UNIVERSIDAD NACIONAL DE LOJA

Facultad de la Energía, las Industrias y los Recursos Naturales No Renovables

CARRERA DE INGENIERÍA EN SISTEMAS

"Análisis y Experimentación con Linked Data aplicado a la Biodiversidad"

"TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERO EN SISTEMAS"

Autor:

- Juan Andrés Maldonado Ajila

Director:

- Ing. María del Cisne Ruilova Sánchez Mg. Sc

BY NC SA

LOJA – ECUADOR

2020

Certificación del Director

Ing. María del Cisne Ruilova Sánchez Mg.Sc
DOCENTE DE LA CARRERA DE INGENIERÍA EN SISTEMAS

CERTIFICA:

Que el egresado Juan Andrés Maldonado Ajila autor del Trabajo de Titulación "Análisis y Experimentación con Linked Data aplicado a la Biodiversidad", ha sido dirigido y orientado bajo mi asesoramiento y reúne a satisfacción los requisitos exigidos en una investigación de este nivel, por lo cual autorizo su presentación y sustentación.

Loja, 30 de Octubre de 2019

Ing. María del Cisne Ruilova Sánchez Mg.Sc
DIRECTOR DEL TRABAJO DE TITULACIÓN

Autoría

Yo, **JUAN ANDRÉS MALDONADO AJILA** declaro ser autor del presente trabajo de titulación y eximo expresamente a la Universidad Nacional de Loja y a sus representantes jurídicos de posibles reclamos o acciones legales por el contenido de la misma.

Adicionalmente acepto y autorizo a la Universidad Nacional de Loja, la publicación del trabajo de Titulación en el Repositorio Institucional – Biblioteca Virtual.

Firma:

Cedula: 1104474828

Fecha: 06 de febrero de 2020

CARTA DE AUTORIZACIÓN DE TESIS POR PARTE DEL AUTOR, PARA LA CONSULTA, REPRODUCCIÓN PARCIAL O TOTAL Y PUBLICACIÓN ELECTRÓNICA DEL TEXTO COMPLETO

Yo, **JUAN ANDRÉS MALDONADO AJILA**, declaro ser el autor de la tesis titulada: "**ANÁLISIS Y EXPERIMENTACIÓN CON LINKED DATA APLICADO A LA BIODIVERSIDAD"**, como requisito para optar al grado de: **INGENIERO EN SISTEMAS**; autorizo al Sistema Bibliotecario de la Universidad Nacional de Loja para con fines académicos, muestre al mundo la producción intelectual de la Universidad, a través de la visibilidad de su contenido de la siguiente manera en el repositorio Digital Institucional:

Los usuarios pueden consultar el contenido de este trabajo en el RDI, en las redes de información del país y del exterior, con las cuales tenga convenio la Universidad.

La Universidad Nacional de Loja, no se responsabiliza por el plagio o copia de la tesis que realice un tercero.

Para constancia de esta autorización, en la ciudad de Loja, a los seis días del mes de febrero del dos mil veinte.

Firma:

Autor: Juan Andrés Maldonado Ajila
Cedula: 1104474828
Dirección: San Sebastián (ciudad de Loja).
Correo Electrónico: jamaldonadoa@unl.edu.ec
Teléfono: 072570591 **celular:** 0967760403

DATOS COMPLEMENTARIOS

Directora de Tesis: Ing. María del Cisne Ruilova Sánchez Mg.Sc.
Tribunal de Grado: Ing. Óscar Miguel Cumbicus Pineda Mg.Sc.
Ing. Roberth Gustavo Figueroa Díaz Mg.Sc.
Ing. Marlon Santiago Viñán Ludeña Mg.Sc.

Agradecimiento

Agradezco primeramente a Dios por permitirme cumplir con mi objetivo y acompañarme continuamente en todas las acciones que he realizado.

Agradezco a mis padres y a mi familia en general por darme todo el apoyo necesario para poder cumplir esta meta, al igual que brindarme toda la orientación indispensable para culminar exitosamente con esta actividad.

Agradezco a la ingeniera María del Carmen Cabrera por su guía y colaboración en el desarrollo del presente trabajo de titulación.

Agradezco a mis profesores por toda la enseñanza impartida en todo el tiempo que estuve en la Universidad, así como, agradecer a mi tutora la ingeniera María del Cisne Ruilova por otorgarme su tiempo y conocimientos, guía para culminar con éxito el desarrollo de mi trabajo de titulación.

Juan Maldonado

Dedicatoria

A mis Padres John y Nancy por el apoyo brindado alrededor de todo el tiempo de estudio que se ha invertido. Además de todas las orientaciones y consejos que siempre me dieron en toda mi formación personal.

A mi Familia en general por estar al pendiente de mí en todos los momentos de complejidad que se presentaban y me enseñaban a través del ejemplo a cumplir mis metas, logrando superar las dificultades con éxito.

Juan Maldonado.

Índice de Contenidos

Certificación del Director II
Autoría III
CARTA DE AUTORIZACIÓN DE TESIS POR PARTE DEL AUTOR, PARA LA CONSULTA, REPRODUCCIÓN PARCIAL O TOTAL Y PUBLICACIÓN ELECTRÓNICA DEL TEXTO COMPLETO IV
Agradecimiento V
Dedicatoria VI
a. Título 1
b. Resumen 2
Summary 3
c. Introducción 4
d. Revisión de Literatura. 6
1. Web Semántica. 6
1.1. Antecedentes. 6
1.2. Características. 7
1.3. Arquitectura. 7
1.4. Aplicaciones 10
2. Linked Data. 10
2.1. Linked Open Data. 11
2.2. Principios de Linked Data. 11
2.3. Esquema de 5 Estrellas LOD. 11
2.4. Trabajos Relacionados con Linked Data. 12
3. Biodiversidad. 13
3.1. Acercamiento de LOD. 13
3.2. Problemas Principales con LOD. 13
3.3. Trabajos Relacionados. 14
3.4. Universidad Nacional De Loja. 15
4. Análisis y Experimentación Linked Data hacia la Biodiversidad. 16
4.1. Metodología de Generación y Publicación Linked Data. 16

4.2. Motores de Representación RDF. 17

4.3. Metodologías para construir Ontologías. 19

4.4. Herramientas. 23

e. Materiales y Métodos 25

1. Contexto 25

2. Proceso 25

2.1. Objetivo 1: Analizar el estado del arte vinculado a la Web Semántica y Linked Data en el campo de Biodiversidad. 26

2.2. Objetivo 2: Especificar un modelo de Linked Data y Web Semántica para implementarla en el campo de Biodiversidad. 26

2.3. Objetivo 3: Evaluar el modelo de Linked Data en la Web Semántica implementado en el campo de Biodiversidad. 26

3. Recursos. 27

3.1. Científicos 27

3.2. Técnicos 28

4. Participantes. 29

f. Resultados 30

1. Objetivo 1: Analizar el estado del arte vinculado a la Web Semántica y Linked Data en el campo de Biodiversidad. 30

1.1. Revisión Sistemática de Literatura. 30

1.2. Beneficios y Limitaciones. 32

1.3. Análisis de los contenidos. 34

2. Objetivo 2: Especificar un modelo de Linked Data y Web Semántica para implementarla en el campo de Biodiversidad. 35

2.1. Identificar y señalar el ambiente de experimentación. 35

2.2. Exploración del entorno local para aplicación de experimentación. 36

2.3. Analizar la estructura, desarrollo de vocabularios, ontologías de datos Linked Data. 36

2.4. Verificar el esquema estructural y la organización de los datos. 60

3. Objetivo 3: Evaluar el modelo de Linked Data en la Web Semántica implementado en el campo de Biodiversidad. 61

3.1. Determinar los criterios orientados a la selección de las herramientas para la evaluación. 61

3.2. Evaluación de la Ontología. 61

3.3. Realizar la validación del modelo. 63

g. Discusión 65

1. Desarrollo de la propuesta alternativa 65

Objetivo Nro. 1: Analizar el estado del arte vinculado a la Web semántica y Linked Data en el campo de Biodiversidad. 65

Objetivo Nro. 2: Especificar un modelo de Linked Data y Web semántica para implementarla en el campo de Biodiversidad. 67

Objetivo Nro. 3: Evaluar el modelo de Linked Data en la Web Semántica implementado en el campo de Biodiversidad 69

h. Conclusiones. 72

i. Recomendaciones. 73

j. Bibliografía 74

k. Anexos 82

ANEXO 1: REVISIÓN SISTEMÁTICA DE LITERATURA 83

ANEXO 2: CRITERIOS DE SELECCIÓN. 112

ANEXO 3: RDF STORES 115

ANEXO 4: CUADRO METODOLOGÍAS PARA ONTOLOGÍAS. 117

ANEXO 5. DOCUMENTACIÓN BLOG JARDÍN BOTÁNICO 119

ANEXO 6. VOCABULARIO – ONTOLOGÍA LDJBRE 124

ANEXO 7. DESCRIPCIÓN LDJBRE 146

ANEXO 8. VERIFICACIÓN RDF 152

ANEXO 9. ENTREVISTA. 153

ANEXO 10. EVALUACIÓN DE LA ONTOLOGÍA. 155

ANEXO 11. MÉTODO EXPERIMENTO. 161

ANEXO 12. LICENCIA CREATIVE COMMONS. 170

ANEXO 13. CERTIFICADO TRADUCCIÓN FINE TUNED ENGLISH 171

ANEXO 14. LISTA DE ACRÓNIMOS..... 172

Índice de Figuras

Figura 1. Estructura de la Web Semántica, Maza [3]. 7
Figura 2. Tecnologías Web Semántica primeros tres niveles..... 8
Figura 3. Tecnologías Web Semántica nivel de Lógica, prueba y confianza. 9
Figura 4. Metodología Generación y Publicación Linked Data 16
Figura 5. Flujo de Procesos hacia Linked Data. 35
Figura 6. Representación esquema de metodología Generación Linked Data. 37
Figura 7. Representación gráfica del Modelo conceptual JBRE. 42
Figura 8. Representación gráfica del Vocabulario LDJBRE. 46
Figura 9. Representación de Clases LDJBRE – Herramienta Protegé. 47
Figura 10.Representación gráfica de propiedades LDJBRE – Herramienta Protegé... 47
Figura 11.Representación gráfica de relaciones LDJBRE – Herramienta Protegé. 48
Figura 12.Representación gráfica limpieza de Datos – Herramienta LODRefine. 49
Figura 13.Representación gráfica Fragmento ontología LDJRE – Herramienta Sublime Text. 50
Figura 14. Representación gráfica LOD Refine – Submenu Exportación. 51
Figura 15. Representación gráfica en la Edición RDF-Schema vista LODRefine. 52
Figura 16. Representación gráfica LOD Refine con Enlaces hacia DBPedia. 53
Figura 17.Representación gráfica LOD Refine con la Vista de Tripletas RDF. 53
Figura 18.Representacion gráfica interfaz Virtuoso OpenLink..... 54
Figura 19.Representación gráfica Consulta Sparql – Herramienta Virtuoso. 55
Figura 20.Representación gráfica de resultados Sparql – Herramienta Virtuoso..... 55
Figura 21.Representación gráfica de resultados generales interfaz Servidor Fuseki. . 56
Figura 22. Representación gráfica interfaz Fuseki-Consulta por defecto. 57
Figura 23.Representación gráfica Fuseki-Resultado consulta pregunta 1..... 57
Figura 24. Representación gráfica Fuseki-Resultado consulta pregunta 2..... 58
Figura 25. Representación gráfica Fuseki-Resultado consulta pregunta 3..... 58
Figura 26. Representación gráfica Fuseki Resultado consulta pregunta 4. 58
Figura 27. Representación gráfica Resultado pregunta 4 servidor Fuseki relación inversa. 59
Figura 28.Representación gráfica Resultado pregunta 5 servidor Fuseki. 59

Figura 29.Representación gráfica de Visualización herramienta Pubby. 60
Figura 30.Representación gráfica Validación Owl – Herramienta OPPS..... 61
Figura 31. Representación gráfica resultados verificación CPQ1 - Protégé 63
Figura 32. Representación gráfica fragmento Vista evaluación automática validación-grafo. 64
Figura 33.Representación gráfica de fragmento de grafo – vista de W3C validador. .. 64
Figura 34. Representación gráfica de Clases Vista-OntoGraf – Herramienta Protege. 68
Figura 35.Representación gráfica clases vista OWLViz – Herramienta Protege. 69
Figura 36. Representación gráfica Vista Fuseki número total de triples. 71
Figura 37.Representación gráfica - vista RDF Stores [65]. 115
Figura 38. Representación gráfica Página de Informes JBRE..... 119
Figura 39.Representación gráfica de la vista de las Publicaciones JBRE. 120
Figura 40.Representación gráfica de la Portada Documento Informe JBRE 2017. ... 121
Figura 41.Representación gráfica del documento Flora Nativa JBRE. 122
Figura 42. Representación gráfica Extracto Validación W3C. 152
Figura 43. Representación gráfica verificación CPQ1 - Protégé 157
Figura 44. Representación gráfica verificación CPQ2 – Protégé..... 158
Figura 45. Representación gráfica verificación CPQ3 – Protégé..... 158
Figura 46. Representación gráfica verificación CPQ4 – Protégé..... 159
Figura 47. Representación gráfica verificación CPQ5 – Protégé..... 160
Figura 48. Escala de Medición con resultados finales..... 163
Figura 49.Representación gráfica de ejemplares categorizados por uso del JBRE... 166
Figura 50.Representación gráfica de las secciones jardín botánico “Reinaldo Espinosa”. 167

Índice de Tablas

TABLA I. ETAPAS DE EVOLUCIÓN DE LA WEB. 6
TABLA II. SUBLENGUAJES DE ONTOLOGÍA - OWL. 9
TABLA III. ESTRELLAS DE LINKED OPEN DATA - LOD. 11
TABLA IV. ETAPAS METODOLOGÍA GENERACION DE LINKED DATA. 17
TABLA V. PROCESOS METODOLOGÍA DE REVISIÓN SISTEMÁTICA DE LITERATURA. 31
TABLA VI. GRUPO DE PROYECTOS DE LA REVISIÓN SISTEMÁTICA DE LITERATURA. 32
TABLA VII. FORMATO DE ENTREGA DE DATOS HACIA USUARIOS. 38
TABLA VIII. NIVELES DE FORMALIDAD DE UNA ONTOLOGÍA. 40
TABLA IX. FASE ESPECIFICACIÓN DE METHONTOLOGY. 40
TABLA X. VOCABULARIOS SOBRE BIODIVERSIDAD. 41
TABLA XI. ESTRUCTURA DE LA MUESTRA DEL ARCHIVO DATOS.CSV. 41
TABLA XII. CLASES Y PROPIEDADES VOCABULARIO LDJBRE. 42
TABLA XIII. SELECCION DE VOCABULARIOS UTILIZADOS. 43
TABLA XIV. URIS REUTILIZADAS PARA LA ONTOLOGÍA. 43
TABLA XV. TÉRMINOS REUTILIZADOS DE LOS VOCABULARIOS. 44
TABLA XVI. TÉRMINOS NUEVOS DE LA ONTOLOGÍA. 44
TABLA XVII. PROPIEDADES DE LA ONTOLOGÍA LDJBRE. 45
TABLA XVIII. PATRON URIS MODELAMIENTO RDF. 51
TABLA XIX. ENLACES HACIA LOS RECURSOS DE DBPEDIA. 52
TABLA XX. FASE DE EXPLOTACIÓN DE INFORMACIÓN. 56
TABLA XXI. RESULTADOS LISTA DE CHEQUEO DE ERRORES COMUNES. 62
TABLA XXII. EVALUACIÓN ONTOLOGÍA CASO DE PRUEBA 1. 62
TABLA XXIII. RESPUESTAS A PREGUNTAS REVISION SISTEMATICA DE LITERATURA. 66
TABLA XXIV. CLASE UNIVERSITY VOCABULARIO LDJBRE. 68
TABLA XXV. PROPIEDAD BIOCLIMATE VOCABULARIO LDJBRE. 68
TABLA XXVI. HERRAMIENTAS GENERACIÓN Y PUBLICACIÓN DE LINKED DATA. 70
TABLA XXVII. FÓRMULAS EMPLEADAS EN LA EVALUACIÓN DEL MODELO. 70
TABLA XXVIII. RESUMEN MÉTRICAS EVALUACIÓN MODELO JARDÍN BOTÁNICO. 71

TABLA XXIX. PALABRAS CLAVES PARA LA GENERACIÓN DE LA CADENA DE BÚSQUEDA. 84
TABLA XXX. CADENAS DE BUSQUEDA RSL..... 84
TABLA XXXI. SELECCIÓN DE DOCUMENTOS RSL..... 85
TABLA XXXII. DESCRIPCIÓN RESULTADOS ES01. 88
TABLA XXXIII. DESCRIPCIÓN RESULTADOS ES02. 89
TABLA XXXIV. DESCRIPCIÓN RESULTADOS ES03. 89
TABLA XXXV. DESCRIPCIÓN RESULTADOS ES04. 90
TABLA XXXVI. DESCRIPCIÓN RESULTADOS ES05. 91
TABLA XXXVII. DESCRIPCIÓN RESULTADOS ES06. 91
TABLA XXXVIII. DESCRIPCIÓN RESULTADOS ES07. 92
TABLA XXXIX. DESCRIPCIÓN RESULTADOS ES08. 93
TABLA XL. DESCRIPCIÓN RESULTADOS ES09. 93
TABLA XLI. DESCRIPCIÓN RESULTADOS ES10. 94
TABLA XLII. DESCRIPCIÓN RESULTADOS ES11. 95
TABLA XLIII. DESCRIPCIÓN RESULTADOS ES12. 95
TABLA XLIV. DESCRIPCIÓN RESULTADOS ES13. 96
TABLA XLV. DESCRIPCIÓN RESULTADOS ES14. 96
TABLA XLVI. DESCRIPCIÓN RESULTADOS ES15. 97
TABLA XLVII. DESCRIPCIÓN RESULTADOS ES16. 98
TABLA XLVIII. DESCRIPCIÓN RESULTADOS ES17..... 98
TABLA XLIX. DESCRIPCIÓN RESULTADOS ES18. 99
TABLA L. DESCRIPCIÓN RESULTADOS ES19. 99
TABLA LI. DESCRIPCIÓN RESULTADOS ES20. 100
TABLA LII. DESCRIPCIÓN RESULTADOS ES21. 101
TABLA LIII. DESCRIPCIÓN RESULTADOS ES22. 101
TABLA LIV. DESCRIPCIÓN RESULTADOS ES23. 102
TABLA LV. DESCRIPCIÓN RESULTADOS ES24. 102
TABLA LVI. DESCRIPCIÓN RESULTADOS ES25. 103
TABLA LVII. DESCRIPCIÓN RESULTADOS ES26. 103
TABLA LVIII. DESCRIPCIÓN RESULTADOS ES27. 104
TABLA LIX. DESCRIPCIÓN RESULTADOS ES28. 104
TABLA LX. DESCRIPCIÓN RESULTADOS ES29. 104
TABLA LXI. DESCRIPCIÓN RESULTADOS ES30. 105
TABLA LXII. DESCRIPCIÓN RESULTADOS ES31. 105

TABLA LXIII. DESCRIPCIÓN RESULTADOS ES32. 105
TABLA LXIV. DESCRIPCIÓN RESULTADOS ES33 106
TABLA LXV. DESCRIPCIÓN RESULTADOS ES34 107
TABLA LXVI. DESCRIPCIÓN RESULTADOS ES35 107
TABLA LXVII. RESPUESTA RSL – PREGUNTA 2. 110
TABLA LXVIII. CRITERIOS DE SELECCIÓN REVISIÓN SISTEMATICA DE LITERATURA. 112
TABLA LXIX. METODOLOGÍAS PARA LA CREACIÓN DE ONTOLOGÍAS PARTE 1. 117
TABLA LXX. METODOLOGÍAS PARA LA CREACIÓN DE ONTOLOGÍAS PARTE 2. 118
TABLA LXXI. DATOS PUBLICADOS POR EL BLOG PRINCIPAL DEL JARDÍN BOTÁNICO. 123
TABLA LXXII. LISTA DE CHEQUEO ERRORES COMUNES EN ONTOLOGÍAS. ... 155
TABLA LXXIII. EVALUACIÓN ONTOLOGÍA CASO DE PRUEBA 1. 156
TABLA LXXIV. EVALUACIÓN ONTOLOGÍA CASO DE PRUEBA 2. 156
TABLA LXXV. EVALUACIÓN DE ONTOLOGÍA CASO DE PRUEBA 3. 156
TABLA LXXVI. EVALUACIÓN DE ONTOLOGÍA CASO DE PRUEBA 4. 157
TABLA LXXVII. EVALUACIÓN DE ONTOLOGÍA CASO DE PRUEBA 5. 157
TABLA LXXVIII. DEFINICIÓN DE OBJETIVOS DE EXPERIMENTACIÓN. 161
TABLA LXXIX. VARIABLES DEL EXPERIMENTO. 162
TABLA LXXX. CATEGORIAS DE USO DEL JARDÍN BOTÁNICO “REINALDO ESPINOSA”. 165
TABLA LXXXI. SECCIONES DEL JARDÍN BOTÁNICO “REINALDO ESPINOSA”. .. 167
TABLA LXXXII. CARACTERÍSTICAS DEL EXPERIMENTO APLICACIÓN LINKED DATA. 168
TABLA LXXXIII. RESULTADOS FINALES DE LA APLICACIÓN DEL EXPERIMENTO. 169
TABLA LXXXIV. TABLA DE ACRÓNIMOS EMPLEADOS. 172

a. Título

“Análisis y Experimentación con Linked Data aplicado a la Biodiversidad”

b. Resumen

La Universidad Nacional de Loja posee un área fuerte a potenciar dentro del área de Biodiversidad, es pertinente destacar la información referente al Jardín Botánico “Reinaldo Espinosa”, fuente de muchas investigaciones, hasta la actualidad muy alejadas de las tecnologías o muy poco utilizadas, por lo tanto, se muestra la oportunidad de exponer mediante un enfoque tecnológico el potencial de compartir dicha información, lograr un impacto y contribución importante en esta área de conocimiento, a través de la incursión de adoptar las tecnologías relacionadas a Web Semántica y Linked Data.

El presente Trabajo de Titulación está orientado a: “Definir un modelo de Web Semántica y Linked Data enfocado en el campo de Biodiversidad”, tal que, el resultado esperado de los diferentes procesos desarrollados, es un modelo prototipo que utilice las tecnologías de Web Semántica y Linked Data, para resaltar la importancia que posee en la conservación de la biodiversidad. En la zona de estudio explorada, el jardín botánico “Reinaldo Espinosa” se denota problemas como la dispersión del conocimiento, debido a que las investigaciones realizadas se encuentran esparcidas en diferentes repositorios; al igual que, la dificultad en el proceso de búsqueda de información, causado por la falta de recursos digitales disponibles para consultar.

El modelamiento principal de los recursos del jardín botánico “Reinaldo Espinosa” así como los procesos subyacentes luego de la generación del modelo son abordados y ejecutados dentro de la Metodología de Generación y Publicación de Linked Data, para la fase de modelado en la creación de la ontología, se contempló emplear la metodología Methontology, conjuntamente con las diferentes herramientas como Open Link Virtuoso, Pubby, CmapTools, Protege, LodRefine, Tomcat y Apache Jena Fuseki, se aplicaron recursos como el método científico, método experimental, trascendentales para el desarrollo del modelo prototipo.

El principal aporte del Trabajo de Titulación es el desarrollo de un modelo prototipo que abarque la realidad de los datos del jardín botánico “Reinaldo Espinosa”, así como un primer vocabulario. Dentro de las líneas futuras se destaca la contribución de investigadores para robustecer el modelo, el desarrollo de una herramienta como un portal de biodiversidad, aplicación semántica capaz de consumir los datos del modelo a posterior.

Summary

Universidad Nacional de Loja has a strong area to enhance within the area of Biodiversity, it is pertinent to highlight the information regarding the Reinaldo Espinosa Botanical Garden, source of many investigations, so far from technology or very little used, therefore, the opportunity is shown to expose through a technological approach the potential to share such information, achieve an important impact and contribution in this area of knowledge, through the incursion of adopting technologies related to Semantic Web and Linked Data.

This Degree Work is oriented to: “Define a Semantic Web and Linked Data model focused on the field of Biodiversity”, such that, the expected result of the different processes developed, is a prototype model that uses Semantic Web technologies and Linked Data, to highlight the importance it has in the conservation of Biodiversity. In the explored study area, the “Reinaldo Espinosa Botanical Garden” denotes problems such as the dispersion of knowledge, because the research carried out is scattered in different repositories; like, the difficulty in the process of searching for information, caused by the lack of available digital resources to consult.

The main modeling of the resources of Reinaldo Espinosa Botanical Garden as well as the underlying processes after the generation of the model are addressed and executed within the Linked Data Generation and Publication Methodology, for the modeling phase in the creation of the ontology, It was contemplated to use the Methontology methodology, together with the different tools such as Open Link Virtuoso, Pubby, CmapTools, Protege, LodRefine, Tomcat and Apache Jena Fuseki, resources were applied such as the scientific method, experimental method, transcendental for the development of the model prototype.

The main contribution of the Degree Work is the development of a prototype model that covers the reality of the data of the Reinaldo Espinosa Botanical Garden, as well as a first vocabulary. Within the future lines, the contribution of researchers to strengthen the model, the development of a tool such as a Biodiversity portal, a semantic application capable of consuming the data of the model afterwards, stands out.

c. Introducción

El presente Trabajo de Titulación se enmarca en la temática central de la Web Semántica y Linked Data, la incorporación de llegar a un contexto más organizado con los datos en la red, representar los contenidos técnicamente y gestionar la información dándole una mayor estructura al publicarla, proporcionar un aporte positivo al área del conocimiento relacionado a la Biodiversidad, a través del modelo que represente al jardín botánico “Reinaldo Espinosa”, cuya utilidad significativa va dirigida hacia editores de datos, investigadores y estudiantes que se relacionan con esta área u otras disciplinas afines.

Se abarca las diferentes conceptualizaciones englobando la terminología alrededor del tópico de la Web Semántica y Linked Data, describiendo diferentes tecnologías acordes a trabajos relacionados, que aporten significativamente en la elaboración de una estructura fuerte, vocabulario idóneo de los datos del jardín botánico “Reinaldo Espinosa”, con el fin de contribuir en la conservación de la biodiversidad. Con lo antes mencionado se logró cumplir con los objetivos que contemplan “Analizar el estado del arte vinculado a la Web Semántica y Linked Data en el campo de Biodiversidad”, así como Especificar un modelo de Linked Data y Web Semántica para posteriormente Evaluar el modelo de Linked Data en la Web Semántica aplicado en el campo de Biodiversidad.

Cabe mencionar que la aplicación de este enfoque tecnológico aporta grandes beneficios, sin embargo, también acarrea varios limitantes, siendo uno de estos el aporte económico, es decir, el factor de financiamiento para cubrir la realización de proyectos de Linked Data es alto; otra limitación circunda en la accesibilidad de los datos, en vista que, algunos proyectos no han podido garantizar la accesibilidad con el paso del tiempo, resultado de datos perdidos, con nodos caídos, es decir sin acceso; también se puede mencionar la poca valoración que se otorga al momento de aplicar los procesos de modelamiento, tarea compleja de ejecutar que pasa desapercibida y su nivel de importancia no es el adecuado; por último se destaca, hacer posible el enfoque de Linked Data demanda un conocimiento técnico importante, en vista que la denominación de Web Semántica (Web de los datos) es conocida como una extensión más técnica de la actual Web (encaminada hacia una Web Inteligente) integra un equipo

multidisciplinario difícil de integrar en ramas de conocimiento del dominio de biodiversidad.

El documento desarrollado contiene secciones de Revisión de Literatura con el preámbulo de los conceptos de Web Semántica, Linked Data, biodiversidad donde se describen las herramientas, estructura y trabajos relacionados con estas principales temáticas; sección de Materiales y Métodos donde se describe el contexto de la metodología tratada, el contenido del jardín botánico “Reinaldo Espinosa”, los recursos empleados, los participantes involucrados; seguido de la sección de Resultados generado de acuerdo a los objetivos planteados; en la sección Discusión se aborda el desarrollo de la propuesta alternativa con la exposición del modelo que se ha creado; sección de Conclusiones donde se presentan los puntos concretos basados en la experimentación y análisis de los datos con relación a los objetivos del Trabajo de Titulación; sección de Recomendaciones en donde se plasma la descripción de los puntos de desenlace con mayor prioridad de acuerdo a la experiencia desarrollada; la información bibliográfica se manifiesta en la sección Bibliografía; al igual que la información complementaria en Anexos.

d. Revisión de Literatura.

La Web actual desde sus inicios ha experimentado una serie de cambios, producto de tecnologías de innovación que evolucionan, crean más necesidades en las personas y ofrecen una gran variedad de beneficios. Producto de la globalización y la amplitud de la red por todo el mundo, en su afán de comunicación y transmisión de información hacia toda la población, surge el término de Web Semántica, cuyo problema a resolver radica en la capacidad de dar a las máquinas mayor comprensión para interpretar y extraer información existente en la Web; problema que se ahondará más adelante.

1. Web Semántica.

Conocido como el creador de la Web, Tim Berners-Lee describe a la Web Semántica como la extensión de la Web (actual) en la cual los datos se entregan con un significado adecuado a su definición, para facilitar que las computadoras y las personas trabajen en cooperación; es decir, una red de datos que pueda procesarse directa o indirectamente por máquinas, Bustán, González [1].

1.1. Antecedentes.

La Web tiene un crecimiento continuo, el manejo de datos llegó a convertirse en la actualidad en un sistema global de información masivo, accedido por la mayoría de personas a nivel mundial, las cuales de manera directa e indirecta al transcurrir la historia han contribuido con sus diferentes etapas de evolución, y son indistintamente el punto de apoyo para su crecimiento. Entonces existen varias etapas por las cuales ha pasado la Web hasta la actualidad, por lo que destacados en Gómez, Romero, Orellana [2], Maza [3] se puede observar en la TABLA I un resumen.

TABLA I.

ETAPAS DE EVOLUCIÓN DE LA WEB.

Año	Nombre	Descripción
1991-2003	Web 1.0	Web que permite solo lectura.
2004	Web 2.0	Web con capacidad de leer y escribir.
2006	Web Semántica	Web de datos, extensión de la web basada en el significado, apoyado en lenguajes universales y accesibilidad fácil.
2010-2020	Web 3.0	Mejora de la web 2.0 con capacidad de procesar y razonar contenido, hacer deducciones lógicas a partir de la información.
2020-2030	Web 4.0	Cooperación de inteligencias tanto de los usuarios como de las cosas para su comunicación y toma de decisiones.

Elaborado por el autor.

1.2. Características.

La Web Semántica posee varios puntos de distinción que mejoran el contenido de la Web, Bustán, Gonzales [1] menciona algunas características de las cuales se destacan:

- Información con mejor estructura para ser localizada y encontrada de forma más sencilla, a través de datos (modelos, formatos) de conocimiento compartido.
- Su estandarización permite que los contenidos logren ser procesados y utilizados por cualquier Software, optimizando la integración e interoperabilidad entre aplicaciones.
- La incorporación de contenido semántico a las páginas que se alojan en Internet, permitiendo una mejor organización de la información, asegurando búsquedas con significado más precisas.
- Permitir la identificación de elementos y recursos disponibles en la web.
- Minimización de la problemática relacionada a la apertura de datos.

1.3. Arquitectura.

Para que la Web Semántica tenga tanta importancia en la organización y procesamiento de la información se requiere de varios componentes que reflejen un estándar, formato que permita la anotación descriptiva del significado de los elementos que se exponen en la Web, de tal manera, basados en los proyectos Bustán [1], Gómez [2], Fernández [4], Quezada [5] se muestra su estructura en la Figura 1 y los detalles de las tres primeras capas en la Figura 2 .

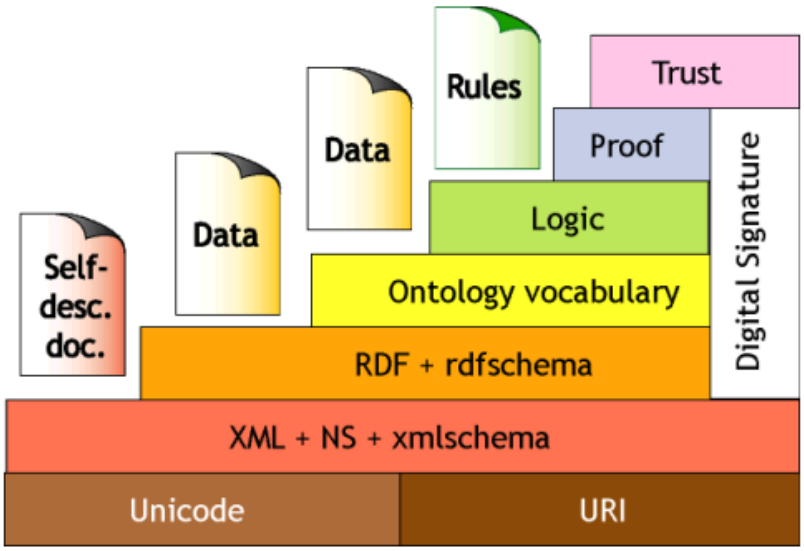


Figura 1. Estructura de la Web Semántica, Maza [3].

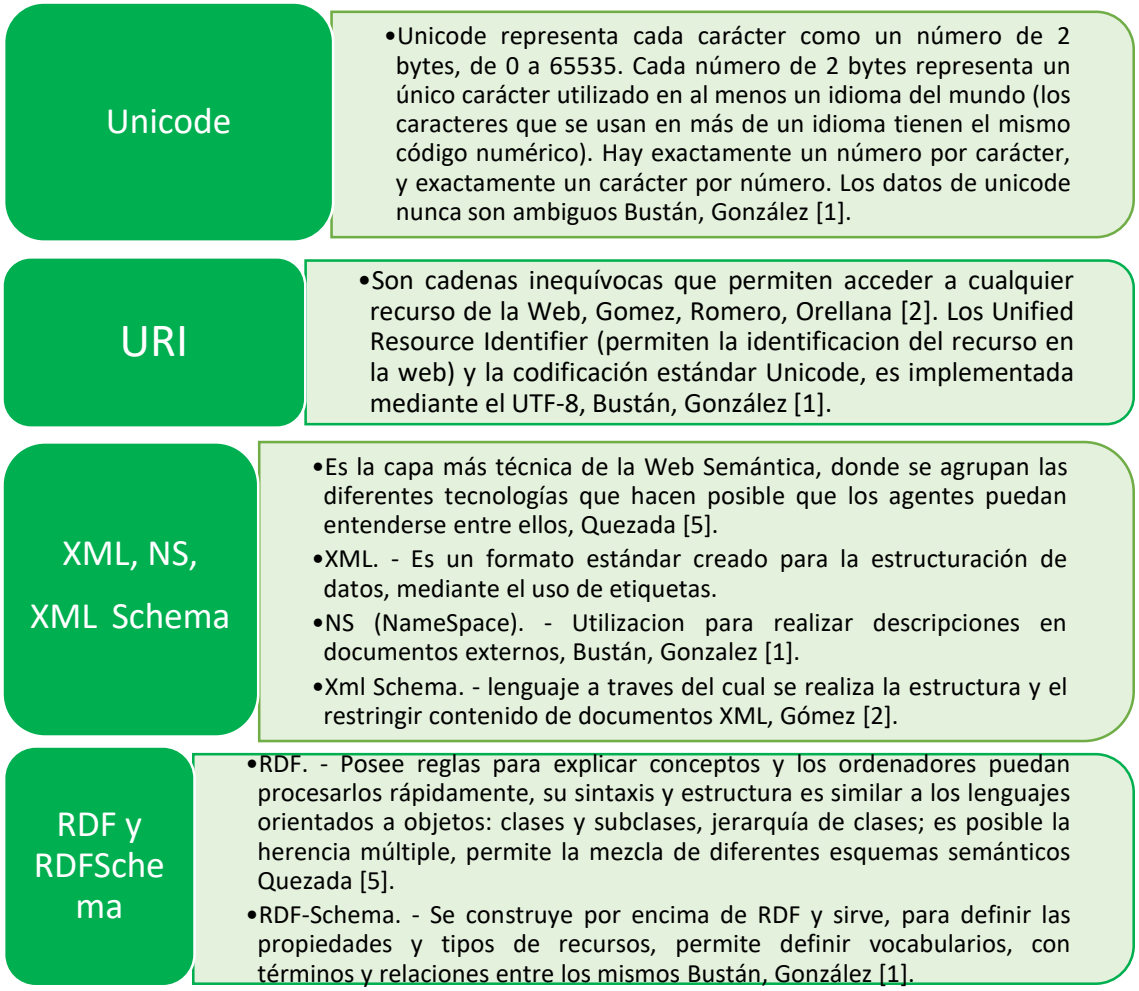


Figura 2. Tecnologías Web Semántica primeros tres niveles.
Elaborado por el Autor.

