontología basada en metadatos LOM, para luego definir el patrón de expresión para las búsquedas y finalmente construir la ontología considerando el nivel del alumno quien busca los recursos para aprender a programar, y también el estilo de aprendizaje del mismo

Para la prueba del software se realizaron entrevistas con los alumnos del primer semestre grupo A de la Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano,



La investigación está dividida en siete capítulos, los cuales se detallan a continuación:

En el Capítulo I contiene el planteamiento del problema, el objetivo principal y objetivos secundarios que se ha propuesto alcanzar en la investigación.

El Capítulo II hace referencia a los antecedentes de la investigación, sustento teórico donde se realizó las investigaciones bibliográficas sobre aplicaciones web y el glosario de términos.

En el Capítulo III hace referencia al tipo de investigación, la delimitación de la población, ubicación del estudio y el tratamiento de los datos.

El Capítulo IV Se muestra los resultados y discusiones de los estudios realizados, en los grupos de control y experimental.

El Capítulo V Se muestra sus respectivas conclusiones a la que se llegó en la investigación.

El Capítulo VI Se muestra las recomendaciones de parte del autor hacia los que participan en la investigación como a futuros investigadores.

El Capítulo VII muestra la bibliografía utilizada en la investigación.

1.1. PLANTEAMIENTO DEL PROBLEMA

En la Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano, todos los semestres ingresan nuevos estudiantes que deben hacer frente al



difícil reto de aprender a programar, muchos de ellos incluso concluyen la Carrera sin haber asimilado satisfactoriamente los conceptos básicos, intermedios y avanzados de programación. Se propone desarrollar un Prototipo de aplicación basado en Web Semántica que permita apoyar a los estudiantes de los diferentes cursos básicos de programación para que puedan encontrar información apropiada al nivel de conocimientos que tienen.

1.2. FORMULACIÓN DEL PROBLEMA

¿Es posible desarrollar un prototipo basado en Web Semántica para buscar Manuales de Programación para la Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano?

1.3. OBJETIVOS DE LA INVESTIGACIÓN

1.3.1. Objetivo general

Desarrollar una Prototipo basado en Web Semántica para buscar manuales de Programación para la Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano

1.3.2. Objetivos específicos

- Establecer un modelo de aplicación basado en Web Semántica.
- Diseñar una ontología para modelar las búsquedas



- Definir una instancia de búsqueda
- Implementar la Instancia de Búsqueda

1.4. HIPÓTESIS DE LA INVESTIGACIÓN

El Prototipo basado en Web Semántica busca Manuales de Programación para la Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano adecuadamente.

1.5. JUSTIFICACIÓN DE LA INVESTIGACIÓN

En la Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano, todos los semestres ingresan nuevos estudiantes que deben hacer frente al difícil reto de aprender a programar, muchos de ellos incluso concluyen la Carrera sin haber asimilado satisfactoriamente los conceptos básicos, intermedios y avanzados de programación. Se desarrolló un Prototipo de aplicación basado en Web Semántica que permita a los estudiantes de los diferentes cursos básicos de programación encontrar información apropiada al nivel de conocimientos que tienen.



CAPITULO II

REVISIÓN DE LA LITERATURA

2.1. ANTECEDENTES

Antecedentes Internacionales

(Gomez A., 2014) “OASEARCH: Modelo de aplicación basado en web semántica para la búsqueda de objetos de aprendizaje mediante perfilado de consultas”. El objetivo principal de esta investigación es desarrollar un prototipo de aplicación informática (motor de búsqueda) que demuestre la factibilidad de crear, de una manera desacoplada de la aplicación, consultas predefinidas que obedecen a conceptos específicos de búsqueda que luego el usuario puede utilizar para recuperar objetos de aprendizaje relevantes en un Sistema Educativo Basado en la Web Semántica.

(Rodriguez G., 2014) “Extracción Semántica de Información basada en Evolución de Ontologías”. El objetivo principal de esta investigación es desarrollar una nueva metodología de anotación y recuperación semántica basado en modelos ontológicos formales, cubriendo todo el ciclo de vida de las anotaciones semánticas teniendo en cuenta las posibles actualizaciones de información y facilitando la adaptación continua del dominio de aplicación a través de la evolución de la ontología subyacente.

Antecedentes Nacionales

(Gomez M., 2014) “Diseño de un modelo para la recuperación de documentos basado en ontologías en el dominio de la Ingeniería Informática”. El objetivo principal de esta investigación es diseñar un modelo basado en ontologías en el dominio de la Ingeniería



Informática, que facilite la recuperación de documentos almacenados en repositorios de contenido digital.

(Espinoza F., 2014) “Diseño de una herramienta para la anotación semántica automática de documentos basados en ontologías en el dominio de la Ingeniería Informática”. El objetivo principal de esta investigación es desarrollar una herramienta que, usando una ontología en el dominio de la Ingeniería Informática, permita la anotación semántica automática de documentos como soporte a la recuperación de conocimiento.

(Coronado Altamirano, 2017) “Desarrollo de una herramienta para la anotación semántica automática de documentos pdf basado en ontologías”. El objetivo principal de esta investigación es desarrollar una herramienta que permita realizar anotaciones semánticas automáticas sobre documentos PDF basado en ontologías.

(Criado Fernández, 2009) “Procedimiento semi-automático para transformar la web en web semántica.” El objetivo principal de esta investigación es desarrollar un procedimiento semi-automático o automático de anotación en OWL basado en la cooperación de los usuarios. Las aplicaciones semánticas.

2.2. BASE TEÓRICA

2.2.1. ¿Qué es la web?

En 1945, Vannevar Bush identificó los problemas en la gestión de grandes colecciones de documentos y los vínculos que establecemos entre ellos. La propuesta de Bush fue considerar esto como un problema científico, y entre las ideas que propuso



estaba la de externalizar y automatizar el almacenamiento y manejo de los vínculos de asociación que hacemos en nuestras lecturas. Veinte años después, Ted Nelson propuso usar una computadora para implementar la idea, usando estructuras de hipertexto e hipermedia para unir partes de documentos. A finales de los sesenta, Douglas Engelbart y el proyecto Augment proporcionaron el ratón y nuevos medios de interacción y los aplicaron en particular a la edición y navegación de hipertexto. El comienzo de los años setenta trajo consigo el trabajo de Vinton Cerf y la aparición de Internet, que conectaba ordenadores de todo el mundo. A finales de los años ochenta, Tim Berners-Lee propuso un nuevo avance: una arquitectura para distribuir hipermedia en Internet, que ahora conocemos como WWW. (Hendler et al., 2020)

La Web proporciona una infraestructura de hipertexto que vincula documentos a través de Internet, es decir, conecta documentos que no están en la misma máquina. La arquitectura web incluye dos partes importantes: los clientes web, el más conocido es el navegador web, y los servidores web, que sirven documentos y datos a los clientes siempre que lo requieran. Para que esta arquitectura funcione, tiene que haber tres componentes esenciales iniciales. Primero, direcciones que nos permitan identificar y ubicar el documento en la Web; segundo, protocolos de comunicación que permiten a un cliente conectarse a un servidor, enviar una solicitud y obtener una respuesta; y tercero, lenguajes de representación para describir el contenido de las páginas, los documentos que se van a transferir. Estos tres componentes comprenden una arquitectura web básica. (Maynard et al., 2016)

2.2.2. Comunicación con Datos.

Para buscar información en la web, esta debe estar organizada en un buscador o un punto de acceso, desde donde se puede iniciar el proceso de navegación, si surgieran



cambios en la información, se tendría que actualizar de forma manual, todas las páginas en las cuales figura la información, lo que se convierte en un proceso tedioso; la información sobre horarios de vuelos, menús de restaurantes, las noticias que se generan en cada momento, deben ser alimentadas en páginas web, y estas luego son archivadas en forma cronológica para que pueda ser recuperada en cualquier momento que se necesite. Todo este proceso que se hace usualmente en forma manual, podría hacerse en forma automática. (Hendler et al., 2020)

2.2.3. La web semántica

Los archivos binarios a menudo contienen metadatos legibles por máquina, como la velocidad de obturación en una imagen JPEG1 o el título del álbum en un archivo de música MP3, los agentes de software automatizados no pueden interpretar el contenido textual de los sitios web tradicionales. Los sitios web que proporcionan semántica a los agentes de software forman la Web Semántica, una extensión de la Web convencional introducida a principios de la década de 2000. La Web Semántica es un aspecto importante de Internet. Para ser utilizado para una colección de tecnologías detrás de la mensajería instantánea, voz sobre IP, wikis, blogs, foros, portales de redes sociales y sindicación web. La próxima generación de la Web se denomina Web 3.0, que es un término general para la personalización, los contenidos semánticos y las aplicaciones web más sofisticadas hacia la inteligencia artificial, incluidos los contenidos generados por computadora. (Sikos, 2015)



2.2.4. La evolución de la Web

En los primeros días de la Web, se podía ver como un conjunto de sitios web que ofrecen una colección de documentos web, y el objetivo era llevar el contenido a sus audiencias. Actuaba como una calle de sentido único: la gente leía todo lo que había allí afuera, con el objetivo de obtener información que pudieran usar de varias maneras. Hoy, la Web se ha vuelto mucho más interactiva. Primero, cada vez más de lo que ahora se conoce como contenido generado por el usuario ha surgido en la Web, y también se crearon una serie de nuevas compañías de Internet en torno a esta tendencia. Más específicamente, en lugar de solo leer el contenido, las personas ahora usan la Web para crear contenido y también para interactuar entre sí mediante el uso de sitios de redes sociales y plataformas web. Y ciertamente, incluso si no creamos ningún contenido nuevo o participamos en ningún sitio de redes sociales, aún podemos disfrutar mucho de la Web: podemos chatear con nuestros amigos, podemos comprar en línea, podemos pagar nuestras facturas en línea y se puede ver un juego de tenis en línea. En segundo lugar, la vida de un desarrollador web también ha cambiado mucho. En lugar de ofrecer contenido web simplemente estático, el desarrollador web de hoy en día es capaz de crear sitios web que pueden ejecutar transacciones comerciales complejas, desde pagar facturas en línea hasta reservar una habitación de hotel y boletos aéreos. Tercero, cada vez más sitios web han comenzado a publicar contenido estructurado para que diferentes entidades comerciales puedan compartir su contenido para atraer y realizar más transacciones en línea. (Yu, 2014)



2.2.5. Beneficios de la Web Semántica

La Web Semántica es una red de datos que proporciona un marco común que permite que los datos se compartan y reutilicen a través de los límites de la aplicación, la empresa y la comunidad. La tecnología web semántica proporciona una serie de lenguajes estandarizados Lenguaje de ontología web para definir el significado del vocabulario, un marco de descripción de recursos, un nuevo modelo de datos para la web, formato de intercambio de reglas y un lenguaje de consulta de Protocolo SPARQL y un lenguaje de consulta RDF, así como diferentes sistemas de razonamiento en la correspondiente lógica de descripción o sistemas de conocimiento basados en reglas. La Web Semántica ha sido ampliamente aceptada en aplicaciones con acceso a datos pesados o trabajos intensivos de gestión del conocimiento. En términos generales, el éxito de estas aplicaciones se basa en dos pasos clave, construir ontologías reutilizables para el dominio, y usar un vocabulario ontológico para anotar datos en el dominio. (Jeff & Pan, 2014)

2.2.6. Arquitectura de la Web Semántica

La arquitectura estándar de la Web Semántica sigue el modelo de capas definido por Tim Berners-Lee en 2004. El objetivo de esta arquitectura es proporcionar a los usuarios y agentes software capacidades de procesamiento, razonamiento y deducción sobre los contenidos de la Web.

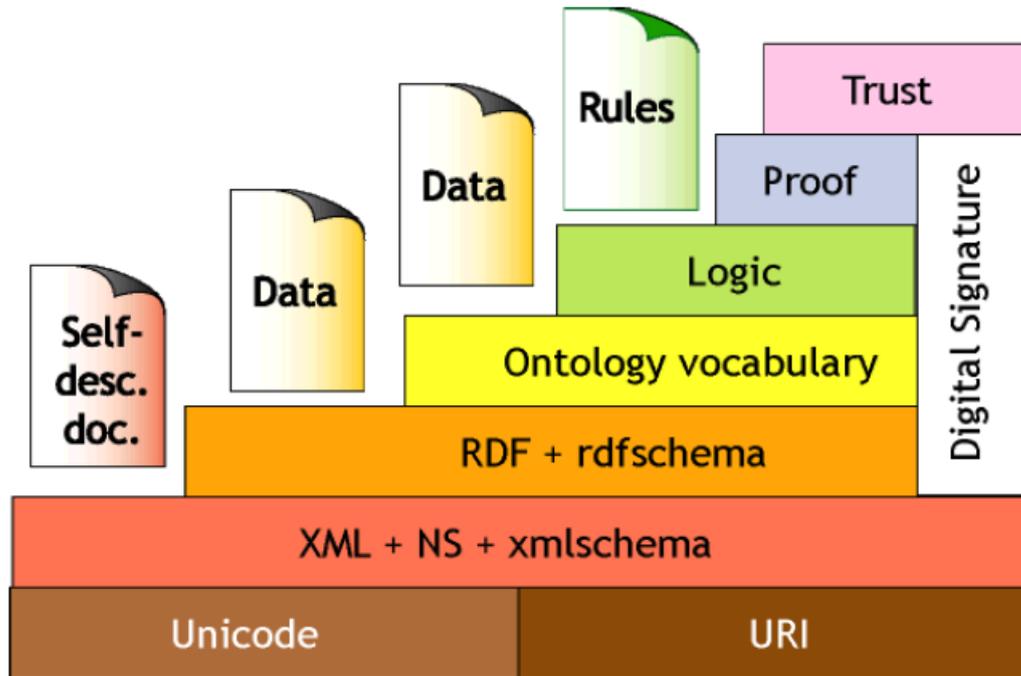


Figura 1: Arquitectura de la Web Semántica

FUENTE: Building Ontologies with Basic Formal Ontology (Arp et al., 2015)

La función que desempeña cada capa dentro de la arquitectura de la Web Semántica es la siguiente.

2.2.6.1. Unicode + URI

Unicode es un sistema internacional estándar que proporciona un número único para cada carácter, sin importar la plataforma ni el programa. Esto permite representar caracteres de cualquier idioma con una codificación unificada. Uniform Resource Identifier (URI) es un sistema de direccionamiento e identificación de recursos. El sistema que usamos actualmente para acceder a los recursos de la Web (URL) es una parte de URI.



2.2.6.2. XML+NS+XML SCHEMA

La combinación XML+NS+XML en un esquema donde eXtended Markup Language (XML) es un sistema que permite definir lenguajes de marcas para usos específicos. Name Spaces (NS) permite combinar diversos lenguajes de marcado creados con XML en un mismo documento. XML Schema sirve para definir tipos de documentos complejos en los que se pueden especificar tipos de datos, listas de componentes y restricciones similares a las del diccionario de datos típico de una base de datos.

2.2.6.3. RDF + rdfschema

Resource Description Framework (RDF) es un modelo de representación de metadatos que, entre otras cosas, permite representar recursos digitales tales como sitios o páginas web. RDF está concebido para representar cualquier clase de recursos (no solamente páginas publicadas en la web). RDF Schema, por su parte, es una extensión de RDF que aporta un lenguaje con mayor capacidad para representar relaciones semánticas complejas.

2.2.6.4. Ontology vocabulary

Una ontología es una especificación formal de un dominio del conocimiento que, en su expresión más simple, se identifica con una taxonomía. Una taxonomía consiste en una jerarquía de conceptos y sus relaciones del tipo clase-subclase. Una ontología formaliza la relación de clase, añade otras relaciones y especifica propiedades para individuos y clases. Ontology-vocabulary se refiere a una ontología concreta sobre un dominio concreto del conocimiento.



2.2.6.5. Logic.

En este contexto, logic se refiere al estudio de las reglas formales que permiten determinar si un razonamiento se sigue de sus premisas. La lógica estudia, por tanto, la estructura de los razonamientos válidos. Se espera que los ordenadores del futuro puedan efectuar razonamientos sobre los recursos y servicios de la Web combinando los conocimientos expresados en las ontologías, los hechos declarados en los metadatos y la aplicación de reglas lógicas.

2.2.6.6. Proof

En este contexto, Proof (prueba) significa demostración matemática. Se considera que un ordenador alcanza la máxima fiabilidad en sus razonamientos cuando es capaz de realizar demostraciones o, lo que es lo mismo a efectos prácticos, cuando es capaz de justificar el motivo por el cual tomó o aconsejó tomar una decisión.

2.2.6.7. Trust (+ Digital Signature)

La última capa, Trust (confianza) debe servir para otorgar confianza a las transacciones en la Web a través que se llevarán a cabo no solamente entre usuarios y sitios web sino también entre programas de software; y todo ello tanto en el plano C2B (consumer to business) como en el B2B (business to business). La Digital Signature (firma digital) proporcionará soporte específico a esta capa, tal como muestra el diagrama.

2.2.6.8. Web Semántica Confiable.

Esta capa tiene la función de evaluar las pruebas ofrecidas por la capa anterior para comprobar de forma exhaustiva si las fuentes de información son confiables.



2.2.6.9. Firma y cifrado (Firma digital).

El objetivo de esta capa es definir el ámbito de confianza para las capas “Prueba” y “Web Semántica Confiable” de manera que permita a los ordenadores y agentes software verificar la seguridad de la información adjuntada y que ésta se envié por una fuente confiable.

La tecnología más importante en el contexto de la Web Semántica son las ontologías, de hecho, constituyen el medio principal para lograr el objetivo de la Web Semántica al facilitar la definición formal de las entidades y conceptos presentes en los diferentes dominios, la jerarquía que les sustenta y las diferentes relaciones que los unen entre sí. De esta manera se garantiza una representación del conocimiento consensuada y reutilizable que puede ser compartida y utilizada automáticamente por cualquier sistema informático. (Barrière, 2016)

2.2.7. Ontologías

La palabra ontología se introdujo originalmente en filosofía para estudiar la naturaleza de la existencia. En informática, la ontología se refiere a la conceptualización a través de un modelo de datos para describir una parte de nuestro mundo, como una organización, un proyecto de investigación, un evento histórico, nuestros colegas, amigos, etc., de manera legible por computadora, formalmente definir un conjunto de clases y conceptos, propiedades y atributos, tipos de relación y entidades como individuos, instancias. Los lenguajes de ontología más avanzados como OWL, admiten los siguientes componentes: Clases que son grupos abstractos, conjuntos o colecciones de objetos, tipos de objetos. Las clases generalmente representan grupos o clases cuyos miembros comparten propiedades comunes. La jerarquía de clases se expresa como nivel



superior, superclase o clase principal y clases de nivel inferior, subclase o clase secundaria. Por ejemplo, una empresa puede representarse como una clase con subclases como departamentos y empleados. Atributos como aspectos, propiedades, características o parámetros que presentan objetos y clases. Individuos como instancias u objetos. Por ejemplo, si nuestro dominio cubre compañías, cada empleado es un individuo. (Sikos, 2015)

2.2.8. Tipos de Ontologías

Una ontología, tal como la concebimos, es un artefacto de representación destinado a representar universales, clases definidas y relaciones entre ellos. Aquí nos interesan específicamente los universales, las clases definidas y las relaciones descubiertas y pertinentes a la investigación científica. Existen algunas distinciones entre tipos de ontologías, incluidas ontologías de dominio, ontologías de nivel superior, ontologías de referencia y ontologías de aplicación.

2.2.8.1. Ontología filosófica Histórica.

La ontología es una rama de la filosofía que tiene sus orígenes en la antigua Grecia en el trabajo de filósofos como Parménides, Heráclito, Platón y Aristóteles. "Ontología" deriva de las palabras griegas ontos (que significa "existencia" o "ser") y logos (que significa "cuenta racional" o "conocimiento"). En este sentido, la ontología se ocupa del estudio de lo que es, de los tipos y estructuras de objetos, propiedades, eventos, procesos y relaciones en la realidad. Desde la perspectiva filosófica, la ontología busca proporcionar una clasificación definitiva y exhaustiva de las entidades en todas las esferas del ser. Los filósofos han enfatizado a este respecto el papel de ciertos tipos básicos o preferidos de entidades, por ejemplo, simples absolutos, que se consideran



verdaderamente reales, en contraste con otras entidades menos privilegiadas, que se consideran construidas sobre su base. (Arp et al., 2015)

2.2.8.2. Ontología filosófica contemporánea

A veces llamada "metafísica analítica", ha permitido en ocasiones el estudio de entidades de las que se ocupan las ciencias especiales (física, química, biología, psicología y otras), pero todavía persigue abrumadoramente un enfoque más general, uno dirigido a proporcionar una descripción y explicación de los tipos de objetos y relaciones que son comunes a todos los dominios científicos. Ejemplos de tales características de la realidad comunes o neutrales al dominio incluyen unidad y pluralidad; causa y efecto; identidad, tanto a la vez como a lo largo del tiempo; estructura compositiva, determinada a través de relaciones parte-todo o miembros-conjunto; ubicación espacial, temporal y espacio-temporal; etcétera. Donde el biólogo estudia las células, el químico estudia las moléculas y el físico estudia la energía y los electrones; el ontólogo filosófico, en cambio, está interesado en dar cuenta de lo que es común a las células, moléculas y electrones (por ejemplo, que son todas entidades o cosas que existen como portadoras de ciertas propiedades o cualidades) y de las relaciones en el que este tipo de entidades se enfrentan entre sí. Tales relaciones pueden extenderse a través de los límites disciplinarios habituales de las ciencias especiales y a través de granularidades, como en el caso de las relaciones entre entidades de nivel micro, como átomos o moléculas, y entidades de nivel macro, como organismos y planetas. El objetivo de la ontología filosófica es, en esencia, proporcionar una descripción clara, coherente y rigurosamente elaborada de las estructuras básicas de toda la realidad así concebida. Ontología filosófica y taxonomía

Una característica importante de la ontología filosófica es la clasificación de diferentes tipos de entidades. Esto implica no solo diferenciar entre categorías básicas de entidades



(por ejemplo, entre objetos como un corazón y procesos como el latido de un corazón), sino también reconocer tipos más específicos de entidades que caen dentro de estas categorías (como la diferencia entre un objeto individual, como un ser humano, y un agregado o grupo de objetos, como una población de seres humanos). El resultado es un conjunto de representaciones de los tipos de objetos que existen, en general o en un campo de investigación dado, organizados según ciertos principios de clasificación. (Rebstock et al., 2016)

2.2.8.3. Ontologías formales versus materiales

Una ontología formal es de dominio neutral. Contiene sólo los términos más generales, como "objeto" y "proceso", que se aplican a todas las disciplinas científicas. Por lo tanto, corresponde al tipo de interés ontológico que identificamos anteriormente como predominante entre los filósofos. Una ontología material (o "dominio") es específica de un dominio. Contiene términos, como "celda" o "carburador", que se aplican solo en un subconjunto de disciplinas. Una ontología formal es una representación de las categorías de entidades y de las relaciones dentro y entre ellas. Aquí adaptamos el término "categoría" de Aristóteles y lo usamos para significar: universal neutral en el dominio. Las categorías son aquellos universales cuyas instancias se encuentran en cualquier dominio de la realidad. Por tanto, las categorías son también universales muy generales. La categoría más obvia es entidad, es decir: cualquier cosa que exista de alguna manera. No importa qué ciencia uno esté considerando, estudia entidades y, por lo tanto, la entidad de categoría se aplica al tema de esa ciencia. Un ejemplo de relación formal-ontológica es parte de, ya que toda ciencia está comprometida con la existencia de al menos algunas entidades que tienen o son partes. (Arp et al., 2015)