Se describe a RDF como un marco de trabajo para representar información en la Web, descritos como conjuntos de sentencias, donde se pueden visualizar a manera de grafos dirigidos y etiquetados (nodos y arcos), Piedra [6]. Estos componentes facilitan la semántica en la Web, de manera que la notación de las relaciones entre los recursos proporciona un modelado de datos considerado como estándar.

- **Ontologías.**

Añade vocabulario para describir propiedades y clases: tales como relaciones entre clases, cardinalidad, igualdad y caracterización de las propiedades. El lenguaje OWL es un mecanismo para desarrollar temas o vocabularios específicos para asociar recursos descrito en Quezada [5]. OWL proporciona tres sublenguajes mostrados en la TABLA II, cada uno con mayor nivel de expresividad que el anterior, diseñados para ser usados por comunidades específicas de programadores y usuarios, ampliado en Fernández [4].

TABLA II.

SUBLENGUAJES DE ONTOLOGÍA - OWL.

Nombre	Descripción
OWL Lite	Lenguaje con más bajo nivel de complejidad formal en su notación, apunta hacia los usuarios que poseen interés en requerir una clasificación jerárquica y restricciones simples, Fernández [4].
OWL DL (Lógica de Descripciones)	Lenguaje más estructurado, con restricciones en las construcciones del lenguaje OWL, su base radica en la ayuda a aquellos usuarios que necesiten la máxima expresividad mientras conservan completamente el factor de procesamiento (todas las conclusiones poseen garantía de ser computables) y resolución (todos los cálculos se resolverán en un tiempo finito) Fernández [4].
OWL FULL	Lenguaje con la prioridad de ofrecer a los usuarios el máximo nivel de expresividad y libertad sintáctica de RDF, sin embargo, sin ofrecer garantías computacionales. Usado básicamente para agregar más significado a un vocabulario RDF o OWL, Fernández [4].

Elaborado por el Autor.

Los niveles para describir los recursos explícitamente comprenden las tecnologías de Unicode, URI, XML, RDF, RDFSchema facilitando la apertura de una mayor organización de información, aclarar cualquier tipo de ambigüedad y enfocarse en un contexto predefinido del conocimiento al etiquetar; el desarrollar vocabularios y utilizar terminología estándar para abstraer información corresponden al nivel de manipulación de datos; finalmente en el nivel de reglas se encuentran las últimas capas de Lógica, Prueba y Confianza (Figura 3) donde se proporciona el factor de entendimiento para la máquina y su procesamiento debido a la inferencia de conocimiento mediante axiomas.

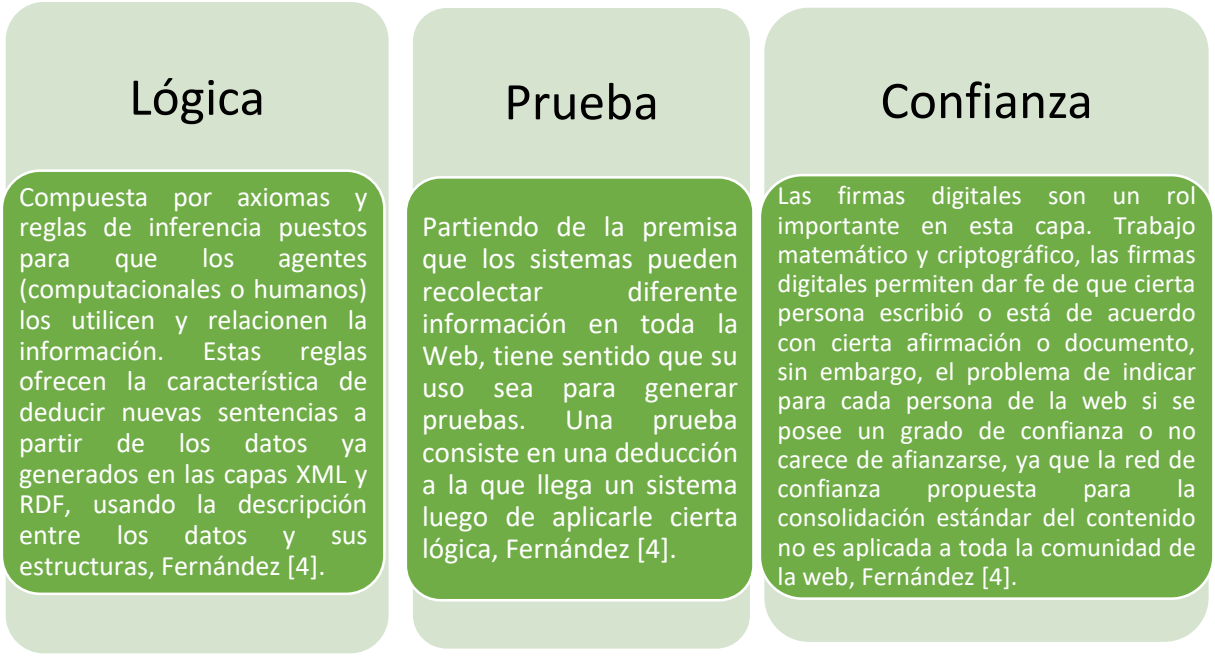


Figura 3. Tecnologías Web Semántica nivel de Lógica, prueba y confianza.
Elaborado por el Autor.

1.4. Aplicaciones.

Las aplicaciones de la Web Semántica poseen mucha variedad, por lo tanto, tratado en Quezada [5] destacamos dos:

- Tecnologías semánticas aplicadas a la gestión del conocimiento: generación de portales web basados en ontologías de manera que mejoran la accesibilidad a la información contenida en la organización.
- Gobierno Semántico: basada en la construcción de infraestructura (tanto de software, modelos, servicios) en la oferta de servicios inteligentes por parte de las administraciones públicas, ejemplo de ello se encuentran los proyectos SemanticGov, HOPS y OntoGov.

A pesar de los esfuerzos tanto de los editores de datos, organizaciones y organismos de enfoque en desarrollo de la Web, el cambio de Web 2.0 a 3.0 no se ve reflejado homogéneamente en el mundo, existen diferencias remarcadas en distintos países que no se encuentran en esa transición, sin embargo, podemos resaltar los ejemplos más conocidos en la aplicación de la Web Semántica, Quezada [5].

- RSS. – Vocabulario RDF basado en XML que permite la catalogación de Información y encontrar con mayor precisión la misma de acuerdo a las preferencias de los Usuarios.
- FOAF. – Utilizado en proyectos que permiten crear páginas Web para describir personas y vínculos, es un vocabulario RDF que permite tener información personal de forma práctica y simplificada para ser procesada, compartida y reutilizada.

2. Linked Data.

El término se refiere a un estilo de publicación y el mecanismo de interconexión de datos estructurados en la Web. El objetivo detrás de Linked Data es que el valor y la utilidad de los datos aumenta cuanto más se interrelaciona con otros datos, Maza [3]. Su finalidad es publicar y enlazar datos estructurados en la Web utilizando RDF, lenguaje para representar información sobre recursos mediante tripletas. Mientras más interconectados estén los datos mayor es su utilidad en la Web ampliado en Gómez, Romero, Orellana [2].

Se habla de Linked Data cuando ha sido el mecanismo por el cual, la Web Semántica a través de la aplicación de las tecnologías que la conforman modelan el contenido de un

dominio de información (área de conocimiento), mediante la vinculación hacia otros conjuntos de datos que tiene una relación que refuerce el conocimiento modelado.

2.1. Linked Open Data.

LOD es un movimiento creciente para que las organizaciones puedan hacer que sus datos existentes puedan estar en un formato legible por la máquina, para permitir a los usuarios crear y combinar conjuntos de datos y para hacer sus propias interpretaciones de los datos disponibles en formatos digeribles, descrito en Maza [3].

2.2. Principios de Linked Data.

Dado los trabajos en Bustán [1], Gómez [2], Maza [3] los cuatro principios en los que se basa la Web de Datos Enlazados son:

1. Utilizar URIs como identificador único de los recursos en la Web.
2. Utilizar el protocolo HTTP para nombrar y resolver la ubicación de los datos identificados mediante las URIs.
3. Representar los datos en RDF y utilizar SPARQL como lenguaje de consulta de los datos generados.
4. Incluir enlaces a otras URIs para permitir la localización de más datos enlazados.

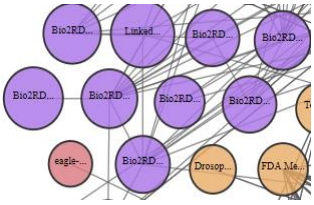
2.3. Esquema de 5 Estrellas LOD.

Tim Berners-Lee [7] propuso un despliegue de 5 estrellas de LOD basándose en la búsqueda de la liberación parcial o total de los datos, permitir con mayor facilidad el uso y reutilización de los mismos, la clasificación de estrellas se muestra en la TABLA III.

TABLA III.

ESTRELLAS DE LINKED OPEN DATA - LOD.

5 Estrellas de Linked Open Data			
Disponibilidad de los datos en la Web	**Estrella**	**Ejemplo**	**Ilustración**
Cualquier Formato	★	PDF	
Forma estructurada	★ ★	XLS	
Formato no propietario	★ ★ ★	CSV	

Mediante el uso de URIs	★★★★	RDF	
Enlazados con otros datos de referencia	★★★★★	LOD	

Elaborado por el Autor.

2.4. Trabajos Relacionados con Linked Data.

LOD es una red mundial de datos organizadas por burbujas temáticas relacionada entre sí a través de internet, mediante la identificación de datos existentes disponibles bajo licencias abiertas, aptos para su conversión en RDF bajo los principios de Linked Data formando así la nube de Linked Data, mencionado en Gómez, Romero, Orellana [2].

Actualmente LOD se encuentra en crecimiento donde se desarrollan varios proyectos en todo el mundo abarcando diversos dominios y áreas de aplicación, de los cuales tenemos algunos relacionados al ámbito universitario e investigativo:

- CourseWare. – Repositorio semántico que contiene y publica datos vinculados RDF, el mismo que posee más de 50 millones de ítems, recomendados para todos los involucrados con la enseñanza ampliado en Quezada [5].
- Gobierno de Estados Unidos. – Maza [3] menciona en el año de 2014 diferentes portales de acuerdo a los niveles de estrellas de Linked Data (ver TABLA III), poseen conjuntos datos en diferentes formatos. Los principales son los portales de sitios como: la ciudad de Austin, Nueva Orleans, Oakland, Chicago, Seattle, San Francisco, estado de Illinois, Colorado, Columbia, New York, Washington, Utah, Nebraska y muchos otros más son los casos de éxitos que se mencionan acorde a la aplicación de Linked Data.
- Proyecto Linked Open Data University of Munster. – LODUM es creado con el objetivo de disponer datos de la investigación de la Universidad de Munster de Alemania como datos vinculados abiertos para mejorar la transparencia y la visibilidad de la Universidad permitiendo la accesibilidad a los datos públicos, ampliado en Quezada [5].
- Gobierno de Chile. – es otro de los países que poseen portales accesibles a LOD como sus datos de la Biblioteca del Congreso nacional de Chile, gobierno

abierto, Consejo de Transparencia alineados a la propuesto de Linked Data y Gobiernos Abiertos mencionado en Maza [3].

- The Open University. – (data.open.ac.uk) es una plataforma que se desarrolló como parte del proyecto CSAC Lucero para extraer, conectar y exponer los datos de los recursos almacenados en los repositorios institucionales de la Universidad y situarlos a disposición abierta para que estos sean reutilizados, redactado en Quezada [5].

3. Biodiversidad.

La Biodiversidad es esencial para la vida en la tierra y motiva muchos esfuerzos en la recopilación de datos sobre especies, que son utilizados por investigadores para el estudio de los seres vivos, Braun, Almonacid [8]. Los datos de colecciones biológicas sin embargo experimentan un cambio, en el caso de la biodiversidad producto de una serie de impactos y degeneración del ambiente en las últimas temporadas, ocasionada por diferentes factores; una solución práctica radica en realizar un análisis de esa información, dada la oportunidad de la digitalización, foco importante para el empleo de las ciencias informáticas, más aun acercándonos a la propuesta de LOD y obtener mayores beneficios en la exposición, reutilización, explotación de información aplicada a la biodiversidad.

3.1. Acercamiento de LOD.

La comunidad informática de la biodiversidad ha hecho grandes progresos hacia el logro de este objetivo mediante la creación de vocabularios comunes compartidos (Darwin Core), en Braun, Almonacid [8] la visión de ofrecer beneficios significativos para la ciencia que estudia la biodiversidad y sus disciplinas asociadas.

Si bien es cierto se han generado logros en los estándares de datos, formatos y mecanismos de intercambio por parte de las organizaciones de normalización como la Organización Internacional de Normalización (ISO), el Open Geospatial Consortium (OGC) y el Worl Wide Web Consortium (W3C), la total interoperabilidad no está garantizada plenamente entre los sistemas de información de biodiversidad tratados en Braun, Almonacid [8], Chavan, Penev [9].

3.2. Problemas Principales con LOD.

La actualidad de los datos primarios sobre biodiversidad radica en una serie de factores que mencionamos a continuación y ratificados en Braun [8], Chavan [9], Wieczorek [10], Piedra [11], tratados como problemas o barreras que todavía no están totalmente solventados:

- Ausencia de incentivos para los editores de datos.
- Rápido descubrimiento de recursos de datos de biodiversidad.
- Procesos de publicación de recursos de datos sobre biodiversidad dispersos y heterogéneos.
- Vacío de conocimientos sobre biodiversidad.
- Falta de integración de información.

3.3. Trabajos Relacionados.

Producto de la realización de la Revisión Sistemática de Literatura (ver Anexo 1) y revisión bibliográfica, se muestra los principales trabajos que se aplican los enfoques de Linked Data y Biodiversidad con mayor relevancia:

- Bio2rdf tiene como objetivo centralizar datos sobre bioinformática y dejarlos disponibles en RDF sobre la Web. Promueve la visión de los datos enlazados dentro de la comunidad de la bioinformática mostrando su potencial al poder ser integrados y utilizados transparentemente por diferentes investigadores en el mundo, Gómez, Romero, Orellana [2], Nolín [12].
- Portal de Biodiversidad Global Biodiversity Information Facility (GBIF) cuyo objetivo es estructurar nodos nacionales en una red accesible vía web de manera libre y gratuita, mencionado en Braun, Almonacid [8].
- Biodiversity international a través del vocabulario Darwin Core, Wieczorek [10] pretende obtener beneficios en nombre de los recursos Fito genéticos europeos y la comunidad de bancos de genes y, al mismo tiempo, compartir un recurso genético rico en disciplinas específicas de información.
- Theophrastus es un proyecto que trata de una plataforma de anotación semántica en el cual un concepto básico es la noción de entidades de interés, que pueden ser consideradas como categorías. Por ejemplo, en el caso del dominio marino, las entidades de interés incluyen especies de peces, zonas acuáticas, países, etc. Cada categoría está asociada a una base de conocimiento (KB accesible a través de un punto final SPARQL[1]) y a una

[1] Lenguaje de consulta, acrónimo recursivo del inglés Protocol and RDF Query Language.

consulta de plantilla SPARQL, el sistema tiene por objeto ayudar a los biólogos en sus investigaciones sobre las especies y la biodiversidad, ampliado en Fafalios, Papadakos [13].

3.4. Universidad Nacional De Loja.

"La Universidad Nacional de Loja (UNL), es una Institución de Educación Superior, laica, autónoma, de derecho público, con personería jurídica y sin fines de lucro, generando propuestas alternativas a los problemas nacionales, con responsabilidad social; reconociendo y promoviendo la diversidad cultural y étnica y la sabiduría popular, apoyándose en el avance científico y tecnológico, en procura de mejorar la calidad de vida del pueblo ecuatoriano", UNL [14].

La UNL encargada de la formación profesional de cada estudiante que se integra a la institución ofrece una apertura en las diferentes líneas de Investigación Científica, Desarrollo Tecnológico e Innovación, que direccionan y promueven la generación de conocimiento científico, base para la innovación tecnológica y la transferencia de conocimientos, ampliado en UNL [15]. Una de ellas a destacar es la Biodiversidad y Biotecnología, entonces de manera preliminar destacamos el interés que la institución, hacia el TT toma de manera positiva en la generación de conocimiento, es decir, modelar datos vinculados aplicados en información de biodiversidad, constituye una buena opción de contribuir con la investigación científica de esta área de conocimiento.

Se señala que la Universidad Nacional de Loja posee el jardín botánico "Reinaldo Espinosa", la información que alberga corresponde al dominio de recursos de biodiversidad que no ha sido explotado en tecnologías asociadas a la Web Semántica, Linked Data o computabilidad, material digital utilizado para investigaciones (trabajos publicados oficialmente) hasta la actualidad.

3.4.1. Jardín Botánico "Reinaldo Espinosa".

Uno de sus objetivos centrales es: "Establecer programas de Investigación Científica, que permitan generar información para el manejo y conservación de la flora nativa y endémica de la región Sur del Ecuador"; al igual que "Impulsar la conservación ex-situ de las especies nativas y endémicas de la región Sur del Ecuador", y otros más en proyección de la formación e incentivación de los estudiantes, interés colectivo por la conservación, JBRE [16].

La visión que tiene es liderar la conservación ex-situ de la flora nativa y endémica de la región sur del Ecuador; se complementa con la misión de “Impulsar la Investigación y Educación Ambiental para lograr la conservación ex-situ de la flora nativa y endémica de la región sur del Ecuador”, JBRE [16]; ofrece una amplia apertura a todos los estudiantes en el campo de investigación y formación, evidenciado tanto en proyectos de titulación como prácticas pre-profesionales.

4. Análisis y Experimentación Linked Data hacia la Biodiversidad.

La problemática ya expuesta en párrafos previos se resume en el desafío actual de la heterogeneidad de los tipos de datos y formatos que son generalmente difícil de descubrir/acceder desde múltiples fuentes, por lo que estudiantes, investigadores, y comunidad en general dedicada a difundir el conocimiento de biodiversidad se enfrenta a un gran desafío.

4.1. Metodología de Generación y Publicación Linked Data.

El procedimiento de generar datos enlazados tomando en cuenta los principios de Linked Data, toma varias recomendaciones donde la mayoría de autores coinciden en el desarrollo de los siguientes pasos, señalados en Bustán [1], Gómez [2], Maza [3], Zarate [8], Robertson [17]; plasmados en la Figura 4 y TABLA IV.

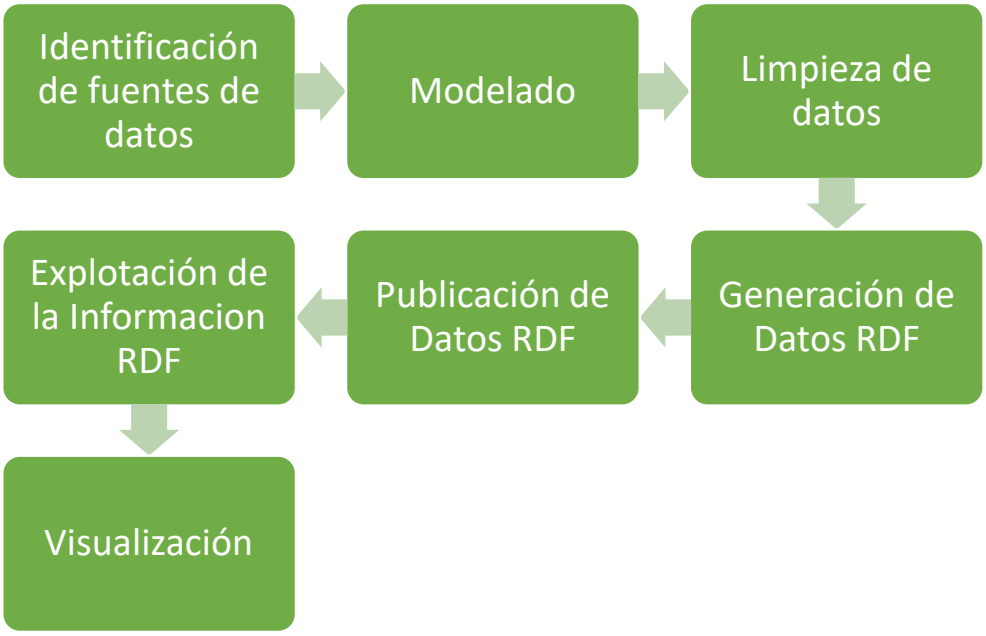


Figura 4. Metodología Generación y Publicación Linked Data
Elaborado por el Autor.

TABLA IV.

ETAPAS METODOLOGÍA GENERACION DE LINKED DATA.

Etapa	Concepto
Identificación de fuentes de datos.	Proceso en el cual se determinan cuáles serán los datos a transformar, es decir, se responden preguntas como: ¿Dónde están los datos?, ¿En qué formato?, ¿quién va a consumir los datos?; una consideración importante, también es el tiempo en el que los datos cambian o actualizan, tratado en Bustán, González [1], Gómez, Romero, Orellana [2].
Modelado.	En este paso se genera un vocabulario que describe cualquier tipo de información que englobe el dominio de nuestro conjunto de datos identificado a transformar, por tal razón, Linked Data propone la reutilización de vocabularios; por ejemplo, entre los vocabularios más utilizados tenemos: FOAF, Dublín Core, AIISO, tratado en Bustán, González [1], Gómez, Romero, Orellana [2].
Limpieza de Datos.	La desambiguación se la realiza de acuerdo al vocabulario utilizado, se evita crear confusión en sus diferentes términos, se verifica que se utiliza el título indicado entre varios homónimos o similares. Este proceso permite corregir inconsistencias, duplicación de datos, datos incompletos, datos irrelevantes, datos incorrectos, de modo que, es la desambiguación de los datos, ampliado en Bustán, González [1], Piedra [6].
Generación de Datos RDF	Ya identificados las fuentes de datos, posibles transformaciones y modificaciones necesarias para que el conjunto de datos se encuentre preparado, se debe generar datos RDF, detallado en Bustán, González [1], Gómez, Romero, Orellana [2].
Publicación de datos RDF.	La publicación de la información por medio de un triple store, es decir, una base de datos para almacenamiento de tripletas RDF. Este paso consiste en que haya un punto de consulta (SPARQL Endpoint, consultas a través del lenguaje de consulta SPARQL) visto en Quezada [5], de las tripletas RDF generadas en el contexto de este trabajo.
Explotación de Datos RDF	El consumo de los datos se realiza mediante la consulta de la información a través de un punto de búsqueda SPARQL, para poder conseguir la información dentro de lo publicado, visto en Bustán, González [1], Gómez, Romero, Orellana [2]. Los principales objetivos de las consultas SPARQL ampliado en Quezada [5] son: • Extraer información en forma de URIs y literales. • Extraer sub-estructuras RDF. • Construir nuevas estructuras de RDF partiendo de las consultas.
Visualización.	Generados los datos RDF y publicados, se requiere del mecanismo que haga posible la visualización de la información ante el usuario más fácil y legible. Es decir, mostrar una interfaz para observar los grafos RDF más amigables al usuario, estudiado en Bustán, González [1].

Elaborado por el Autor.

4.2. Motores de Representación RDF.

La metodología de generación y publicación de Linked Data determina que se requiere un repositorio para almacenar los diferentes archivos de RDF que se generen. Estos sistemas proporcionan la opción de almacenar e interactuar con tripletas, de manera que se detallan a continuación los más comunes de ellos tratados en Bustán [1], Gómez [2], Zarate [8], García [18], Semantic Web [19], cabe destacar que existen variedad de opciones (ver Anexo 3 RDF Stores).

Se destaca a continuación los principales sistemas de almacenamiento debido a su popularidad y utilización más común en la realización de proyectos Linked Data donde se evidencia su efecto positivo.

4.2.1. 4-Store.

Escrito en C, diseñado para el desarrollo de aplicaciones web semánticas, es un RDF/SPARQL Store, permite hacer gran cantidad de consultas sobre la información almacenada por los usuarios, descrito en Gómez, Romero, Orellana [2], Semantic Web [19] y caracterizado por:

- Soporte de datos RDF.
- Licencia disponible GPL
- Sistema operativo de trabajo basado en UNIX.
- Soporte para SPARQL.
- Catalogado con buen desempeño, escalabilidad y estabilidad de los datos.

4.2.2. Sesame.

Arquitectura desarrollada para al almacenamiento y consultas de grandes cantidades de metadatos en RDF y RDF Schema de manera eficiente. Entre sus principales características visto en Gómez, Romero, Orellana [2], tenemos:

- Para su implementación es necesario un espacio de trabajo para JAVA, para bibliotecas de programación se requiere JRE (Java Runtime Enviroment), SLF4J (Simple Logging Facade for Java), para Sesame como servidor un contenedor de Java Servlet (generalmente recomendado Apache Tomcat 7.0 o superiores).
- Posee un subsistema que define interfaces e implementación de todas las entidades básicas RDF denominado RDF Modelo.
- Tiene un conjunto de parsers y escritores para varios formatos RDF.
- HTTP Server es una serie de Servlets en java que implementan un protocolo para acceder a los repositorios de Sesame a través de HTTP.

4.2.3. AllegroGraph.

Es una base de datos comercial, para almacenar y consultar datos RDF (desarrollada por Franz Inc.), la versión libre está limitada al almacenamiento de menos de 50 millones de tripletas tratado en Gómez, Romero, Orellana [2] caracterizado por :

- Disponibilidad para sistemas Windows, Linux y Mac OS X en plataformas de 32 y 64 bits.

- Dispone de una arquitectura de protocolo REST, esencialmente un súper conjunto de clientes Sesame HTTP.
- En su nivel más bajo, sobre esta capa se encuentra la base de datos.

4.2.4. Open Link Virtuoso.

Almacenamiento híbrido para una amplia gama de modelos de datos entre los que destacamos RDF, XML y documentos de Texto libre. Las principales características que se menciona en Bustán [1], Gómez [2], García [18] son:

- Data Access Drivers: Punto de entrada a los servicios del motor.
- Security Manager: modulo responsable de proteger los datos y su transmisión.
- Query Manager: provee análisis léxico/sintáctico de las consultas.
- Meta Data Manager: provee al Query Manager información respecto de las entidades sobre las que trabajara el plan de ejecución.
- Concurrency Manager: permite que los clientes y servicios puedan abrir múltiples sesiones concurrentes.
- Load Data I/O Manager: maneja la lectura y escritura desde fuentes de datos externas.
- Replication Manager: controla la migración de datos, su transformación y la sincronización a través de dos o más servidores.

4.2.5. Apache JENA - Fuseki.

Framework para la construcción de aplicaciones de tipo Web Semántica, implementa APIs para trabajar con bloques de construcción como RDF y OWL. Su arquitectura proporciona un servidor SPARQL denominado Fuseki mencionado en Bustán [1], Gómez [2], Zarate [8] que en conjunto provee:

- Modelo abstracto que traslada operaciones de alto nivel a bajo nivel.
- Permite un almacenamiento persistente a través de las operaciones sobre las sentencias en grafos RDF.
- Provee un conjunto de utilidades en línea de comandos para el procesamiento de RDF, sin embargo, no posee por sí mismo características de cliente servidor.

4.3. Metodologías para construir Ontologías.

En la actualidad se encuentran diversas metodologías para generar ontologías apropiadas a los proyectos que se desarrollan en diferentes áreas del conocimiento (ver

Anexo 4), de las cuales se menciona algunas destacadas en Maza [3], Guzmán [20], Villalon [21], Arciniega [22], por su uso común con una reseña positiva.

4.3.1. Metodología Cyc.

Originaria de 1984, por parte de la Corporación de Tecnología en Computación y Microelectrónica, pieza del proyecto de Inteligencia Artificial, define a una ontología como la extracción del conocimiento común implícito en múltiples fuentes, la base de conocimiento es propietaria, aunque una pequeña versión fue liberada, esta metodología busca definir un vocabulario común para el conocimiento automatizado, consta de varios pasos, tratado en Maza [3], Guzmán [20]:

- Codificación manual del conocimiento. - Consiste en la extracción manual del conocimiento común percibido en diferentes fuentes (de manera implícita).
- Utilización de herramientas de software para la adquisición de conocimiento. - Adquirir nuevo conocimiento mediante la delegación de herramientas de procesamiento de lenguaje natural o aprendizaje computacional.

4.3.2. Metodología de Grüninger y Fox.

Dentro de esta metodología las cuestiones de competencia hacen referencias a consultas a las cuales la ontología debería responder, así como la proyección de que ámbito se propone enfrentar, visto en Maza [3], Guzmán [20] sus pasos son:

- Definición de escenarios motivadores: identificación de las posibles aplicaciones en las que la ontología será usada.
- Formulación de preguntas en lenguaje natural: conocidas como cuestiones de competencia, para determinar el ámbito de la ontología.
- Especificación de la Terminología: realizadas las preguntas se define conceptos principales, relaciones, propiedades, etc.
- Formalización de las interrogantes.
- Especificación de axiomas formales.
- Verificación de la ontología.

4.3.3. Metodología Kactus.

La ontología se construye sobre una base de conocimiento mediante un proceso de abstracción, al igual que despliega diferentes puntos de vista en cuanto a sus posibles modelos en Maza [3], Guzmán [20] mencionan sus etapas como:

- Especificación de la aplicación. – Referencia del dominio que asumirá la aplicación, los objetos de interés y tareas que realizará la ontología, define el punto de vista de modelado en diversos tipos como: dinámico-estático, funcional-causal.
- Diseño preliminar apuntalado en categorías ontológicas top-level destacadas.
- Conformación del refinamiento y estructura ontológica.

4.3.4. Metodología Methontology.

Toma la creación de ontologías como un proyecto informático, es decir, abarca actividades propias para la generación de la ontología, también incluye actividades para la planificación del proyecto, calidad del resultado, documentación, tratado en Maza [3], Guzmán [20], Villalon [21], sus pasos son:

- Especificación. – Consiste en determinar el alcance, objetivos, propósito, nivel de formalidad y usuarios finales de la ontología.
- Conceptualización. – Crea glosario de términos que pertenecen al dominio, definirlos y creación de taxonomía.
- Formalización. – Se transforma el modelo conceptual en un modelo formal o semicomputable.
- Implementación. – Se codifica la ontología utilizando un lenguaje formal (ontolingua, XOL, OIL, OWL, y otros).
- Mantenimiento. – Fase de actualización y corrección de la ontología.

4.3.5. Metodología On-to-Knowledge.

Proyecto de las Tecnologías de la Sociedad de la Información (IST) por el cual se han desarrollado herramientas y métodos que soporten la administración de conocimiento, apoyado en una ontología compartible y usable, sus fases vistas en Maza [3], Guzmán [20], Villalon [21] son:

- Estudio de factibilidad. – Determina los factores que podrían influir en éxito o fracaso, consta de actividades de identificar problemas y oportunidades, seleccionar áreas más prometedoras y soluciones potenciales, entre otras.
- Comienzo. – Se elabora el documento para la especificación de requerimientos, desarrollo de una taxonomía informal, y analizar las fuentes de conocimiento.
- Refinamiento. – Se refiere a la taxonomía informal de acuerdo a los requerimientos y objetivos.

- Evaluación. – Prueba de utilidad de la ontología desarrollada, si satisface los requerimientos, capacidad de respuesta, funcionamiento.
- Mantenimiento. – Realizar modificaciones de acuerdo a los cambios identificados y actualizar de acuerdo a reglas específicas.

4.3.6. Metodología de Uschold y King.

Permite la creación de ontologías en base a otras ya existentes, tratadas a profundidad en Maza [3], Guzmán [20], sus pasos son:

- Identificar el propósito. – Determinar el propósito por el cual se crea la ontología.
- Capturar los conceptos y relaciones. – Los conceptos y los términos.
- Codificar la ontología. – Se documenta y evalúa la ontología, así como la utilización de otras ontologías para crear una nueva.

4.3.7. Metodología NeÓn.

Permite la construcción de redes ontológicas y reutilización de recurso a través de nueve escenarios que se pueden combinar entre sí, visto en Guzmán [20], Villalon [21].

- Escenario 1. Desarrollo de redes de ontologías desde la especificación hasta la implementación.
- Escenario 2. Elaboración de redes de ontologías mediante la reutilización y reingeniería de recursos no ontológicos.
- Escenario 3. Desarrollo de redes de ontologías mediante reutilización de recursos ontológicos.
- Escenario 4. Desarrollo de redes de ontologías mediante reutilización y reingeniería de recursos ontológicos.
- Escenario 5. Elaboración de redes de ontologías mediante reutilización y mezcla de recursos ontológicos.
- Escenario 6. Reutilización, mezcla y reingeniería de recursos ontológicos para obtener redes de ontologías.
- Escenario 7. Realización de redes de ontologías a través de la reutilización de patrones de diseño ontológicos.
- Escenario 8. Desarrollo de redes de ontologías mediante reestructuración de recursos ontológicos.
- Escenario 9. Desarrollo de redes de ontologías mediante la localización de recursos ontológicos.

4.4. Herramientas.

Para realizar una experimentación con buenos resultados además de las diferentes metodologías inmersas en la generación de Linked Data, creación de ontologías que en los apartados anteriores se mencionan, es necesario describir las herramientas que complementan una buena implementación, por lo que aplicados en Bustán [1], Gómez [2], Maza [3], Zarate [8], Robertson [17], Guzmán [20], Arciniega [22] se detallan las más referenciadas.

4.4.1. Open Refine.

Herramienta que permite trabajar con grandes cantidades de datos en la extracción, transformación y procesamiento, permite trasladar datos desde múltiples fuentes, para reformatearlos, realizar la limpieza y cargarlos en otra base de datos, ampliado en Bustán [1], Zarate [8], Arciniega [22] caracterizado por:

- Exploración de datos.
- Abarca múltiples formatos como: TSV, CSV, XLS, JSON.
- Permite enlazar los recursos con otros Datasets.
- Disponible para sistemas operativos como: Windows, Linux, Mac Os.

4.4.2. Protegé.

Herramienta utilizada en el desarrollo de Ontologías y Sistemas fundamentados en el conocimiento. Origen remontado en la Universidad de Stanford, las características mencionadas en Bustán [1], Maza [3], Guzmán [20], Arciniega [22] son:

- Basado en Java y de código abierto.
- Compatible con lenguaje de ontologías OWL 2 y especificaciones RDF de W3C.
- Gran cantidad de plugins en la creación de ontologías.

4.4.3. W3C RDF VALIDATION.

Servicio online gratuito desarrollado por el consorcio W3C para el control de documentos RDF, ampliado en Bustán, González [1], Gómez, Romero, Orellana [2], sus características son:

- Permitir detección de errores y visualizar los resultados generados.
- Los documentos RDF se pueden pegar directamente en el formulario desplegado, o a través de la especificación del URI respectivo.

4.4.4. Ontology Pitfall Scanner (OOPS).

Herramienta basada en la Web para identificar peligros potenciales que desembocan en errores de modelado de ontologías, diseñada para proveer soporte a los diferentes desarrolladores durante la actividad de validación de la ontología, visto en Bustán, González [1], Arciniega, Faggioni [22].

4.4.5. Pubby.

Generador de grafos RDF sobre el protocolo HTTP para su utilización sobre un punto de consulta SPARQL, así permitir el acceso tanto a servidores SPARQL remotos como locales, detallado en Bustán, González [1], Arciniega, Faggioni [22], con características como:

- Generación de una salida en HTML.
- Implementa convenio de contenido.
- Compatible con contenedores servlets, donde se destacan Apache Tomcat y Jetty.

4.4.6. OntoEdit.

Herramienta gráfica que soporta el desarrollo y mantenimiento de ontologías. El modelo interno se basa mediante el uso de clases, relaciones, axiomas, slots y atributos. Se le pueden añadir nuevas funcionalidades también a través de plugins que le aportan interoperabilidad, permitiendo trabajar con otras herramientas como Sesame por ejemplo, ampliado en Maza [3], Guzmán [20].

e. Materiales y Métodos

De acuerdo con el Reglamento de Régimen Académico que rige a las Instituciones de Educación Superior de Ecuador, en el artículo 21, numeral 3, se estipula que un Trabajo de Titulación (TT) se basará en procesos de investigación e intervención, CES [23].Por otro lado, todo TT deberá consistir en una propuesta innovadora que contenga como mínimo una investigación exploratoria y diagnóstica, además, de acuerdo con el artículo 72 del mismo reglamento, la investigación a nivel de grado es de carácter exploratorio y descriptivo, Ludeña [24].