Una ontología material, por el contrario, consiste en representaciones de los universales materiales (es decir, no formales) que se instancian en algún dominio específico de la realidad, como la genética, la anatomía, la biología vegetal, el cáncer, etc. Cuando una ontología formal contendrá representaciones de universales compartidas por muchas ontologías, cada ontología material contendrá representaciones utilizadas solo en la ontología. Así, los universales representados en una ontología de la biología celular, por ejemplo, no se superpondrán con los representados en una ontología de la cosmología o de la arquitectura. Esta distinción resultará de gran importancia para el proyecto de diseño de ontología en apoyo de la ciencia impulsada por la información. La ontología específica de dominio o material se analiza con más detalle en la siguiente sección. (Rebstock et al., 2016)

2.2.8.4. Ontología de dominio

Un dominio es una porción delimitada de la realidad correspondiente a una disciplina científica como la biología celular o la microscopía electrónica, o un área de conocimiento o interés como la Gran Guerra o la recolección de sellos o permisos de construcción. No todo lo que es parte de una entidad dentro de un dominio dado también es parte de ese dominio. Así, todo ser humano tiene moléculas como partes, pero las moléculas no forman parte del dominio de, por ejemplo, la geografía humana o el derecho de los derechos humanos (un tema que se tratará bajo el título de “granularidad” en el capítulo 3). Cada ontología de dominio consta de una taxonomía (una jerarquía estructurada por la relación es junto con otras relaciones como parte de, contenido en, adyacente a, has agente, precedido por, etc., junto con definiciones y axiomas que gobiernan cómo sus términos y relaciones deben ser comprendido. Por tanto, una ontología de dominio es una taxonomía que se ha mejorado para incluir más información



sobre los universales, las clases y las relaciones que representa. Una ontología de dominio proporciona una representación estructurada y controlada de las entidades dentro del dominio relevante, una que puede usarse, por ejemplo, para anotar datos pertenecientes a entidades en ese dominio con el fin de hacer que los datos sean más fácilmente accesibles y compartibles por los seres humanos y procesable por computadoras. (Arp et al., 2015)

2.2.8.5. Clasificación por el conocimiento que contienen

A partir del contenido de la ontología, Riichiro Mizoguchi establece las siguientes categorías ontológicas

Ontologías del dominio: expresan conceptualizaciones específicas a un dominio en particular, es decir, describen los conceptos y sus relaciones respecto a un dominio específico.

Ontologías de tarea: describen el vocabulario específico de los conceptos de manera que se pueda utilizar el conocimiento del dominio para realizar tareas específicas.

Ontologías generales: proporcionan descripciones generales acerca de objetos, eventos, relaciones temporales, relaciones causales, modelos de comportamiento y funcionalidades. (Allemang & Hendler, 2012)

2.2.8.6. Clasificación por el tipo de estructura

Se clasifican en:

Ontologías terminológicas o lingüísticas: especifican los términos que son usados para representar conocimiento en un dominio determinado. Se suelen utilizar para unificar el vocabulario de un dominio concreto.



Ontologías de información: especifican la estructura de los registros de almacenamiento de una base de datos. Estas ontologías proporcionan un marco estándar para el almacenamiento de la información.

Ontologías para modelar conocimiento: especifican conceptualizaciones de conocimiento. Estas ontologías tienen una estructura interna mucho más rica que las anteriores, que las hace interesantes para los desarrolladores de sistemas basados en conocimiento. Estas ontologías se ajustan al uso particular del conocimiento que describen. (Hendler et al., 2020)

2.2.8.7. Clasificación por motivación

Mediante esta clasificación se categorizan las ontologías a partir del motivo por el que éstas se elaboran.

Ontologías para la representación del conocimiento, también conocidas como representacionales: permiten especificar las conceptualizaciones que subyacen en los formalismos de representación de conocimiento por lo que también se denominan meta-ontologías (del inglés, meta-level ontologies o top-level ontologies). (Hendler et al., 2020)

2.2.8.8. Ontologías genéricas:

Definen conceptos considerados genéricos y fundacionales del conocimiento como las estructuras parte/todo, la cuantificación, los procesos o los tipos de objetos. También se conocen por el nombre de ontologías abstractas o superteorías, debido a que permiten definir conceptos abstractos y pueden ser reutilizables en diferentes dominios.



Ontologías genéricas o de alto nivel: describen conceptos generales que normalmente son independientes del dominio o problema particular. Se consideran, generalmente, de utilidad en todos los dominios o aplicaciones.

Ontologías de dominio: definen un vocabulario para describir conceptos y relaciones de un dominio específico. Estos conceptos son normalmente especializaciones de términos introducidos en la ontología de alto nivel.

Ontologías de tareas: describen el vocabulario relacionado con una actividad, tarea genérica o artefacto especializando los términos introducidos en la ontología de alto nivel.

Ontologías de aplicación: describen conceptos dependientes de un dominio y de una tarea en particular, los cuales son frecuentemente especializaciones de ontologías relacionadas. Estos conceptos normalmente suelen corresponder a funciones realizadas por las entidades de dominio mientras desempeñan una determinada actividad.

Ontologías de grano grueso y de grano fino: con esta clasificación Guarino distingue las ontologías según la exactitud con la que reflejan la conceptualización. (Rebstock et al., 2016)

Las ontologías de grano grueso, que desarrollan una axiomatización más rica, pueden consistir en un conjunto mínimo de axiomas escritos en un lenguaje de expresividad mínima, para apoyar un conjunto de servicios específicos a una amplia comunidad de usuarios que ya han acordado una conceptualización subyacente. Este tipo de ontologías también son llamadas ontologías “off-line”, porque son sólo accesibles temporalmente por los usuarios para fines de referencia. Por otro lado, las ontologías de grano fino se desarrollan adoptando un dominio más rico y un conjunto más amplio de



axiomas y relaciones conceptuales relevantes que permiten definir con bastante precisión los conceptos a los que se refiere. También pueden ser llamadas ontologías “online”, debido a que su función es proporcionar soporte al núcleo del sistema. (Rebstock et al., 2016)

2.2.8.9. Por el grado de formalidad de la ontología

Existe otro criterio que puede ser utilizado para clasificar ontologías basadas en el grado de formalidad de la ontología

Ontologías descriptivas: Poli define estas ontologías como un recolector de elementos del mundo, ya sea de un dominio específico de análisis o en general. El autor define el mundo como resultado de un complejo entramado de conexiones de dependencia y formas de independencia entre los elementos que lo componen. Es decir, en el mundo existen cosas materiales, plantas y animales, así como los productos de los talentos y actividades de animales y humanos, pensamientos, sensaciones y decisiones, así como el completo espectro de actividades mentales y reglas que se organizan en relación a la dependencia o independencia entre los mismos elementos. En este ámbito, las ontologías descriptivas están relacionadas con esta recopilación de información.

Ontologías formales: según Poli, este tipo de ontologías tienen la función de destilar, filtrar y organizar los resultados de una ontología descriptiva, ya sea en su entorno local o global. Según esta interpretación, la ontología formal es formal en el sentido descrito por Husserl. El filósofo alemán fue el primero en introducir la ontología formal en la filosofía. Husserl define las ontologías formales basándose en la mereología, la teoría de la dependencia y la topología. Para él, una ontología formal se basa en interconexiones de las cosas, con objetos y propiedades, parte y todo que representan



categorías puras que caracterizan aspectos y tipos de realidad que todavía no tienen nada que ver con el uso de ningún formalismo específico. (Arp et al., 2015)

2.2.9. Elementos de una ontología

Las ontologías proporcionan un vocabulario común de un área y definen, a diferentes niveles de formalismo, el significado de los términos y relaciones entre ellos. El conocimiento en ontologías se formaliza principalmente usando cinco tipos de componentes: clases, atributos, relaciones, axiomas e instancias.

2.2.9.1. Clases.

Se suele usar tanto el término “clases” como “conceptos”. Un concepto representa cualquier entidad que se puede describir, tiene asociado un identificador único, puede poseer diferentes atributos y establecer relaciones con otros conceptos. Las clases en la ontología se suelen organizar en taxonomías. Algunas veces, la noción de ontología se diluye en el sentido que las taxonomías se consideran ontologías completas.

2.2.9.2. Atributos.

Los atributos representan la estructura interna de los conceptos. Atendiendo a su origen, los atributos se clasifican en específicos y heredados. Los específicos son los propios del concepto al que pertenecen, mientras que los heredados vienen dados por las relaciones taxonómicas en las que el concepto desempeña el rol de hijo y, por tanto, hereda los atributos del padre. Los atributos se caracterizan por el rango en el cual pueden tomar valor.



2.2.9.3. Relaciones.

Las relaciones representan un tipo de interacción entre los conceptos del dominio. Se definen formalmente como cualquier subconjunto del producto cartesiano de n conjuntos, esto es: “ $R: C_1 \times C_2 \times \dots \times C_n$ ”. Las relaciones normalmente suelen ser binarias entre dos conceptos. No todas las relaciones tienen el mismo significado. Existen relaciones binarias de especialización como “is-a” o de composición como “part-whole”, que se pueden modelar con distintas propiedades como la simetría, reflexividad, transitividad, asimetría, etc.

2.2.9.4. Axiomas.

Los axiomas son expresiones que son siempre ciertas, es decir, modelan las “verdades” que siempre se cumplen en el modelo y en el dominio. Pueden ser incluidas en una ontología con muchos propósitos, tales como definir el significado de los componentes ontológicos, definir restricciones complejas sobre los valores de los atributos, argumentos de relaciones, etc. Estos axiomas verifican la corrección de la información especificada en la ontología y pueden ayudar en la generación de nuevo conocimiento.

2.2.9.5. Instancias.

Las instancias son las ocurrencias en el mundo real de los conceptos. En una instancia todos los atributos del concepto tienen asignado un valor concreto.

2.2.10. Lenguajes para la representación de ontologías

Los lenguajes ontológicos surgieron como medio para proporcionar capacidad de representación del conocimiento. El primer lenguaje para la definición de ontologías para

la Web fue SHOE. Posteriormente surgieron otros lenguajes con funciones similares, entre los que se pueden destacar, como más representativos, RDF, DAML+OIL y, por último, OWL (del inglés, "Web Ontology Language").

A continuación, se proporciona una breve descripción de estos lenguajes para la representación de ontologías.

2.2.10.1. Simple HTML Ontology Extensions (SHOE)

SHOE (Luke et al., 1997) fue el primer lenguaje de etiquetado utilizado para definir ontologías en la Web. Las ontologías y las etiquetas son incrustadas en archivos HTML. Este lenguaje permitía definir clases, relaciones entre clases, así como reglas de inferencia expresadas en forma de cláusulas de Horn. Una de las principales limitaciones de este lenguaje es que no dispone de ningún mecanismo para expresar negaciones o disyunciones. Aunque el proyecto fue abandonado a medida que se desarrollaron nuevos lenguajes de representación de ontologías como OIL y DAML (McGuinness, 2000), este lenguaje disponía de diferentes herramientas, API y editores de anotaciones que, entre otras funciones, permitían la serialización de este lenguaje en XML.

2.2.10.2. Resource Description Framework (RDF)

El lenguaje RDF es un modelo de datos que utiliza tripletas para representar recursos y las relaciones que pueden ser establecidas entre ellos. Una tripleta está compuesta por dos nodos (sujeto y objeto) unidos por un arco (predicado).

Sujeto: identifica el recurso (persona, lugar o cosa) que la sentencia describe. Un recurso RDF puede ser cualquier cosa en un modelo de datos (documento, usuario, producto, etc.). Cada recurso está identificado de forma única a través de una URI.



Predicado: representa una propiedad del sujeto. Al igual que ocurre con los sujetos, cada propiedad se identifica con una URI única.

Objeto: especifica el valor del predicado para un sujeto. En RDF un objeto puede ser otro recurso o un literal. En el caso de que sea un recurso, el objeto definirá otra URI que identifique al recurso. Por su parte, los literales son una cadena simple de caracteres u otro tipo de datos primitivo definido por XML.

En términos de RDF, un literal puede contener marcado XML pero no es interpretado por RDF. El modelo RDF distingue a los literales de los recursos restringiendo a los literales ser sujeto o predicado en una declaración.

RDF provee un mecanismo capaz de describir recursos e identificarlos unívocamente. Para identificar globalmente los recursos y las propiedades se utilizan la URI. Una URI es una cadena de caracteres que identifica unívocamente.

2.2.11. El modelo RDF

La Web Semántica define un modelo de información llamado RDF. Es el medio por el cual la Web Semántica facilita la "catalogación por asociación". En su núcleo está la noción de una declaración que describe una asociación. Una declaración tiene varias características. Identifica de qué se trata. Expresa una relación con otra cosa. Tiene un valor de verdad, es decir, cada declaración es verdadera o falsa. Aquí hay algunos ejemplos de declaraciones: Las aves tienen plumas. La tierra órbita alrededor del sol. Cero Absoluto es 0 grados Kelvin. Una declaración RDF está destinada a ser consumida por el software. En otras palabras, está escrito de tal manera que una computadora puede identificar inequívocamente cada uno de sus componentes para que pueda relacionarlos con otras declaraciones. Para evitar la ambigüedad, los componentes de las declaraciones



consisten principalmente en identificadores únicos. Una declaración RDF se puede modelar de dos maneras, sin cambiar su significado o su valor de verdad. Puede representar una declaración RDF como un triple, que parece una oración. Un triple consiste en un sujeto seguido de un predicado seguido de un objeto. También puede representar como un segmento gráfico que consta de dos nodos con un borde dirigido entre ellos. En declaraciones RDF, identificadores únicos se utilizan para identificar al sujeto, predicados y, a veces para los valores de objeto de triples RDF. Al igual que los URI y las URL, un IRI identifica de forma exclusiva algo. Pero a diferencia de una URL, no se requiere que este IRI resuelva nada. Por lo tanto, si conecta este IRI a un navegador web, es posible que no recupere nada. (Hopkins & Powell, 2015)

2.2.12. El Lenguaje de Ontologías Web (OWL)

Está diseñado para ser usado en aplicaciones que necesitan procesar el contenido de la información en lugar de únicamente representar información para los humanos; facilita una mejor forma de interpretar el contenido Web, mejor que los mecanismos admitidos por XML, RDF, y el esquema RDFS proporcionando un vocabulario adicional junto con una semántica formal. OWL tiene tres sub lenguajes, con un nivel de expresividad en forma creciente: OWL Lite, OWL DL, y OWL Full. Una funcionalidad clave de OWL es la capacidad de definir clases de restricción. Las clases sin nombre se definen en base a restricciones en los valores para propiedades particulares de la clase. Usando este mecanismo, OWL puede usarse para modelar situaciones en las que los miembros de una clase particular deben tener ciertas propiedades. En RDFS, las restricciones de dominio y rango pueden permitirnos hacer inferencias sobre todos los miembros de una clase (como jugadas para relacionar a un jugador de béisbol con un equipo). En OWL, uno puede usar declaraciones de restricción para diferenciar el caso



entre algo que se aplica a todos los miembros de una clase versus algunos miembros, e incluso insistir en un valor particular para una propiedad específica de todos los miembros de una clase. (Allemang & Hendler, 2012)

2.2.13. Información de modelado en software

En los sistemas de software, el modelado de información es de suma importancia. Los atributos de los objetos del mundo real que elige capturar en su software y la forma en que elige representarlos determinan en gran medida las operaciones que su sistema puede realizar y las preguntas que puede responder. En un sistema semántico esto es aún más importante, porque al modelar conceptos y objetos con suficiente rigor, es posible reutilizar datos de una aplicación en otra. Todo software debe diseñarse con un modelo de información de algún tipo. El carácter del modelo de información depende de los objetivos del programa y del contexto para el que está diseñado. Por ejemplo, considere una aplicación simple que administra una colección de direcciones de correo. Un enfoque es modelar la información como un conjunto de objetos relacionados. Puede esperar una clase Persona y una clase Dirección, cada una con varios miembros de datos apropiados. Un enfoque alternativo, es usar un modelo relacional y pensar en una tabla de Persona y una tabla de Dirección, cada una con columnas apropiadas para valores, y una tabla de enlaces para definir las relaciones entre ellos. (Hebeler & Fisher, 2009)



2.3. MODELOS DE RECUPERACIÓN DE LA INFORMACIÓN

2.3.1. Modelo booleano

El modelo booleano se basa en la teoría de conjuntos y el álgebra de Boole. En éste modelo inicial el usuario especifica en su consulta una expresión booleana formada por una serie de términos ligados mediante operadores booleanos (comúnmente AND, OR y NOT). Dada la expresión lógica de la consulta, el sistema devolverá aquellos documentos que la satisfacen y que conformarán el

Conjunto de documentos relevantes. El sistema simplemente particiona los documentos de la colección en dos conjuntos, aquéllos que cumplen la condición especificada (relevantes), y aquellos que no la cumplen (no relevantes), sin ordenación interna alguna, de forma similar a lo que ocurriría con una base de datos tradicional.

Las desventajas importantes asociada al modelo está dado por la dificultad que conlleva la formalización de la necesidad de información del usuario en forma de expresión booleana, sobre todo cuando se trata de usuarios inexpertos y de necesidades complejas. (Workman, 2015)

2.3.2. Modelo Vectorial

Este modelo plantea un marco formal diferente en que se permite tanto la asignación de correspondencias parciales, como la existencia de grados de relevancia en base a los pesos de los términos en consultas y documentos. En éste modelo las consultas y documentos son representados mediante vectores dentro de un espacio multidimensional de finido por los propios términos, de tal forma que cada uno de los términos diferentes del sistema, definen una dimensión. Desde un punto de vista



geométrico, si ambos vectores, consulta y documento, están próximos, es factible asumir que el documento es similar a la consulta, el documento es posiblemente relevante

Conceptualmente, este modelo utiliza una matriz documento-término que contiene el vocabulario de la colección de referencia y los documentos existentes. En la intersección de un término y un documento se almacena un valor numérico de importancia del término t en el documento d ; tal valor representa su poder de discriminación. Así, cada documento puede ser visto como un vector que pertenece a un espacio n -dimensional, donde n es la cantidad de términos que componen el vocabulario de la colección. En teoría, los documentos que contengan términos similares estarán a muy poca distancia entre sí sobre tal espacio. De igual forma se trata a la consulta, es un documento más y se la mapea sobre el espacio de documentos. Luego, a partir de una consulta dada es posible devolver una lista de documentos ordenados por distancia (los más relevantes primeros). Para calcular la semejanza entre el vector consulta y los vectores que representan los documentos se utilizan diferentes fórmulas de distancia, siendo la más común la del coseno. El siguiente ejemplo se muestra un documento y una consulta.

2.3.3. Modelo probabilístico

Fue propuesto por Robertson y Spark-Jones, El modelo probabilístico formaliza el proceso de recuperación en términos de teoría de probabilidades. A partir de una expresión de consulta se puede dividir una colección de N documentos en cuatro subconjuntos distintos: REL conjunto de documentos relevantes, REC conjunto de documentos recuperados, RR conjunto de documentos relevantes recuperados y NN el conjunto de documentos no relevantes no recuperados. El resultado ideal de a una consulta se da cuando el conjunto REL es igual REC. Como resulta difícil lograrlo en



primera intención, el usuario genera una descripción probabilística del conjunto REL y a través de sucesivas interacciones con el SRI se trata de mejorar la performance de recuperación. Dado que una recuperación no es inmediata dado que involucra varias interacciones con el usuario y que estudios han demostrado que su performance es inferior al modelo vectorial, su uso es bastante limitado.

Según el principio de orden de probabilidades, el rendimiento óptimo de un sistema se consigue cuando los documentos son ordenados de acuerdo a sus probabilidades de relevancia. El modelo parte de las siguientes suposiciones:

a. Todo documento es, bien relevante, bien no relevante para la consulta.

b. El hecho de juzgar un documento dado como relevante o no relevante no aporta información alguna sobre la posible relevancia o no relevancia de otros documentos. Existen múltiples medidas de similitud utilizadas en éste modelo, siendo el más conocido el sistema Okapi, cuyo esquema de pesos se encuentra entre los más efectivos y, junto al vectorial if-idf es punto de referencia para el desarrollo y evaluación de nuevos modelos y nuevos esquemas de pesos. (Maynard et al., 2016)

2.3.4. Modelos para documentos estructurados

Los modelos clásicos responden a consultas, buscando sobre una estructura de datos que representa el contenido de los documentos de una colección, únicamente como listas de términos significativos. Un modelo de recuperación de documentos estructurados utiliza la estructura de los mismos a los efectos de mejorar la performance y brindar servicios alternativos al usuario (por ejemplo, uso de memoria visual, recuperación de elementos multimedia, mayor precisión sobre el ámbito de la consulta y demás). La estructura de los documentos a indexar está dada por marcas o etiquetas,



siendo los estándares más utilizados el SGML (Standard General Markup Language), el HTML(HyperText Markup Language), el PDF (Portable Document Format), el XML (eXtensible Markup Language) y LATEX.

Al poseer la descripción de parte de la estructura de un documento es posible generar un grafo sobre el que se navegue y se respondan consultas de distinto tipo, por ejemplo:

- a. Por estructura: ¿Cuáles son las secciones del segundo capítulo?
- b. Por metadatos o campos: Documentos de "Editorial UNL" editados en 2020
- c. Por contenido: Término "agua" en títulos de secciones
- d. Por elementos multimedia: Imágenes cercanas a párrafos que contengan Bush

Existen dos modelos en esta categoría "nodos proximales" y "listas no superpuestas". Ambos modelos se basan en almacenar las ocurrencias de los términos a indexar en estructuras de datos diferentes, según aparezcan en algún elemento de estructura (región) o en otro como capítulos, secciones, subsecciones y demás. En general, las regiones de una misma estructura de datos no poseen superposición, pero regiones en diferentes estructuras sí se pueden superponer. Es necesario mencionar que algunos motores de búsqueda de Internet ya utilizan ciertos elementos de la estructura de un documento, por ejemplo, los títulos, a efectos de realizar tareas de ranking, resumen automático, clasificación y otras.

La expansión de estos lenguajes de demarcación, especialmente en servicios sobre Internet, hace que se generen y publiquen cada vez más documentos semiestructurados. Es necesario entonces, desarrollar técnicas que aprovechen el valor agregado de los



nuevos documentos. Si bien, en la actualidad éstas no se encuentran tan desarrolladas como los modelos tradicionales, consideramos su evolución como una cuestión importante en el área de RI, especialmente a partir de investigaciones con enfoques diferentes que abordan la problemática (Uschold et al., 2018)

2.4. HERRAMIENTAS DE DESARROLLO

2.4.1. RDF/XML

RDF es el modelo de datos para metadatos, se basa en tripletes modelo entidad-atributo-valor. Un ejemplo puede ser: Objeto (curso), atributo (alumnos matriculados) y el valor (34). Este mecanismo para describir recursos es utilizado como recomendación de la W3C.

La colección de varios tripletes relacionados, se puede describir en un grafo dirigido y etiquetado, como se puede observar en la figura un ejemplo de RDF/XML Figura 3. RDF/XML es la sintaxis normativa propuesta por la W3C, para expresar/serializar un grafo RDF como un documento XML. (Gayo et al., 2017)

2.4.2. OWL

Web Ontology Language por sus siglas en inglés, es un lenguaje de marcación semántica para la publicación de ontologías en la Web. Se desarrolla como una extensión del lenguaje RDF propuesto por la W3C para los procesos que se refieren a web semántica.



Dependiendo del nivel de exactitud que se desee emplear el lenguaje OWL se subdivide en, OWL LITE, OWL FULL y OWL DL, ésta última que se usará en este proyecto pues nos garantiza integridad de información, al generar clases y subclases en la ontología.