Manifestado lo anterior se resalta a la investigación exploratoria, que reside en la incursión de la manipulación de los datos y la falta de mecanismos que proporcionen información estructurada acerca de biodiversidad referenciada hacia el jardín botánico “Reinaldo Espinosa”, es decir; la aplicación de tecnologías de Web Semántica y Linked Data basadas en el conocimiento adquirido por la RSL conjuntamente con la experimentación con los datos del jardín. En el ámbito descriptivo se documentan los diferentes materiales y métodos que son tomados en cuenta para obtener el objetivo principal de Definir un modelo de Web Semántica y Linked Data enfocado en el campo de Biodiversidad, así como el estado actual de la incursión de las tecnologías, describiendo todos sus componentes y conceptos principales abordados en la sección de Revisión de Literatura.

En base a lo descrito anteriormente se describen los materiales y métodos que se emplearon en el presente TT.

1. Contexto

El proyecto de investigación se desarrolla en un entorno virtual enfocado hacia la Universidad Nacional de Loja, basado en el jardín botánico “Reinaldo Espinosa” donde se desenvolverá el escenario de experimentación aplicado a biodiversidad, que posteriormente servirá de guía para el desarrollo de modelos a otras universidades y grupos de investigación.

2. Proceso

Para alcanzar el objetivo general del presente proyecto de investigación se empleó el siguiente proceso para cada uno de los objetivos específicos:

2.1. Objetivo 1: Analizar el estado del arte vinculado a la Web Semántica y Linked Data en el campo de Biodiversidad.

Para el cumplimiento del primer objetivo y poder adquirir los conocimientos asociados al estado del arte vinculado a la Web Semántica y Linked Data en la biodiversidad se realizaron las siguientes tareas:

2.1.1. Revisión Sistemática de Literatura (ver Anexo 1).

2.1.2. Revisión bibliográfica (ver Revisión de literatura).

2.1.3. Análisis de Documentos (ver Resultados).

2.1.4. Beneficios y Limitaciones (Ver Resultados).

2.1.5. Documentar actividades (Preparativo del reporte técnico).

2.2. Objetivo 2: Especificar un modelo de Linked Data y Web Semántica para implementarla en el campo de Biodiversidad.

El desarrollo del modelo prototipo, así como la ontología elaborada, se desenvolvieron a través del cumplimiento de las siguientes tareas:

2.2.1. Identificación del ambiente de experimentación aplicado a Biodiversidad (ver Resultados - Objetivo 2).

2.2.2. Exploración del entorno de prueba para aplicación de la Web Semántica y Linked Data (ver Resultados sección 2.2 y Anexo 5 – Documentación JBRE).

2.2.3. Análisis de la estructura, desarrollo de vocabularios, ontologías y publicación de datos Linked Data (ver Resultados sección 2.3.).

2.2.4. Verificar el esquema estructural y la organización de los datos (véase Resultados sección 2.4 y la Figura 30).

2.2.5. Documentar actividades (Redacción del reporte técnico).

2.3. Objetivo 3: Evaluar el modelo de Linked Data en la Web Semántica implementado en el campo de Biodiversidad.

Para la evaluación de la ontología, así como el modelo prototipo generado, y cumplir con el tercer objetivo planteado se atravesó por las siguientes tareas:

2.3.1. Determinar los criterios orientados a la selección de las herramientas para la evaluación (ver TABLA XX y sección de Resultados 3.1.).

2.3.2. Realizar la evaluación de la ontología (ver Resultados en la sección 3.2.).

2.3.2. Realizar la evaluación del modelo (ver Resultados sección 3.3., Figura 33 y TABLA XXVIII).

2.3.3. Documentar actividades (Completar el reporte técnico).

3. Recursos.

Para dar respuesta a la pregunta de investigación y cumplir los objetivos planteados se usaron los siguientes recursos:

3.1. Científicos

3.1.1.Método Científico: Procedimiento que busca "Descubrir la existencia de procesos objetivos y sus conexiones internas y externas para generalizar y profundizar en los conocimientos así adquiridos para demostrarlos con rigor racional y comprobarlos con el experimento y técnicas de su aplicación" Castán [25], realizado en sus diferentes etapas, primeramente la fase de Observación - Definición y Planteamiento del problema donde se planteó la pregunta de investigación (¿Cuál será el impacto tecnológico de utilizar Linked Data en la conservación de la biodiversidad?) abordada en la Revisión sistemática de Literatura (ver Anexo 1), continuando con la fase de Supuestos contrastables - formulación de la hipótesis base fundamental en la obtención del objetivo principal: "Definir un modelo de Web Semántica y Linked Data enfocado en el campo de Biodiversidad", seguido de la fase de Comprobación (ver Resultados – Objetivo 1 y Discusión), para finalmente en la fase de Conclusiones: generalización de los resultados describir sus desafíos principales y problemas más significativos (ver sección Conclusiones), además de otorgar las directrices pertinentes para que sean empleados en monitoreo del medio ambiente, como análisis de especies en general y proyectos futuros.

3.1.2. Experimento: "Es la investigación empírica que manipula el factor o variable independiente del entorno o fenómeno estudiado, midiendo el efecto que tiene sobre otra variable o factor dependiente", Bocco [26]; el proceso experimental desarrollado contempló 5 diferentes actividades (ver Anexo 11) enfocadas en la evaluación del modelo prototipo.

3.1.3. Observación Activa: Permitió al investigador obtener una mayor comprensión sobre el problema de estudio, se realizó el análisis y determinó con eficacia los beneficios de aplicar Linked Data en un ámbito de biodiversidad (ver Discusión), a través de la identificación del entorno de exploración (ver Resultados sección 2.1), así como trabajos relacionados (ver Revisión de Literatura sección 3.3).

3.1.4. Metodología RSL: Realizado como parte principal del cumplimiento del objetivo número uno del presente TT, tomando como base las propuestas de Bárbara Kitchenham y una guía rápida para RSL tratada en Kitchenham [27], Velásquez [28] respectivamente, se considera las etapas de Planificación, Ejecución y Reporte (Ver Anexo 1) que detalla los resultados obtenidos de las investigaciones encontradas en respuesta a las preguntas de investigación formuladas.

3.2. Técnicos

3.2.1. Metodología de Generación y Publicación de Linked Data: comprende 7 fases iniciando con *Identificación de fuentes de datos* donde se exploró de manera inicial el entorno y se determinó donde se encontraba la información (vea sección Resultados 2.3), continuando con el *Modelado* se realizó la elaboración de mapas conceptuales que representen la ontología creada y desarrollada con la metodología Methontology (vea sección Resultados 2.3.1.2), avanzando con la *Limpieza de Datos* tras la elaboración de la ontología la herramienta LOD Refine, dispuso las alternativas de refinar los datos, así como encontrar posibles errores de escritura, espacios innecesarios, para la obtención final del archivo RDF a publicar (vea sección Resultados 2.3.1.2), continuando con la *Generación de datos RDF* donde se obtuvo el modelo prototipo con los nexos entre elementos a través de las similitudes encontradas con los recursos de la DBpedia (ver Resultados), para llegar a la *Publicación de Datos RDF* por medio de un triple store, este paso consiste en que haya un punto de consulta (SPARQL Endpoint, consultas a través del lenguaje SPARQL) ampliado en Quezada [5], las tripletas RDF generadas se puede evidenciar en la sección Resultados 2.3.1.5., casi al culminar la *Explotación de la información RDF* se la realizó en la sección Resultados 2.3.1.6, donde se ejecutaron las consultas para satisfacer las preguntas de la TABLA XX, mostrando finalmente la *Visualización* mediante un enfoque más amigable ante los usuarios, la información se expone como una aplicación web de FrontEnd (ver Figura 29) donde se visualiza en formato HTML los datos RDF que se almacenaron en el servidor Apache Jena-Fuseki.

3.2.2. Metodología Methontology: Abarca 5 etapas iniciando con la *Especificación* mostrado en la Elaborado por el Autor.
TABLA IX con mayor detalle, la *Conceptualización* fase de la creación del vocabulario ldjbre[2] modelado en la sección Resultados 2.3.1.2, la *Formalización* corresponde a la

[2] Ontología Linked Data Jardín Botánico Reinaldo Espinosa acrónimo definido como ldjbre

etapa realizada a través de un modelo conceptual en un modelo formal o semi-computable (ver Figura 7 y Figura 8) explicado dentro de la sección 2.3.1.2, continuando con la *Implementación* etapa de la codificación en un lenguaje formal OWL a través de la herramienta Protege (ver Figura 34, finalmente la *Actualización y corrección de la ontología* incluye el proceso de mantenimiento de la ontología (ver sección Resultados 2.3.1.2, corrección de la misma frente a conceptos, relaciones, y su actualización permanente producto de su validación, se muestra en la Figura 30 el resultado.

3.2.3. Método Triangular: proceso de evaluación de ontologías donde considera los parámetros de evaluación en tres aspectos fundamentales: lista de errores comunes al elaborar una ontología, conocida como lista de chequeo, el planteamiento de casos de prueba y el uso de razonadores, la verificación de los casos de prueba permite evaluar el cumplimiento de los requerimientos de las ontologías expresados en las preguntas de competencia. Los razonadores facilitan revisar las propiedades lógico-formales durante todo el ciclo de vida, López [29].

4. Participantes.

El proyecto de investigación se ejecutó por el estudiante, dirigido por la Directora del proyecto de TT, docente de la Universidad Nacional de Loja, Ing. María del Cisne Ruilova Sánchez Mg. Sc., así como la intervención del técnico encargado del jardín botánico “Reinaldo Espinosa”.

f. Resultados

En esta sección se abarca los diferentes hallazgos encontrados, relacionados a los objetivos previamente definidos en el anteproyecto de Trabajo Titulación, a continuación, se describen las fases desarrolladas y los resultados generados.

1. Objetivo 1: Analizar el estado del arte vinculado a la Web Semántica y Linked Data en el campo de Biodiversidad.

El objetivo principal es incursionar en la investigación de proyectos enfocados al campo de Biodiversidad, cuyo eje base de desarrollo sea la aplicación de la Web Semántica y Linked Data, para obtener los conocimientos respectivos enfocados a generar el modelo aplicado a la biodiversidad, tomando en cuenta los aspectos de cada contribución realizada hasta la actualidad.

1.1. Revisión Sistemática de Literatura.

En esta primera actividad de RSL se lleva a cabo el análisis de los conceptos y terminologías técnicas acordes al estado del Arte pertinentes a la Web Semántica y Linked Data en el campo de la Biodiversidad aplicando la metodología propuesta por Bárbara Kichenham [27], la cual facilitó de manera significativa la recopilación y extracción de trabajos que sistemáticamente fueron analizados para una mayor comprensión global del desarrollo del presente TT.

Producto del proceso de revisión sistemática de literatura resumido en la TABLA V, el resultado obtenido arroja un total de 35 trabajos relacionados (ver Anexo 1 – sección 2.2), que proveen un amplio contexto de literatura que esclarece los diferentes conceptos vinculados a la temática principal del presente TT, y en base a los cuales se dará respuesta a las preguntas de investigación planteadas (ver Anexo 1 – sección 1.2).

Ante lo mencionado, la aplicación de la Web Semántica y Linked Data en la biodiversidad trata de resolver el principal problema de dispersión del conocimiento, como lo menciona Amanqui [30], ya que los diferentes proyectos que generan y organizan conocimiento de biodiversidad se encuentra distribuidos en variados dominios, evidenciando el uso de diferentes metodologías no compartidas por los proyectos, aislamiento de información, falta de integración manifestado en Horn[31],

carencia de formatos estándares, son algunos de los problemas que surgen y son mencionados en Zarate [8], Fafalios [13], Zvoleft [32] .

TABLA V.

PROCESOS METODOLOGÍA DE REVISIÓN SISTEMÁTICA DE LITERATURA.

Etapas	**Actividades**
Planificación	1.Justificación de la revisión.
	2.Formulación de preguntas de Investigación
	3.Estrategias de Búsqueda.
	4.Fuentes Bibliográficas Seleccionadas.
	5.Cadenas de Búsqueda.
	6.Criterios de Inclusión.
	7.Criterios de Exclusión.
Ejecución	1.Búsqueda de Documentos.
	2.Selección de Documentos
Reporte	1.Hallazgos.
	2.Discusión.
	3.Conclusiones.

Elaborado por el Autor.

La solución posible radica en aplicar una perspectiva de Linked Data y Web Semántica, es decir, la combinación de datos bajo un formato estándar (RDF), construcción de una ontología a partir del conocimiento generado por cualquier trabajo, con la difusión de la construcción de dicha ontología expuesto en Domínguez [33], para extraer e integrar datos a partir de diferentes bases de datos, solución que se pretende exponer de manera estándar tratados con mayor amplitud en Zarate [8], Chavan [9], Amanqui [34], y disminuir la barrera de interoperabilidad que hasta la actualidad se mantiene.

En la TABLA VI se exponen un grupo de proyectos que aportaron con el desarrollo de nuevos formatos e iniciativas novedosas de gestión entre conjuntos de datos, todos los trabajos encontrados en el proceso de Revisión Sistemática de Literatura se encuentra documentados en el Anexo 1 enmarcados en la aplicación del dominio de Biodiversidad.

Por otro lado el Marco de Desarrollo de la Generación de Linked Data y su respectiva publicación relacionado hacia biodiversidad, se identifican tres puntos fundamentales, como son los procesos de Modelado, Generación y Visualización; tratadas en las respuestas de las preguntas de investigación de la Revisión Sistemática de Literatura(ver Anexo 1–sección Discusión; también se identifican puntos de conclusión para la difusión de información (ver Anexo 1–sección Conclusiones).

TABLA VI.

GRUPO DE PROYECTOS DE LA REVISIÓN SISTEMÁTICA DE LITERATURA.

TITULO DE ESTUDIO	PROYECTO	DESCRIPCION
Beyond INSPIRE: An Ontology for Biodiversity Metadata Records	BIOME	Proyecto con el objetivo de mostrar la ontología BIOME, que establece los requisitos mínimos para la gestión de los conjuntos de datos de biodiversidad. Aunque todavía necesita validación, en lugar de comenzar con una ontología muy rica, se utilizan otras más simples junto con una plataforma (Dendro) para gestionar los conjuntos de datos, Silva [35].
Transformando Datos de Biodiversidad en Linked Data	Biodiversidad Global Biodiversity Information Facility (GBIF).	Portal, cuyo objetivo es estructurar nodos nacionales en una red accesible vía web de manera libre y gratuita, Zarate [8]. Aporte de métricas valiosas, y datos de espacio territorial concreto.
Hacia la Interoperabilidad Semántica para el Manejo Inteligente y Sostenible de Territorios de Alta Biodiversidad usando SmartLand-LD.	SmartLand.	Gestión sostenible e inteligente de los territorios de alta biodiversidad, el pilotaje de la iniciativa se ejecuta en tres provincias del sur de Ecuador, que abarcan los ecosistemas de la Costa, los Andes y la Amazonía montañosa, Piedra [11]. Proyecto potenciador de unión de enfoques tanto de desarrollo sustentable y sostenible como tecnológico.
Darwin core: An evolving community-developed biodiversity data standard.	Biodiversity international	Notación del nombre de los recursos Fito genéticos europeos y la comunidad de bancos de genes y, al mismo tiempo, compartir un recurso genético rico en disciplinas específicas de información, Wieczoreck [10]. Aporte idóneo de formato de compartición de datos relacionado a biodiversidad.

Elaborado por el Autor.

1.2. Beneficios y Limitaciones.

Culminado el proceso de RSL se pone en evidencia los puntos asociados a los beneficios que se identificaron en los diferentes documentos, tales como:

- Integración automatizada de datos medioambientales.
- Transformación de datos ambientales en la generación de modelos de biodiversidad.
- Generación de estudios de usabilidad para la mejora continua de la aplicación y procesos para la conservación de recursos de biodiversidad.
- Descripción de datos mediante una forma estructurada, legible por el ser humano y máquina.
- Identificación rápida de la existencia de datos y recursos, hacia la atención de la comunidad académica e investigativa.

- Adaptación de vocabularios e integración de ontologías en la descripción de recursos de biodiversidad.
- Fácil acceso a los formatos estándares y la información detallada de sus metadatos en la utilización para descripción de recursos.
- Desarrollo de herramientas que facilitan la movilización e intercambio de datos sobre biodiversidad.
- Generación de recomendaciones hacia los gestores y desarrolladores en la manipulación de datos sobre biodiversidad.
- Desarrollo de estudios evolutivos y ecológicos que potencien la utilización y generación de modelos de biodiversidad.
- Mapeo eficiente y conversión de datos hacia el formato estándar RDF en base a un modelo previamente consensuado.

A continuación, se detallan las limitaciones que se identifican alrededor del desarrollo y ejecución de proyectos que aplican tecnologías de Web Semántica y Linked Data, descubiertas en los artículos investigados.

- Propagación de errores producto de contar con metadatos sin depuración y obsoletos.
- Falta de mecanismos de sincronización en la búsqueda de servicios y recuperación de datos.
- Carencia de un modelo para la publicación y descubrimiento de datos por parte de sus autores a pequeña escala.
- Falta de accesibilidad hacia el amplio espectro de datos primarios sobre biodiversidad.
- Costo elevado en infraestructura para la creación y mantenimiento de los datos por parte de editores de mediana y pequeña escala.
- Ausencia de una definición formal de las relaciones entre los términos referentes hacia biodiversidad (muy general), ya que posee muchos subdominios y ramificaciones.
- Tiempo considerable en la ejecución y validación de recuperar datos de biodiversidad, difícil tarea por parte de expertos de biodiversidad.
- El conocimiento a nivel de expertos de biodiversidad es difícil de recuperar y conseguir en los equipos multidisciplinarios que exigen los diferentes proyectos, con su respectivo seguimiento.

- Baja eficacia en las consultas, al buscar información y utilizar el lenguaje de consulta con querys SPARQL simples.

1.3. Análisis de los contenidos.

El objetivo principal de esta actividad está dirigida a realizar un análisis de todo el contenido obtenido por la búsqueda de información generada por la revisión bibliográfica y la revisión sistemática de literatura, desarrollada en torno a las temáticas de Web Semántica, Linked data y su aplicación a biodiversidad, para poder construir un esquema que agrupe conceptos que expliquen el objeto de estudio y sus diversas ramificaciones concernientes a herramientas, tecnologías, principios y trabajos relacionados.

Con la recopilación de trabajos acordes al objeto de estudio, y un análisis de literatura adecuado para los mismos (ver Anexo 1 – sección 2.3), se evidencia los problemas de la Web derivados de los diferentes cambios en tecnologías y plataformas generados por el avance masivo que posee el internet hasta la actualidad, la estructura y futuro de la Web toman un giro distinto, producto de nuevas exigencias, se optó por un nuevo enfoque basado en la inclusión de semántica explícita, para ganar un punto en la distribución de información organizada (computabilidad), visto en Zarate [8], necesaria para poder llegar al objetivo de la Web Semántica y alcanzar un nivel de "comprensión" hacia la máquina, dando soporte a la humanidad en tareas más complejas.

Producto del análisis realizado se define la aplicación del conjunto de varias tecnologías, componentes del marco de Web Semántica, donde se identifican 7 capas trascendentales descritas en 3 niveles importante de detallado. El primer nivel "Descripción establecida en el documento" con capas de estructura equivalentes a la agrupación de Unicode – URI – XML+ NS + XMLSchema – RDF + RDFSchema, ampliadas teóricamente en el apartado 1.3 (ver Revisión de Literatura).

El segundo nivel "Datos" donde el conjunto de capas relacionado a vocabularios de Ontologías y la Lógica define el comportamiento de los datos y las relaciones inherentes de la expresividad de sus conceptos (véase Ontologías). El tercer nivel "Reglas" encontradas las capas de Pruebas y Confianza enfatizadas en los objetivos de aportar con las pruebas de inferencia, razonamiento computacional capaz de establecer si la información de la Web de un sitio tiene nivel de confianza o carece del mismo, documentado en la Figura 3.

Concerniente a Linked Data se destacan cuatro principios definidos, al igual que la estructura de niveles referente a su compartición conocido como el esquema de estrellas LOD, llevado con minuciosidad en la Revisión de Literatura sección 2.3., se termina por hacer un flujo de procesos base que los proyectos poseen mostrado en la Figura 5, se representan los datos que se manejan, los mecanismos de transformación y almacenamiento para generar el archivo RDF que posteriormente serán sometidos a consultas en la Web para observarlos por navegadores, aplicaciones.

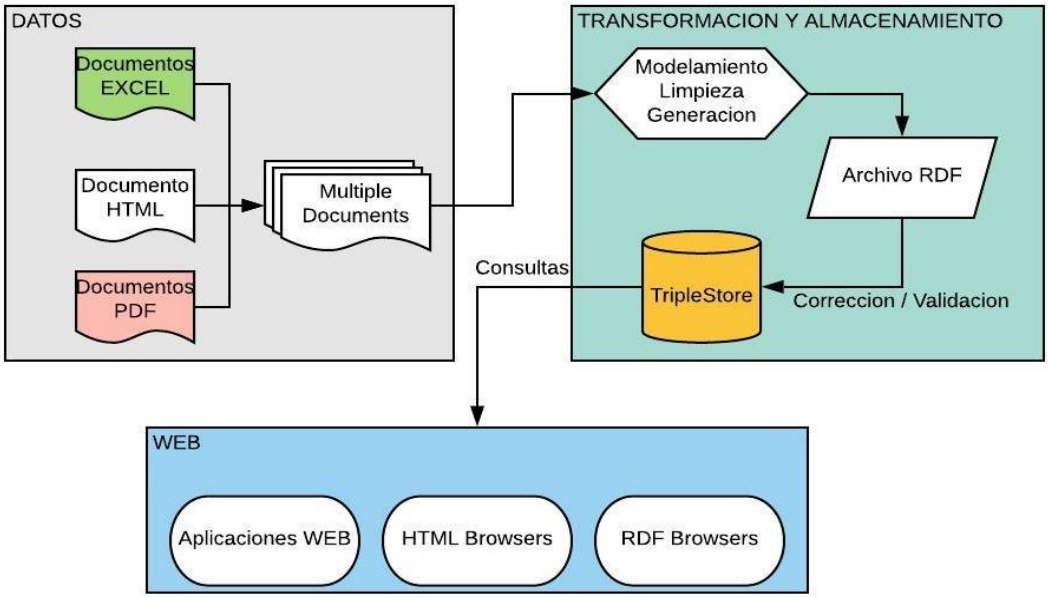


Figura 5. Flujo de Procesos hacia Linked Data.
Elaborado por el Autor.

2. Objetivo 2: Especificar un modelo de Linked Data y Web Semántica para implementarla en el campo de Biodiversidad.

En esta fase se genera el modelo que abarque los datos de biodiversidad comprendidos entre los recursos pertenecientes al jardín botánico “Reinaldo Espinosa”, para que puedan ser vistos con la propuesta de Linked Data a través de la aplicación de la metodología de Generación y Publicación de Linked Data (ver Revisión de Literatura sección 4.1) y Methontology (ver sección 4.3.4).

2.1. Identificar y señalar el ambiente de experimentación.

Para llegar a exponer mediante Linked Data los datos asociados a un área de conocimiento determinada y lograr la integración de la información, dentro del contexto

universitario se expone el escenario circundante a la Universidad Nacional de Loja, para que mediante un estándar permita su accesibilidad, minimizar la barrera de compartición, poder analizar los datos asociados al jardín botánico que posee, y en un ambiente investigativo, exploratorio se enmarque dentro del área de Biodiversidad.

2.2. Exploración del entorno local para aplicación de experimentación.

La búsqueda general en la Web no permite la ágil recopilación de información concreta que el usuario requiere en el dominio de biodiversidad, no es diferente con el jardín botánico "Reinaldo Espinosa". El portal de la Universidad, página principal de referencia del jardín botánico, aporta con descripciones breves, pero la información de mayor precisión radica en un enlace externo dirigido a su blog, donde la colección de datos que presenta se enfoca en formatos de lectura, es decir, archivos PDF, que describen las principales especies, objetivos y su historia tratado en JBRE [16], pero inoperable para un análisis inmediato (manipulación de recursos digitales) de las diferentes plantas albergadas, y sumado a esto la falta de recursos computables, puesto que la información recogida se centra en la dirección web única: https://jbreinaldoespinosa.wordpress.com.

La información obtenida a través de la búsqueda virtual y consulta de las plataformas que dispone el jardín botánico, así como el acercamiento presencial con una entrevista al técnico del jardín botánico "Reinaldo Espinosa" documentado en el Anexo 9; permitieron ratificar que los documentos publicados en su blog son manejados en PDF, archivos solo de lectura.

2.3. Analizar la estructura, desarrollo de vocabularios, ontologías de datos Linked Data.

Para la realización del vocabulario que soporte los datos de la colección Biológica del JBRE y sus diferentes especies, la metodología aplicada para la generación de Linked Data, junto con Methontology en la creación de la ontología (ver Recursos Técnicos sección 3.2) muestran los siguientes resultados.

2.3.1. Metodología de Generación y Publicación de Linked Data.

La aplicación de la metodología abarca 7 etapas, las cuales se pueden ver reflejadas en el esquema de la Figura 8, y donde se detallan a continuación las actividades que fueron realizadas para poder generar tanto la ontología como el desarrollo del modelo prototipo que represente los recursos del jardín botánico "Reinaldo Espinosa".

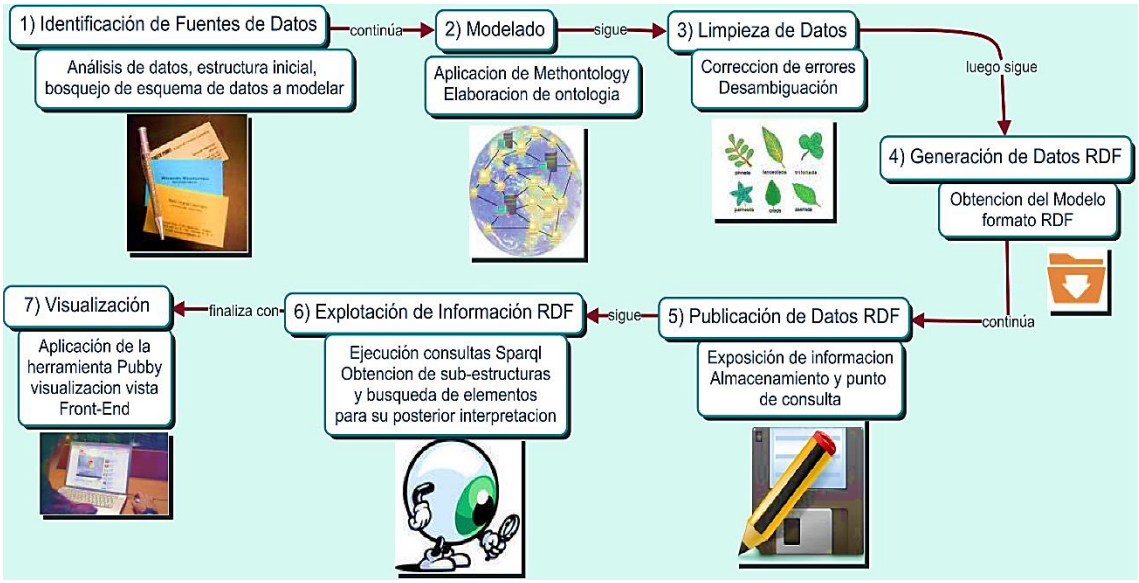


Figura 6. Representación esquema de metodología Generación Linked Data.

2.3.1.1. Identificación de fuentes de datos.

Proceso en el cual se determinan cuáles serán los datos a transformar, es decir, se responden preguntas como: ¿Dónde están los datos?, ¿En qué formato?, ¿quién va a consumir los datos?; una consideración importante, también es el tiempo en el que los datos cambian o actualizan. Primeramente, se considera la exploración de los datos concernientes a la información que se pueda encontrar del jardín botánico, de manera centrada hacia los diferentes recursos digitales que estén a disposición de los usuarios, publicados en la red.

¿Dónde están los datos?

- Alojados en el portal web, blog de Wordpress[3] con la información del jardín botánico en General.

¿En qué formato?

- Se encuentran en documentos con formato .PDF, lo que significa que su estado actual en cuanto a la información publicada es solo de lectura para los investigadores.

¿Quién va a consumir los datos?

- Los datos a modelar están proyectados para los investigadores, editores de datos que tengan como objetivo procesar, modelar o gestionar conocimiento a partir de colecciones de datos aplicados a la biodiversidad, en concreto con la información del jardín botánico.

a. Términos del dominio.

[3] Web JBRE: https://jbreinaldoespinosa.wordpress.com

- Universidad: Institución educativa de tercer nivel.
- Jardín Botánico: organización de conjuntos de elementos asociados a la flora y fauna de un determinado espacio geográfico.
- Sección: sectores que tiene distribuido el jardín botánico, puede ser: Arboretum, Plantas andinas, Medicinales y Ornamentales, Orquideario, plantas Xerofíticas, Bonsái, etc.
- Nombre Común: denominación simple que define el nombre popular de la planta.
- Nombre Científico: notación técnica de la especie, correspondiente a la planta descrita de manera científica.
- Categoría uso: Tipo de uso que cotidianamente se da a las plantas, puede ser: cereal, hortalizas, tubérculos, frutales, etc.
- Familia: Describe la familia a la que pertenece la planta.

b. Datos que se van a entregar.

Obtenidos los datos que se pretenden otorgar en la definición del dominio, la selección de los datos que se van a mostrar son:

- Datos Informativos generales del jardín botánico “Reinaldo Espinosa”.
- Nombres y datos informativos de las secciones que tiene el jardín botánico “Reinaldo Espinosa”.
- Plantas que pertenecen al JBRE.
- Nombre común, científico y familia a la que pertenece cada especie.
- Tipo de uso asociado a cada elemento del jardín botánico “Reinaldo Espinosa”.
- Relación identificada con la Universidad Nacional de Loja.
- Nexos encontrados en la DBPedia[4] con las especies del jardín botánico.

c. Formatos de entrega de datos.

Para los formatos entregados al usuario se toma en cuenta la propuesta de Tim Berners Lee [7], para lo cual la TABLA VII muestra la definición de formatos.

TABLA VII.

FORMATO DE ENTREGA DE DATOS HACIA USUARIOS.

ESTRUCTURA	TIPO DE ARCHIVO	EXTENSION
RDF/XML	rdf+xml	.rdf
Notación triples	Text/plain	.nt
CSV	Text/csv	.csv
HTML	Text/html	.html

Basado en Bustán, Gonzales [1].

[4] Proyecto para la extracción de datos de Wikipedia para proponer una versión Web semántica.

d. Posible Consumo de datos.

Los datos de la colección de plantas del JBRE de la UNL, están proyectados para el consumo posible de:

- Administrador: persona encargada del JBRE, con interacción directa hacia el almacenamiento de datos RDF, para el procesar y actualizar la información.
- Posibles Usuarios: lo integran las personas como estudiantes, docentes, investigadores y a futuro público en general que podrá acceder al dataset de la colección biológica.

e. Extracción de los datos.

La información sobre el jardín botánico "Reinaldo Espinosa" se obtuvo mediante el acceso al blog correspondiente JBRE[5] y sus diferentes especies accediendo a los archivos que expone el portal como: flora nativa, guía de plantas útiles, donde se estructuró la información perteneciente a los siguientes datos:

Información de Colección

- Información del Jardín Botánico.
- Plantas que poseen.
- Nombre Común de la especie.
- Nombre Científico de la especie.
- Familia a la que pertenece la especie.
- Categoría correspondiente al tipo según su uso cotidiano.
- Sección en la que se encuentra distribuida.
- Información del Personal: nombre, apellido, categoría laboral.

2.3.1.2. Modelado.

En este paso se realizan distintas actividades, como producto de analizar metodologías se aplica Methontology, en cuanto a los vocabularios y su reutilización, se genera uno nuevo que describe cualquier tipo de información que englobe el dominio de nuestro conjunto de datos, por tal razón Linked Data promueve la reutilización de vocabularios LOV[6], siempre que sea posible; por ejemplo, entre los vocabularios más utilizados globalmente tenemos: FOAF, Dublín Core, AIISO, entre otros.

- **Metodología Methontology.**

[5] Jardín Botánico Reinaldo Espinosa: https://jbreinaldoespinosa.wordpress.com/
[6] Linked Open Vocabularies: https://lov.linkeddata.es/dataset/lov/vocabs

Methontology - Especificación: Una vez identificados los datos a modelar, se aplica la metodología (ver Materiales y Métodos sección 3.2.2) donde se describe las características concernientes a la determinación del alcance, objetivos, propósito, nivel de formalidad y posibles usuarios (ver TABLA IX), para el nivel de formalidad se toma en cuenta lo tratado en Maza [3], Barchini [36] y se muestra en la TABLA VIII su respectivo detalle.

TABLA VIII.

NIVELES DE FORMALIDAD DE UNA ONTOLOGÍA.

Nivel de Formalidad	Descripción
Altamente Informal	Expresada en lenguaje natural.
Semi-informal	Expresada en lenguaje natural restringido y estructurado
Semi-formal	Expresado en lenguaje artificial definido formalmente
Rigurosamente formal	Términos meticulosamente determinados con semántica formal, teoremas y pruebas.

Elaborado por el Autor.

TABLA IX.

FASE ESPECIFICACIÓN DE METHONTOLOGY.

Dominio	Biodiversidad e Información asociada a las diferentes especies que posee el Jardín Botánico "Reinaldo Espinosa"
Alcance	• Representación de las especies que conforman el jardín botánico. • Detallar la información referente a el jardín botánico mediante la exposición de clases, propiedades.
Objetivo	• Difundir la información de biodiversidad del jardín botánico "Reinaldo Espinosa" a través de la representación de las diferentes especies que posee, con la intención de promover la investigación.
Propósito	• Estructurar la información de una muestra de datos con las diferentes especies que se encuentran en el jardín botánico. • Determinar las relaciones entre los diferentes conceptos que describan los recursos de la ontología.
Nivel de Formalidad	• Semi-formal.
Posibles Usuarios	• Personas con interés en la información de biodiversidad del jardín botánico "Reinaldo Espinosa". • Investigadores interesados en información de biodiversidad con un enfoque de tecnologías semánticas y compartición de datos en formatos RDF,TTL.

Elaborado por el Autor.

El dominio del conjunto de datos gira entorno a la biodiversidad, corresponde a las diferentes especies que posee el jardín botánico "Reinaldo Espinosa", describe una variedad de recursos, debido a esto, los principales vocabularios de conexión acordes a la temática que se han investigado se muestran en la TABLA X.

TABLA X.

VOCABULARIOS SOBRE BIODIVERSIDAD.

Vocabulario	Descripción
Botany[7]	Vocabulario derivado de Biological Taxonomy Vocabulary, utilizado para llenar la necesidad de describir las plantas con un mayor nivel de detalle del que permite el vocabulario básico.
Audubon Core[8]	Conjunto de vocabularios diseñados para representar metadatos para colecciones y recursos multimedia de biodiversidad. La unión de los vocabularios abordan cuestiones como la gestión de los medios y las colecciones, las descripciones de su contenido, su cobertura taxonómica, geográfica y temporal, y las formas adecuadas de recuperarlos, atribuirlos y reproducirlos.
Darwin Core[9]	Incluye un glosario de términos destinados a facilitar el intercambio de información sobre la diversidad biológica al proporcionar identificadores, etiquetas y definiciones. Darwin Core se basa principalmente en los taxones, su presencia en la naturaleza como se documenta mediante observaciones, muestras e información relacionada.

Elaborado por el Autor.

Methontology - Conceptualización: se organiza y construye una especificación semi-informal a través de un conjunto de tablas, diagramas (clases, propiedades, modelos conceptuales) con el objetivo de lograr mayor comprensión por parte de los expertos que deseen conocer y comprender el desarrollo de la ontología.

a. Identificación de clases y propiedades.

Al analizar los datos del J.B.R.E, tomando en cuenta la muestra del documento datos.CSV, se diseñó el modelo conceptual mostrado en la Figura 7, que describe e identifica las clases y propiedades del vocabulario LDJBRE[10] parte de la estructura del modelo prototipo, manifestado en la TABLA XI y TABLA XII.

TABLA XI.

ESTRUCTURA DE LA MUESTRA DEL ARCHIVO DATOS.CSV.

Elemento	Identificación	Descripción
Nombre Común	Instancia Literal	Nombre con el cual se identifica comúnmente a la especie
Nombre Científico	Instancia Literal	Identificador de la especie por su denominación científica.
Categoría de uso	Propiedad	Datos asociados al hábitat climatológica de la vegetación en su uso común.
Familia a la que pertenece	Clase	Nombre de representación a la clase de familia que pertenezca

[7] Vocabulario revisado en: http://ontologi.es/biol/botany#
[8] Extraído de: https://www.tdwg.org/standards/ac/
[9] Consultado de: https://www.tdwg.org/standards/dwc/
[10] LDJBRE: acrónimo de Linked Data Jardín Botánico Reinaldo Espinosa

Sección	Instancia Literal	Descripción de la zona en la que se encuentra ubicada la especie.

Elaborado por el Autor.

TABLA XII.

CLASES Y PROPIEDADES VOCABULARIO LDJBRE.