2.4.3. PROTÉGÉ

Es un editor de ontologías de código abierto desarrollado por la Universidad de Stanford¹, que ofrece una representación gráfica de las ontologías para un mejor entendimiento utilizando librerías gráficas en JAVA.

2.4.4. JOWL

Es un plugin JQuery para navegar y visualizar documentos en formato XML/RDF. Tiene diferentes componentes, entre los que destacan un navegador de ontologías en vista de árbol, esto permite al usuario tener un control de lo que se busque en la estructura RDF.

Además, un visualizador de contenidos de las entidades de la estructura ontológica, es decir un detalle de todas las relaciones que una entidad pueda (Rebstock et al., 2016)

2.5. DEFINICIÓN DE TERMINOS BÁSICOS

2.5.1. XML

Proporciona una sintaxis superficial para documentos estructurados, pero no impone restricciones semánticas en el significado de estos documentos.



2.5.2. XML Schema

Es un lenguaje que se utiliza para restringir la estructura de los documentos XML, además de para ampliar XML con tipos de datos.

2.5.3. RDF

Es un modelo de datos para objetos ("recursos") y relaciones entre ellos, proporcionando una semántica simple para éste. Este tipo de modelo de datos puede ser representado en una sintaxis XML.

2.5.4. RDF Schema

Es un vocabulario utilizado para describir propiedades y clases de recursos RDF, con una semántica para la generalización y jerarquización tanto de propiedades como de clases.

2.5.5. OWL

Añade más vocabulario para describir propiedades y clases: entre otros, relaciones entre clases (por ejemplo, desunión), cardinalidad (por ejemplo, "uno exacto"), igualdad, más tipos de propiedades, características de propiedades (por ejemplo, simetría), y clases enumeradas.



CAPITULO III

MATERIALES Y MÉTODOS

3.1. TIPO Y DISEÑO DE INVESTIGACIÓN

3.1.1. Tipo del problema de Investigación

La presente investigación será cualitativa.

3.1.2. Diseño del problema de Investigación

El diseño cualitativo que utilizaremos será el de Investigación/Acción. Este diseño se aplica cuando una problemática de una comunidad necesita resolverse y se pretende lograr el cambio (Hernandez S., 2014)

3.2. POBLACIÓN Y MUESTRA DE INVESTIGACIÓN

3.2.1. Población de la investigación

La población tomada en cuenta para el desarrollo del presente estuvo conformada por los alumnos del 1er semestre Grupo A 2019-I de la Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano



Tabla 1: Alumnos del 1er semestre Grupo A 2019-I EPIS-UNAP

Condición	Cantidad
Alumno	27
TOTAL	27

FUENTE: Registro Académico

3.2.2. Muestra

Por la naturaleza de la investigación para el presente trabajo, el cálculo de muestra no es necesario porque la población total no supera las 27 unidades. (Sampieri, 2018)

3.2.3. Método de recolección de datos

El método utilizado para la recolección de los datos ha sido la encuesta a los alumnos del 1er semestre Grupo A 2019-I de la Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano. (Sánchez H. y Reyes, 1998)



CAPITULO IV

RESULTADOS Y DISCUSION

4.1. ANÁLISIS DEL PROBLEMA

Para lograr el objetivo de esta investigación, se desarrolló un prototipo de motor de búsqueda para manuales de programación que usa una representación ontológica de Concepto de búsqueda para devolver los resultados.

Un Concepto de búsqueda es una construcción en base a los metadatos de un Objeto de aprendizaje. Como ejemplo Estudiantes con déficit de atención. Este concepto de búsqueda se construye en base a metadatos de un objeto de aprendizaje, podemos considerar Interactividad=Alto, Recurso=Video, Duración < 3 minutos.

Para este ejemplo, una aplicación busca todos los objetos para estudiantes con déficit de atención, para esto se debe definir los objetos aptos para estudiantes con déficit de atención y esto se hace al darle semántica a los datos y objetos usando Ontologías.

Entonces se debe contar con un mecanismo ontológico para representar los metadatos de un objeto de aprendizaje.

El estándar de metadatos que usaremos para esta aplicación es el estándar LOM (Learning Object Metadata) o metadatos para objetos de aprendizaje.

La ontología que emplearemos es LOM2OWL, Esta ontología está implementada en el lenguaje OWL y proporciona una herramienta que permita hacer descripciones semánticas de los objetos de aprendizaje. Estas descripciones sirven para mejorar la

búsqueda, reutilización y uso de objetos de aprendizaje mapeando cada uno de los elementos que constituyen el estándar LOM.

Una vez que hemos elegido la Ontología en la que se basarán los metadatos de los objetos de aprendizaje empezamos a diseñar la estructura de datos para contener la ontología y definir diferentes Conceptos de búsqueda.

4.2. DESCRIPCIÓN DEL PROTOTIPO DE LA APLICACIÓN.

La aplicación se construyó para buscar conceptos previamente definidos que son creados fuera de la lógica de la aplicación. Está basado en las ontologías que desacoplan el modelo de conocimiento de la aplicación usando los metadatos asociados a los objetos de aprendizaje.

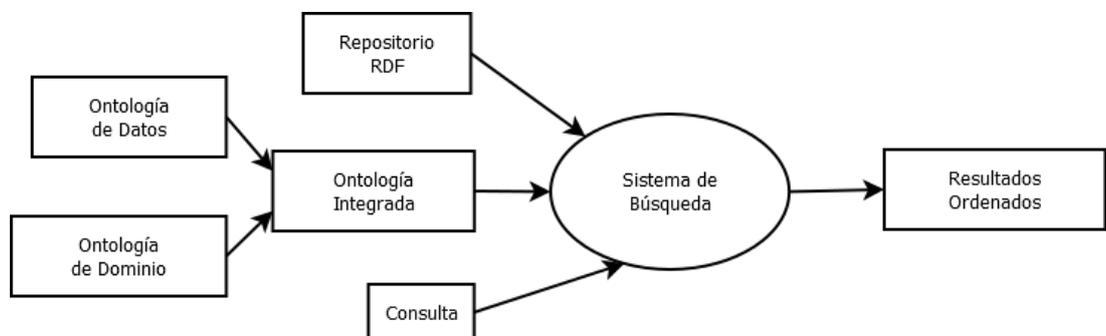


Figura 2: Entradas y salidas de información del Sistema

Elaboración propia

4.3. EL MODELO DE DATOS

4.3.1. Ontología basada en metadatos LOM

Cuando se trabaja con ontologías Web, se recomienda reusar las ontologías que ya han sido definidas, en este trabajo se usará la ontología LOM2OWL para modelar los datos de los objetos.

Las clases y relaciones básicas que son utilizadas por la ontología LOM2OWL se muestran en la siguiente figura.

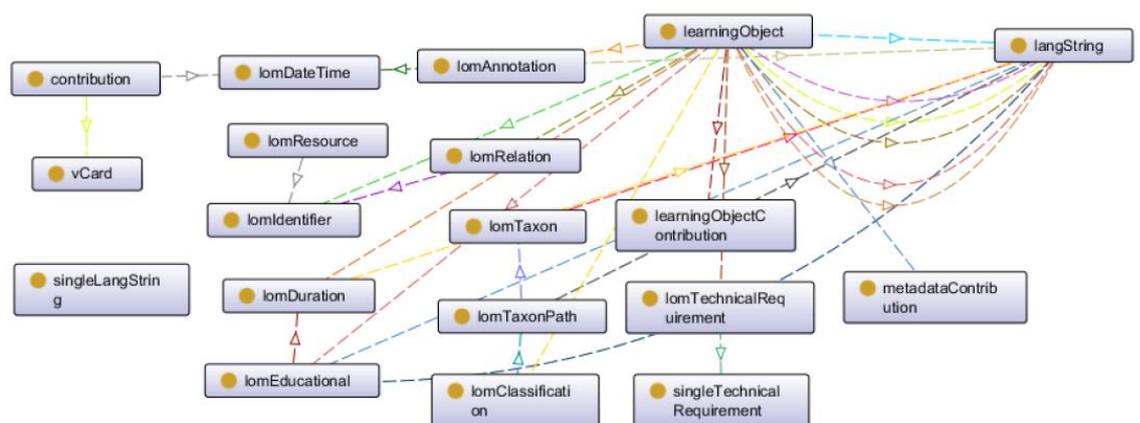


Figura 2: Clases y relaciones básicas de la ontología LOM2OWL

Fuente: Building Ontologies with Basic Formal Ontology (Arp et al., 2015)

4.3.2. Definición del patrón de expresión para las búsquedas

El patrón de expresión es la forma en que los usuarios externos deben definir los conceptos utilizando OWL para la búsqueda de objetos. Esta definición de patrón de expresión, es la construcción de una Ontología de Aplicación

La edición de la Ontología Integrada se realizó utilizando la herramienta Protege [61] la cual facilita la creación de todas las estructuras OWL además de las entidades de



prueba y la inclusión de razonadores. Cada uno de estos elementos fue necesario para poder expresar, mediante la Ontología Integrada, el tipo de búsqueda “Estilos de aprendizaje” que corresponde al patrón de expresión que se denominará “Clasificación por Aspectos”.

Como se dijo anteriormente, para definir un primer patrón de expresión que soportaría la aplicación y al mismo tiempo crear un tipo específico de búsqueda que sirviera como ejemplo, se partió del trabajo de Mejía y compañía en el cual se define un modelo que permite inferir el Estilo de Aprendizaje al que obedece un Objeto de aprendizaje basado en una combinación de sus metadatos LOM.

Específicamente, este modelo hace una clasificación de los objetos de aprendizaje en los diferentes estilos de acuerdo a los tipos de recursos que estos contienen. Los tipos de recursos determinan el estilo de aprendizaje y esta relación se expresa en una tabla que surge de un cuestionario a diversos expertos temáticos en el que se les pregunta qué tanta importancia le da un estudiante que tiene un determinado estilo de aprendizaje al hecho de que ese objeto de aprendizaje contenga un tipo de recurso determinado. Los resultados se expresan en una tabla de la siguiente manera:

Tabla 2: Relación entre los recursos LOM y las dimensiones de aprendizaje

Tipos de Recursos	Estilos de Aprendizaje							
	Procesamiento		Percepcion		Comprension		Entrada	
	Activo	Reflexivo	Intuitivo	Sensorial	Secuencial	Global	Visual	Verbal
Ejercicio	G	G	VG	G	G	I	G	VG
Simulacion	VG	I	I	VG	I	G	VG	I
Cuestionario	G	I	I	I	G	I	I	I
Diagrama	I	G	G	G	G	VG	VG	I
Figura	I	G	G	G	G	VG	VG	I
Gráfico	I	G	G	G	G	VG	VG	I
Indice	I	I	I	I	G	VG	I	G
Diapositiva	I	G	G	G	G	VG	VG	G
Tabla	I	I	I	G	G	G	G	G
Texto Narrativo	G	VG	VG	I	I	I	I	VG
Examen	G	I	I	I	G	I	I	I
Experimento	VG	G	G	VG	G	I	VG	I
Problema	VG	G	G	VG	I	I	G	VG
Planteamiento del problema	G	I	I	I	G	I	I	I
Autoevaluación	I	VG	VG	I	I	I	I	VG

Elaboración propia

Esta tabla expresa el conocimiento de los expertos temáticos sobre cómo los alumnos responden a un objeto para enseñanza de acuerdo a su estilo de aprendizaje, es decir, un recurso se adapta mejor a un estudiante de acuerdo al tipo de aprendizaje.

Se definió un Patrón de expresión para la Clasificación por Aspectos el que definirá el alcance de la aplicación

4.3.3. Construcción de la ontología.

Se ha modelado el patrón con tres clases, la clase búsqueda, la clase concepto de búsqueda, y la clase categorías que se asocian a conceptos de búsqueda.

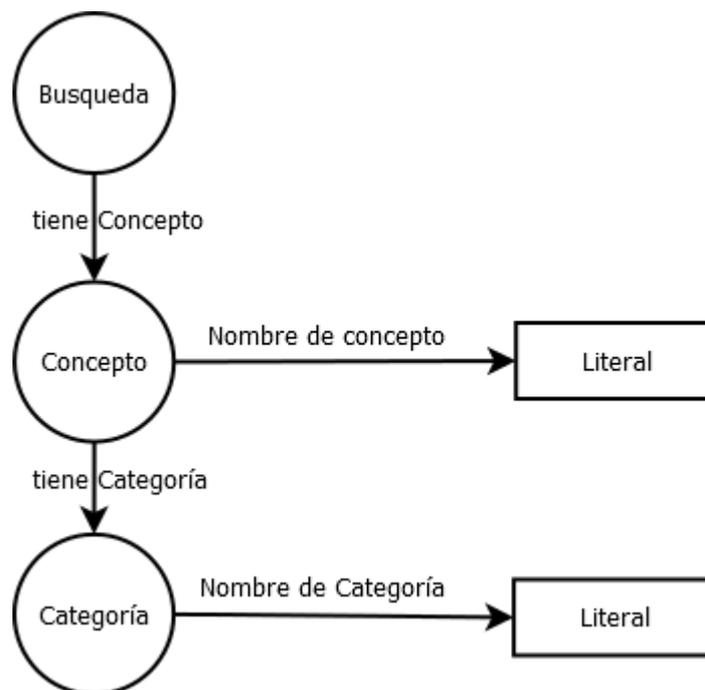


Figura 3: Clases y relaciones básicas de la ontología LOM2OWL

Elaboración propia

Con esta representación la aplicación de búsqueda obtiene un listado de los conceptos disponibles y luego las categorías asociadas a cada concepto y estos conceptos se obtienen de la información que se obtiene de otra ontología que define el modelo estudiante. La ontología que modela al estudiante y las características asociadas a él están en la siguiente figura:

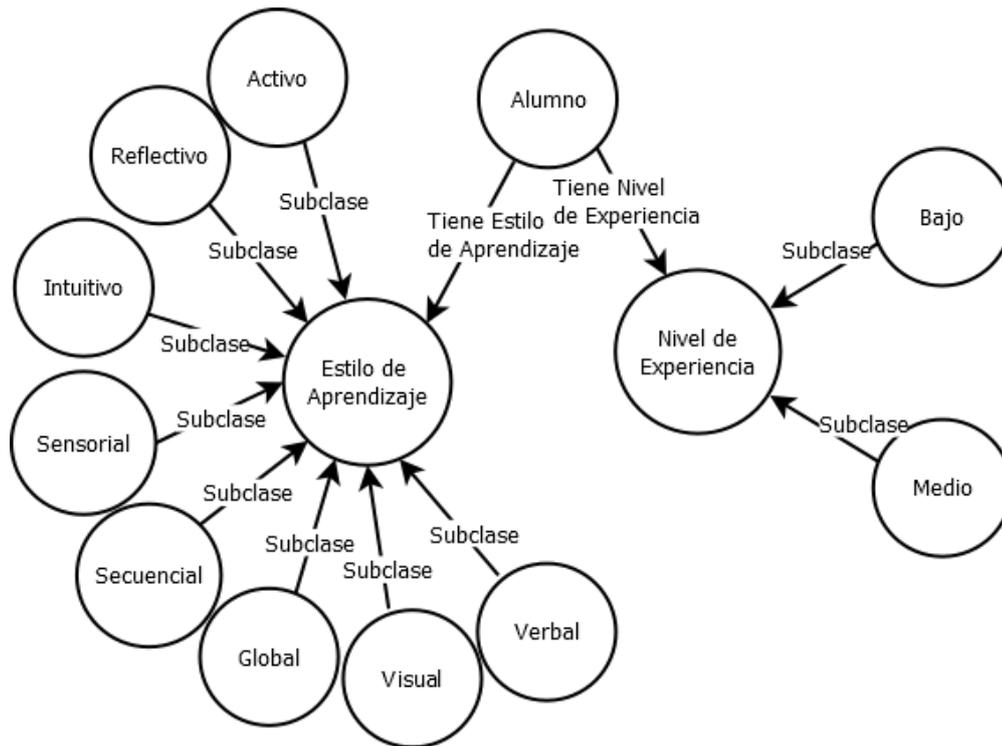


Figura 4: Ontología del alumno
Elaboración propia

Esta ontología modela que un alumno posee una característica denominada Estilo de Aprendizaje que tiene como subclases los estilos de aprendizaje y también el Nivel de experiencia.

El problema que surge ahora es cómo alinear estas dos ontologías para lograr que cada característica del estudiante se vea reflejada en la ontología de búsqueda. Para esto, se propone el uso de una relación que asocie la unión de cada característica con los conceptos de búsqueda.

4.4. ELABORACIÓN DE UN MODELO DE APLICACIÓN BASADO EN WEB SEMÁNTICA.

El modelo de aplicación basado en Web Semántica está compuesto por cinco módulos, tal como lo muestra la figura siguiente.

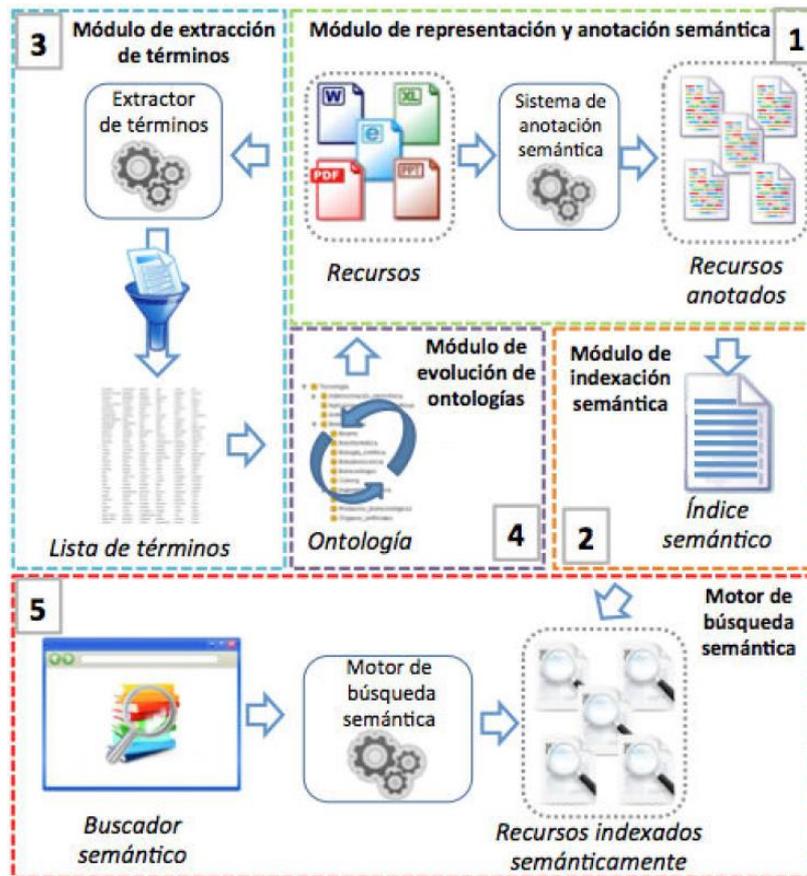


Figura 5: Modelo de aplicación basado en Web Semántica.

FUENTE: Building Ontologies with Basic Formal Ontology (Arp et al., 2015)

El modelo está compuesto de 5 módulos:

4.4.1. Módulo de representación y anotación semántica

Este módulo crea un conjunto de anotaciones en base a conceptos y propiedades de la ontología. Estas anotaciones se usarán durante el proceso de anotación semántica para identificar las expresiones lingüísticas que se asociarán con los conceptos definidos en el dominio ontológico.



4.4.2. Módulo de indexación

Este módulo obtiene las anotaciones para indexar cada anotación con información que es obtenida del dominio ontológico. Este módulo usa las relaciones taxonómicas de la ontología para indexar nueva información a cada anotación obtenida.

4.4.3. Módulo de extracción de términos

Este módulo identifica los términos más significativos de un recurso determinado. Usando patrones lingüísticos y datos estadísticos para extraer una lista de términos conformados por palabras simples o términos compuestos. Estos términos serán definidos luego como conceptos en la ontología.

4.4.4. Módulo de evolución de ontologías

Para cada término candidato, este módulo encontrará una categoría de unión entre el término y otro concepto definido en la ontología del prototipo.

4.4.5. Modulo motor de búsqueda semántico

Este módulo recupera toda la información asociada con consultas en base a palabras clave, usando las anotaciones y los índices semánticos guardados en el prototipo para hacer la búsqueda.

4.5. PRUEBA DE LA INSTANCIA DE BÚSQUEDA

Los alumnos del primer semestre grupo A 2019-I de la Escuela Profesional de ingeniería de Sistemas de la Universidad Nacional del Altiplano, luego de realizar las búsquedas al realizar las encuestas sobre el Prototipo basado en Web Semántica para



buscar Manuales de Programación para la Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano, los resultados se presentan a continuación.

RESULTADOS DE LA ENCUESTA AL USUARIO

1. Consideras que la aplicación disminuye el tiempo de búsqueda de información

Tabla 3: Disminución de tiempo de búsqueda de información

Disminución de tiempo de búsqueda de información	Nro	Porc
Totalmente de acuerdo	5	19
De acuerdo	16	59
Ni de acuerdo ni en desacuerdo	6	22
En desacuerdo	0	0
Totalmente en desacuerdo	0	0

Elaboración propia

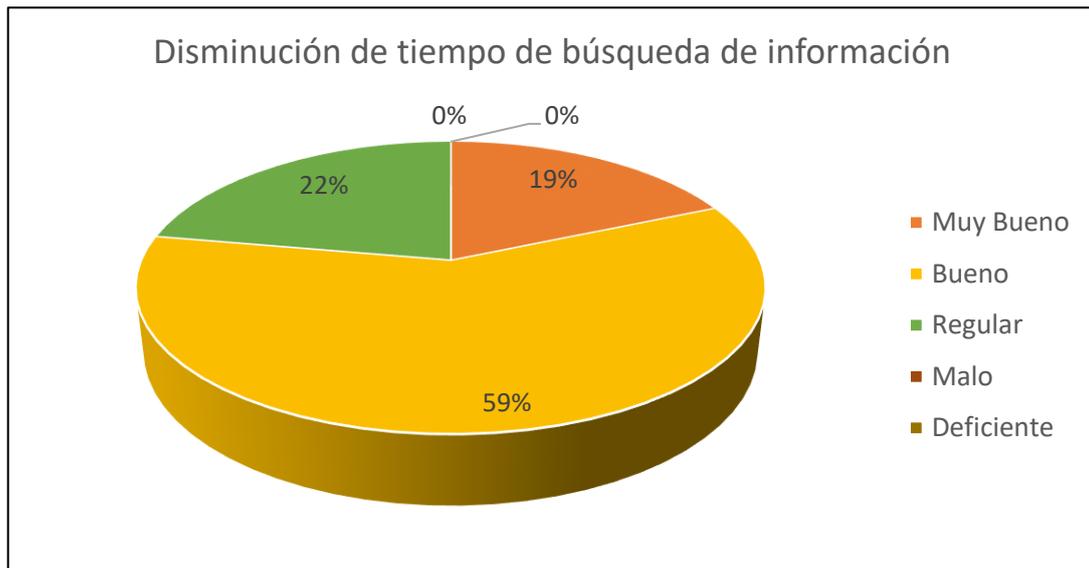


Figura 6: Disminución de tiempo de búsqueda de información

Elaboración propia.

De acuerdo a la encuesta aplicada a los usuarios, se aprecia que: el 59% de los usuarios está de acuerdo en que la aplicación disminuye el tiempo de búsqueda, seguido

del 22 % que no está ni de acuerdo ni en desacuerdo, el 19% de los usuarios están totalmente de acuerdo y ninguno está en desacuerdo y tampoco totalmente en desacuerdo.