Clases	Propiedades	Clases	Propiedades
Datos Bioclimáticos	Superficie	Persona	Nombre
	Altura		Apellido
	Tipo de clima	Rol Laboral	Nombre
	Bioclima		Descripción
Especie	Categoría uso	Universidad	Dirección
	Sección		Categoría
	Nombre científico		Email
	Nombre Común		Nombre
	Identificador (Id)		Página Web
Jardín Botánico	Página web		
	Nombre		

Elaborado por el Autor.

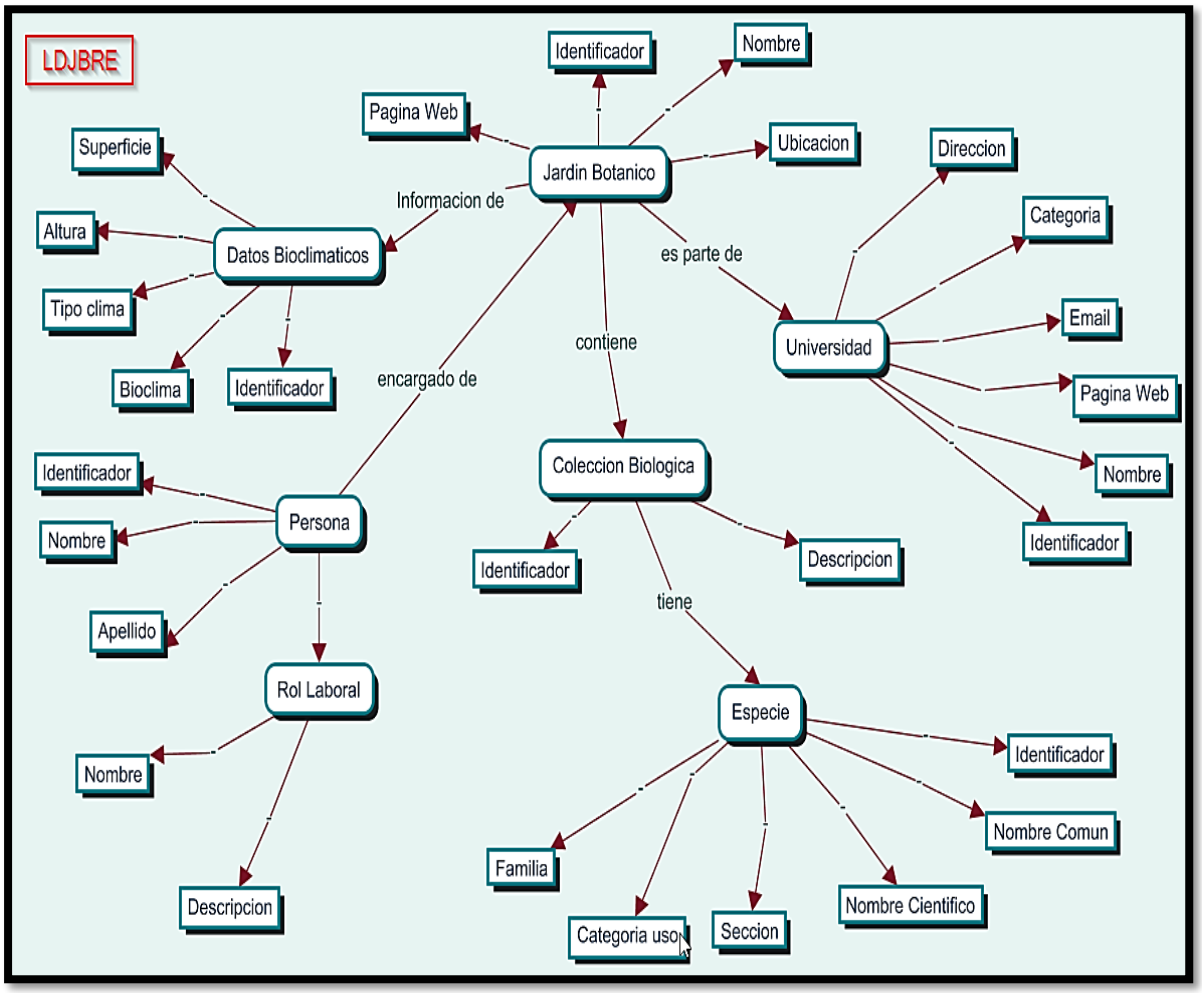


Figura 7. Representación gráfica del Modelo conceptual JBRE.

b. Vocabularios seleccionados.

La TABLA XIII muestra los vocabularios utilizados, la TABLA XIV presenta los términos a reutilizar con sus URIs y prefijos correspondientes, empleados en la generación de la nueva ontología que representa la información del jardín botánico “Reinaldo Espinosa”.

TABLA XIII.

SELECCION DE VOCABULARIOS UTILIZADOS.

Prefijo	URI	Uso
Foaf	http://xmlns.com/foaf/0.1/	Descripción de la persona encargada del JBRE, su rol laboral.
Dbo	http://dbpedia.org/ontology/	Definición de la clase universidad.
Dwc	http://rs.tdwg.org/dwc/terms/	Los términos utilizados son aquellos nombres, lugares y eventos taxonómicos que documentan los datos biológicos de la colección.
Vcard	http://www.w3.org/2006/vcard/ns#	Definición de metadatos de la ontología ldjbre y Universidad.
Dc	http://purl.org/dc/elements/1.1/	Descripción rol laboral de la persona.
Org	http://www.w3.org/ns/org	Descripción de una organización

Elaborado por el Autor.

TABLA XIV.

URIS REUTILIZADAS PARA LA ONTOLOGÍA.

Prefijo	Termino	Uri
Foaf	Homepage	http://xmlns.com/foaf/0.1/homepage
	Person	http://xmlns.com/foaf/0.1/person
	Name	http://xmlns.com/foaf/0.1/name
	familyName	http://xmlns.com/foaf/0.1/familyName
dbo	Place	http://dbpedia.org/ontology/Place
	Location	http://dbpedia.org/ontology/location
Dwc	scientificName	http://rs.tdwg.org/dwc/terms/scientificName
	vernacularName	http://rs.tdwg.org/dwc/terms/vernacularName
	Taxón	http://rs.tdwg.org/dwc/terms/Taxon
	Family	http://rs.tdwg.org/dwc/terms/Family
Vcard	Street-address	http://www.w3.org/2006/vcard/ns#street-addres
	Category	http://www.w3.org/2006/vcard/ns#category
	Mbox	http://www.w3.org/2006/vcard/ns#mbox
Dc	Description	http://purl.org/dc/elements/1.1/description
	isPartOf	http://purl.org/dc/elements/1.1/isPartOf
org	Identifier	http://www.w3.org/ns/org#identifier
	Location	http://www.w3.org/ns/org#location

Elaborado por el Autor.

En la TABLA XV muestra las clases y términos de los vocabularios reutilizados adaptados a los datos de la colección del jardín botánico junto a su descripción y tipo correspondiente.

TABLA XV.

TÉRMINOS REUTILIZADOS DE LOS VOCABULARIOS.

Prefijo vocabulario	Termino	Tipo	Descripción
Foaf	Homepage	property	Página web de Inicio
	Person	Class	Define la clase persona.
	Name	property	Nombre de alguna cosa.
	familyName	property	Describe los apellidos de la persona.
dbo	Place	property	Lugar de ubicación de una entidad
	Location	property	Identifica la ubicación de un objeto
Dwc	scientificName	property	El nombre científico completo, con autoría e información de la fecha si se conoce.
	vernacularName	property	Nombre común o vernáculo
	Taxón	Class	Un grupo de organismos considerados por los taxónomos para formar una unidad homogénea.
	Family	property	Nombre de la familia a la que pertenece el taxón.
Vcard	Street-address	property	Dirección referente de la Universidad
	Category	property	Categoría en la que se ubica a la universidad, ejemplo: A,B,C.
	Mbox	property	Correo de la Universidad
Dc	Description	property	Describe el rol laboral de la persona
	isPartOf	property	Se utiliza para indicar que el JBRE es parte de la Universidad.
org	Identifier	Property	Identificador de la organización.
	Location	property	Ubicación de la organización con respecto a las personas.

Elaborado por el Autor.

c. Modelado de la estructura.

Teniendo en cuenta los términos de los vocabularios a utilizar, en la TABLA XVI describe los nuevos elementos que se crean en el vocabulario.

TABLA XVI.

TÉRMINOS NUEVOS DE LA ONTOLOGÍA.

Termino	Descripción
Bioclimatic data	Datos bioclimáticos del jardín botánico “Reinaldo Espinosa”
Labor Role	Identifica la posición o cargo técnico de la persona encargada.
Biological collection	Permite identificar el conjunto de datos a los que pertenecen las especies
University	Institución Educativa de tercer nivel
Species	Elemento constituyente de la colección biológica.
Botanical Garden	jardín botánico

Elaborado por el Autor.

Methontology - Formalización e Implementación: Se combinan estas dos etapas de manera que se evidencia el tratamiento de la metodología como un proceso iterativo y trabajo conjunto para proyectos de software, consiste en obtener los resultados de la

formalización para transformar el modelo conceptual en un modelo semi-computable, usando representaciones lógicas, grafos conceptuales, esquemas. Seguido de la unión de los procesos para su implementación, centrada en el uso de las herramientas expuestas en la Revisión de Literatura (vea Herramientas sección 4.4.) de traducción de lenguajes ontológicos, lo que implica realizar la codificación de la ontología utilizando un lenguaje formal, que en este caso será OWL, en base al uso de la herramienta Protégé.

Definidas las clases, se describe sus propiedades, teniendo en cuenta la definición del vocabulario como prefijo: "ldjbre" con la URI creada localmente identificada: http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre# que representa "Linked Data Jardín Botánico Reinaldo Espinosa", mostrada en la TABLA XVII.

TABLA XVII.

PROPIEDADES DE LA ONTOLOGÍA LDJBRE.

Termino	Descripción	Dominio	Rango
Surface	Superficie del jardín botánico	ldjbre:Bioclimatic data	Xsd:string
Height data	Altura sobre el nivel del mar del jardín botánico	ldjbre:Bioclimatic data	xsd:string
Bioclimate	Tipo de bioclima que pertenece el jardín botánico	ldjbre:Bioclimatic data	Xsd:string
Type of climate	Categoría del clima según Koppen	ldjbre:Bioclimatic dataldjbre	Xsd:string
Description	Permite describir el rol laboral	ldjbre:Labor Role	Xsd:string
Idspecies	Identificador para cada una de las especies	ldjbre:Species	Xsd:sttring
Hascategory	Identifica la categoría de la persona en un ambiente laboral	foaf:person	ldjbre:Labor Role
Section	sección de las especies	ldjbre:Species	Xsd:string
Idbotanic	Identificador de la clase del jardín botánico	ldjbre:Botanic_Garden	Xsd:string
category use	Tipo de uso catalogado por el jardín	ldjbre:Species	Xsd:string
Id_collection	Identificador de la clase colección biológica	ldjbre:Biological_collection	Xsd:string
Id_person	Identificador de la clase persona	ldjbre:Person	Xsd:string
Is charge of	Relaciona una persona con el cargo del jardín	foaf:person	ldjbre:Botanic Garden
Information of	Relaciona los datos bioclimáticos con el jardín	ldjbre:Botanic Garden	ldjbre:Bioclimatic data
contains	Relaciona al jardín con la colección biológica	ldjbre:Botanic garden	ldjbre:Biological collection

have	Relaciona la colección biológica con las especies	ldjbre:Biological collection	ldjbre:Species
Id_bioclimate	Identifica la clase de datos bioclimáticos	ldjbre:Bioclimate_data	Xsd:String

Elaborado por el Autor.

El desarrollo de las clases y propiedades acordes a nuestro dominio se reflejan en la Figura 8 donde se expone la representación conceptual con las propiedades descritas, y las relaciones modeladas.

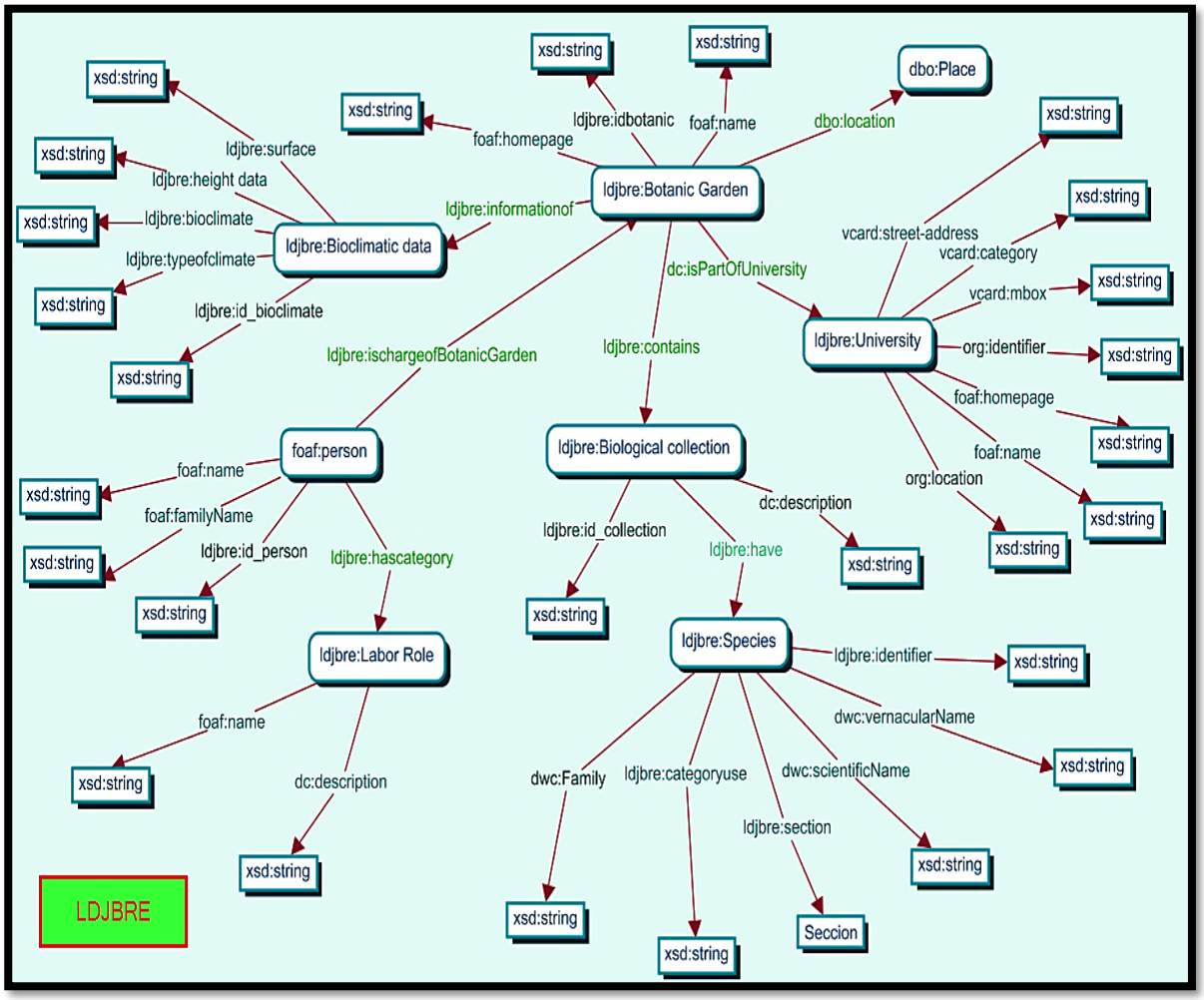


Figura 8. Representación gráfica del Vocabulario LDJBRE.

Se presenta el modelo construido mediante la herramienta PROTÉGÉ donde se muestran las clases en la Figura 9.

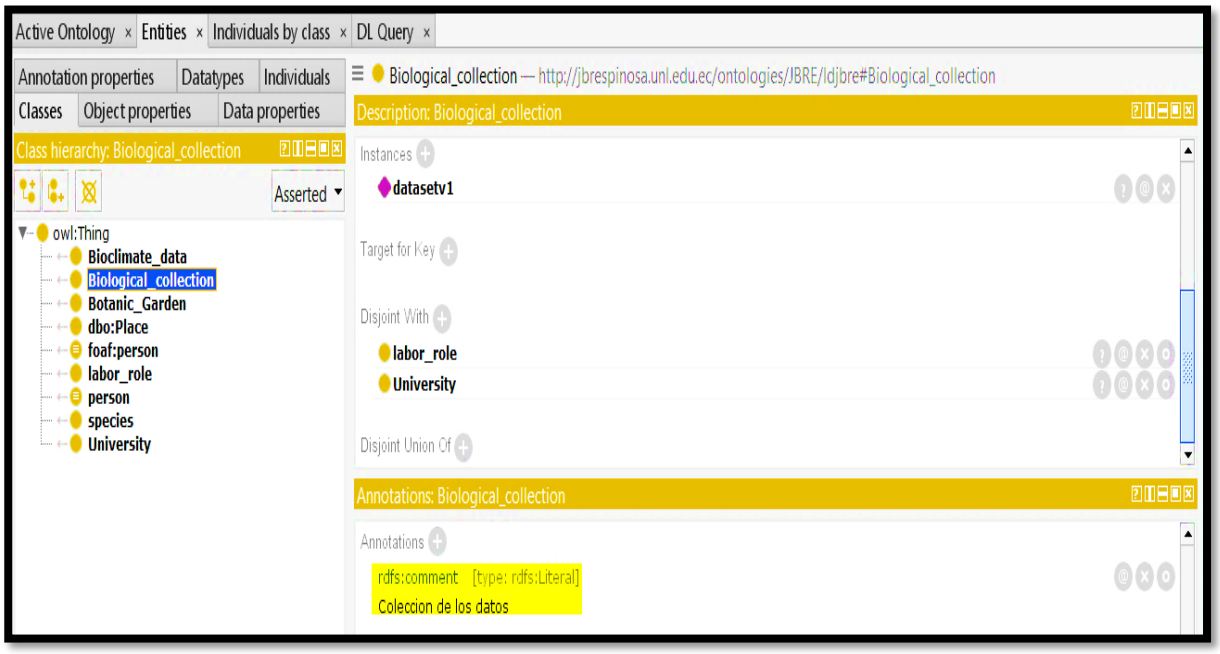
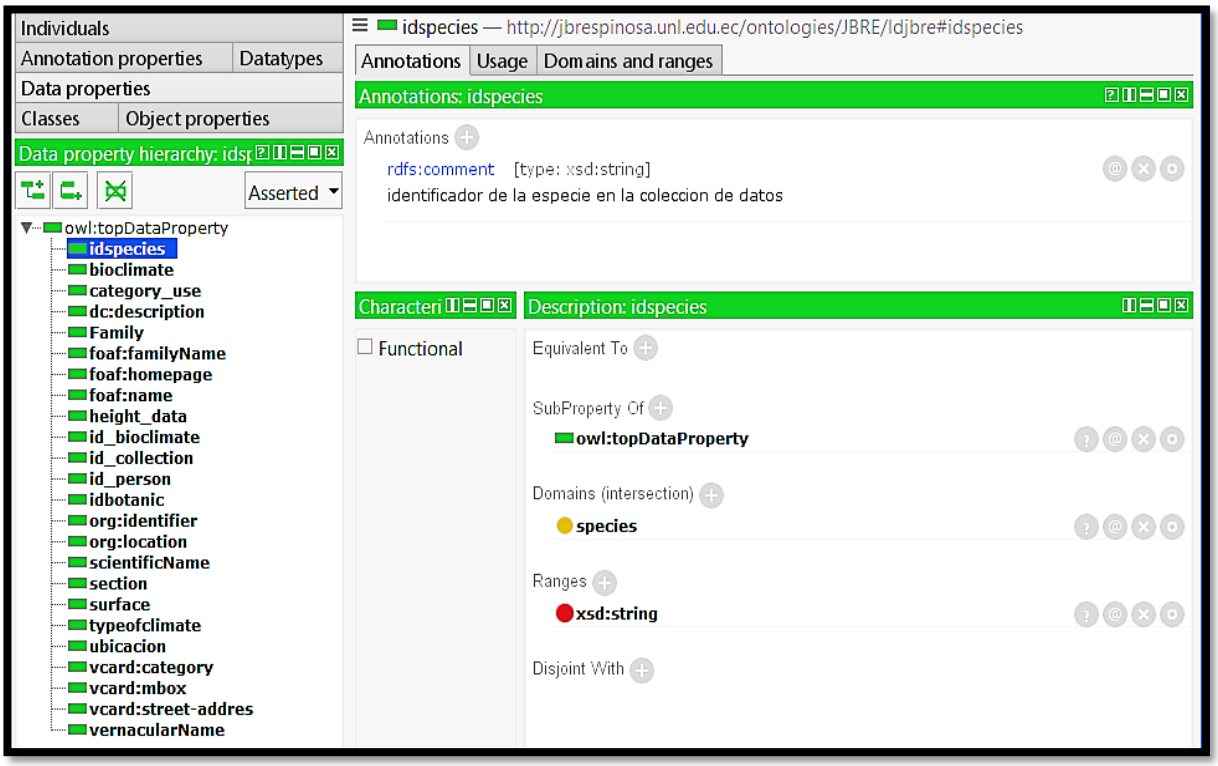


Figura 9. Representación de Clases LDJBRE – Herramienta Protegé.

Se muestra las propiedades en la

Figura 10 con su rango y dominio, al igual que las relaciones que poseen en la Figura



11.

Figura 10.Representación gráfica de propiedades LDJBRE – Herramienta Protegé.

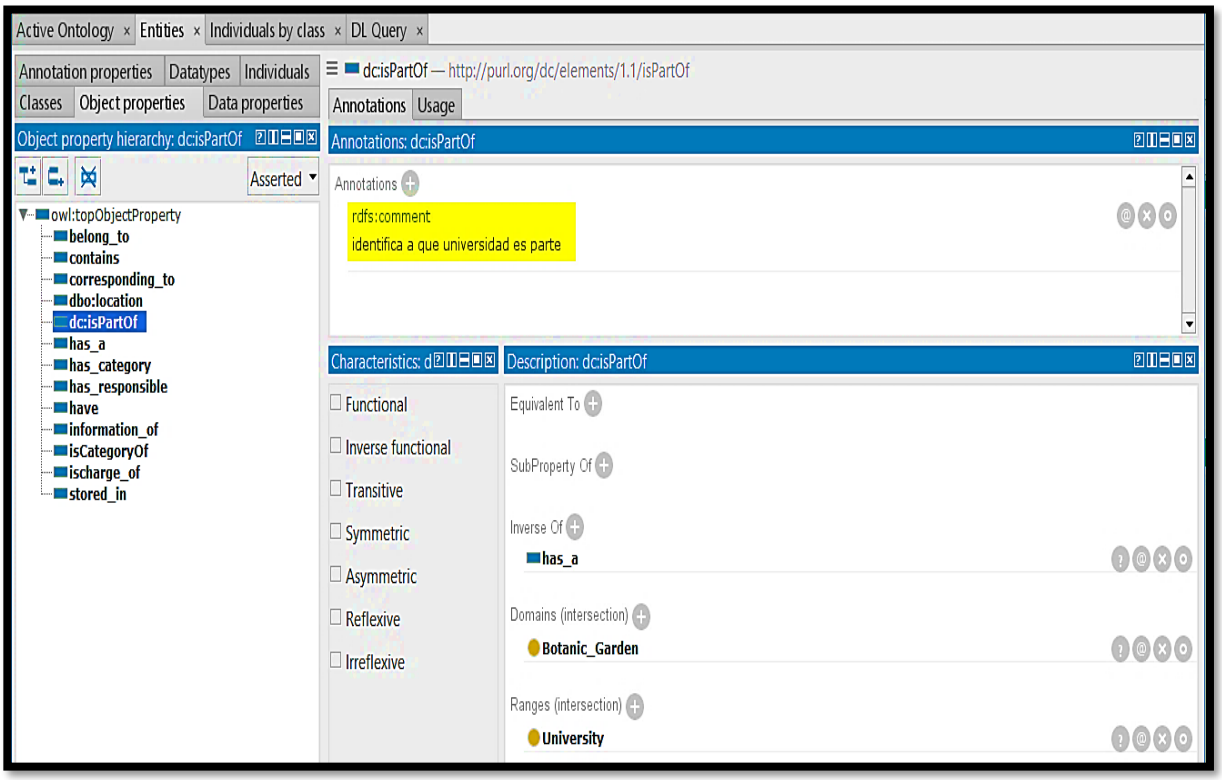


Figura 11.Representación gráfica de relaciones LDJBRE – Herramienta Protegé.

Methontology - Actualización y Corrección de la ontología: se señala que cada uno de los cambios o actualizaciones que se han desarrollado en la Ontología del jardín botánico (LDJBRE) han sido efectuados conforme se han desarrollado las actividades previas (descritas en la sección 2.3.1.2 Modelado) con la corrección, identificación y creación de clases, propiedades y relaciones, que a su vez forman parte de la metodología de generación y publicación de Linked Data, parte fundamental para poder transformar, enfocar, obtener datos RDF aplicados al entorno de biodiversidad, conocida generalmente como la etapa de mantenimiento.

2.3.1.3. Limpieza de Datos.

Realizada la elaboración de la ontología, en esta epata se aborda la depuración de los datos, se hace evidente minimizar el factor de ruido o barrera originaria del mapeo de datos que se realizó. El saneamiento de los datos comprendió el aplicar los filtros necesarios para detectar:

- Espacios en blanco innecesarios.
- Términos repetidos.
- Escritura mal empleada.

Se analizaron todos los campos mapeados como Nombre Común, Familia, Nombre Científico, entre otros revisados para refinar y corregir la información, se utilizó la herramienta LOD REFINE en el momento de ejecutar esta tarea (ver en la Figura 12).

El proceso de limpieza abarca el ámbito correspondiente a la desambiguación, evitar inconsistencias en los datos, verificando términos que se encuentre descritos correctamente, completos; contrastados con la guía de información del Jardín Botánico revisando su catálogo de inventarios de especímenes, así como el adquirir los recursos equivalentes almacenados en el repositorio de la DBpedia, comparando efectivamente la terminología empleada.

Dentro de los errores corregidos se detectó en el campo Familia y Nombre Científico se mapeaban datos con espacios en blanco innecesarios, dentro de algunas inconsistencias se percibió que algunas especies se repetían caso de "Ciprés" y "Eucalipto" con misma Familia, categoría de uso, pero diferente nombre científico, que se fue resuelto adoptando el sistema de clasificación APGIII llevado en el jardín y eliminando el registro similar no reconocido globalmente.

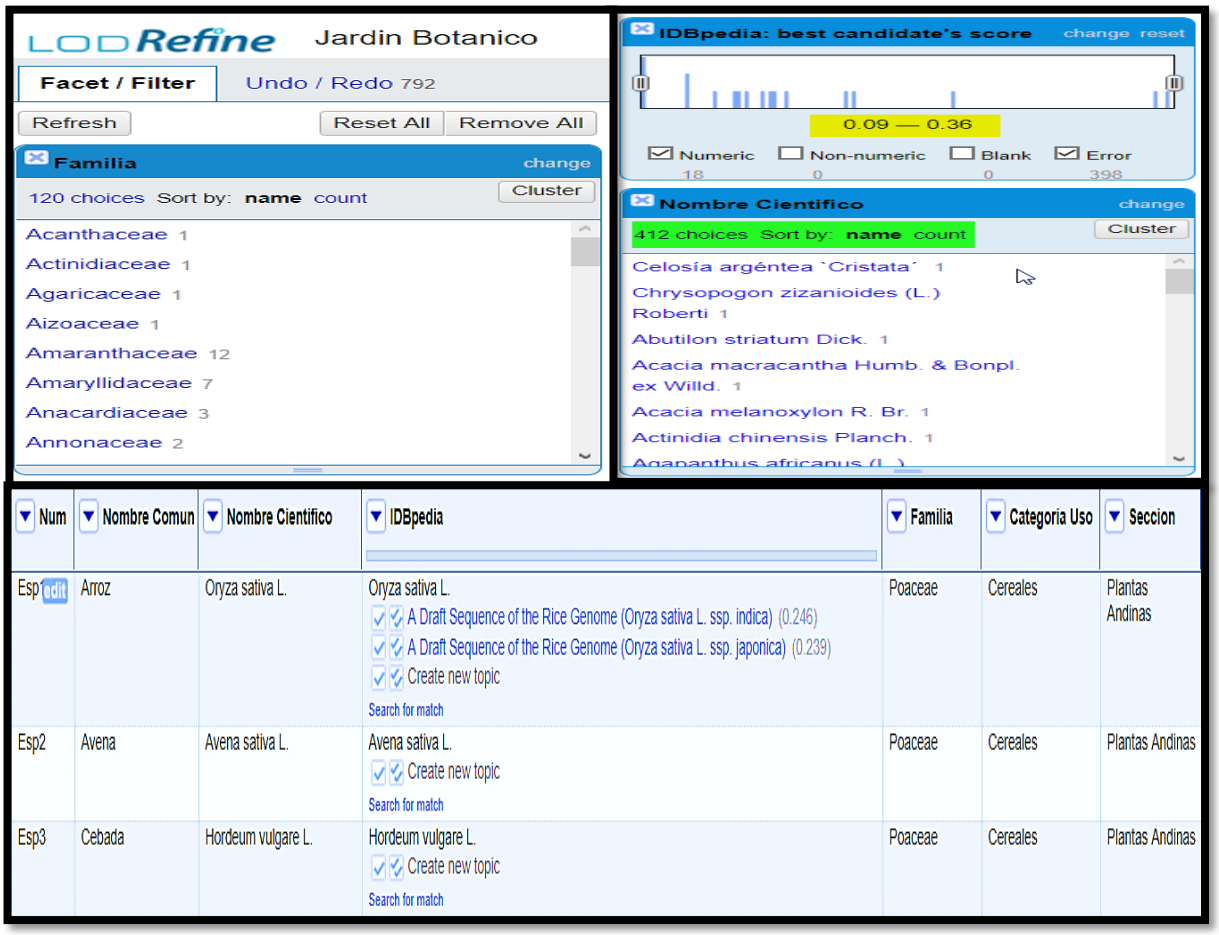


Figura 12.Representación gráfica limpieza de Datos – Herramienta LODRefine.

2.3.1.4. Generación de Datos RDF.

Una vez identificados las fuentes de datos, modificaciones necesarias para el conjunto de datos se encuentre preparado, se generan datos RDF; por lo tanto, diseñado el vocabulario LDJBRE, se procede a la exportación del mismo en formato RDF/XML; en la Figura 13 se muestra un fragmento del vocabulario elaborado y en el Anexo 6 la ontología completa.

```
<?xml version="1.0"?>
<rdf:RDF xmlns="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#"
     xml:base="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre"
     xmlns:cc="http://creativecommons.org/ns#"
     xmlns:dc="http://purl.org/dc/elements/1.1/"
     xmlns:ns="http://www.w3.org/2006/vcard/ns#"
     xmlns:dbo="http://dbpedia.org/ontology/"
     xmlns:dwc="http://rs.tdwg.org/dwc/terms/"
     xmlns:org="http://www.w3.org/ns/org#"
     xmlns:owl="http://www.w3.org/2002/07/owl#"
     xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
     xmlns:xml="http://www.w3.org/XML/1998/namespace"
     xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
     xmlns:foaf="http://xmlns.com/foaf/0.1/"
     xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
     xmlns:skos="http://www.w3.org/2004/02/skos/core#"
     xmlns:terms="http://purl.org/dc/terms/"
     xmlns:ldjbre="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#">
    <owl:Ontology rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre">
        <owl:versionIRI rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/
        ldjbre/1.0"/>
        <terms:creator rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Juan
        Maldonado</terms:creator>
        <terms:license rdf:datatype="http://www.w3.org/2001/XMLSchema#anyURI">
        https://creativecommons.org/licenses/by-nc-sa/4.0/</terms:license>
        <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-schema#Literal">
        Ontologia Jardin Botanico reinaldo Espinoza</rdfs:comment>
        <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Vocabulario
        ldjbre</rdfs:label>
    </owl:Ontology>
```

Figura 13.Representación gráfica Fragmento ontología LDJRE – Herramienta Sublime Text.

El proceso de generación de datos RDF a través de la utilización del vocabulario LDJBRE, se realizó mediante el esquema de URIS HTTP que utiliza el diseño "/" tratados en Sauermann, Cyganiak [37] para el conjunto de información del dataset, que explica al "slash namespace" como esquema recomendado por el W3C en la elaboración de la estructura de las tripletas que identifica el entorno de trabajo con el jardín botánico como temática principal. La TABLA XVIII muestra el patrón URIs para los elementos del jardín botánico mapeados por la herramienta LOD REFINE.

TABLA XVIII.

PATRON URIS MODELAMIENTO RDF.

Elemento	Patrón	Ejemplo
Vocabulario	http://(uri-base)/(type-file)/(acronym botanical garden)/(vocabulary-name#)	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University
Recursos	http://(uri-base)/(botanical-garden)/resource/	http://localhost:8080/jardinbotanico/resource/

Elaborado por el Autor.

La nomenclatura de los recursos toma en consideración la incursión del servidor local (dirección localhost:8080) de Apache Tomcat como uri-base, el cual se anticipa utilizar en la visualización e integración de la herramienta Pubby mencionada en la Revisión de Literatura (sección Herramientas 4.4.).

LOD Refine proporciona la generación de datos RDF, exportándolos en diferentes formatos como TTL, CSV, XSL, RDF/XML visto en la Figura 14; al igual que la visualización del esquema actualmente desarrollado, y la inclusión de vocabularios para la edición u corrección de las tripletas mostrado en la Figura 15.

Figura 14. Representación gráfica LOD Refine – Submenu Exportación.

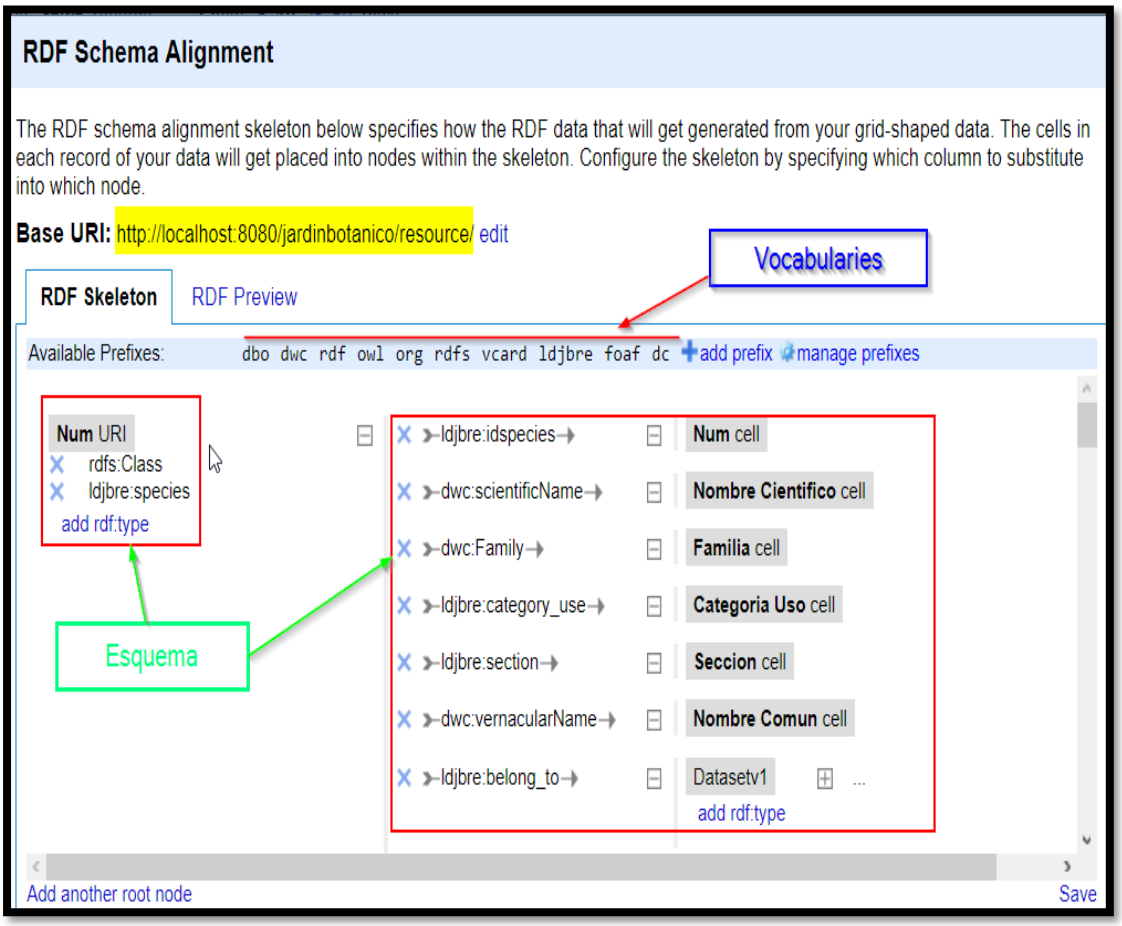


Figura 15. Representación gráfica en la Edición RDF-Schema vista LODRefine.

La búsqueda de recursos en la DBPedia sobre elementos que podamos realizar el enriquecimiento de datos fue realizada de forma semiautomática y manual (ver Figura 16 y Figura 17), producto de la ambigüedad de los nombres de especímenes que varían en la web tras la publicación de diversos autores; la TABLA XIX muestra un ejemplo de los enlaces definidos.

TABLA XIX.

ENLACES HACIA LOS RECURSOS DE DBPEDIA.

Recurso	Tipo	Enlace hacia recurso
http://localhost:8080/jardinbotanico/resource/Esp1	Owl:sameAs	http://dbpedia.org/resource/Rice
http://localhost:8080/jardinbotanico/resource/Jardin_Botanico_Reinaldo_Espinosa	Owl:sameAs	http://es-la.dbpedia.org/page/resource/Jardín_botánico_Reinaldo_Espinosa

Elaborado por el Autor.

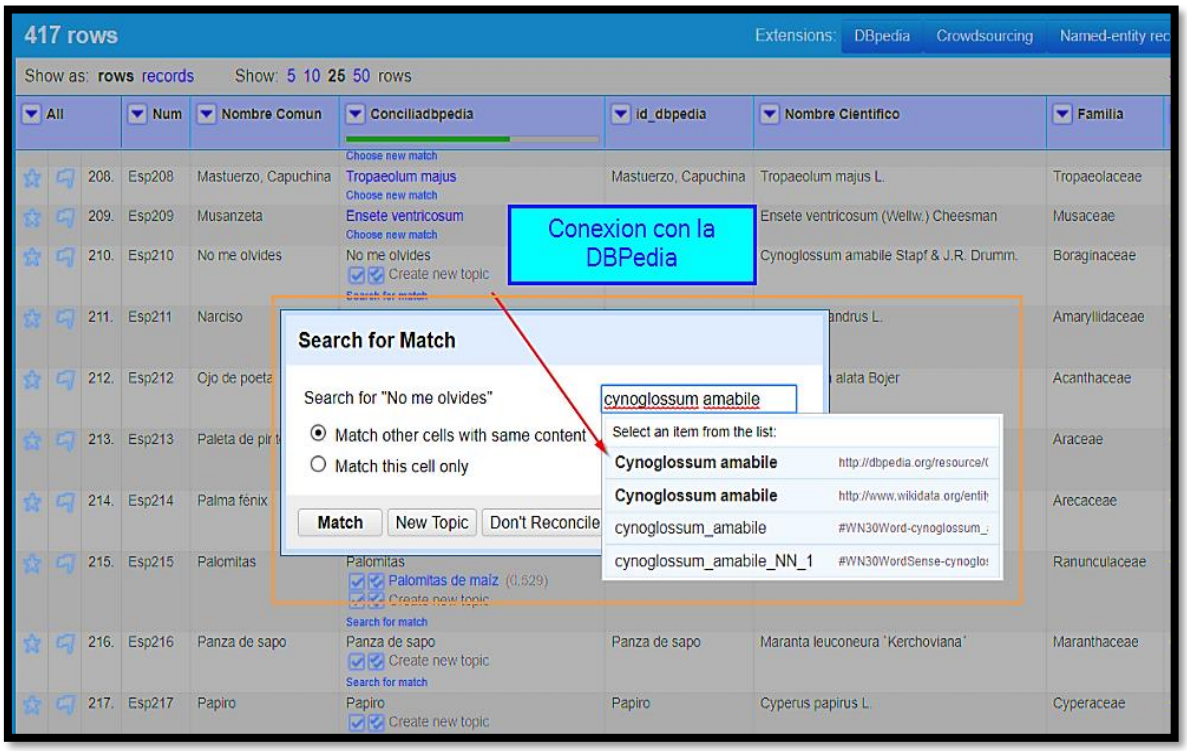


Figura 16. Representación gráfica LOD Refine con Enlaces hacia DBPedia.

Figura 17.Representación gráfica LOD Refine con la Vista de Tripletas RDF.

2.3.1.5. Publicación de Datos RDF.

Con la generación de las tripletas de los recursos del jardín botánico, se almacenan en una herramienta store, véase la Figura 18 para visualizar su interfaz, Virtuoso Open Source y el servidor de apache Jena Fuseki proporciona el Endpoint sparql, con la capacidad de realizar las consultas necesarias para la evaluación del modelo sobre las tripletas RDF almacenadas; inclusive con la posibilidad de realizar acciones CRUD en los datos.

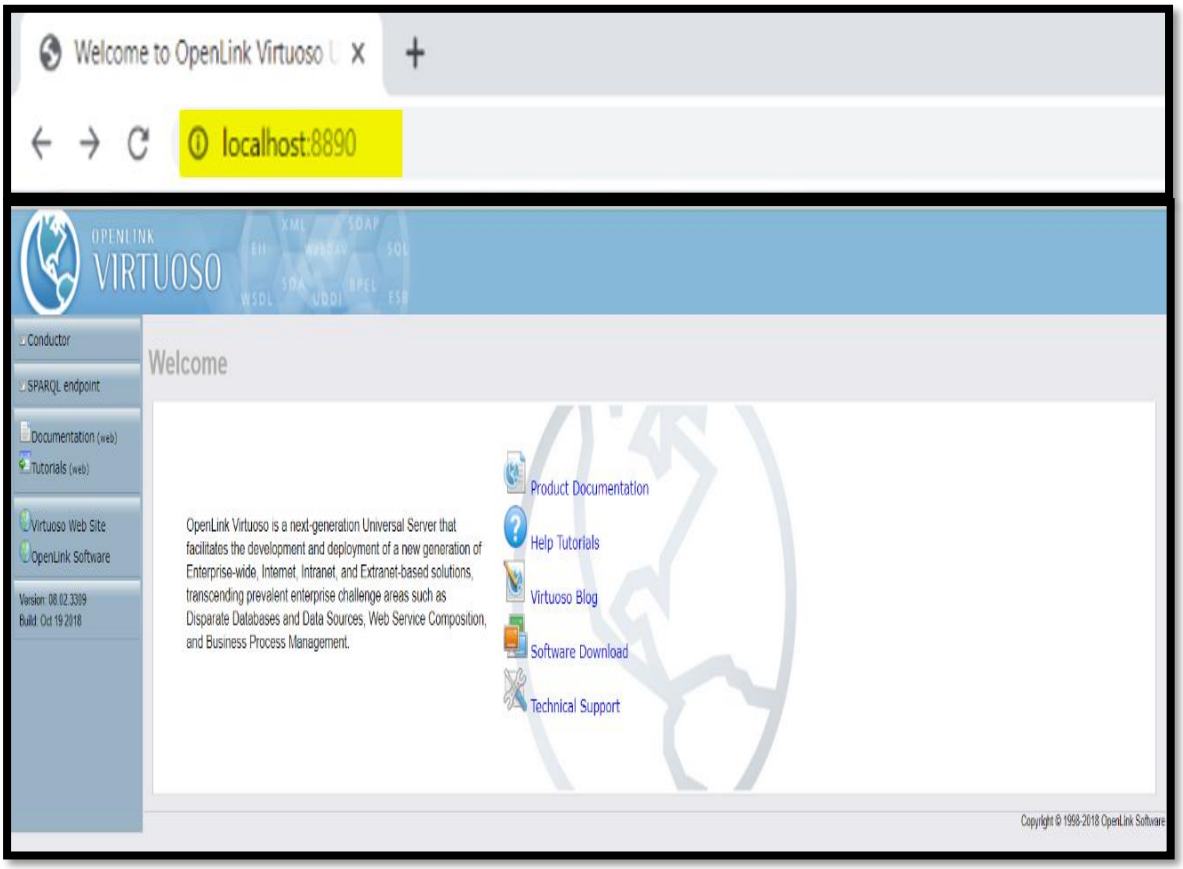


Figura 18.Representacion gráfica interfaz Virtuoso OpenLink.

Para realizar las consultas empleando el lenguaje Sparql la Figura 19 muestra un ejemplo relacionado con la información de ID, nombre científico y uso respecto a las especies almacenadas, un fragmento del resultado de la consulta se visualiza en la Figura 20.

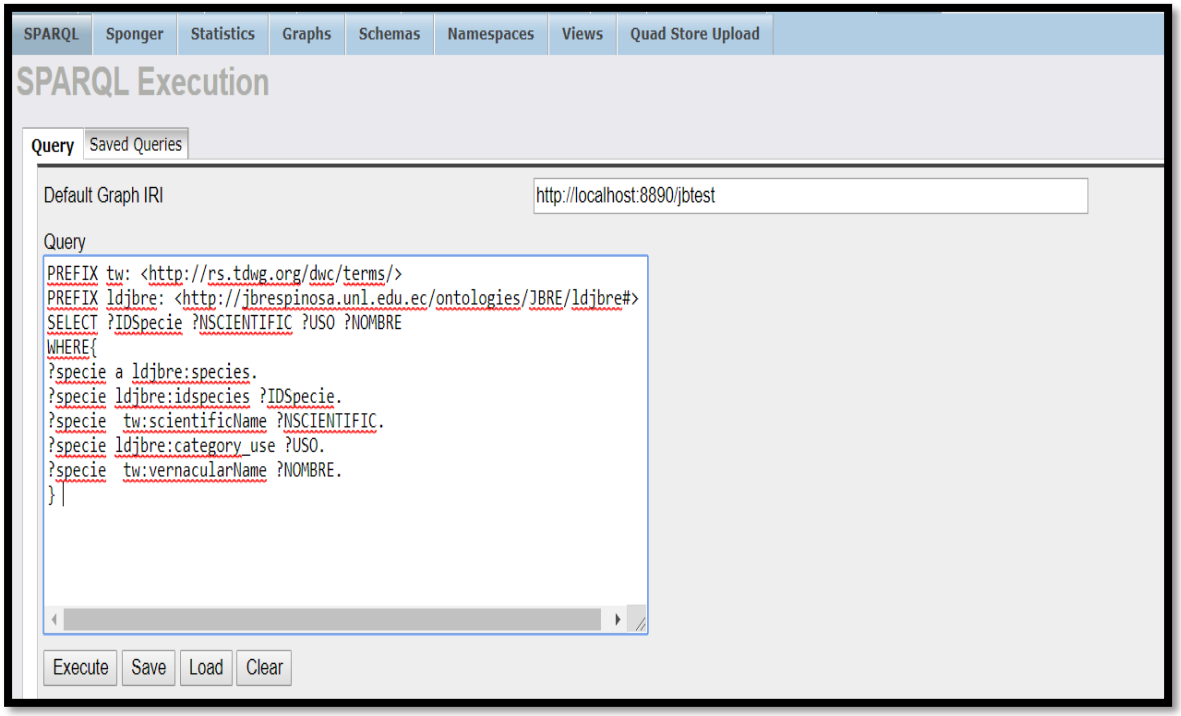


Figura 19.Representación gráfica Consulta Sparql – Herramienta Virtuoso.

IDSpecie	NSCIENTIFIC	USO	NOMBRE
"Esp150"	"Zantedeschia elliottiana (Will. Watson) Engl."	"Ornamentales"	"Cartucho amarillo"
"Esp1"	"Oryza sativa L."	"Cereales"	"Arroz"
"Esp10"	"Panicum miliaceum L."	"Cereales"	"Mijo"
"Esp100"	"Avena fatua L."	"Pastos"	"Avena forrajera"
"Esp101"	"Urochloa brizantha (Hochst. ex A.Rich) R. Webster"	"Pastos"	"Braquiaria"
"Esp102"	"Bromus catharticus M. Vahl."	"Pastos"	"Cebadilla"
"Esp103"	"Centrosema pubescens Benth."	"Pastos"	"Centrosema"
"Esp104"	"Panicun maximum Jacq."	"Pastos"	"Chilena"
"Esp105"	"Pennisetum purpureum Shumach"	"Pastos"	"Elefante"

Figura 20.Representación gráfica de resultados Sparql – Herramienta Virtuoso.

2.3.1.6. Explotación de la información RDF.

La información a explotar se basa en la ejecución de consultas en el Sparql - Endpoint proporcionado por Fuseki (ver Figura 21 y Figura 22), para obtener información mejor estructurada; por lo tanto, se han planteado preguntas que evidencien como se puede sacar beneficios en este punto, ver la TABLA XX.

TABLA XX.

FASE DE EXPLOTACIÓN DE INFORMACIÓN.

Núm.	Pregunta	Resultado
1	¿Cuantas especies se encuentran en el dataset y la información de su ID, su nombre común, científico, categoría de uso y sección?	Se muestran 417 especies en total, su información ver Figura 23
2	¿Cuál es la información de la especie 307 identificada en el jardín botánico, su nombre común, científico, categoría de uso y sección?	Científicamente denominada como "Populus nigra Duroy" en la categoría de uso de "Maderables" comúnmente conocido con el nombre de "Álamo negro" encontrado en la sección de "Arboretum" ver Figura 24
3	¿Cuál es la persona encargada del jardín botánico, rol de desempeño?	La persona con el rol de “Director” del jardín botánico es “Aguirre Mendoza Zhofre“ ver Figura 25
4	¿Qué relación posee el jardín botánico y La Universidad Nacional de Loja?	La relación que existe es “isPartOf” lo que explica que es parte de la Universidad Nacional de Loja, ver Figura 26. Así como su inversa “has_a “ ver Figura 27.
5	¿Cuáles son las URIs de los recursos de especies equivalentes?	Dada la muestra de 417 especies con la ayuda de owl:sameAs podemos observar las URIs de los recursos asociados a la DBPedia en la Figura 28

Elaborado por el Autor.

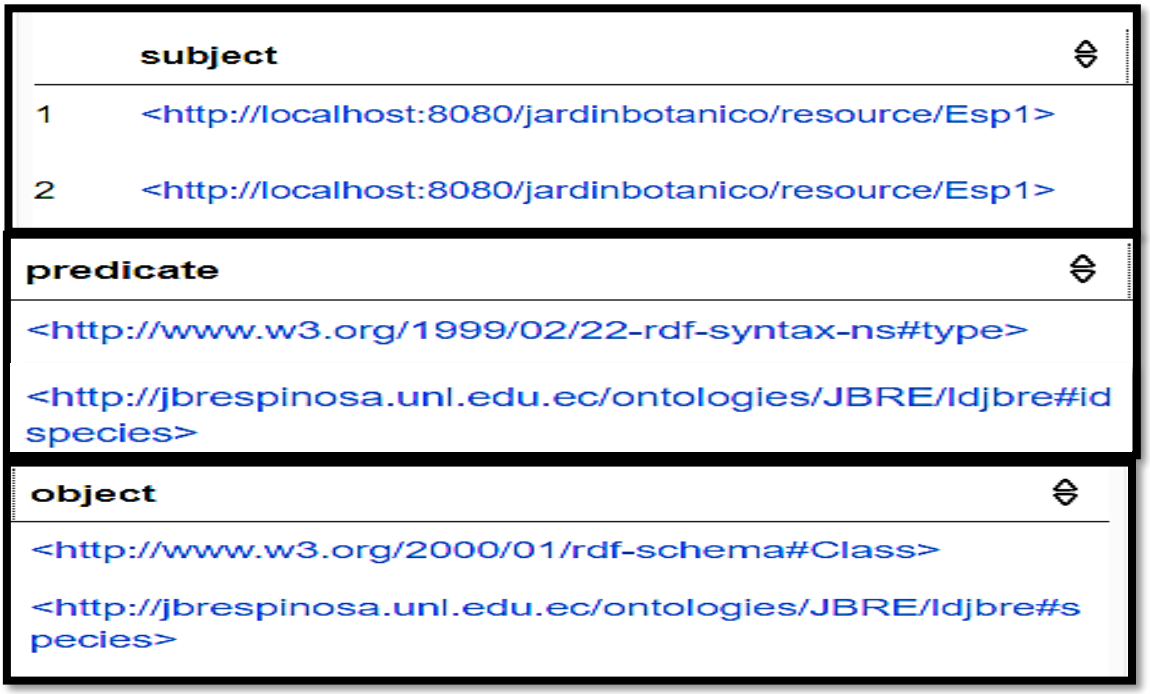


Figura 21.Representación gráfica de resultados generales interfaz Servidor Fuseki.

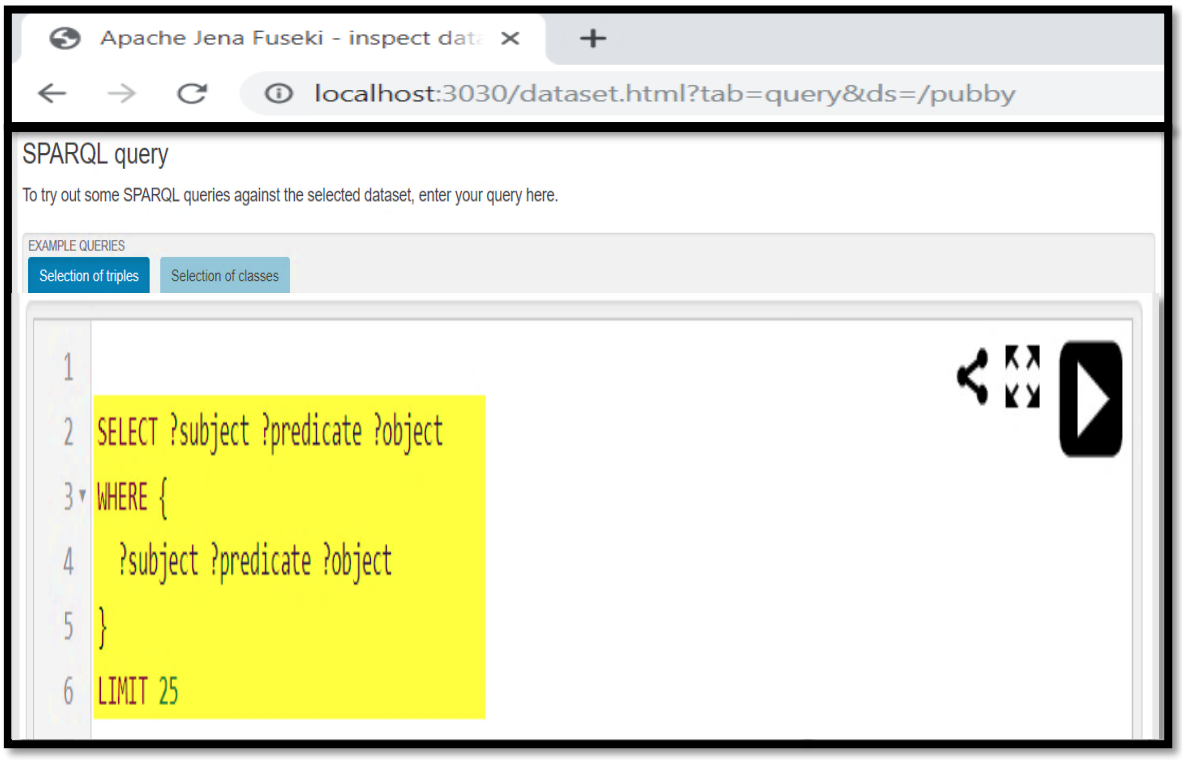


Figura 22. Representación gráfica interfaz Fuseki-Consulta por defecto.

QUERY RESULTS

Table | Raw Response

Showing 1 to 50 of 417 entries

Search:

	IDSpecie	NSCIENTIFIC	USO	NOMBRE
1	"Esp1"	"Oryza sativa L."	"Cereales"	"Arroz"
2	"Esp2"	"Avena sativa L."	"Cereales"	"Avena"
3	"Esp3"	"Hordeum vulgare L."	"Cereales"	"Cebada"
4	"Esp4"	"Secale cereale L."	"Cereales"	"Centeno"
5	"Esp5"	"Amaranthus caudatus L."	"Cereales"	"Kiwicha"
6	"Esp6"	"Chenopodium quinoa Willd."	"Cereales"	"Quinua"
7	"Esp7"	"Sorghum bicolor L."	"Cereales"	"Sorgo"
8	"Esp8"	"Triticum vulgare L."	"Cereales"	"Trigo"
9	"Esp9"	"Zea mays L."	"Cereales"	"Maíz"
10	"Esp10"	"Panicum miliaceum L."	"Cereales"	"Mijo"

Figura 23.Representación gráfica Fuseki-Resultado consulta pregunta 1.

QUERY RESULTS

Table | Raw Response

Showing 1 to 1 of 1 entries

Search:

	NSCIENTIFIC	USO	NOMBRE	SECCION	IDSpecie
1	"Populus nigra Duroy"	"Maderables"	"Alamo negro"	"Arboretum"	"Esp307"

Showing 1 to 1 of 1 entries

Figura 24. Representación gráfica Fuseki-Resultado consulta pregunta 2.

QUERY RESULTS

Table | Raw Response

Showing 1 to 1 of 1 entries

Search: Show 50

	IDbotanic	URIEncargado	Nombre	Role	Descripcion
1	"JBREv1"	<http://localhost:8080/jardinbotanico/resource/person1>	"Aguirre Mendoza Zhofre"	"Director"	"Director del JBRE"

Showing 1 to 1 of 1 entries

Figura 25. Representación gráfica Fuseki-Resultado consulta pregunta 3.

```
PREFIX ldjbre: <http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#>
PREFIX local: <http://localhost:8080/jardinbotanico/resource/>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?IDbotanic ?Recurso ?Relaciones
WHERE{
?pe a ldjbre:Botanic_Garden.
?pe ldjbre:idbotanic ?IDbotanic.
?pe foaf:name ?Recurso.
?Detalle ?Relaciones local:Universidad_Nacional_de_Loja.
  Filter regex(?IDbotanic,"JBREv1")
}
```

QUERY RESULTS

Table | Raw Response

Showing 1 to 1 of 1 entries

Search: Show 50 entries

	IDbotanic	Recurso	Relaciones
1	"JBREv1"	"Reinaldo Espinoza"	<http://purl.org/dc/elements/1.1/isPartOf>

Showing 1 to 1 of 1 entries

Figura 26. Representación gráfica Fuseki Resultado consulta pregunta 4.

```
PREFIX ldjbre: <http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#>
PREFIX local: <http://localhost:8080/jardinbotanico/resource/>
PREFIX org: <http://www.w3.org/ns/org#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT  ?IDU ?Recurso ?Relaciones
WHERE{
?pe a ldjbre:University.
?pe org:identifier ?IDU.
?pe foaf:name ?Recurso.
?pe ?Relaciones local:Jardin_Botanico_Reinaldo_Espinosa.
  Filter regex(?IDU,"UNL")
}
```

QUERY RESULTS

Table | Raw Response

Showing 1 to 1 of 1 entries

Search: Show 50 entries

	IDU	Recurso	Relaciones
1	"UNL"	"Universidad Nacional de Loja"	ldjbre:has_a

Showing 1 to 1 of 1 entries

Figura 27. Representación gráfica Resultado pregunta 4 servidor Fuseki relación inversa.

QUERY RESULTS

Table | Raw Response

Showing 1 to 50 of 417 entries

Search: Show 50 entries

	IDSpecie	NSCIENTIFIC	NOMBRE	Recurso	Equivalencias
1	"Esp1"	"Oryza sativa L."	"Arroz"	<http://localhost:8080/jardinbotanico/resource/Esp1>	<http://dbpedia.org/resource/Rice>
2	"Esp2"	"Avena sativa L."	"Avena"	<http://localhost:8080/jardinbotanico/resource/Esp2>	<http://dbpedia.org/resource/Oat>
3	"Esp3"	"Hordeum vulgare L."	"Cebada"	<http://localhost:8080/jardinbotanico/resource/Esp3>	<http://www.wikidata.org/entity/Q11577>
4	"Esp4"	"Secale cereale L."	"Centeno"	<http://localhost:8080/jardinbotanico/resource/Esp4>	<http://www.wikidata.org/entity/Q12099>
5	"Esp5"	"Amaranthus caudatus L."	"Kiwicha"	<http://localhost:8080/jardinbotanico/resource/Esp5>	<http://dbpedia.org/resource/Kiwicha>
6	"Esp6"	"Chenopodium quinoa Willd."	"Quinua"	<http://localhost:8080/jardinbotanico/resource/Esp6>	<http://dbpedia.org/resource/Quinua>
7	"Esp7"	"Sorghum bicolor L."	"Sorgo"	<http://localhost:8080/jardinbotanico/resource/Esp7>	<http://dbpedia.org/resource/Sorghum_bicolor>

Figura 28.Representación gráfica Resultado pregunta 5 servidor Fuseki.

2.3.1.7. Visualización.

Una vez realizada las consultas, la representación de los recursos en forma de tablas que muestran los resultados, aunque expuestas de manera estructurada, no pueden ser tan amigables para los usuarios, debido a esto la herramienta Pubby proporciona un aspecto Front-End de los Sparql-Endpoint tanto locales como remotos; la Figura 29 muestra como es el aspecto de interpretación al Endpoint: <http://localhost:3030/pubby/query> provisto por el servidor Fuseki.

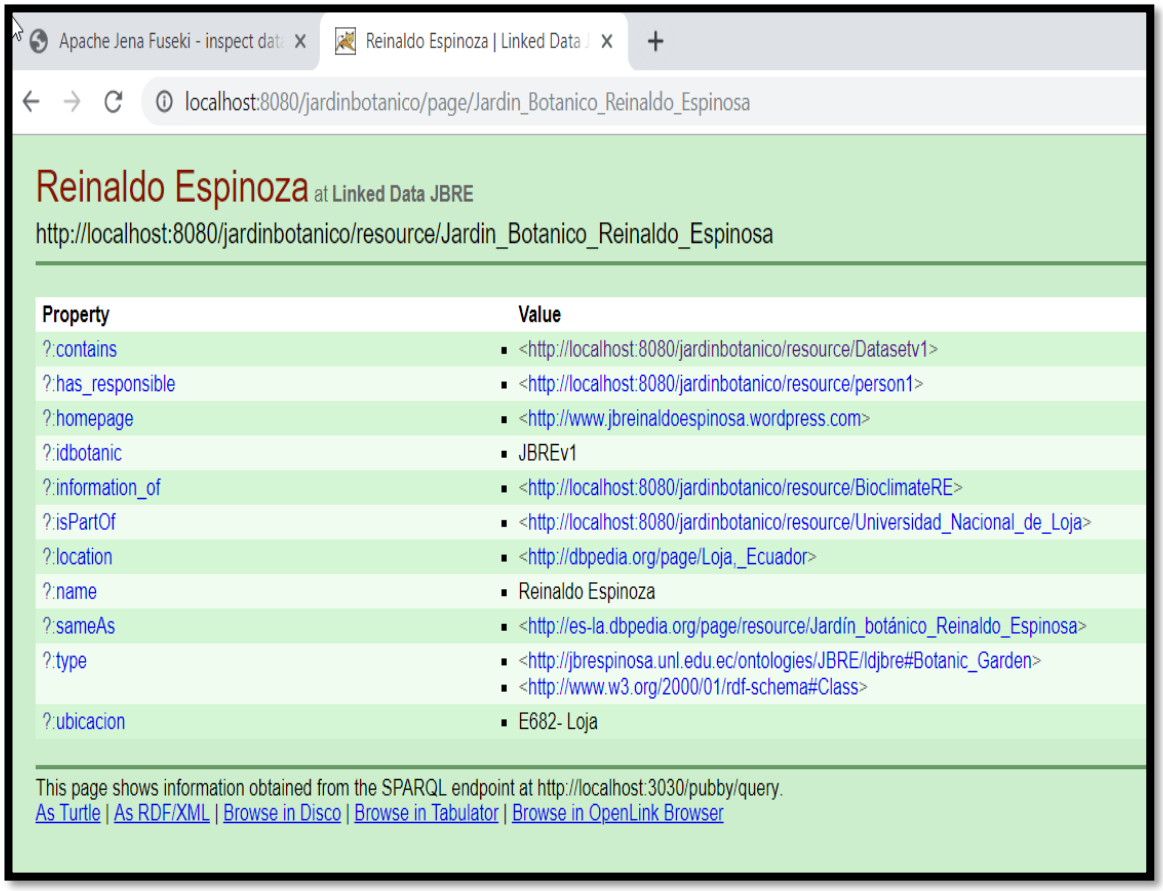


Figura 29.Representación gráfica de Visualización herramienta Pubby.

2.4. Verificar el esquema estructural y la organización de los datos.

Para poder verificar el esquema estructural, se procede a la validación de la ontología generada por la herramienta Protegé, el resultado se puede evidenciar en la Figura 30, utilizando la herramienta OPPS (Ontology Pitfall Scanner), donde se detectan varias observaciones a tomar en cuenta, el resultado sin errores críticos que haya que corregir de forma inmediata, amerita una evaluación exitosa y de manera satisfactoria.

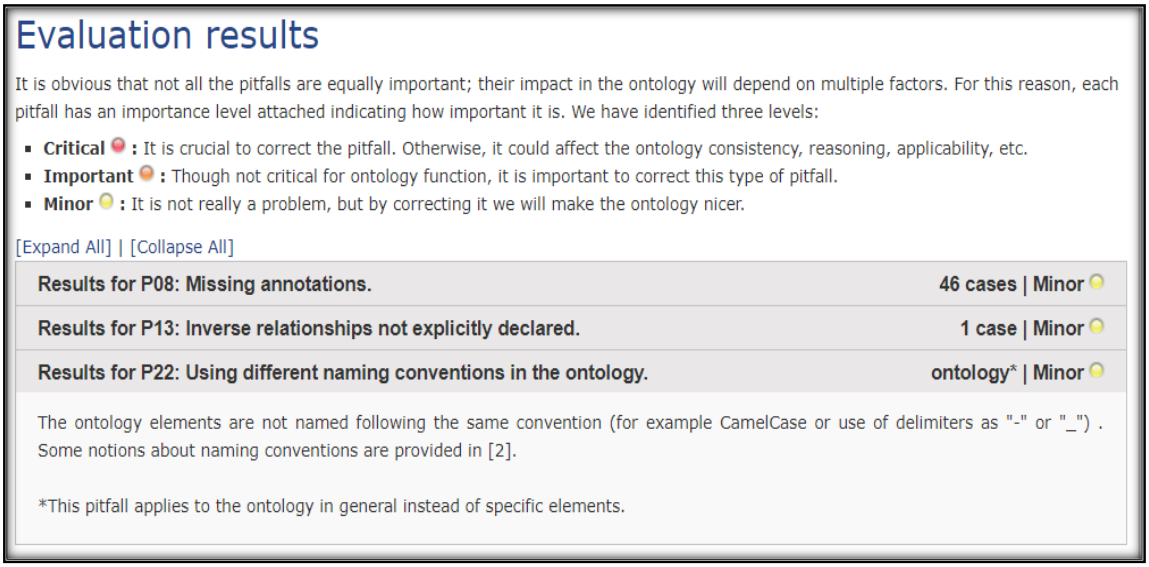


Figura 30.Representación gráfica Validación Owl – Herramienta OPPS.

3. Objetivo 3: Evaluar el modelo de Linked Data en la Web Semántica implementado en el campo de Biodiversidad.

En esta fase se evalúa el modelo generado en la etapa anterior, con el objetivo de explotar la información almacenada en la base de datos de triples, verificar el potencial de la búsqueda de información bajo el enfoque de Web Semántica y Linked Data de manera estructurada en el Sparql EndPoint; paso fundamental para poder determinar en base a información concreta, cual es el impacto de poseer datos con calidad en la conservación de biodiversidad referentes a la muestra obtenida del jardín botánico “Reinaldo Espinosa”.

3.1. Determinar los criterios orientados a la selección de las herramientas para la evaluación.

Para la evaluación de la ontología se empleó el método del triángulo (ver Materiales y métodos sección 3.2.3.), en referencia al modelo a través de las herramientas donde se almacenan los recursos modelados previamente, el método experimental (ver Materiales y métodos sección 3.1.2.) fue el proceso de evaluación en el servidor Fuseki de apache Jena y la verificación visual en la herramienta Pubby para su despliegue (perspectiva Front-End).

3.2. Evaluación de la Ontología.

La evaluación de la ontología se basó en el método triangular (ver Materiales y métodos sección 3.2.3.) donde los resultados obtenidos se muestran a continuación.

3.2.1. Lista de Chequeo de Errores.

TABLA XXI.
RESULTADOS LISTA DE CHEQUEO DE ERRORES COMUNES.

Tipo de Error	Elemento a Verificar	Evaluación
I	Una misma clase es definida como subclase y superclase al mismo tiempo en diferentes niveles de la taxonomía.	Negativo
II	Uso excesivo de la relación es-un.	Negativo
III	Existencia de más de un concepto principal.	Positivo
IV	Existencia de clases incompletas que provocan ambigüedad por no estar correctamente documentadas	Negativo
V	Falta de conocimiento disjunto. No se declara la disyunción entre conceptos provocando una incorrecta formalización del conocimiento.	Negativo
VI	Falta de exhaustividad. Se declaran subclases sin tener en cuenta la división completa de los conceptos en partes.	Negativo
VII	Existencia de términos repetidos.	Positivo
VIII	Poca especificación o delimitación de las propiedades que provoca un pobre razonamiento.	Negativo
IX	No se corresponden los elementos del dominio con los conceptos declarados o no se corresponde el conocimiento del dominio con los conceptos, relaciones y axiomas declarados.	Negativo
X	Existencia de redundancia entre las extensiones disjuntas de un concepto.	Negativo
XI	No se tiene en cuenta la traducción de los conceptos de la taxonomía a otros idiomas.	Positivo
XII	Falta de estandarización. Los nombres de los términos no siguen un estándar.	Negativo

Basado en López [29].

Se detectaron errores de tipo III, VII y XI; este último contempla el factor de internacionalización (solo tomado en cuenta el idioma inglés), para finalmente obtener como correcto la elaboración final del diseño de la ontología.

3.2.2. Verificación de los requisitos de la ontología.

Para demostrar la validez de la ontología se plantearon 5 casos de pruebas que engloban preguntas de competencia a través de la ejecución del razonador HermiT (evaluación a detalle en Anexo 10), el planteamiento de la estructura de los casos de prueba se describe en la TABLA XXII.

TABLA XXII. EVALUACIÓN ONTOLOGÍA CASO DE PRUEBA 1.

Caso de Prueba Pregunta de Competencia 1 (CPQ1)	
Pregunta de Competencia	¿Cuantas especies se encuentran representadas en la ontología LDJBRE?
Escenario de prueba	Mediante protege se encuentran almacenadas las instancias de la clase species con un total de 416.
Resultado esperado	Al aplicar el razonador deberá desplegar la lista de 416 instancias que sean de la clase species.
Resultado obtenido	Satisfactorio. El resultado generado muestra un listado de 416 Instancias de clase species.

Para utilizar el razonador HermiT se crearon las instancias de las clases con sus propiedades y relaciones de manera que los detalles de la verificación en relación al primer caso de prueba se encuentran en la Figura 31 (evaluación completa Anexo 10), generados por la herramienta Protégé.

- Verificación Pregunta de Competencia 1 (CPQ1).

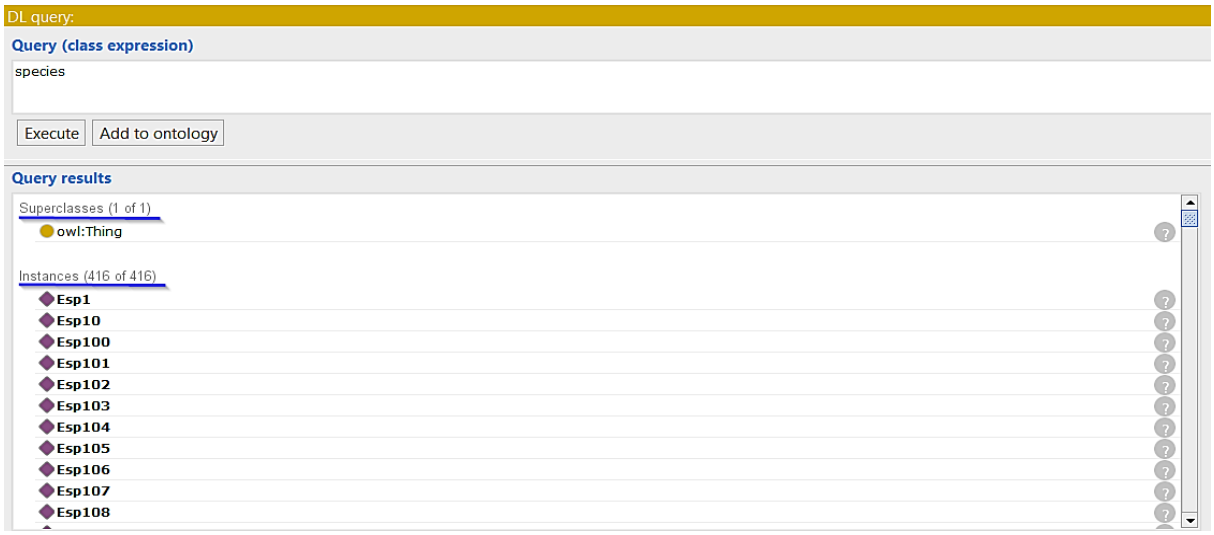


Figura 31. Representación gráfica resultados verificación CPQ1 - Protégé

La ontología apunta a la representación de los datos del jardín botánico "Reinaldo Espinosa", razón por la cual se planteó la pregunta ¿Cuantas especies se encuentran representadas en la ontología LDJBRE?, el razonador nos muestra la superclase correspondiente, seguido de 416 instancias almacenadas representativas a la clase "species", obteniendo un desenlace muy satisfactorio.