2. Consideras que la aplicación ayuda en la búsqueda de información.

Tabla 4: La aplicación ayuda en la búsqueda de información

Disminución de tiempo de búsqueda de información	Nro	Porc
Totalmente de acuerdo	4	15
De acuerdo	17	63
Ni de acuerdo ni en desacuerdo	6	22
En desacuerdo	0	0
Totalmente en desacuerdo	0	0

Elaboración propia

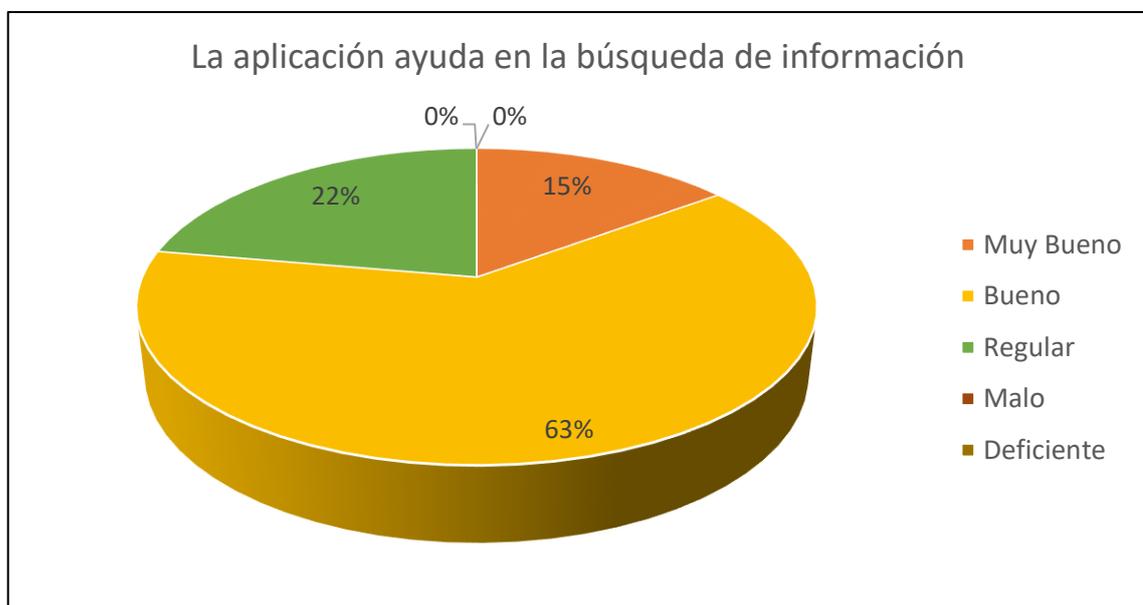


Figura 7: La aplicación ayuda en la búsqueda de información

Elaboración propia.

De acuerdo a la encuesta aplicada a los usuarios, se aprecia que: el 63% de los usuarios está de acuerdo en que la aplicación ayuda en la búsqueda de información,

seguido del 22 % que no está ni de acuerdo ni en desacuerdo, el 15% de los usuarios están totalmente de acuerdo y ninguno está en desacuerdo y tampoco totalmente en desacuerdo.

3. Consideras que la aplicación es de utilidad.

Tabla 5: La aplicación es de utilidad

La aplicación es de utilidad	Nro	Porc
Totalmente de acuerdo	1	4
De acuerdo	21	78
Ni de acuerdo ni en desacuerdo	5	18
En desacuerdo	0	0
Totalmente en desacuerdo	0	0

Elaboración propia

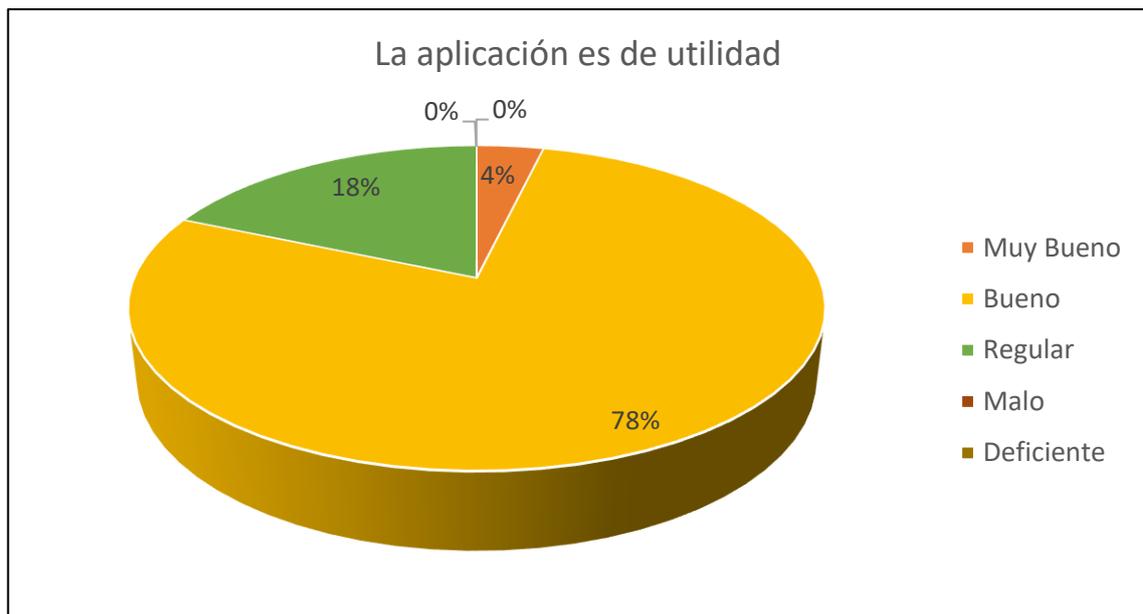


Figura 8: La aplicación es de utilidad

Elaboración propia.

De acuerdo a la encuesta aplicada a los usuarios, se aprecia que: el 78% de los usuarios está de acuerdo en que la aplicación ayuda en la búsqueda de información,

seguido del 19 % que no está ni de acuerdo ni en desacuerdo, el 4% de los usuarios están totalmente de acuerdo y ninguno está en desacuerdo y tampoco totalmente en desacuerdo.

4. Consideras que la aplicación tiene un buen desempeño

Tabla 6: La aplicación tiene un buen desempeño

La aplicación tiene un buen desempeño	Nro	Porc
Totalmente de acuerdo	2	8
De acuerdo	19	70
Ni de acuerdo ni en desacuerdo	6	22
En desacuerdo	0	0
Totalmente en desacuerdo	0	0

Elaboración propia

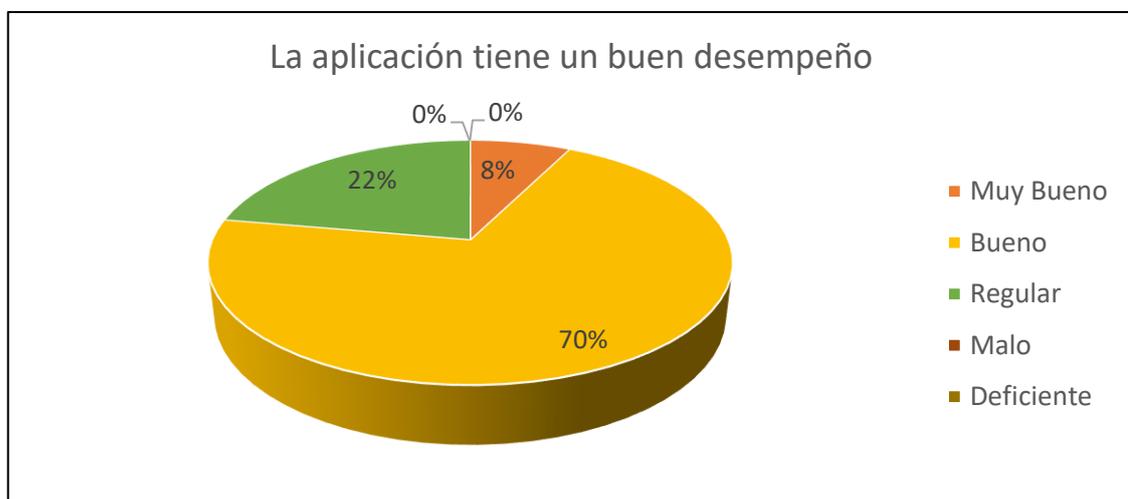


Figura 9: La aplicación tiene un buen desempeño

Elaboración propia.

De acuerdo a la encuesta aplicada a los usuarios, se aprecia que: el 70% de los usuarios está de acuerdo en que la aplicación tiene un buen desempeño, seguido del 22 %

que no está ni de acuerdo ni en desacuerdo, el 7% de los usuarios están totalmente de acuerdo y ninguno está en desacuerdo y tampoco totalmente en desacuerdo.

5. Considera que la aplicación es de fácil uso

Tabla 7: La aplicación es de fácil uso

La aplicación es de fácil uso	Nro	Porc
Totalmente de acuerdo	6	22
De acuerdo	13	48
Ni de acuerdo ni en desacuerdo	8	30
En desacuerdo	0	0
Totalmente en desacuerdo	0	0

Elaboración propia

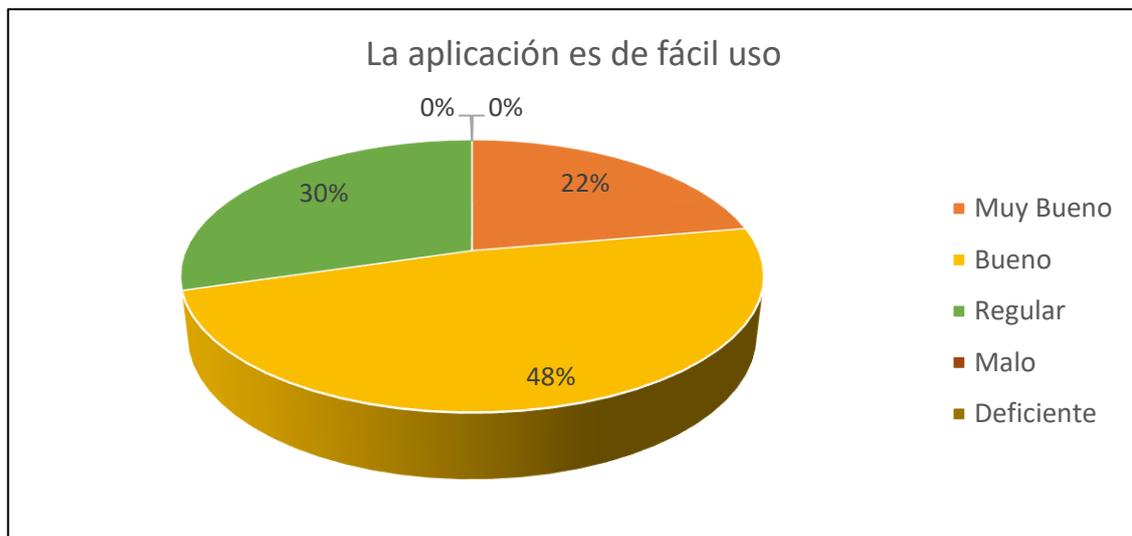


Figura 10: La aplicación es de fácil uso

Elaboración propia.

De acuerdo a la encuesta aplicada a los usuarios, se aprecia que: el 48% de los usuarios está de acuerdo en que la aplicación es de fácil uso, seguido del 30 % que no

está ni de acuerdo ni en desacuerdo, el 22% de los usuarios están totalmente de acuerdo y ninguno está en desacuerdo y tampoco totalmente en desacuerdo.