3.2.3. Conclusiones de Evaluación.

Producto de la ejecución del método del triángulo se puede concluir que:

- EL error más importante en el que se incurrió corresponde a la existencia de más de un concepto principal, superado al realizar consultas Sparql.
- Los casos de prueba planteados verifican de manera satisfactoria la validez de los requerimientos para los que fue creada la ontología, tal que, se puede representar los recursos del Jardín Botánico de manera eficiente.

3.3. Realizar la validación del modelo.

A través del validador del W3C[11] se realiza la aprobación del modelo generado con el mapeo de las diferentes especies, y en formato RDF que permita compartir, buscar información a través de URIs válidos, además se representa un fragmento del grafo respectivo de la validación mostrado en la Figura 32 y Figura 33.

[11] Validador web ejecutado en: https://www.w3.org/RDF/Validator

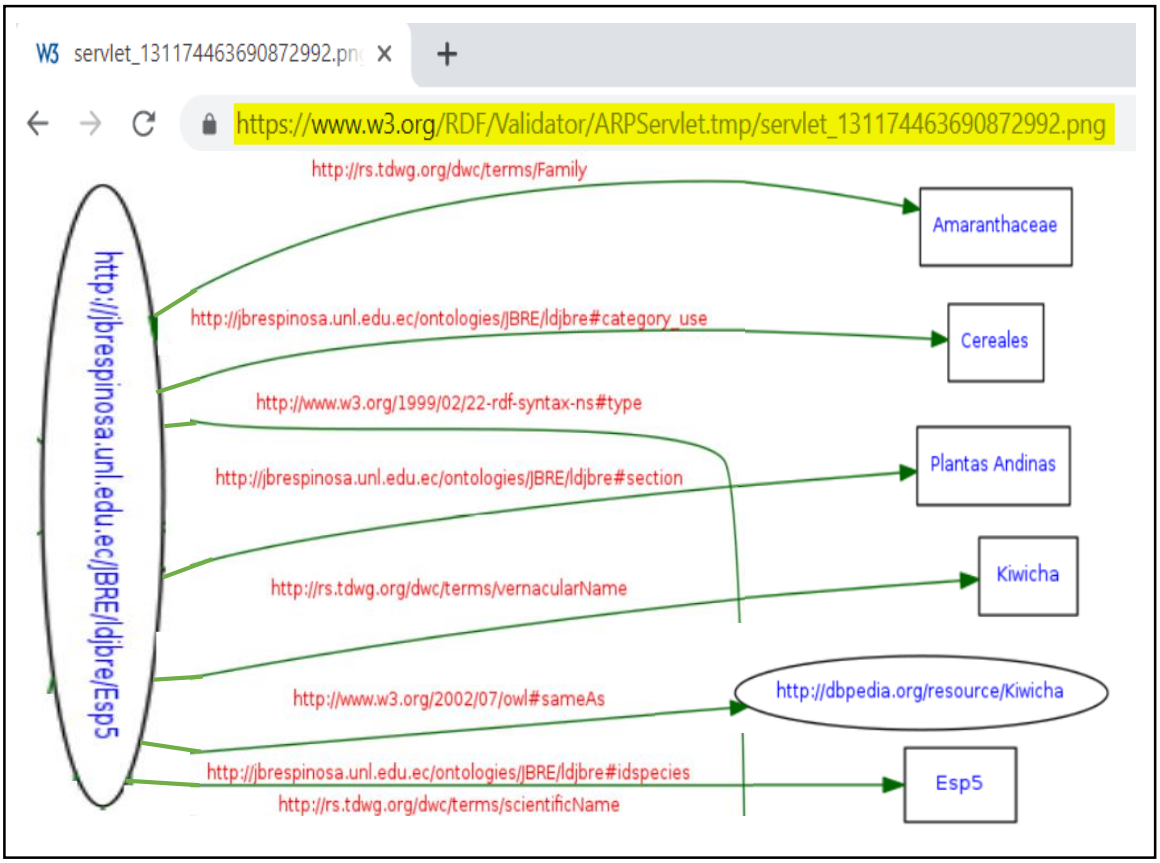


Figura 32. Representación gráfica fragmento Vista evaluación automática validación-grafo.

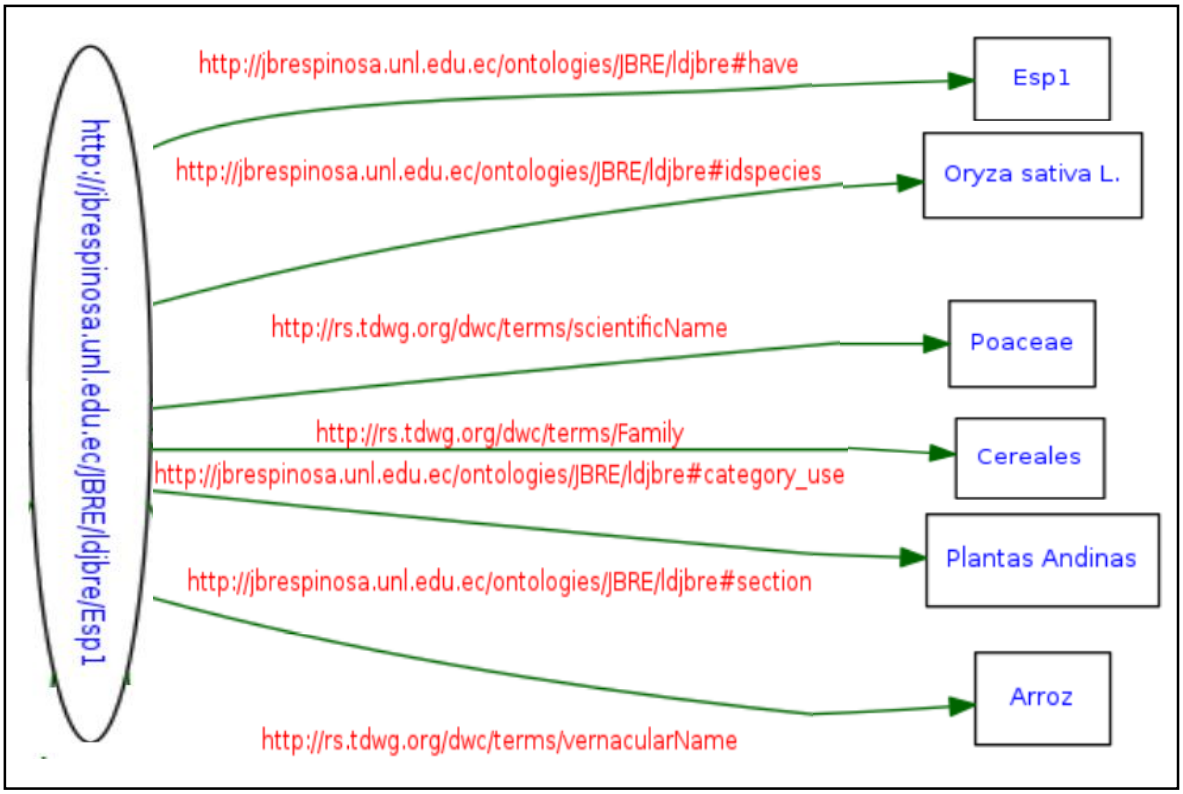


Figura 33.Representación gráfica de fragmento de grafo – vista de W3C validador.

g. Discusión

Dentro del dominio de Biodiversidad y la conservación de la vegetación ambiental, la tecnología ofrece la oportunidad de potenciar y difundir la información de esta área de conocimiento a través de herramientas como Linked Data y Web Semántica; tal que se describe el desarrollo de una alternativa solución de manera experimental que, especifique un modelo basado en el enfoque de las herramientas antes mencionadas y explique los beneficios que se generan de dicha propuesta.

1. Desarrollo de la propuesta alternativa

El presente proyecto alberga una investigación de carácter exploratorio y descriptivo, cuyo objetivo es definir un modelo Linked Data para la generación de datos de biodiversidad de las diferentes especies del jardín botánico "Reinaldo Espinosa" de la UNL, entendible tanto para las personas como computadores, cuyo objetivo es de estructurar la información en formato RDF y brindar a la sociedad una perspectiva diferente de exponer la información y los datos de la colección biológica de la UNL.

Objetivo Nro. 1: Analizar el estado del arte vinculado a la Web semántica y Linked Data en el campo de Biodiversidad.

Para cumplir este objetivo se llevó a cabo, una Revisión Sistemática de Literatura (ver Anexo 1), conjuntamente con la investigación bibliográfica, permitió responder las preguntas planteadas, al igual que formar una idea con bases sólidas del alineamiento al enfoque de aplicar el enfoque de Web Semántica y Linked Data a entornos de biodiversidad, así como identificar el aporte que llega a desencadenar en la conservación del ambiente.

A través de la recolección de información, selección de documentos, el respectivo análisis de literatura (ver anexo 1 sección 2.3), organización de los contenidos obtenidos (ver Revisión de Literatura sección d.), se identificó el amplio contexto de herramientas, tecnologías, plataformas, arquitecturas y marcos de trabajo que son utilizados hasta la actualidad con su aplicación vinculada hacia la biodiversidad, la TABLA XXIII muestra las preguntas planteadas, así como las respuestas elaboradas producto de la investigación realizada.

TABLA XXIII.

RESPUESTAS A PREGUNTAS REVISION SISTEMATICA DE LITERATURA.

Preguntas Revisión Sistemática de Literatura	Respuesta
¿QUÉ TIPO DE CONTRIBUCIONES Y BENEFICIOS SE HAN REALIZADO A TRAVÉS DE LA WEB SEMÁNTICA Y LINKED DATA APLICADAS A BIODIVERSIDAD?	Han sido varias las contribuciones aplicadas a biodiversidad, en las que se destacan proyectos que muestran plataformas y portales de biodiversidad como Bio2RDF tratado en Nolin [12], GBIF[12] mencionado en Zarate [8], Robertson [17], Bustan [1], EBI visto en Jupp [38], EOL[13], SpeciesLink[14], que aplican las tecnologías de Web Semántica y Linked Data para la generación de colecciones de datos disponibles para investigadores, estudiantes y editores de datos que buscan la operatividad de los datos para futuros análisis de manera más simple con la ayuda de la máquina. En cuanto a los beneficios de aplicar Linked Data, se señala el tratamiento de los datos de biodiversidad, su manipulación por la máquina y el refuerzo de criterio en el análisis realizado por una persona, descrito en Zarate [8], así como el desarrollo de estudios más concretos, el determinar que la situación taxonómica está en desacuerdo con un porcentaje del 13% mencionado en Robertson [17], no hubiese sido posible sin la aplicación de este enfoque. Cabe mencionar las limitaciones existentes dentro de los artículos que no hacen posibles la total integración de los datos, donde se mencionan el vacío de conocimiento, el aislamiento de grupos de investigación en diferentes silos de información, tratados en Chavan [9], Piedra [11], debido a esto no se puede determinar cuan dispersos están los conjuntos de información acerca de biodiversidad.
¿CUÁLES SON LAS TECNOLOGIAS DE WEB SEMANTICA Y LINKED DATA ACTUALMENTE APLICADAS?	Hay una amplia gama de tecnologías, herramientas y plataformas aplicadas hasta la actualidad, donde las tecnologías eje de los artículos resultado de la RSL son XML, RDF, SPARQL, OWL; apoyadas en las distintas herramientas como Protege, GraphDB, Jena, Virtuoso, Open Refine, coinciden en Peñuela [39], Hallo [40], Orciuoli [41], Boonlamp [42], García [18] . Sin embargo, cabe destacar que los vocabularios son distintos en cada investigación, por lo que destacamos entre los más expuestos a Dublín Core, XML Schema, FOAF, tratados en Yoshiura [43], Moussallem [44], Piedra [45], y Darwin Core, con más extensiones de vocabularios RDF detallados en Vandenbussche [46], Baskauf [47], Wieczorek [10].

Elaborado por el Autor.

Se identificó los diferentes beneficios, adicionalmente las limitaciones existentes de los artículos investigados, descritos en resultados Objetivo 1 (ver sección 1) donde se evidencia que existe una gran operatividad, accesibilidad y más opción de compartir datos al aplicar tecnologías de Web Semántica y Linked Data, ya que permiten ver la vegetación de manera eficaz, punto importante para la conservación de biodiversidad.

[12] Global Biodiversity Information Facility, portal: http://gbif.org
[13] Encyclopedia of life, portal: http://www.eol.org
[14] Portal de Biodiversidad: http://splink.org.br

Objetivo Nro. 2: Especificar un modelo de Linked Data y Web semántica para implementarla en el campo de Biodiversidad.

Los datos del jardín botánico “Reinaldo Espinosa” muestran la principal falencia de no tener recursos digitales manipulables, expuesta en la identificación de problemas correspondiente a Revisión de Literatura (ver sección 3.2.), también se menciona sobre la carencia de disponibilidad de recursos de biodiversidad, que tengan el enfoque idóneo para que los investigadores, editores de datos y académicos en general tengan una apreciación de los datos de las especies que este jardín posee de manera estructurada, capaz de ser computable para la máquina (que pueda realizar análisis a través de la identificación de especies y familias a las que pertenecen) para determinar un punto de impacto, o una valoración rápida generadora de indicadores ambientales que contribuyan a la conservación ambiental.

Algunas ilustraciones de la información hallada en el portal web del JBRE se encuentran en el Anexo 5, donde se resume que acorde las estrellas de Linked Data y su nivel de exposición en la Web (ver Revisión Literatura – sección 2.3) corresponde a un primer nivel, con documentos PDF de solo lectura. De acuerdo a la dispersión de los elementos relacionados con las especies, en concordancia con los informes que publica el jardín botánico, se menciona que repercute en alentar la investigación de cualquier persona en la búsqueda eficiente de la Taxonomía, estudio de la familia de cualquier espécimen, tarea que resulta difícil de realizar, al igual que identificar especies asociadas (relacionadas) entre sí.

De manera progresiva la estructura que se propone en base a una muestra elaborada de 417 especies, identificadas en los documentos antes mencionados, mediante la creación de un esquema conceptual detallado en resultados Objetivo 2 (ver sección 2.3.1.2), la Figura 34 muestra una vista de las clases con la estructura de la colección biológica, así como un fragmento de las clases y propiedades del vocabulario LDJBRE (vea Anexo 7 descripción completa) en la TABLA XXIV y TABLA XXV elaborado para la representación de los recursos del jardín botánico “Reinaldo Espinosa”, para consolidar una mayor estructura y organización de la información a través de la aplicación de tecnologías de Web Semántica y Linked Data.

TABLA XXIV.

CLASE UNIVERSITY VOCABULARIO LDJBRE.

Clase University vocabulario ldjbre	
rdf:Class	ldjbre:University
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University
rdf:subClassOf	owl:Thing
rdfs:comment	Institución a la que es parte el jardín botánico

Elaborado por el Autor.

TABLA XXV.

PROPIEDAD BIOCLIMATE VOCABULARIO LDJBRE.

Propiedad bioclimate ldjbre	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#bioclimate
rdf:subPropetyOf	owl:TopDataProperty
rdfs:comment	Bioclima
rdfs:domain	ldjbre:Bioclimate_data
rdfs:range	rdfs:Literal

Elaborado por el Autor.

En la Figura 34 y Figura 35 se observan de una manera gráfica las clases que se han creado en Protégé producto del modelado antes realizado (véase Resultados sección 2.3.1.2.).

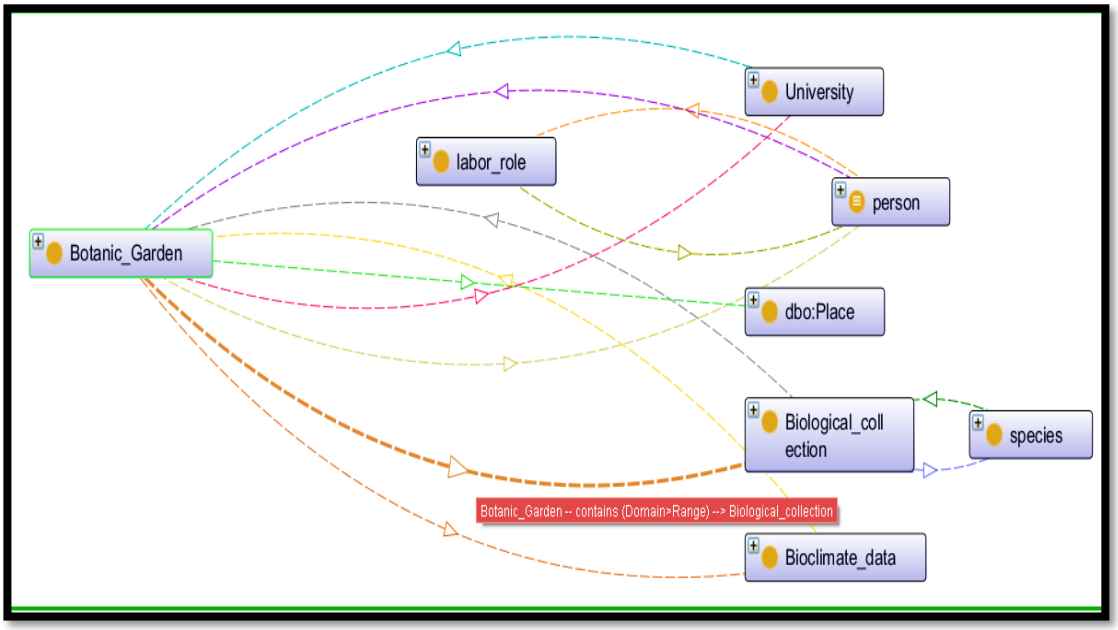


Figura 34. Representación gráfica de Clases Vista-OntoGraf – Herramienta Protege.

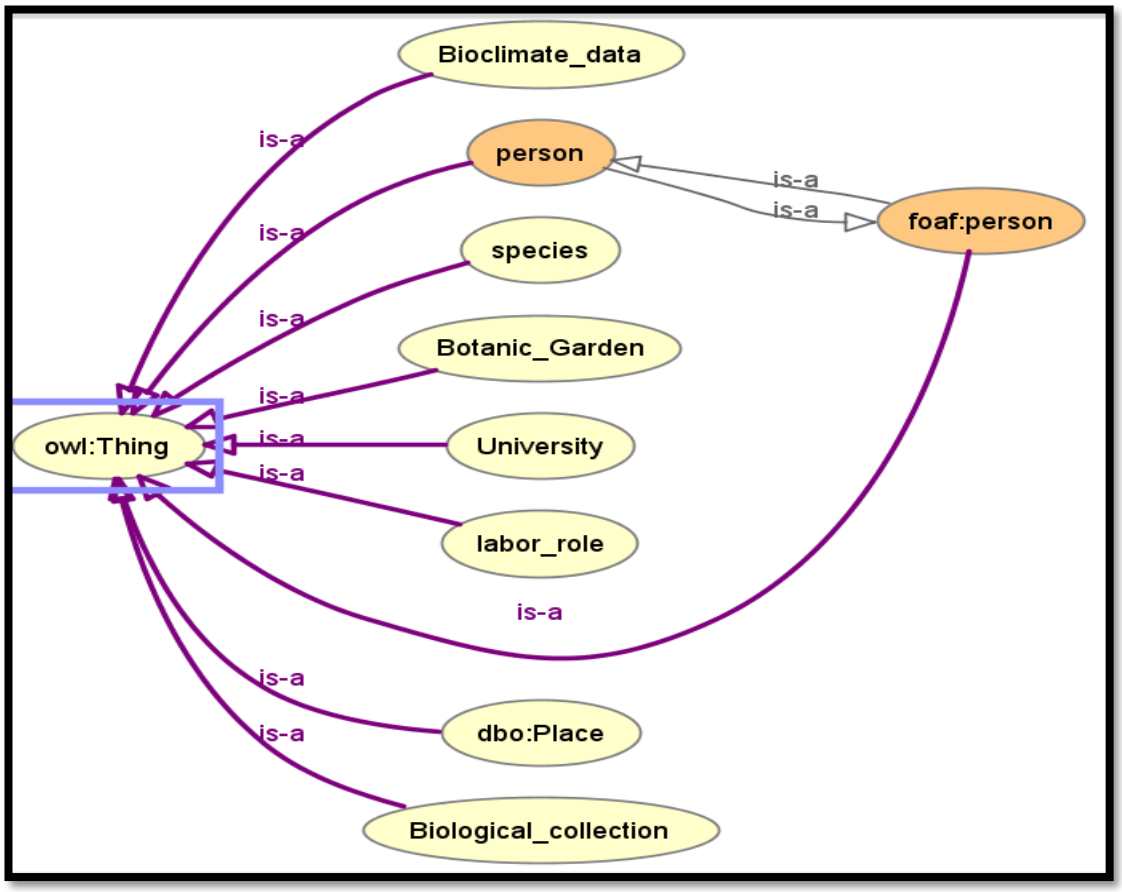


Figura 35.Representación gráfica clases vista OWLViz – Herramienta Protege.

Objetivo Nro. 3: Evaluar el modelo de Linked Data en la Web Semántica implementado en el campo de Biodiversidad

El punto principal de la investigación se fija en la base fundamental de extraer información completa, correcta y oportuna acorde al tiempo, para poder determinar el objetivo planteado, en el presente TT se evaluó el modelo prototipo generado, mediante la validación tanto de forma automática como de modo manual; también, se evaluó la ontología aplicando el método del triángulo, se logró corroborar que la información que se consulta en el Sparql-Endpoint sea correcta, completa y adecuada, para que al publicar la información tenga un impacto tecnológico alto.

Para realizar la evaluación del modelo se exponen las diferentes herramientas que se utilizaron en la Metodología de Generación y Publicación de Linked Data (ver sección 3.2.1.) en la TABLA XXVI, tomando en cuenta el trabajo Bustán [1] con la especificación de las fases acordes a la metodología en relación con su concepto de utilización.

Realizada la evaluación de la ontología se exponen los errores de tipo III, VII y XI (ver TABLA XXI), donde ejecutadas las consultas Sparql fueron satisfactoriamente solventadas, además se identificó al detectar el error de tipo XI una posible proyección a futuro de incorporar la traducción de la ontología hacia otros idiomas.

Referente a la evaluación ejecutada la TABLA XXVII muestra las fórmulas empleadas, así como la TABLA XXVIII muestra el resumen de los resultados de las métricas (ver proceso experimental sección 3.1.2), obtenidas al realizar las consultas en el Sparql Endpoint para adquirir información mayor estructurada.

TABLA XXVI.

HERRAMIENTAS GENERACIÓN Y PUBLICACIÓN DE LINKED DATA.

FASE GENERACIÓN Y PUBLICACIÓN LINKED DATA	**UTILIZACIÓN**	**HERRAMIENTA**
Identificación De Fuentes De Datos	Conceptualización	CMaps Tools
Limpieza De Datos Generación De Datos RDF	Limpieza y Generación	LOD Refine
Modelado	Busqueda Desarrollo de Vocabularios	LOV PROTÉGÉ
Limpieza De Datos	Validación	OOPS
Generación De Datos RDF	Validación	W3C RDF VALIDATION
Publicación De Datos RDF Explotación De Información RDF	Almacenamiento	Apache Jena Fuseki
Explotación De Información RDF Visualización	Visualización	Pubby

Elaborado por el Autor.

TABLA XXVII.

FÓRMULAS EMPLEADAS EN LA EVALUACIÓN DEL MODELO.

Variables	**Formulas**
Disponibilidad	$D = 1 - \frac{Solicitudes\ erroneas}{solicitudes\ totales}$
Validez Sintáctica	$VS = \frac{datos\ incorrectos}{cantidad\ de\ propiedades}$
Precisión Semántica	$PS = \frac{valores\ incorrectos}{cantidad\ de\ propiedades}$
Correctitud	$Co = \frac{enlaces\ incorrectos}{enlaces\ presentes}$
Impacto	$I = 1 - \frac{(VS + PS + Co)}{3}$

Elaborado por el Autor.

TABLA XXVIII.

RESUMEN MÉTRICAS EVALUACIÓN MODELO JARDÍN BOTÁNICO.

Criterio: Evaluación Automática			
Herramienta	Archivo	Resultado	Conclusión
Validador RDF W3C	Jardin-Botanico.RDF	Successfully	Validación correcta.
Criterio: Evaluación Manual			
Variable	Escala de Medición	Rango de Valores	Conclusión
Calidad (D: Disponibilidad)	D=1 - (2/20) D= 1 – 0.1 D= 0.9	D=90%, equivalente a una Evaluación Correcta.	El 90% de los datos son alcanzables y listos para su uso.
Impacto (Validez sintáctica: VS, Precisión Semántica: PS, Correctitud: Co).	VS = 94/423 VS = 0.2222 PS = 0/423 PS = 0. Co = 54/417 Co = 0.1295 I = 1 – (VS+PS+Co)/3 I = 1 – (0.3517/3) I = 0.8828.	El Impacto de la contribución del modelo, corresponde al promedio de las medidas de validez sintáctica, precisión semántica y correctitud, se obtuvo un resultado igual a 88.28% equivalente a "Alto".	Dado el comportamiento de las métricas VS, PS, Co y su valor promedio de error, comprueba la calidad del documento de 88.28% , que evidencia el impacto Alto que posee al exponer Información del jardín botánico.

Elaborado por el Autor.

La información del dataset almacenado en el servidor Apache Jena Fuseki se visualiza en la Figura 36, donde se destaca que 4642 triples es el número total de tuplas generadas, que contiene el modelo con la descripción de los recursos del jardín botánico "Reinaldo Espinosa".

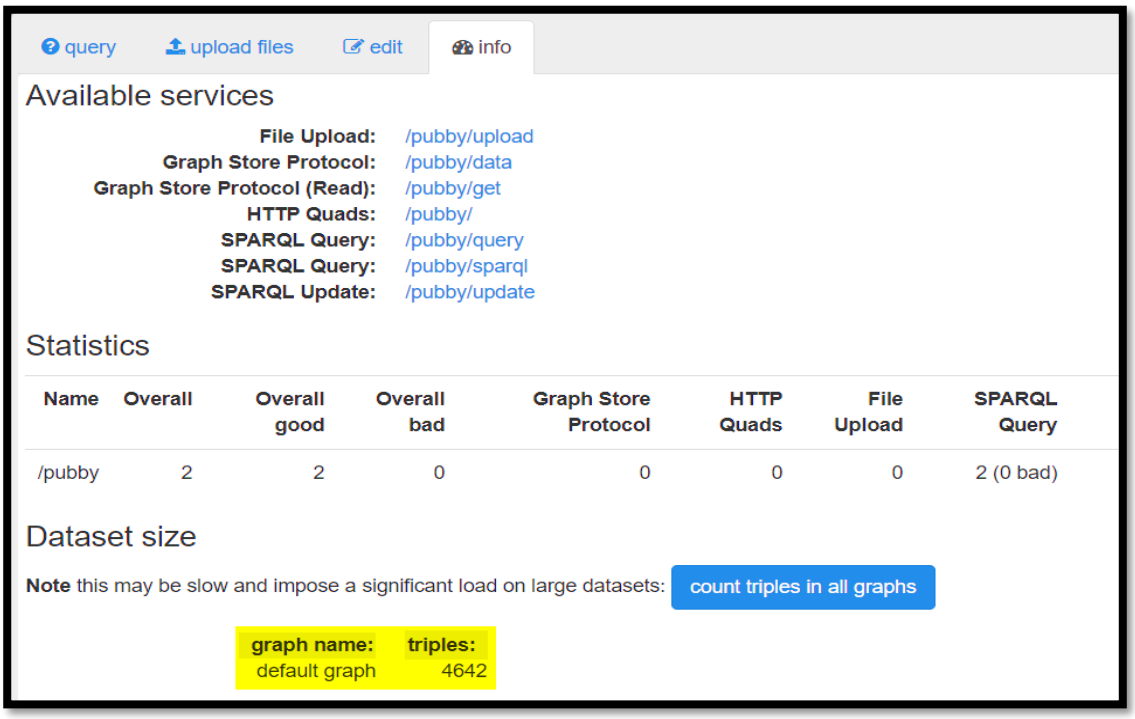


Figura 36. Representación gráfica Vista Fuseki número total de triples.

h. Conclusiones.

Cumplidos los objetivos planteados en el presente TT, bajo la premisa de aplicar tecnologías de Linked data y Web Semántica en el jardín botánico “Reinaldo Espinosa” se establecen las siguientes conclusiones:

- El problema principal de Linked Data en el jardín botánico es la integración total de los datos, ya que todavía existe dispersión de conocimientos, en donde se evidenció producto de la exploración la carente aplicación de tecnologías, así como la falta de recursos digitales.
- Los resultados generados producto de la RSL a través del análisis de 35 trabajos, permitieron identificar y utilizar las metodologías (Generación y Publicación de Linked Data, Methontology), herramientas conjuntas (Apache Jena Fuseki, Pubby, Virtuoso) para el desarrollo del modelo prototipo del jardín botánico “Reinaldo Espinosa” con datos reales, exponiendo los beneficios de modelar los recursos con más estructura al describir sus características, plasmadas en propiedades y relaciones.
- LOD Refine se constituye en la herramienta primordial referente a la depuración, limpieza y desambiguación de datos permitiendo eliminar términos repetidos, espacios en blanco, considerados como ruido al momento de realizar la conversión efectiva de la información al estándar RDF.
- La ontología LDJBRE permite realizar una asignación correcta con los elementos del jardín botánico “Reinaldo Espinosa”, evaluado a través del método del triángulo se logró cumplir con sus requerimientos, destacando los errores comunes de no poseer un concepto principal, lenguaje basado en un idioma, barreras superadas en consultas Sparql, evidenciando la utilidad de elaborar una ontología para representar la información en el dominio de biodiversidad.
- La aplicación de tecnologías de Web Semántica y Linked Data integradas a biodiversidad a través del modelo definido en Jardin-Botanico.RDF es una manera eficaz de representar los datos del jardín botánico “Reinaldo Espinosa”, bajo un modelo de datos (RDF) adoptado como un estándar internacional.
- El estudio derivado de trabajar en el jardín botánico “Reinaldo Espinosa” con un enfoque de Linked Data es más eficiente frente al manejo de documentación en archivos PDF, logrando adherirse al paradigma de Open Data y Open Science para su beneficiosa repercusión en la comunidad científica.

i. Recomendaciones.

Con la generación del modelo prototipo, aplicadas las tecnologías de Web Semántica y Linked Data surgen varias posibilidades en diferentes ramas asociadas al campo de experimentación, por lo que se realizan las siguientes recomendaciones:

- La generación de un estudio previo de los vocabularios existentes hasta la actualidad, para seleccionar los más utilizados en el área de conocimiento donde se desea representar la información, ya que se encuentran en constante evolución y desarrollo.
- En la selección de vocabularios a reutilizar en torno al jardín botánico “Reinaldo Espinosa” y tópicos de biodiversidad, revisar el vocabulario Darwin Core por su amplio contexto de aplicación, correcta representación de recursos, integrado en el desarrollo de la ontología elaborada.
- La integración del vocabulario LDJBRE en la adopción del enfoque de desarrollo de aplicaciones semánticas que fomenten la investigación acerca de biodiversidad en la región sur del Ecuador, modelando los datos RDF para llegar a mejorar y compartir más información entre la comunidad científica y el jardín botánico “Reinaldo Espinosa”.
- El impulsar la diligencia de tecnologías de Web Semántica y Linked Data por parte de las entidades educativas para desarrollar aplicaciones semánticas o sistemas inteligentes que fortalezcan la educación, rama de la ciencia poco aplicada y examinada.
- A la Universidad Nacional de Loja generar un esquema organizacional para poder realizar un proceso de generación de datos organizadamente proyectado a su accesibilidad con la Web Inteligente (Web del futuro).
- Como trabajo a futuro la creación de portales de biodiversidad, propios de la Universidad Nacional de Loja, que permiten la creación de conjuntos de datos estructurados con mejor organización bajo el enfoque de Linked Data y Web Semántica para que los usuarios e investigadores empleen el modelo generado, representando los recursos de biodiversidad del jardín botánico “Reinaldo Espinosa”, montando la infraestructura tecnológica necesaria para que este pueda enlazarse permanentemente con otras bases de datos, permitiendo dar una visibilidad internacional por la variedad de especímenes que existe en la región sur de Ecuador.

j. Bibliografía

[1] G. Bustán, A. María, and D. F. J. González, “Principios y tecnologías linked data para la publicación de datos académicos de las diferentes carreras de la Universidad Nacional de Loja.,” 2016.

[2] L. V. Gomez Navas, A. E. Romero Mejia, and G. R. Orellana Quezada, “Publicación de datos enlazados observando los principios de Linked Data,” 2012.

[3] L. A. Maza Puchaicela, “Apertura de datos Públicos de Ecuador en base a los principios de Linked Open Data,” 2014.

[4] D. Fernández, “La naturaleza de la información semántica en la Web,” 2009.

[5] C. M. Quezada Maldonado, “Estudio de la Mineria de contenido Web y la Web Semantica aplicada a la busqueda de Informacion en la web,” 2008.

[6] N. Piedra, “Marco de Trabajo de Interoperabilidad Semántica para la Educación Abierta,” 2016.

[7] T. Berners Lee, “Linked Data - Design Issues,” 2009. [Online]. Available: https://www.w3.org/DesignIssues/LinkedData.html. [Accessed: 13-Jan-2019].

[8] M. Zarate, G. Braun, and S. Almonacid, “Transformando Datos de Biodiversidad en Linked Data,” pp. 364–368, 2017.

[9] V. Chavan and L. Penev, “The data paper: a mechanism to incentivize data publishing in biodiversity science.,” *BMC Bioinformatics*, vol. 12 Suppl 1, no. Suppl 15, 2011.

[10] J. Wieczorek *et al.*, “Darwin core: An evolving community-developed biodiversity data standard,” *PLoS One*, vol. 7, no. 1, 2012.

[11] N. Piedra and J. P. Suárez, “Hacia la Interoperabilidad Semántica para el

Manejo Inteligente y Sostenible de Territorios de Alta Biodiversidad usando SmartLand-LD," *RISTI - Rev. Ibérica Sist. e Tecnol. Informação*, vol. 26, pp. 104–121, 2018.

[12] M.-A. Nolin *et al.*, "Bio2RDF Network Of Linked Data," *Web*, no. January, pp. 1–8, 2008.

[13] P. Fafalios and P. Papadakos, "Theophrastus: On demand and real-time automatic annotation and exploration of (web) documents using open linked data," *J. Web Semant.*, vol. 29, pp. 31–38, 2014.

[14] Universidad Nacional de Loja, "Nosotros | Universidad Nacional de Loja." [Online]. Available: http://www.unl.edu.ec/universidad/nosotros. [Accessed: 21-Jan-2019].

[15] Universidad Nacional de Loja, "Inicio Investigacion | Universidad Nacional de Loja." [Online]. Available: http://www.unl.edu.ec/investigacion/inicio-investigacion. [Accessed: 21-Jan-2019].

[16] jbreinaldoespinosa, "Datos Generales | jbreinaldoespinosa." [Online]. Available: https://jbreinaldoespinosa.wordpress.com/curriculum/. [Accessed: 21-Jan-2019].

[17] T. Robertson *et al.*, "The GBIF integrated publishing toolkit: Facilitating the efficient publishing of biodiversity data on the internet," *PLoS One*, vol. 9, no. 8, 2014.

[18] A. N. García, "Plataforma de Datos Virtuoso: Arquitectura, Tecnologías y Caso de Estudio," p. 49, 2015.

[19] Semantic Web, "4store - semanticweb.org.edu." [Online]. Available: http://semanticweb.org/wiki/4store. [Accessed: 16-Jan-2019].

[20] J. Alberto, G. Luna, M. S. M. López, I. Ingrid, and D. Torres, "Metodologías y métodos para la construcción de ontologías," no. 50, pp. 133–140, 2012.

[21] M. P. Villalon, "METODOLOGÍA NEON APLICADA A LA

REPRESENTACIÓN DEL CONTEXTO," no. September 2010, 2019.

[22] M. Arciniega, M. Belén, and S. Faggioni, "Modelo ontológico para la representación de datos académicos y su publicación con tecnología semántica," 2016.