6. Consideras que la aplicación es amigable

Tabla 8: La aplicación es amigable

La aplicación es amigable	Nro	Porc
Totalmente de acuerdo	3	11
De acuerdo	20	74
Ni de acuerdo ni en desacuerdo	4	15
En desacuerdo	0	0
Totalmente en desacuerdo	0	0

Elaboración propia

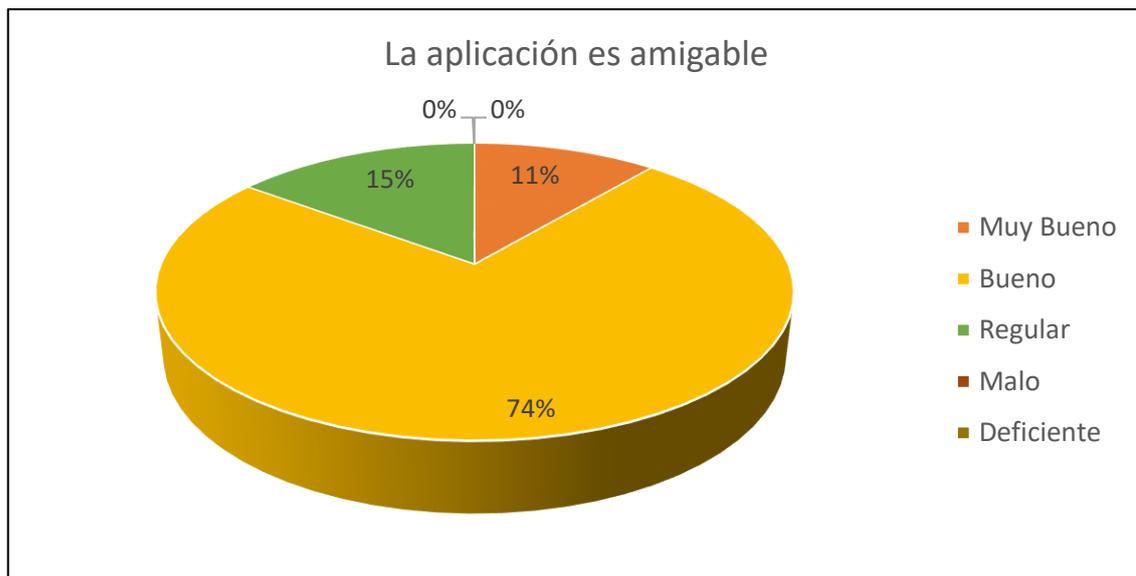


Figura 11: La aplicación es amigable

Elaboración propia.

De acuerdo a la encuesta aplicada a los usuarios, se aprecia que: el 74% de los usuarios está de acuerdo en que la aplicación es amigable, seguido del 15 % que no está



ni de acuerdo ni en desacuerdo, el 11% de los usuarios están totalmente de acuerdo y ninguno está en desacuerdo y tampoco totalmente en desacuerdo.



DISCUSIÓN

A continuación, se realiza la discusión de los resultados considerando los objetivos planteados y a los antecedentes en el marco teórico, contrastando con la evidencia empírica obtenida a través de los instrumentos de recolección de datos.

Para el objetivo general se concluye que se ha logrado desarrollar un prototipo basado en Web Semántica para buscar recursos de Programación para la Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano alcanzando un desempeño aceptable en las búsquedas de información, de acuerdo con los resultados obtenidos por (Gomez A., 2014) cuya investigación concluye con el desarrollo del modelo OASEARCH, basado en web semántica para la búsqueda de objetos de aprendizaje mediante perfilado de consultas.

El uso del modelo RDF, permitió el desarrollo del prototipo basado en Web Semántica, lo que ha permitido recuperar recursos de programación con un desempeño aceptable de forma automática lo que concuerda con los resultados obtenidos por (Rodriguez G., 2014) cuya investigación concluye con el desarrollo de una nueva metodología de anotación y recuperación semántica basado en modelos ontológicos formales, cubriendo todo el ciclo de vida de las anotaciones semánticas teniendo en cuenta las posibles actualizaciones de información y facilitando la adaptación continua del dominio de aplicación a través de la evolución de la ontología subyacente.

Se ha desarrollado el modelo de búsquedas, diseñando una ontología para que los usuarios realicen consultas en la web. Lo que concuerda con los resultados obtenidos por (Coronado Altamirano, 2017) cuya investigación concluye con el desarrollo de una



herramienta para la anotación semántica automática de documentos pdf basado en ontologías.

Se ha logrado la definición e implementación de una instancia de búsqueda que permite realizar consultas basadas en la Web Semántica lo que concuerda con los resultados obtenidos por (Gomez M., 2014) cuya investigación concluye con el diseño de un modelo basado en ontologías en el dominio de la Ingeniería Informática, que facilite la recuperación de documentos almacenados en repositorios de contenido digital.



V. CONCLUSIONES

- Se ha logrado el desarrollado de un prototipo basado en Web Semántica para buscar recursos de Programación para la Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano alcanzando un desempeño aceptable en las búsquedas de información que han realizado los usuarios.
- El modelo de aplicación desarrollado en base a la Web Semántica usando el modelo RDF, ha permitido recuperar recursos de programación con un desempeño aceptable de forma automática.
- El diseño de una ontología para modelar las búsquedas que realizan los usuarios del sistema se ha logrado de forma satisfactoria, permitiendo realizar consultas en la web con un desempeño aceptable.
- La definición de una instancia de búsqueda que permita realizar las consultas basadas en la Web Semántica ha permitido que el usuario encuentre información relacionada a programación de forma satisfactoria.
- La implementación de la Instancia de Búsqueda se ha realizado utilizando el lenguaje de programación Python, usando las distintas librerías que integra, lo que posibilitó una gran flexibilidad en las consultas en la Web.



VI. RECOMENDACIONES

A los investigadores en web semántica se recomienda realizar comparaciones de eficiencia en la recuperación automática de información semántica mediante el uso de corpus alternativo usando diferentes técnicas de cálculo de similitud para el proceso de los datos.

A los desarrolladores de sistemas web basados en ontologías se recomienda evaluar otros gestores de bases de datos para mejorar los tiempos de respuesta, considerando que gran parte del rendimiento de la aplicación depende del rendimiento de la base de datos y de la arquitectura en la que ha sido diseñada.

A los tesisistas que deseen continuar esta investigación se recomienda evaluar nuevas herramientas de recuperación automática de información semántica considerando el estado del arte y las nuevas aplicaciones orientadas al procesamiento de información.

A los estudiantes y egresados de la escuela profesional de Ingeniería de Sistemas se recomienda utilizar técnicas y algoritmos de recuperación de información automática para procesar y revalorare los idiomas Quechua y Aymara.



VII. REFERENCIAS

- Allemang, D., & Hendler, J. (2012). *Semantic Web for the Working Ontologist*. Morgan Kaufmann.
- Arp, R., Smith, B., & Spear, A. D. (2015). *Building Ontologies with Basic Formal Ontology*. MIT Press.
- Barrière, C. (2016). *Natural Language Understanding in a Semantic Web Context*. Springer International Publishing.
- Coronado Altamirano, G. (2017). *Desarrollo de una herramienta para la anotación semántica automática de documentos pdf basado en ontologías*. Pontificia Universidad Católica del Perú.
- Criado Fernández, L. (2009). *Procedimiento semi-automático para transformar la web en web semántica*. Universidad Alfonso X el Sabio.
- Espinoza F. (2014). *Diseño de una herramienta para la anotación semántica automática de documentos basados en ontologías en el dominio de la Ingeniería Informática*. Pontificia Universidad Católica del Perú.
- Gayo, J. E. L., Prud'hommeaux, E., Boneva, I., Kontokostas, D., Ding, Y., & Groth, P. (2017). *Validating RDF Data*. Morgan & Claypool Publishers.
- Gomez A. (2014). *OASEARCH: Modelo de aplicación basado en web semántica para la búsqueda de objetos de aprendizaje mediante perfilado de consultas*. Universidad EAFIT.
- Gomez M. (2014). *Diseño de un modelo para la recuperación de documentos basado en ontologías en el dominio de la Ingeniería Informática*. Pontificia Universidad Católica del Perú.
- Hebeler, J., & Fisher, M. (2009). *Semantic Web Programming*. Programming-Wiley.
- Hendler, J., Gandon, F., & Allemang, D. (2020). *Semantic Web for the Working*



Ontologist: Effective Modeling for Linked Data, RDFS, and OWL. Association for Computing Machinery and Morgan & Claypool Publishers.

Hernandez S. (2014). Metodología de la investigación. In *Mc Graw Hill Education 6a edición*. <https://doi.org/>- ISBN 978-92-75-32913-9

Hopkins, M., & Powell, J. (2015). *A Librarian's Guide to Graphs, Data and the Semantic Web*. Chandos Publishing .

Jeff, Z., & Pan, Z. (2014). *Semantic Web Enabled Software Engineering*. IOS Press.

Maynard, D., Bontcheva, K., & Augenstein, I. (2016). *Natural Language Processing for the Semantic Web*. Morgan & Claypool Publishers. <https://books.google.com.pe/books?id=5fm3DQAAQBAJ>

Rebstock, M., Naujok, K. D., Janina, F., Huemer, C., Paulheim, H., Röder, P., & Tafreschi, O. (2016). *Ontologies-Based Business Integration*. Springer Berlin Heidelberg.

Rodriguez G. (2014). *Extracción Semántica de Información basada en Evolución de Ontologías*. Universidad de Murcia.

Sampieri, R. H. (2018). *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta*. McGraw-Hill Interamericana. <https://books.google.com.pe/books?id=5A2QDwAAQBAJ>

Sánchez H. y Reyes, C. (1998). *Metodología y Diseños en la Investigación Científica*. Mantaro.

Sikos, L. (2015). *Mastering Structured Data on the Semantic Web*. Data-Apress.

Uschold, M., Ding, Y., & Groth, P. (2018). *Demystifying OWL for the Enterprise*. Morgan & Claypool Publishers.

Workman, M. (2015). *Semantic Web: Implications for Technologies and Business Practices*. Springer International Publishing.

Yu, L. (2014). *A Developer's Guide to the Semantic Web*. Springer.



ANEXOS

UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO
FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y
SISTEMAS
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS
ENCUESTA

Introducción:

La información es netamente académica y nos ayudará con la investigación del desarrollo de un prototipo basado en web semántica para buscar manuales de programación para la escuela profesional de Ingeniería de Sistemas de la Universidad Nacional del Altiplano

Instrucciones:

La presente encuesta consta de 10 preguntas, solicito a Ud. Expresar su opinión personal, marcando con una X una de las alternativas que cree conveniente.

Edad: _____

Fecha: _____

1. Consideras que la aplicación disminuye el tiempo de búsqueda de información
 - (1) Totalmente en desacuerdo
 - (2) En desacuerdo
 - (3) Ni de acuerdo ni en desacuerdo
 - (4) De acuerdo
 - (5) Totalmente de acuerdo

2. Consideras que la aplicación ayuda en la búsqueda de información
 - (1) Totalmente en desacuerdo
 - (2) En desacuerdo
 - (3) Ni de acuerdo ni en desacuerdo
 - (4) De acuerdo
 - (5) Totalmente de acuerdo



3. Consideras que la aplicación es de utilidad
 - (1) Totalmente en desacuerdo
 - (2) En desacuerdo
 - (3) Ni de acuerdo ni en desacuerdo
 - (4) De acuerdo
 - (5) Totalmente de acuerdo

4. Consideras que la aplicación tiene un buen desempeño
 - (1) Totalmente en desacuerdo
 - (2) En desacuerdo
 - (3) Ni de acuerdo ni en desacuerdo
 - (4) De acuerdo
 - (5) Totalmente de acuerdo

5. Consideras que la aplicación es de fácil uso
 - (1) Totalmente en desacuerdo
 - (2) En desacuerdo
 - (3) Ni de acuerdo ni en desacuerdo
 - (4) De acuerdo
 - (5) Totalmente de acuerdo

6. Consideras que la aplicación es amigable
 - (1) Totalmente en desacuerdo
 - (2) En desacuerdo
 - (3) Ni de acuerdo ni en desacuerdo
 - (4) De acuerdo
 - (5) Totalmente de acuerdo