[23] CES, "Reglamento de Régimen Académico," no. 289, pp. 1–53, 2013.

[24] J. D. Ludeña Jiménez, "'Estudio de factibilidad para la migración de los servicios de Tecnología de Información hacia Cloud Computing, para el Gobierno Autónomo Descentralizado Municipal de Loja,'" Universidad Nacional de Loja, 2019.

[25] Y. Castán, "Introduccion Al Metodo Cientifico Y Sus Etapas," p. 6, 2014.

[26] M. Genero Bocco, J. A. Cruz Lemus, and M. G. Piattini Velthuis, *Metodos de Investigacion en ingenieria del Software*. 2014.

[27] B. Kitchenham, "Procedures for Performing Systematic Reviews," 2004.

[28] J. D. Velasquez, "Una Guía Corta para Escribir Revisiones Sistemáticas de Literatura Parte 3," *Dyna*, vol. 82, no. 189, pp. 9–12, 2015.

[29] Y. A. López Rodríguez, Y. Hidalgo Delgado, and N. Silega Martínez, "Un método práctico para la evaluación de ontologías," no. May, 2018.

[30] F. K. Amanqui, K. J. Serique, S. D. Cardoso, J. L. D. Santos, A. Albuquerque, and D. A. Moreira, "Improving biodiversity data retrieval through semantic search and ontologies," *Proc. - 2014 IEEE/WIC/ACM Int. Jt. Conf. Web Intell. Intell. Agent Technol. - Work. WI-IAT 2014*, vol. 1, pp. 726–731, 2014.

[31] T. Horn, "Integrating Biodiversity Data into Botanic Collections," *Biodivers. Data J.*, vol. 4, p. e7971, 2016.

[32] A. Zvoleft and J. Ahumada, "UNDERSTANDING THE LINK BETWEEN POPULATION DYNAMICS AND BIODIVERSITY CONSERVATION THROUGH REMOTE SENSING AND GRIDDED POPULATION DATA

INTEGRATION," pp. 2560–2563, 2015.

[33] M. I. C. Domínguez, C. M. Q. Castro, A. V Ramírez, I. G. González, and R. M. Rodríguez, "Ontology-based system to support decision-making in the environmental management business process [Sistema basado en ontología para el apoyo a la toma de decisiones en el proceso de gestión ambiental empresarial]," *CEUR Workshop Proc.*, vol. 1797, no. November, pp. 54–66, 2016.

[34] F. K. Amanqui *et al.*, "A model of provenance applied to biodiversity datasets," *Proc. - 25th IEEE Int. Conf. Enabling Technol. Infrastruct. Collab. Enterp. WETICE 2016*, pp. 235–240, 2016.

[35] J. R. da Silva, J. A. Castro, C. Ribeiro, J. Honrado, Â. Lomba, and J. Gonçalves, "Beyond INSPIRE: An Ontology for Biodiversity Metadata Records," pp. 597–607, 2014.

[36] G. E. Barchini and M. M. Álvarez, "Dimensiones e indicadores de la calidad de una ontología," *Rev. Mex. Física*, vol. 52, no. 1, pp. 69–74, 2010.

[37] L. Sauermann and R. Cyganiak, "Cool URIs for the Semantic Web," 2008. [Online]. Available: https://www.w3.org/TR/cooluris/. [Accessed: 31-May-2019].

[38] S. Jupp *et al.*, "The EBI RDF platform: Linked open data for the life sciences," *Bioinformatics*, vol. 30, no. 9, pp. 1338–1339, 2014.

[39] A. Meroño Peñuela *et al.*, "Semantic technologies for historical research: A survey," *Jane's Def. Ind.*, vol. 6, no. AUG., pp. 539–564, 2015.

[40] M. Hallo, S. Lujan-Mora, and A. Mate, "Publishing a scorecard for evaluating the use of open-access journals using linked data technologies," *IEEE CITS 2015 - 2015 Int. Conf. Comput. Inf. Telecommun. Syst.*, no. Dl, pp. 1–5, 2015.

[41] F. Orciuoli, "Supporting seamless learning with semantic technologies and situation awareness," *Proc. - 2014 Int. Conf. Intell. Netw. Collab. Syst. IEEE*

INCoS 2014, pp. 726–730, 2014.

[42] L. Boonlamp, “A linked data approach to planning collaboration amongst local governments in Thailand,” *Proceeding 2017 2nd Int. Conf. Inf. Technol. INCIT 2017*, vol. 2018-Janua, pp. 1–5, 2018.

[43] V. T. Yoshiura *et al.*, “Towards a health observatory conceptual model based on the semantic web,” *Procedia Comput. Sci.*, vol. 138, pp. 131–136, 2018.

[44] D. Moussallem, M. Wauer, and A. C. N. Ngomo, “Machine Translation using Semantic Web Technologies: A Survey,” *J. Web Semant.*, vol. 51, pp. 1–19, 2018.

[45] N. Piedra, E. Tovar, R. Colomo-Palacios, J. Lopez-Vargas, and J. A. Chicaiza, “Consuming and producing linked open data: The case of OpenCourseWare,” *Program*, vol. 48, no. 1, pp. 16–40, 2014.

[46] P. Y. Vandenbussche, G. A. Atemezing, M. Poveda-Villalón, and B. Vatant, “Linked Open Vocabularies (LOV): A gateway to reusable semantic vocabularies on the Web,” *Semant. Web*, vol. 8, no. 3, pp. 437–452, 2017.

[47] S. J. Baskauf and C. O. Webb, “Darwin-SW: Darwin Core-based terms for expressing biodiversity data as RDF,” *Semant. Web*, vol. 7, no. 6, pp. 629–643, 2016.

[48] J. D. Velásquez Henao, “Una Guía Corta para Escribir Revisiones Sistemáticas de Literatura. Parte 2,” *Dyna*, vol. 82, no. 190, pp. 9–12, 2015.

[49] N. Villanueva-Rosales, L. G. Chavira, N. Del Rio, and D. Pennington, “EScience through the integration of data and models: A biodiversity scenario,” *Proc. - 11th IEEE Int. Conf. eScience, eScience 2015*, pp. 171–176, 2015.

[50] T. Jones *et al.*, “Trends in access of plant biodiversity data revealed by Google Analytics,” *Biodivers. Data J.*, vol. 2, p. e1558, 2014.

[51] G. Skevakis, K. Makris, P. Arapi, and S. Christodoulakis, “Elevating natural history museums’ cultural collections to the linked data cloud,” *CEUR Workshop Proc.*, vol. 1091, no. Sda, pp. 40–51, 2013.

[52] M. J. Costello, W. K. Michener, M. Gahegan, Z. Q. Zhang, and P. E. Bourne, “Biodiversity data should be published, cited, and peer reviewed,” *Trends Ecol. Evol.*, vol. 28, no. 8, pp. 454–461, 2013.

[53] G. Munoz, W. D. Kissling, and E. Van Loon, “Biodiversity Observations Miner: A web application to unlock primary biodiversity data from published literature,” *Biodivers. Data J.*, no. in revision, 2018.

[54] L. A. M. Zalani, K. S. Jye, S. Balakrishnan, and S. K. Dhillon, “Biodiversity Databases and Tools,” *Ref. Modul. Life Sci.*, pp. 1–13, 2018.

[55] V. Proença *et al.*, “Global biodiversity monitoring: From data sources to Essential Biodiversity Variables,” *Biol. Conserv.*, vol. 213, pp. 256–263, 2017.

[56] P. Haase *et al.*, “The next generation of site-based long-term ecological monitoring: Linking essential biodiversity variables and ecosystem integrity,” *Sci. Total Environ.*, vol. 613–614, pp. 1376–1384, 2018.

[57] R. G. Mateo, K. Mokany, and A. Guisan, “Biodiversity Models: What If Unsaturation Is the Rule?,” *Trends Ecol. Evol.*, vol. 32, no. 8, pp. 556–566, 2017.

[58] M. Mountantonakis and Y. Tzitzikas, “High performance methods for Linked Open Data connectivity analytics,” *Inf.*, vol. 9, no. 6, pp. 1–33, 2018.

[59] F. Michel, C. Faron-Zucker, and F. Gandon, “Integration of Web APIs and Linked Data Using SPARQL Micro-Services — Application to Biodiversity Use Cases,” *Biodivers. Inf. Sci. Stand.*, vol. 2, p. e25481, 2018.

[60] M. Diepenbroek *et al.*, “Terminology supported archiving and publication of environmental science data in PANGAEA,” *J. Biotechnol.*, vol. 261, no. July, pp. 177–186, 2017.

[61] A. Lausch, A. Schmidt, and L. Tischendorf, “Data mining and linked open data - New perspectives for data analysis in environmental research,” *Ecol. Modell.*, vol. 295, pp. 5–17, 2015.

[62] S. D. Cardoso, K. J. Serique, F. K. Amanqui, J. L. C. Dos Santos, and D. A. Moreira, “A gazetteer for biodiversity data as a linked open data solution,” *Proc. Work. Enabling Technol. Infrastruct. Collab. Enterp. WETICE*, pp. 435–440, 2014.

[63] C. S. Parr *et al.*, “TraitBank: Practical semantics for organism attribute data,” *Semant. Web*, vol. 7, no. 6, pp. 577–588, 2016.

[64] S. J. C. Janssen *et al.*, “Towards a new generation of agricultural system data, models and knowledge products: Information and communication technology,” *Agric. Syst.*, vol. 155, pp. 200–212, 2017.

[65] Creative Commoons, “Sobre las licencias - Creative Commons.” [Online]. Available: https://creativecommons.org/licenses/?lang=es. [Accessed: 21-Feb-2019].

[66] DB-Engines, “DB-Engines Ranking - popularity ranking of RDF stores.” [Online]. Available: https://db-engines.com/en/ranking/rdf+store. [Accessed: 16-Jan-2019].

[67] MarkLogic Corp., “Feature-Rich Database to Integrate Data Better and Faster - MarkLogic.” [Online]. Available: https://www.marklogic.com/product/marklogic-database-overview/database-features/. [Accessed: 21-Jan-2019].

[68] Amazon Inc., “Amazon Neptune: base de datos de gráficos rápida y de confianza desarrollada para la nube.” [Online]. Available: https://aws.amazon.com/es/neptune/. [Accessed: 21-Jan-2019].

[69] Stardog Union, “Platform | Stardog.” [Online]. Available: https://www.stardog.com/platform/#features. [Accessed: 21-Jan-2019].

[70] Stardog Union, “Features | Stardog.” [Online]. Available:

https://www.stardog.com/features/. [Accessed: 21-Jan-2019].

[71] Stardog Union, “Estudio stardog.” [Online]. Available: https://www.stardog.com/studio/. [Accessed: 21-Jan-2019].

[72] Eclipse RDF4J, “About – Eclipse RDF4J.” [Online]. Available: http://rdf4j.org/about/. [Accessed: 21-Jan-2019].

[73] R. Leonardo, R. Coronel, P. Alarcón, and R. Torres, “Calidad De Datos En Contextos De Linked Data,” 2014.

k. Anexos

Lista de Anexos.

ANEXO 1: REVISIÓN SISTEMÁTICA DE LITERATURA..... 83.
ANEXO 2: CRITERIOS DE SELECCIÓN..... 112.
ANEXO 3: RDF STORES..... 115.
ANEXO 4: CUADRO METODOLOGÍAS PARA ONTOLOGÍAS..... 117.
ANEXO 5. DOCUMENTACIÓN BLOG JARDÍN BOTÁNICO.....119.
ANEXO 6. VOCABULARIO – ONTOLOGÍA LDJBRE.....124.
ANEXO 7. DESCRIPCIÓN LDJBRE..... 146.
ANEXO 8. VERIFICACIÓN RDF.....152.
ANEXO 9. ENTREVISTA..... 153.
ANEXO 10. EVALUACIÓN DE LA ONTOLOGÍA ..... 155.
ANEXO 11. MÉTODO EXPERIMENTO..... 161.
ANEXO 12. LICENCIA CREATIVE COMMONS..... 170.
ANEXO 13. CERTIFICADO TRADUCCIÓN FINE TUNED ENGLISH..... 171.
ANEXO 14. LISTA DE ACRÓNIMOS..... 172.

ANEXO 1: REVISIÓN SISTEMÁTICA DE LITERATURA

Tiene como fin resumir, compilar, criticar y sintetizar la investigación existente sobre un área temática o fenómeno de interés usando un proceso de búsqueda, análisis y síntesis que pueda clarificar el estado del arte, dar soporte a nuevas investigaciones e identificar variables importantes acordes al área de dominio de la investigación, Velásquez [48].

1. Planificación.

1.1 Justificación de la Revisión

La revisión sistemática tiene como meta principal obtener una aproximación al estado del arte sobre la Web Semántica y Linked Data aplicada a la Biodiversidad. Buscando un resultado final concreto de una síntesis de literatura apropiada para la realización del TT.

1.2 Formulación de la Pregunta de Investigación

Se busca la respuesta a las siguientes interrogantes:

- **P1. ¿QUÉ TIPO DE CONTRIBUCIONES Y BENEFICIOS SE HAN REALIZADO A TRAVÉS DE LA WEB SEMÁNTICA Y LINKED DATA APLICADAS A BIODIVERSIDAD?**

- **P2. ¿CUÁLES SON LAS TECNOLOGÍAS DE WEB SEMÁNTICA Y LINKED DATA ACTUALMENTE APLICADAS?**

1.3 Diseño de la búsqueda

a. Estrategias de búsqueda

✓ Los documentos que serán utilizados en esta revisión bibliográfica serán solo los obtenidos en las fuentes bibliográficas que se detallan en la sección de resumen de fuentes y estrategia de búsquedas.

✓ La búsqueda se basará en palabras específicas que den contestación a las preguntas de investigación, como: Web Semántica, Linked Data, Biodiversidad; así como, sinónimos aplicados a la temática de estudio como RDF, ontologías, modelos.

b. Fuentes bibliográficas seleccionadas

- Google Scholar: https://scholar.google.com/
- Scopus: https://www.scopus.com/
- ScienceDirect: https://www.sciencedirect.com/
- IEEExplorer: https://ieeexplore.ieee.org/Xplore/home.jsp

c. Cadenas de búsqueda

Para la construcción de la cadena de búsqueda apropiada se realizó una investigación preliminar que facilitó la identificación de términos clave para establecer la cadena final que se consultará en las fuentes bibliográficas pertinentes, la TABLA XXIX muestra los estudios previos.

TABLA XXIX.

PALABRAS CLAVES PARA LA GENERACION DE LA CADENA DE BÚSQUEDA.

TITULO DE ESTUDIO	PALABRAS CLAVE
Named based Approach to Build a HUB for Biodiversity LOD	Linked Open Data (LOD), biodiversity, analysis, databases.
The data paper: a mechanism to incentivize data publishing in biodiversity science.	Biodiversity, conservation, metadata, data publishing.
The EBI RDF platform: Linked open Data for the Life sciences.	Linked Open Data, exploration resources, Semantic Web, web services, datasets.
Linked Data – The Story So Far	Linked Data, Web of Data, Semantic Web, Data exploration.
Added Values of Linked Data in Education: A survey and Roadmap	Linked Data, education, interlinking,
Open Data y Linked data : su impacto en el área de bibliotecas y documentación	Linked Open data, application, experiments, data, semantic web.

Elaborado por el Autor.

Para la obtención de literatura pertinente acorde al tema de consulta y responda a las preguntas de RSL, se elaboró dos cadenas de búsqueda. Con el objetivo de responder la primera pregunta planteada (ver sección 1.2) P1 se creó la cadena de búsqueda siguiente: (((((((((Semantic Web) OR Linked Data) OR Linked Open Data) AND applied*) OR model*) OR experiments*) OR conservation*) OR analysis*) AND Biodiversity).

En respuesta a la pregunta P2 se generó la siguiente cadena de búsqueda: ((“Semantic Web“) and (“Linked Data”)) and (“technologies”) and (“models”)*

TABLA XXX.

CADENAS DE BUSQUEDA RSL.

Fuentes Bibliográficas	Cadena de Búsqueda
Google Scholar, Scopus, SciendeDirect, IEEExplorer.	(((((((((Semantic Web) OR Linked Data) OR Linked Open Data) AND applied*) OR model*) OR experiments*) OR conservation*) OR analysis*) AND Biodiversity).
Google Scholar, Scopus, SciendeDirect, IEEExplorer	Semantic Web and Linked Data and technologies and models*.

Elaborado por el Autor.

d. Fuentes y Estrategias De Búsqueda.

Motores de búsqueda: Google Scholar, Scopus, ScienceDirect, IEEExplorer.

Idioma: español o inglés.

e. Criterios de inclusión.

Los criterios de inclusión de los documentos a obtener comprenderán:

- **Tipos de estudio:** Tesis, artículos científicos, libros, revistas científicas.
- **Año de Publicación:** artículos con vigencia de 6 años atrás.
- **Criterio Investigador:** artículos con relación explícita a la cadena de búsqueda, que sobresalga en la investigación palabras clave similares, y que se considere necesario incluir para la RSL.

f. Criterios de exclusión

Se excluirán todos los documentos que no contesten las preguntas de la RSL, que no cumplan con los criterios de inclusión, es decir; se descartaran los estudios que acorde a su contenido no posean mayor grado de relevancia, términos equivalentes a la temática central, relación directa con las preguntas de investigación a resolver y cadenas de búsqueda.

2. Ejecución.

2.1 Búsqueda de documentos

Luego de establecer los parámetros para la RSL, precedemos a realizar la búsqueda de literatura en las fuentes bibliográficas establecidas, aplicando las cadenas de búsqueda correspondientes con cada motor de búsqueda.

2.2 Selección de documentos

Se realiza la búsqueda de literatura, aplicando los criterios de exclusión e inclusión, y acorde al criterio del investigador en la inclusión de literatura de trabajos importantes que no cumplen con los años de publicación de vigencia, pero son fundamentales; se obtuvieron un total de 35 trabajos relacionados que responderán las preguntas de la RSL, tales como:

TABLA XXXI.

SELECCIÓN DE DOCUMENTOS RSL.

Código	Titulo	Términos	Año de Publicación	Buscador
ES01	EScience through the integration of data and	Semantic web, Biodiversidad, model web.	2015	IEEExplorer

	models: A biodiversity scenario [49].			
ES02	The data paper: a mechanism to incentivize data publishing in biodiversity science [9].	Linked Data, Biodiversidad	2011	Google Scholar
ES03	Darwin core: An evolving community-developed biodiversity data standard [10].	Biodiversidad, estándar.	2012	Google Scholar
ES04	Trends in access of plant biodiversity data revealed by Google Analytics [50].	Linked Data, Biodiversity.	2014	Scopus
ES05	Darwin-SW: Darwin Core-based terms for expressing biodiversity data as RDF [47].	Web Semántica, Linked Data, Biodiversidad	2016	Google Scholar
ES06	Elevating Natural History Museums' Cultural Collections to the Linked Data Cloud [51].	Linked Data.	2013	Google Scholar
ES07	Biodiversity data should be published, cited, and peer reviewed [52].	Linked Data, Biodiversidad	2013	Google Scholar
ES08	Biodiversity Observations Miner: A web application to unlock primary biodiversity data from published literature [53].	Linked Data, Biodiversity.	2018	Google Scholar
ES09	Biodiversity Databases and Tools [54].	Biodiversity	2018	Science Direct
ES10	Global biodiversity monitoring: From data sources to Essential Biodiversity Variables [55].	Linked Data	2017	Science Direct
ES11	The next generation of site-based long-term ecological monitoring: Linking essential biodiversity variables and ecosystem integrity [56].	Linked Data, Biodiversidad.	2018	Science Direct
ES12	Biodiversity Models: What If Unsaturation Is the Rule? [57].	Biodiversity, models.	2017	Science Direct
ES13	Transformando Datos de Biodiversidad en Linked Data [8].	Linked Data, Biodiversidad	2017	Google Scholar
ES14	Hacia la Interoperabilidad Semántica para el Manejo Inteligente y Sostenible de Territorios de Alta Biodiversidad usando SmartLand-LD [11].	Computer Science, E nginneering	2018	Scopus
ES15	High Performance Methods for Linked Open Data Connectivity Analytics [58].	Computer Science, Linked Data	2018	Scopus

ES16	Integration of Web APIs and Linked Data Using SPARQL Micro-Services — Application to Biodiversity Use Cases [59].	Computer Science, Biodiversity	2018	Scopus
ES17	Theophrastus: On demand and real-time automatic annotation and exploration of (web) documents using open linked data [13].	Linked Data, Biodiversity.	2014	Science Direct
ES18	Terminology supported archiving and publication of environmental science data in PANGAEA [60].	Linked Data, Biodiversity.	2017	Science Direct
ES19	A Model of Provenance Applied to Biodiversity Datasets [34].	Linked Data, Biodiversity	2016	IEEExplorer
ES20	Data mining and linked open data – New perspectives for data analysis in environmental research [61].	Linked Open Data, Semantic Web.	2015	Science Direct
ES21	Understanding the link between population dynamics and biodiversity conservation through remote sensing and gridded population data integration [32].	Biodiversity.	2015	IEEExplorer
ES22	Improving Biodiversity Data Retrieval through Semantic Search and Ontologies [30]	Semantic, Biodiversity.	2014	IEEExplorer
ES23	A Gazetteer for Biodiversity Data as a Linked Open Data Solution [62].	Semantic web, Biodiversity	2014	IEEExplorer
ES24	Semantic technologies for historical research: A survey [39].	Semantic Web, technologies	2015	Scopus
ES25	Linked Open Vocabularies (LOV): A gateway to reusable semantic vocabularies on the Web [46].	Linked Data	2017	Scopus
ES26	Publishing a scorecard for evaluating the use of open-access journals using linked data technologies [40].	Linked Data, semantic web.	2015	IEEExplorer
ES27	Supporting seamless learning with semantic technologies and situation awareness [41]	Linked Data, Semantic web.	2014	IEEExplorer
ES28	A linked data approach to planning collaboration amongst local governments in Thailand [42]	Linked Data	2018	IEEExplorer
ES29	Towards a health observatory conceptual	Semantic web	2018	Science Direct

	model based on the semantic web [43]			
ES30	Machine Translation using Semantic Web Technologies: A Survey [44].	Linked Data, Semantic web	2018	Science Direct
ES31	Consuming and producing linked open data: the case of OpenCourseWare [45].	Linked open Data.	2014	Google Scholar
ES32	The EBI RDF platform: linked open data for the life sciences [38].	Linked Data	2014	Google Scholar
ES33	TraitBank: Practical semantics for organism attribute data [63].	Biodiversity, semantic Web.	2016	Google Scholar
ES34	Beyond INSPIRE: An Ontology for Biodiversity Metadata Records [35].	Biodiversity, Linked Open Data.	2014	Google Scholar
ES35	Towards a new generation of agricultural system data, models and knowledge products: Information and communication technology [64].	Linked Data	2017	Science Direct

Elaborado por el autor.

2.3 Análisis de Literatura

Luego de seleccionar la literatura correspondiente para resolver las preguntas de la RSL, detallaremos los hallazgos en la literatura encontrada.

TABLA XXXII.

DESCRIPCIÓN RESULTADOS ES01.

Estudio	**ES01**
Problema	La creación de modelos de Biodiversidad requiere la integración y manipulación de conjuntos de datos desacoplados a través de diferentes dominios del conocimiento.
Solución	Desarrollo del marco ELSEWeb, cuyo objetivo es facilitar la integración de datos ambientales para la generación de modelos de distribución de especies. Lidia la arquitectura y su desarrollo, incluyendo la interfaz de usuario y las lecciones aprendidas de este esfuerzo colaborativo y multidisciplinario.
Conclusiones	Se establece el marco de trabajo de Tierra, Vida y Web Semántica (ELSEWeb) utilizada para automatizar la recuperación y transformación de los datos ambientales necesarios para generar un modelo de Biodiversidad suficientemente robusto y eficaz.
Contribución	Construcción de la visión de la ciber-infraestructura que permite a los usuarios centrarse en cuestiones científicas de interés en lugar de aprender otros lenguajes de programación y navegar a través de docenas de sitios web en busca de los datos necesarios para probar un modelo, requerido en la investigación.
Beneficios y limitaciones	Encontramos beneficios como: Integración automatizada de datos medioambientales, transformación de los datos ambientales necesarios para generar un modelo de biodiversidad, estudio de usabilidad que aporta una mejora continua de la aplicación y flujo de trabajo con los usuarios. Al igual que limitaciones como: la producción de errores al tener metadatos obsoletos sin depuración, la necesidad de mecanismos de sincronización

	frecuentes para evitar la inconsistencia en las descripciones de los servicios y la recuperación de datos.

TABLA XXXIII.

DESCRIPCIÓN RESULTADOS ES02.

Estudio	**ES02**
Problema	Los datos primarios de Biodiversidad no son de fácil acceso ni detectables. Entre varios problemas, uno es la falta de incentivos a los editores de datos para publicar sus recursos, mecanismos que actualmente falta para el reconocimiento a través de la publicación académica convencional de metadatos enriquecidos, lo que debería garantizar el rápido descubrimiento de los recursos de datos de biodiversidad 'aptos para el uso'.
Solución	Propone el establecimiento del "documento de datos sobre biodiversidad" como un posible mecanismo para ofrecer reconocimiento académico a los esfuerzos y la inversión de los editores de datos en la creación de metadatos ricos y la publicación de los mismos como documentos académicos de referencia.
Conclusiones	Los documentos de datos agilizan la publicación de recursos sobre biodiversidad adecuados para su uso. Sin embargo, la adopción y el establecimiento del documento de datos como posible mecanismo de reconocimiento académico requiere un alto grado de compromiso e inversión por parte de las comunidades interesadas.
Contribución	Mecanismo que trata de ofrecer reconocimiento académico a través del registro de prioridades, la citabilidad y la difusión de los esfuerzos y la inversión de los editores de datos en la creación de metadatos ricos.
Beneficios y Limitaciones	Expone beneficios como: proporcionar una publicación en revistas que ofrezcan crédito académico a los editores de datos, describir los datos de forma estructurada y legible por el ser humano, señalar la existencia de los datos a la atención de la comunidad académica, vínculos estrechos o interconexiones con los ciclos de publicación de datos y publicación académica. Plantea limitaciones como: Carencia de un modelo para la publicación y el descubrimiento de datos hacia sus autores a pequeña escala, falta de esfuerzos definitivos para asegurar el descubrimiento de recursos de datos de regiones mega-diversas, en desarrollo y subdesarrolladas del mundo, falta de acceso a un amplio espectro de datos primarios sobre biodiversidad, el costo de la creación y el mantenimiento de la infraestructura por parte de los editores de datos de pequeña y mediana escala.
Tecnologías y Plataformas	Destaca la plataforma GBIF. Metalenguaje GMP. Formato de Darwin Core (DwC-A).

TABLA XXXIV.

DESCRIPCIÓN RESULTADOS ES03.

Estudio	**ES03**
Problema	Un paso esencial hacia la comprensión de los patrones globales de biodiversidad es proporcionar una visión estandarizada de estas fuentes de datos heterogéneas para mejorar la interoperabilidad. Fundamental para este avance las definiciones de términos comunes pueden ser el mejor camino.
Solución	Se centra en las categorías de términos que definen el estándar de Darwin Core, propone la utilización del mismo, de modo que aumenta en gran medida el valor y la reutilización de los datos de biodiversidad accesibles y de libre disponibilidad para que puedan integrarse e incorporarse efectivamente en otros esfuerzos científicos de "gran desafío".

Conclusiones	Darwin Core cumple una función esencial en la descripción de datos en un formato estándar, basado en las aportaciones de la comunidad y en la utilidad demostrada en Biodiversidad.
Contribución	La creación de un formato estándar para los datos de biodiversidad y el creciente uso de los Archivos Centrales de Darwin (estándar) han ayudado a superar el desafío de la limitada disponibilidad de datos para responder a las preguntas sobre biodiversidad.
Beneficios y limitaciones	Se presentan beneficios como: Adaptar los vocabularios del TDWG, integrar recursos con otros vocabularios y ontologías para su uso en la descripción de los recursos de la biodiversidad, producir y proporcionar fácil acceso a los archivos centrales de Darwin y a los metadatos de recursos; desarrollo de herramientas para simplificar la movilización y el intercambio de datos sobre biodiversidad utilizando el Núcleo Darwin. Se presentan limitantes como: falta de una coordinación entre los datos de la biodiversidad, falta de una ontología bien definida, una definición formal de las relaciones entre los términos en un dominio definido.
Tecnologías y Plataformas	Estándar de Darwin Core. Plataforma GBIF. Tecnologías: XML, CSV. Formatos: RDF, Access to Biological Collections Data (modelo ABCD).

TABLA XXXV.

DESCRIPCIÓN RESULTADOS ES04.

Estudio	**ES04**
Problema	La cantidad de datos sobre Biodiversidad vegetal disponible a través de la web ha aumentado en la última década, pero para que estos datos estén disponibles se requiere una considerable inversión de tiempo y trabajo, consideraciones vitales tanto para las organizaciones como para las instituciones que buscan validar los factores de impacto de estos trabajos en línea.
Solución	Utilizar Google Analytics (GA) para medir el valor de la presencia digital, examina las tendencias de uso utilizando 15 cuentas GA diferentes, distribuidas en 451 instituciones o proyectos botánicos que comprenden más del cinco por ciento de los herbarios del mundo. Los datos de usuario de la muestra revelan: 1) más de 17 millones de sesiones web, 2) en cinco sistemas operativos principales, 3) el tráfico directo y de búsqueda domina con un impacto mínimo de los medios sociales, 4) los tipos de dispositivos móviles y nuevos se han duplicado cada año durante los últimos tres años, 5) los navegadores web, las herramientas que utilizamos para interactuar con la web, están cambiando.
Conclusiones	Herbaria es una amplia gama de servicios informativos que van más allá de la preservación básica de plantas e incluyen: recursos de nomenclatura, literatura, identificación, glosarios necesarios para la jerga botánica e información importante derivada de especímenes.
Contribución	Examinación de los datos de Google Analytics (GA) de varios sitios web relacionados con plantas y hongos. En particular, los cambios que podrían sugerir nuevas direcciones que se deberían tomar para maximizar el valor de la inversión que los museos y los herbarios están haciendo en los esfuerzos de digitalización.
Beneficios y Limitaciones	Tiene beneficios como: proporcionar recomendaciones a los actuales gestores y desarrolladores de información sobre la interfaz de usuario y la experiencia; recursos que permiten además estudios evolutivos y ecológicos que proporcionan una ventaja adicional de un modelo bien establecido que se encuentra en el reino de las plantas. Posee limitaciones como: el tiempo, el conocimiento a nivel de expertos requiere tiempo para adquirirlo, para ser efectivo, todas las páginas requieren la colocación del código de seguimiento.
Tecnologías	Google Analytics.

TABLA XXXVI.

DESCRIPCIÓN RESULTADOS ES05.

Estudio	ES05
Problema	Expresar datos de Biodiversidad en el Marco de Descripción de Recursos (RDF) tiene una serie de ventajas sobre los formatos de texto de campo simple (donde una línea en el archivo contiene los datos para un solo registro).
Solución	Utilizar DSW para facilitar el marcado coherente de los datos sobre Biodiversidad para que los gráficos de los conjuntos de datos creados por los diferentes proveedores pudieran fusionarse fácilmente.
Conclusiones	DSW proporciona un mecanismo para comprobar si sus términos se utilizan de manera coherente con su modelo. Sus propiedades permiten la creación de consultas SPARQL simples que pueden utilizarse para descubrir nueva información sobre recursos enlazados cuyos metadatos son generados por diferentes proveedores.
Contribución	Desarrollo de un modelo consensuado de RDF para la comunidad informática de la biodiversidad, las propiedades de los objetos Darwin-SW pueden ser utilizadas para llevar a cabo tareas de razonamiento en el contexto de la integración de datos de instancias de biodiversidad.
Beneficios y Limitaciones	Aporta con beneficios como: establecer un mapeo estándar de un modelo relacional de consenso para el estándar RDF, permitir la integración de datos y la eficacia de consulta, las propiedades de los términos DSW permiten eliminar los vínculos incoherentes entre recursos, facilita la vinculación de los recursos básicos de los museos (especímenes, taxones) con los recursos externos. Así como limitaciones como: el uso inconsistente de las propiedades de los objetos DSW causado por la des-normalización del modelo DSW inhibe en la eficacia de la consulta, el razonamiento limitado para detectar enlaces inapropiados posee equivalencia en realizar varias consultas SPARQL simples.
Tecnologías y Plataformas	Modelo de Darwin-SW. Plataforma GBIF. Modelo RDF. Lenguaje SPARQL.

TABLA XXXVII.

DESCRIPCIÓN RESULTADOS ES06.

Estudio	ES06
Problema	La abundancia de contenido científico de alta calidad sobre Historia Natural y Biodiversidad es producida de manera distribuida y abierta por Museos de Historia Natural (Museos de Historia Natural (MHN)) utilizando sus propios estándares establecidos y mejores prácticas. El Gestionar la publicación de tal riqueza y variedad de contenidos en la Web, así como el apoyo a los procesos de creación de contenidos distribuidos e interoperables, plantea el reto de que los enfoques de publicación tradicionales no son adecuados para encontrarnos.
Solución	El proyecto Natural Europe ofrece una solución coordinada a retos de nivel europeo que tiene como objetivo mejorar la disponibilidad, la capacidad de descubrimiento y la calidad de vida de las personas, y la pertinencia de los contenidos culturales ambientales para la educación y la formación a lo largo de toda la vida. El uso del aprendizaje, en un contexto multilingüe y multicultural; presenta la arquitectura de la infraestructura semántica desarrollada para la transición de la federación Natural Europe de Repositorios culturales de los MHN a la Web Semántica, así como la metodología.

Conclusiones	Europeana ha seguido la ingesta y conversión de los metadatos del patrimonio cultural de los Museos Nacionales de Salud en Datos Vinculados; al igual que se expone una metodología que puede aplicarse también en otros ámbitos, aprovechando esquemas, y en relación con los vocabularios / taxonomías; proporcionando contenido de biodiversidad y culturalidad.
Contribución	Mejorar la disponibilidad y la pertinencia de los contenidos culturales medioambientales para la educación y la utilización del aprendizaje permanente, en un contexto multilingüe y multicultural a través de una plataforma.
Beneficios y Limitaciones	Beneficios como: la creación de grandes portales nacionales e internacionales del patrimonio cultural, como Europeana, grandes depósitos de datos abiertos, como la nube abierta de datos vinculada, y publicaciones masivas de datos de bibliotecas vinculadas; facilita el ciclo de vida de la gestión de metadatos. También limitaciones como: el acceso transparente a fuentes de datos federadas de CDR que cumplen con diferentes esquemas de Ontología.
Tecnologías y Plataformas	Formatos: RDF(S), SKOS. Tecnologías: SPARQL, OWL, XML. Vocabularios: Catalogue of life (CoL), ABCD. Services/datasets: GeoNames, DBpedia, Uniprot, GEMET.

TABLA XXXVIII.

DESCRIPCIÓN RESULTADOS ES07.

Estudio	**ES07**
Problema	La preocupación por la calidad de los datos impide el uso de bases de datos públicas sobre biodiversidad y los beneficios subsiguientes para la sociedad. La publicación de los datos podría seguir el proceso de publicación bien establecido: con controles de calidad automatizados, revisión por pares y decisiones editoriales. Esto mejoraría la precisión de los datos, reduciría la necesidad de que los usuarios "limpien" los datos y podría aumentar su uso. Los autores y editores obtendrían el crédito debido por una publicación (de datos) revisada por pares a través de métricas de uso y citación.
Solución	El proceso de publicación de datos que justifique el establecimiento de revistas de datos sobre biodiversidad en línea de acceso abierto, la adopción de la publicación de datos por parte de las revistas existentes. Hace hincapié en la importancia del acceso abierto a los datos sobre biodiversidad y reconoce el apoyo financiero que se requiere; principio detrás de muchas políticas gubernamentales e intergubernamentales para la disponibilidad de datos.
Conclusiones	La publicación de datos sobre Biodiversidad mediante metadatos podría convertirse en una ontología de dominio, con una rica semántica que describa las diversas categorías utilizadas y sus interrelaciones; el hecho de que la publicación final sea revisada por pares y publicada al estilo de una revista científica convencional atrae a científicos que contribuyan con impulsar la generación colaborativa de información en esta área de conocimiento.
Contribución	Proceso de publicación de datos realizado a través de revistas de datos que archivan datos en bases de datos integradas gestionadas por centros de datos, como GBIF.
Beneficios y limitaciones	Beneficios: establecimiento de revistas de datos sobre biodiversidad en línea de acceso abierto, la adopción de la publicación de datos por parte de las revistas existentes, desarrollo de nuevos estándares de datos. Limitaciones: gran parte del material suplementario no está archivado permanentemente y puede llegar a ser inaccesible con el tiempo, publicación de datos sobre biodiversidad a través de GBIF por parte de organizaciones gubernamentales, de los cuales sólo fragmentos pueden estar asociados con trabajos de investigación, necesidad de apoyo financiero.

Tecnologías y Plataformas	Plataforma GBIF. Sistema de Información Biogeográfica de los Océanos. Bases de Datos: FishBase, AlgaeBase, Cetabase. Estándar: Darwin Core Archive.

TABLA XXXIX.

DESCRIPCIÓN RESULTADOS ES08.

Estudio	**ES08**
Problema	Una parte considerable de los datos primarios sobre biodiversidad se encuentra digitalmente encerrada en la literatura publicada, que a menudo se almacena como archivos pdf. Los enfoques a gran escala de la ciencia de la Biodiversidad podrían beneficiarse de la recuperación de esta información y hacerla accesible digitalmente y legible por máquinas. No obstante, la cantidad y diversidad de literatura publicada digitalmente plantea muchos desafíos para el descubrimiento y la recuperación de conocimientos.
Solución	La minería de textos ha sido ampliamente utilizada para tareas de descubrimiento de datos en grandes cantidades de documentos. Presenta una nueva herramienta de minería de texto de código abierto, el Biodiversity Observations Miner (BOM). Esta aplicación web, escrita en R, permite el descubrimiento semiautomático de observaciones puntuales de la biodiversidad (por ejemplo, interacciones bióticas, rasgos funcionales) asociadas a los nombres científicos presentes en un corpus de literatura científica.
Conclusiones	La minería de texto es una herramienta efectiva para el descubrimiento de colecciones de datos relacionadas hacia la biodiversidad.
Contribución	Herramienta que facilita el descubrimiento, movilización y reutilización de datos ecológicos almacenados en la literatura científica.
Beneficios y Limitaciones	Beneficios: progreso sustancial para cerrar la brecha computacional en ecología, las observaciones de los rasgos funcionales y las interacciones bióticas de las especies en el campo de la minería de texto son punto de partida para llenar las lagunas de conocimiento que limitan el avance de la ecología. Limitaciones: la necesidad de un acceso fácil y estándar a los datos de biodiversidad en aumento, encontrar información en la literatura puede convertirse en un reto y requerir mucho tiempo.
Tecnologías y Plataformas	Biodiversity Observations Miner (BOM). Herramienta Global Names Recognition and Discovery.

TABLA XL.

DESCRIPCIÓN RESULTADOS ES09.

Estudio	**ES09**
Problema	Existen diferentes tipos de datos biológicos, como la secuencia de ADN, secuencia de proteínas, estructura de las proteínas, biodiversidad y muchos otros campos especializados.
Solución	La base de datos proporciona un método conveniente para almacenar una enorme cantidad de datos y permite a los biólogos o investigadores recolectar los datos, organizarlos sistemáticamente, analizar los datos y compartir la información con otros investigadores de manera eficiente. Los datos biológicos a menudo se componen de datos complicados. La base de datos biológica es una gran biblioteca organizada de información sobre ciencias de la vida, recopilada a partir de experimentos científicos, literatura publicada, tecnología de experimentos de alto rendimiento y análisis computacionales.

Conclusiones	El desarrollo de una base de datos biológica aporta significativamente, ya que permite afrontar el desafío de capacidad de la tecnología para gestionar los datos complejos que engloba la información biológica, bajo un enfoque de tecnologías semánticas, con una base de datos eficaz.
Contribución	Ofrece una descripción de las necesidades de investigación y biodiversidad motivadas por un conjunto de tecnologías con portales frontales que proporcionan dos bases de datos, la base de datos GBIF y la base de datos EOL, que son los mejores ejemplos de bases de datos y herramientas de biodiversidad.
Beneficios y limitaciones	Beneficios: permite documentar la diversidad biológica de la vida, demuestra los cambios en el medio ambiente que han tenido lugar a lo largo del tiempo, revelar las especies exóticas invasoras que son una amenaza para la biodiversidad, predecir el potencial de una especie no autóctona en la región. Limitaciones: las políticas y la toma de decisiones a todos los niveles en la forma en que los seres humanos perciben la biodiversidad, eficacia de sistemas automatizados con la capacidad de ayudar a devolver identificaciones precisas de las especies con anotaciones.
Tecnologías y Plataformas	Bases de datos: Antweb, Fishbase, Georgian Biodiversity. Plataforma GBIF, EOL, IBIS.

TABLA XLI.

DESCRIPCIÓN RESULTADOS ES10.

Estudio	**ES10**
Problema	Los estudios de campo ecológicos y los esquemas de monitoreo intensivo informan a una amplia gama de EBVs, pero los primeros tienden a entregar datos a corto plazo, mientras que la cobertura geográfica de los segundos es limitada. Las Variables Esenciales de Biodiversidad (EBVs) proporcionan una consolidación de información de diversas fuentes de observación de la biodiversidad de manera que, ¿se puede describir cómo eficaz?, en contraste, los esquemas de monitoreo extensivo informan la abundancia de la población EBV, pero entregan datos a largo plazo a través de una extensa red de sitios.
Solución	Clasifica las fuentes de observaciones primarias en cuatro tipos: sistemas de vigilancia amplios e intensivos, estudios de campo ecológicos y teledetección por satélite; caracteriza su cobertura geográfica, taxonómica y temporal. La teledetección por satélite es particularmente adecuada para proporcionar información sobre la función y la estructura de los ecosistemas de los EBV.
Conclusiones	El conjuntos de datos EBV representativos depende tanto de la capacidad de integrar los datos disponibles, a través de la armonización de datos y los esfuerzos de modelado, como del establecimiento de nuevos programas de monitoreo para abordar las brechas críticas de datos.
Contribución	Variables Esenciales de Biodiversidad a fin de proporcionar indicadores de alto nivel para informar a los científicos, administradores y al público sobre el cambio global de la biodiversidad.
Beneficios y Limitaciones	Beneficios: promover enfoques eficientes en función de los costos, evaluar las fuentes de datos subyacentes e identificar los sesgos existentes en la actualidad, proponer métodos para abordar las brechas en los datos, mejor uso posible de los datos generados por los estudios de campo ecológicos y los esquemas de monitoreo. Limitaciones: mayor deterioro en el estado y las tendencias de la biodiversidad, la cobertura geográfica de los programas de monitoreo de biodiversidad existentes es insuficiente, el conjunto actual de indicadores pasa por alto aspectos importantes de la biodiversidad, debido a las lagunas en los datos primarios.
Tecnologías y Plataformas	Datasets LPI. Base de datos TRY, Plataforma GBIF, GloBI.

TABLA XLII.

DESCRIPCIÓN RESULTADOS ES11.

Estudio	ES11
Problema	Los efectos del cambio global sobre la Biodiversidad y el bienestar humano exigen la mejora de los datos ambientales a largo plazo como base para la ciencia, las políticas y la toma de decisiones, incluyendo una mayor interoperabilidad, multifuncionalidad y armonización.
Solución	Basándose en el ejemplo de dos iniciativas mundiales, la Red Internacional de Investigación Ecológica a Largo Plazo (ILTER) y el Grupo de Observación de la Biodiversidad de la Tierra (GEOBON), propone fusionar los marcos en los que se basan estas iniciativas, la integridad de los ecosistemas y las variaciones esenciales de la biodiversidad, para que sirvan de guía mejorada para la investigación y la vigilancia a largo plazo basadas en los sitios en los ecosistemas terrestres, de agua dulce y costeros en el futuro.
Conclusiones	El beneficio de la fusión de los marcos de las dos iniciativas mundiales promete que las variables enumeradas se utilizan comúnmente en el monitoreo de ecosistemas. Para obtener los mejores beneficios de estos datos, sin embargo, aún es necesario establecer circuitos de retroalimentación más eficientes entre las mediciones realizadas en el sitio y los requisitos locales, regionales y globales tanto para la ciencia como para las políticas.
Contribución	Establecimiento de una red de sitios en una misma ubicación dentro de un ecosistema con tareas de investigación y supervisión compartidas que proporcionen datos modulares para usos flexibles y polivalentes.
Beneficios y limitaciones	Beneficios: contribuir a detener la pérdida mundial de biodiversidad para 2020, proporcionar datos ambientales coherentes que permitan realizar análisis causales de la investigación básica de los ecosistemas, promover la comparabilidad e interoperabilidad de los datos entre los distintos emplazamientos y redes de control. Limitaciones: falta de interoperabilidad de los datos recogidos y la creación de interfaces entre los diferentes enfoques adoptados por las distintas comunidades, necesidad de ajustes en la recolección de datos para documentar de forma robusta los cambios de las unidades biológicas en el espacio y el tiempo.
Tecnologías	Frameworks: EI y EBV.

TABLA XLIII.

DESCRIPCIÓN RESULTADOS ES12.

Estudio	ES12
Problema	Mejorar las predicciones de la Biodiversidad es esencial si queremos hacer frente a los retos que plantea el cambio global. Como el conocimiento es clave para los modelos de alimentación, necesitamos evaluar cómo la teoría debatida puede afectar a los modelos. Un importante debate en curso es si las limitaciones ambientales limitan el número de especies que pueden coexistir en una comunidad (saturación), con hallazgos recientes que sugieren que la riqueza de especies en muchas comunidades podría ser insaturada.
Solución	Los modelos de biodiversidad podrían abordar esta cuestión teniendo en cuenta una dualidad: considerar a las comunidades como insaturadas, pero donde la composición de las especies está limitada por diferentes generadores de biodiversidad dependientes de la escala. Identifica una variedad de avances prometedores para incorporar esta dualidad en los enfoques de modelización de la biodiversidad de aplicación común y mejorar sus predicciones espaciales.

Conclusiones	Necesidad de esfuerzos considerables, incluidos nuevos enfoques metodológicos y conceptuales, para iluminar esta área de investigación; ya que actualmente se encuentra en una etapa prematura.
Contribución	El desarrollo de la modelización de la biodiversidad y su avance debe tomar en cuenta la dualidad de las comunidades insaturadas, en las que la composición de las especies se ve limitada por una variedad de factores de biodiversidad dependientes de la escala (principalmente la energía y la heterogeneidad del hábitat).
Beneficios y limitaciones	Beneficios: Considerar la composición final prevista o la riqueza de especies como una distribución de probabilidad, predecir la composición y riqueza de especies a escala local. Limitaciones: el modelo de biodiversidad dentro de una perspectiva comunitaria insaturada, requiere una serie de soluciones innovadoras de modelización, tamaño de especies y región (hábitat) de estudio, falta proporcionar flexibilidad en los modelos para que puedan considerar teorías complementarias o alternativas a las actuales teorías principales.

TABLA XLIV.

DESCRIPCIÓN RESULTADOS ES13.

Estudio	**ES13**
Problema	Las colecciones biológicas contienen información heredada irreemplazable sobre nuestra biosfera, que es esencial para comprender como la Biodiversidad está cambiando en una era de impactos humanos sin precedentes. Dichos análisis solo son prácticos si se digitalizan, integran y hacen disponibles en línea los datos de las colecciones biológicas de todo el mundo.
Solución	Desarrollar una arquitectura para convertir y publicar datos de Biodiversidad utilizando tecnologías de la Web Semántica, en particular los principios establecidos por la iniciativa Linked Open Data (LOD) para compartir y relacionar información.
Conclusiones	Los datos de biodiversidad a través de tecnologías de web semántica adquieren un grado más alto de importancia que los dispersos en su estado anterior. GraphDB es un repositorio semántico que permite almacenar las tripletas generadas por OpenRefine y también permite trabajar con motores de inferencia y de consultas SPARQL que sobresaltan la elección de enfocarlos a través de tecnologías semánticas.
Contribución	Análisis de diferentes tecnologías y formalismos que se requieren para poder ofrecer una fuente de datos enlazada y abierta, con información de biodiversidad.
Beneficios y Limitaciones	Beneficios: diseñar un prototipo para la visualización y consulta de los datos generados y enlazados. Limitaciones: vinculación de repositorios propios.
Tecnologías y Plataformas	Plataformas OpenRefine, GraphDB, JENA. Formatos CSV, JSON, RDF, Darwin Core. Recursos como FOAF, Dublín Core. Lenguaje de consulta SPARQL.

TABLA XLV.

DESCRIPCIÓN RESULTADOS ES14.

Estudio	**ES14**

Problema	Un reto presente en los actuales esfuerzos de desarrollo sostenible es potenciar la interoperabilidad entre variables e indicadores almacenados en sistemas de información, proveedores de servicios de datos autónomos, distribuidos y heterogéneos. El acceso restringido a fuentes de datos y servicios limita las capacidades de un territorio inteligente. sostenible para mejorar los aspectos sociales, ambientales y económicos.
Solución	La construcción de gráficos de conocimiento, la interoperabilidad semántica y las técnicas colaborativas de ciencia abierta, el marco propuesto permite el mapeo e integración de datos sobre indicadores de desarrollo sostenible en territorios de alta biodiversidad.
Conclusiones	La calidad y la información oportuna son vitales para que los gobiernos, el sector académico, las organizaciones internacionales, el sector privado y el público en general puedan tomar decisiones informadas.
Contribución	Enfoque para el manejo inteligente y sostenible de territorios de alta biodiversidad, gestión inteligente de las interacciones entre personas y territorios.
Beneficios y Limitaciones	Beneficios: creación de indicadores que faciliten la planificación a través de múltiples escalas de tiempo, disponibilidad de una arquitectura de datos web abierta que abarca la recopilación, preservación de datos, reutilización, análisis, modelado predictivo y la visualización de datos. Limitaciones: límites semánticos en los que los significados, la terminología y el vocabulario son diferentes, desajustes de coordinación que existen entre las variables e indicadores.
Tecnologías y Plataformas	Vocabularios: DCAT, SDMX, RDF Data Cube, DCMI, Geodata, FOAF, VoID, PROV. Tecnologías: URIs, RDF, XML, SPARQL. Herramienta Silk, Virtuoso.

TABLA XLVI.

DESCRIPCIÓN RESULTADOS ES15.

Estudio	**ES15**
Problema	La carencia de métricas, mediciones (e índices) de conectividad disponibles que involucren más de dos conjuntos de datos, que cubran todo el contenido (por ejemplo, entidades, esquemas, triplicaciones) o "rebanadas" (por ejemplo, triplicaciones para una entidad específica) de conjuntos de datos. Generalmente, no es una tarea fácil encontrar las conexiones entre los conjuntos de datos, ya que existe un gran número de conjuntos de datos LOD y el cierre transitivo y simétrico de las relaciones de equivalencia debe calcularse para las conexiones que no faltan.
Solución	Métodos y algoritmos escalables, para realizar el cálculo del cierre transitivo y simétrico de las relaciones de equivalencia (ya que pueden producir más conexiones entre los conjuntos de datos), para medir la conectividad entre dos o más conjuntos de datos.
Conclusiones	La conectividad entre dos o más conjuntos de datos a través de Entidades comunes, Triples, Literales y elementos de Esquema, aportan un enfoque muy productivo mientras que pueden ocurrir más conexiones debido a las relaciones de equivalencia entre URIs (owl:sameas).
Contribución	Introducción de algoritmos para calcular el cierre transitivo y simétrico de las relaciones de equivalencia, mientras que muestra cómo crear en paralelo índices semánticos para diferentes conjuntos de elementos (entidades, triples, clases, etc.).
Beneficios y Limitaciones	Beneficios: mediciones concernientes a la conectividad entre dos o más conjuntos de datos, detección de conjuntos de datos más conectados.

	Limitaciones: el cálculo del cierre se realiza en el tiempo de consulta, que puede consumir mucho tiempo; se deben realizar varias uniones (y comparaciones) para encontrar la intersección para un subconjunto específico de conjuntos de datos, falta de conexiones para los elementos del esquema.
Tecnologías y Plataformas	Framework MapReduce. Tecnologías: SPARQL, RDF, OWL. Herramientas: AMADA, H2RDF, MAPSIN.

TABLA XLVII.

DESCRIPCIÓN RESULTADOS ES16.

Estudio	**ES16**
Problema	A menudo no logran hacer interoperables las descripciones de los recursos debido al hecho de que se basan en vocabularios propios que carecen de semántica formal. Los principios de Linked Data también buscan la publicación masiva de datos en la Web, pero con el objetivo específico de garantizar la interoperabilidad semántica.
Solución	Utiliza los principios de la arquitectura de micro-servicios para definir una arquitectura SPARQL Micro-Service, destinada a la consulta de APIs Web utilizando SPARQL, pretende solventar la aparición de un ecosistema de micro servicios SPARQL publicados por proveedores independientes para permitir a las aplicaciones basadas en Linked Data extraer fácilmente datos de una gran cantidad de servicios distribuidos, escalables y fiables.
Conclusiones	Los micro servicios SPARQL proporcionan metadatos auto descriptivos procesables por máquina, tales como los argumentos esperados, la forma en que se transmiten al micro servicio (por ejemplo, parámetros de cadena de consulta HTTP) y el gráfico típico que se puede generar.
Contribución	Uso de los microservicios SPARQL en varios casos de uso de la vida real relacionados con el dominio de la biodiversidad.
Beneficios y limitaciones	Beneficios: La reutilización de términos de vocabularios bien adoptados adjunta una semántica claramente definida, vincular los resultados a la URI de un gráfico o un documento que proporcione metadatos y triples de control por separado. Limitaciones: cada micro-servicio de SPARQL se defina junto con sus propios términos a medida.
Tecnologías y Plataformas	Tecnologías: JSON, XML, RDF, SPARQL. Plataforma: Docker, MongoDB, GitHub, GBIF.

TABLA XLVIII.

DESCRIPCIÓN RESULTADOS ES17.

Estudio	**ES17**
Problema	El proceso de identificación y recopilación de información relacionada con una especie es lento y engorroso. Específicamente, se informó de los cuellos de botella de este proceso, no hay un punto central que pueda proporcionar la exploración de especies, géneros, etc. similares.
Solución	Sistema que soporta la anotación automática de documentos (web) a través de la minería de entidades y proporciona servicios de exploración mediante la explotación de Linked Open Data (LOD), el sistema tiene por objeto ayudar a los biólogos en sus investigaciones sobre las especies y la biodiversidad, ha sido diseñado para ser altamente configurable en diferentes aspectos como entidades de interés, fichas de información y sistemas de búsqueda externos.

Conclusiones	El principal cuello de botella expuesto es el rendimiento y la fiabilidad de los puntos finales SPARQL en línea, debido a su tecnología inmadura y constantes cambios.
Contribución	Explota el carácter dinámico y abierto del LOD para configurar las entidades de interés, así como para especificar cómo enlazar, enriquecer y navegar por las entidades identificadas y la información semántica relacionada.
Beneficios y Limitaciones	Beneficios: exploración a través de la minería de entidades y la explotación de Datos Abiertos Vinculados, aumento de la eficiencia y la fiabilidad del servicio de exploración de las entidades. Limitaciones: No es fácil localizar, identificar y resumir todas las especies referenciadas en una fuente de información, el tiempo necesario para recuperar la información relevante es variable, fiabilidad y el rendimiento de los puntos finales SPARQL en línea.
Tecnologías y Plataformas	Tecnologías: XML, RDF, SPARQL. Plataforma: DBpedia endpoint.

TABLA XLIX.

DESCRIPCIÓN RESULTADOS ES18.

Estudio	**ES18**
Problema	Desde sus primeras etapas, los datos se archivaron de manera consistente y se anotaron de acuerdo con la forma en que se produjeron, incluyendo información sobre los investigadores principales, los tipos de medición y observación, los métodos de muestreo y análisis, y los dispositivos, así como las referencias a la literatura.
Solución	Se elabora un sistema que integra un catálogo terminológico (TC), con interfaces que permiten reproducir y trabajar manualmente con terminologías. El tesauro de clasificación de términos de PANGAEA, que forma parte del TC, se utilizó como un vocabulario general que enlaza los distintos dominios y permite realizar exploraciones y simulaciones laterales con diversas facetas.
Conclusiones	La aplicación de terminologías en sistemas tiene un efecto positivo mutuo para los servicios de terminología y los sistemas de información como PANGAEA, ya que mejora el contenido, la fiabilidad y la interoperabilidad.
Contribución	Establecimiento de restricciones para la configuración de las consultas de datos, basadas en la elaboración de un catálogo de términos construidas en su sistema.
Beneficios y limitaciones	Beneficios: vincular los valores de los datos con los términos de la cooperación técnica fomenta la armonización de los datos, generación dinámica de la matriz de datos a descargar a partir del conjunto de resultados. Limitaciones: El tiempo necesario para la revisión y aceptación de los términos varía, la falta de coherencia en el nivel de detalle proporcionado por los proveedores de datos.
Tecnologías y Plataformas	Tecnologías: XML, RDF. Catalogo TC.

TABLA L.

DESCRIPCIÓN RESULTADOS ES19.

Estudio	**ES19**
Problema	Los mecanismos actualmente disponibles no proporcionan suficiente información sobre la procedencia de la información sobre biodiversidad. Un número cada vez mayor de instituciones de investigación de la biodiversidad agregan nuevos especímenes y la información relacionada con ellos a sus colecciones biológicas y ponen esta información a disposición en la Web.

Solución	Modelo de procedencia de la biodiversidad que amplía el modelo de datos PROV del W3C. Los datos de Biodiversidad se asignan a términos de ontologías relevantes, como Dublin Core y GeoSPARQL, almacenadas en bases de datos para triples y consultados utilizando los puntos finales de SPARQL.
Conclusiones	BioProv permite que las aplicaciones que analizan la biodiversidad incorporen datos de procedencia en su información. Definimos un documento de mapeo para los datos de biodiversidad de IBT para generar triples de RDF.
Contribución	Modelo de procedencia: Actividad de Recolección y Catalogación, que utiliza la información de procedencia para que los expertos en biodiversidad puedan realizar consultas y responder a preguntas científicas.
Beneficios y Limitaciones	Beneficios: referencia para evaluar la precisión y el recuerdo de consultas en datos de biodiversidad. Limitaciones: realización de consultas estructuradas más avanzadas, en asociación con investigadores de la biodiversidad.
Tecnologías y Plataformas	Tecnologías: XML, RDF, PROV, SPARQL, GeoSPARQL, OWL, RML. Plataformas: GBIF, INPA, LBA, NYBG.

TABLA LI.

DESCRIPCIÓN RESULTADOS ES20.

Estudio	**ES20**
Problema	El rápido desarrollo de la tecnología de la información y la informática, aumento extremo de la recogida y el almacenamiento de datos digitales. La técnica de la minería de datos ofrece una manera prometedora de extraer conocimientos y patrones de conjuntos de datos grandes, multidimensionales y complejos.
Solución	Visión general de las técnicas de minería de datos existentes y de las herramientas relacionadas, así como ilustrar el potencial de la minería de datos para diferentes áreas de investigación por medio de ejemplos de aplicaciones. Los procesos, tipos y métodos de minería de datos se presentan de manera comprensible y se proporcionaron referencias importantes a las herramientas de minería de datos de código abierto existentes, que se caracterizan por su uso simple e intuitivo en apoyo del análisis de datos complejos.
Conclusiones	La minería de datos ofrece una solución prometedora con numerosos enfoques para el descubrimiento de conocimientos en datos multidimensionales y complejos. Las técnicas de extracción de datos se utilizan actualmente en áreas de investigación analítica como la informática, la física o las matemáticas, y requieren una amplia experiencia en programación.
Contribución	Estudio y comparación de las técnicas tradicionales de minería de datos, referencia a las ventajas del LOD, así como a las herramientas comerciales y de código abierto existentes entre estas.
Beneficios y limitaciones	Beneficios: permitir enfoques globales de minería de datos y el descubrimiento de asociaciones, patrones y conocimientos verdaderamente nuevos, mostrar nuevas perspectivas para resolver problemas complejos y hacer mejores predicciones. Limitaciones: los patrones y el conocimiento sólo pueden ser extraídos de conjuntos de datos cerrados, los datos complejos de los que se dispone actualmente deben ponerse en perspectiva como resultado del impacto de las decisiones económicas, sociales, ambientales y de planificación.
Tecnologías y Plataformas	Tecnologías: XHTML, RDF, SPARQL, OWL. Herramientas: Rapidminer, Rattle, WEKA, Knime, JENA, 4Store.

TABLA LII.

DESCRIPCIÓN RESULTADOS ES21.

Estudio	ES21
Problema	Existen algunos progresos en la contención de la pérdida de bosques en los trópicos, la deforestación sigue amenazando la biodiversidad tropical. Muchas áreas protegidas tropicales existen como "islas" rodeadas de cambios continuos en el uso de la tierra y el desarrollo.
Solución	Encontrar relaciones significativas entre la longitud de la red de carreteras y la pérdida de bosques y entre la longitud de la red de carreteras y el desarrollo agrícola. Para entender las interconexiones entre la cubierta terrestre, la biodiversidad y el crecimiento de la población, utilizamos un conjunto de datos único de la Red de Evaluación y Monitoreo de la Ecología Tropical (TEAM) junto con productos derivados de la cubierta terrestre de Landsat, datos de población de la Población en Cuadrícula del Mundo (GPW), y datos de carreteras del Conjunto de Datos de Acceso Abierto de Carreteras Mundiales (gROADS).
Conclusiones	No se encuentra de manera clara patrones que describan la ocupación de las especies y la pérdida de bosques o la ganancia agrícola, o entre la ocupación de las especies y el cambio de la población humana.
Contribución	Examinar los patrones de cambio de la población y de la LCLUC dentro y alrededor de las Áreas Protegidas tropicales, y los efectos de estos cambios en la biodiversidad.
Beneficios y Limitaciones	Beneficios: Evidenciar los efectos de la expansión de las plantaciones forestales, facilita la migración y la extracción de recursos, ayudando a las economías locales. Limitaciones: Carencia de patrones claros entre la ocupación de las especies y la pérdida de bosques, variantes a la zona de estudio.

TABLA LIII.

DESCRIPCIÓN RESULTADOS ES22.

Estudio	ES22
Problema	la actualidad, los investigadores en el campo de Biodiversidad utilizan bases de datos para extraer nuevos conocimientos y también para compartir sus propios descubrimientos. Sin embargo, cuando estos investigadores necesitan encontrar información relevante en los datos, no es apropiado para ser utilizado en estas grandes cantidades de datos heterogéneos de biodiversidad, lo que conduce a resultados de búsqueda con baja precisión y memoria.
Solución	Nueva arquitectura que aborda este problema utilizando los datos de biodiversidad y su mapeo con términos de ontologías relevantes, como Darwin Core, DBpedia, Ontobio y Catálogo de Vida, se almacenan utilizando formatos de web semántica y se consultan utilizando herramientas de web semántica (como las tiendas triples).
Conclusiones	Una vez que los datos de recolección están en RDF es fácil integrarlo con datos de fuentes diferentes e independientes (si comparten términos ontológicos comunes). La integración de datos y los scripts son muy importantes porque, en muchos casos, después de encontrar datos relevantes, es casi imposible manipularlos de forma manual dado su tamaño.
Contribución	Arquitectura que utiliza un sistema de búsqueda semántica para consultar datos de biodiversidad y formatos, herramientas de web semántica para representarla.

Beneficios y Limitaciones	Beneficios: Mapeo entre los datos de biodiversidad y las ontologías que los describen, reutilización enfocada en Linked Open Data, comparar los términos de ontología, derivados de las palabras clave de entrada del usuario, búsqueda semántica con mejores resultados de medición. Limitaciones: necesidad de mucho esfuerzo, por parte de los expertos en dominios, para averiguar qué datos deben ser devueltos en cada consulta.
Tecnologías y Plataformas	Plataformas: GBIF, INPA, NYGB, SpeciesLink, Protege. Tecnologías: Extensible Observation Ontology (OBOE), SPARQL, OntoBio. Vocabularios: Darwin Core, EnvO, Geonames.

TABLA LIV.

DESCRIPCIÓN RESULTADOS ES23.

Estudio	**ES23**
Problema	La falta de información espacial geográfica precisa en los datos de ocurrencia de especies, especialmente de regiones ricas en diversidad (como la Selva Amazónica), conduce a problemas en muchas actividades de conservación, como la planificación sistemática para la protección de especies en peligro de extinción.
Solución	Presenta un nomenclátor (un directorio geográfico que asocia los lugares con nombres a coordenadas geográficas) para los datos de biodiversidad que está disponible como un recurso de Datos Abiertos Vinculados (usando un Punto Final GeoSPARQL) y muestra cómo se puede utilizar para mejorar los datos de recolección de datos geográficos imprecisos. Compara la eficiencia de Gazetteer con tres recursos abiertamente disponibles, Geonames, WikiMapia y Wikipedia.
Conclusiones	El nomenclátor aporta beneficios como el poder añadir coordenadas geográficas ausentes a los registros de biodiversidad y eliminar información geográfica inexacta.
Contribución	Mejorar la precisión de la información geográfica contenida en los datos biológicos y almacenarla en un Gazetteer (directorio geográfico que asocia lugares con nombres a coordenadas geográficas).
Beneficios y Limitaciones	Beneficios: permitir realizar consultas complejas, que no son posibles con los datos estructurados comunes en los nomenclátores, mejorar eficazmente la precisión de los datos sobre biodiversidad, facilidad de intercambiar datos mediante tecnologías LOD. Limitaciones: falta de registros con datos de latitud y longitud y otros tienen coordenadas incorrectas (o inexactas), registros con ocurrencias antiguas.
Tecnologías y Plataformas	Plataformas: GBIF, SpeciesLink. Tecnologías: RDF, GeoSPARQL, XML, Repositorios: Geonames, Wikipedia, WikiMapia.

TABLA LV.

DESCRIPCIÓN RESULTADOS ES24.

Estudio	**ES24**
Problema	Durante los años noventa del siglo pasado, historiadores e informáticos crearon juntos una agenda de investigación en torno al ciclo de vida de la información histórica. Comprendió las tareas de creación, diseño, enriquecimiento, edición, recuperación, análisis y presentación de información histórica con ayuda de tecnología de la información.

Solución	En este documento se estudia el trabajo conjunto de historiadores e informáticos en el uso de los métodos y tecnologías de la Web Semántica en la investigación histórica.
Conclusiones	El utilizar un enfoque de Web Semántica con información de historiadores aborda un desafío muy complejo para poder ser desplegadas correctamente en la historia. Deben aplicarse a los datos históricos en una tubería compleja, estratificada y adaptada adecuadamente.
Tecnologías y Plataformas	Tecnologías: RDF, URIs, OWL, XML, SPARQL. Vocabularios: RDFa, Dublin Core. Herramientas: AllegroGraph, IBMDB2, OpenLink Virtuoso, OWLIM.

TABLA LVI.

DESCRIPCIÓN RESULTADOS ES25.

Estudio	**ES25**
Problema	Una de las principales barreras para el despliegue de datos vinculados es la dificultad que tienen los editores de datos para determinar qué vocabularios usar para describir la semántica de los datos.
Solución	Informe sistemático que describe los vocabularios abiertos vinculados (LOV), un catálogo de vocabularios reutilizables de alta calidad para la descripción de los datos en la Web.
Conclusiones	LOV recopila y hace indicadores visibles, tales como las interconexiones entre los vocabularios y el historial de versiones de cada vocabulario, junto con el editor anterior y actual (individual u organización).
Tecnologías y Plataformas	Tecnologías: RDF, SPARQL, OWL, URI Vocabularios: Dublín Core, VOAF, VoID, Data Cube, DCAT. Herramientas: Jena Fuseki, RESTful.

TABLA LVII.

DESCRIPCIÓN RESULTADOS ES26.

Estudio	**ES26**
Problema	Las revistas de acceso abierto recopilan, conservan y publican información científica en formato digital, pero sigue siendo difícil no solo para los usuarios sino también para las bibliotecas digitales evaluar el uso y el impacto de este tipo de publicaciones.
Solución	Proceso para desarrollar y publicar un cuadro de mandos en la web semántica basado en la norma ISO 2789: 2013 que utiliza tecnologías de datos vinculados de tal manera que se puede vincular a conjuntos de datos relacionados.
Conclusiones	Las tecnologías de Datos Vinculados componen una oportunidad para enriquecer la información relacionada por los KPI, conectándolos a conjuntos de datos relevantes en toda la web.
Tecnologías y Plataformas	Herramientas: Open Link Virtuoso, CubeViz. Tecnologías: RDF, SPARQL, URI, XML. Vocabularios: Data Cube, BIBO, Dublín Core, FOAF, ORG, SKOS.

TABLA LVIII.

DESCRIPCIÓN RESULTADOS ES27.

Estudio	ES27
Problema	Una experiencia de aprendizaje sin fisuras, para una persona, se caracteriza por la capacidad de mantener la continuidad del aprendizaje, conectar actividades de aprendizaje a través de lugares, tecnologías, entornos sociales e integrar diferentes enfoques de enseñanza y aprendizaje. El aprendizaje ubicuo y el aprendizaje sin fisuras no son sinónimos. El aprendizaje ubicuo podría considerarse una condición necesaria para un aprendizaje sin fisuras. De hecho, la tecnología ubicua apoya a los alumnos de la manera correcta, en el lugar correcto y en el momento correcto, basándose en los contextos personales y ambientales del mundo real, que son características útiles para un aprendizaje sin fisuras.
Solución	Proponer y motivar el uso de los Datos Vinculados (realizados mediante la Pila de Web Semántica) y las técnicas de Concientización de la Situación para respaldar los escenarios de Aprendizaje sin fisuras. En particular, las tecnologías y metodologías de Datos Vinculados y Web Semántica se consideran muy útiles para modelar y respaldar la continuidad de la experiencia perfecta en actividades de aprendizaje heterogéneas (en calidad, tiempo y espacio).
Conclusiones	El acceso ubicuo se ve desde una representación de entorno inteligente en el que hay que explotar la información sobre la realidad actual para mejorar los procesos de aprendizaje.
Tecnologías y Plataformas	Tecnologías: RDF, SPARQL, RDFS, OWL, URIs, OWL2, XML. Vocabularios: SKOS.

TABLA LIX.

DESCRIPCIÓN RESULTADOS ES28.

Estudio	ES28
Problema	Para lograr una interoperabilidad eficiente de los sistemas de información, los datos vinculados desempeñan un papel importante en la relación de datos.
Solución	Desarrollo de un nuevo marco para la organización administrativa del subdistrito (SAO) del gobierno local en Tailandia. El objetivo es publicar datos vinculados entre los gobiernos locales que permitan descubrir, acceder y fomentar la interoperabilidad.
Conclusiones	La exposición de un enfoque de Linked Data en el gobierno local de Tailandia proporciona un beneficio bueno respecto a la interoperabilidad de los datos de la localidad. Siguiendo los enlaces de los resultados se obtendrá información útil relacionada con temas que antes no estaban disponibles.
Tecnologías	Tecnologías: RDF, SPARQL, OWL, URIs. Herramientas: Tripify, Virtuoso, D2R, Drupal.

TABLA LX.

DESCRIPCIÓN RESULTADOS ES29.

Estudio	ES29
Problema	Los Observatorios de la Salud han ganado popularidad mundial y se han establecido en todo el mundo desde la década de 1970. Con el arribo de la Web semántica, las máquinas pueden procesar, reutilizar, integrar y entender el significado (semántica) de la información y el conocimiento en la World Wide Web para realizar tareas complejas.

Solución	Proponer un modelo conceptual de observatorio de salud asentado en la web semántica para ayudar en el diseño, desarrollo e implementación de los procesos de un observatorio de salud. El modelo propuesto se basó en un análisis conceptual que contiene una revisión narrativa de las tecnologías de la Web Semántica, una arquitectura de software de múltiples capas y una revisión integradora de los Observatorios de la Salud.
Conclusiones	El modelo propuesto puede prestar oportunidades para el desarrollo e implementación de nuevos observatorios o para la adecuación de los observatorios de salud existentes.
Tecnologías	Tecnologías: RDF, SPARQL, OWL, Basic Formal Ontology (BFO), RDFa.

TABLA LXI.

DESCRIPCIÓN RESULTADOS ES30.

Estudio	**ES30**
Problema	Recientemente se ha desarrollado un gran número de enfoques de traducción automática para facilitar la migración fluida de contenido en todos los idiomas. Sin embargo, la literatura insinúa que todavía se deben enfrentar muchos obstáculos para lograr mejores traducciones automáticas.
Solución	La revisión sistemática de los enfoques de traducción automática que se basan en las tecnologías de la Web Semántica para traducir textos. En general, su encuesta sugiere que si bien las tecnologías de la Web Semántica pueden mejorar la calidad de los resultados de la traducción automática para diversos problemas, la mezcla de ambos está aún en su infancia.
Conclusiones	Los SWT se han utilizado principalmente para desambiguación en los sistemas MT y que sus capacidades han aumentado constantemente. Considerando el poder de decisión de SWT, no pueden ser inadvertidos por los futuros sistemas de MT.
Tecnologías	Métricas BLUE, MEANT, NIST y otras más.

TABLA LXII.

DESCRIPCIÓN RESULTADOS ES31.

Estudio	**ES31**
Problema	OpenCourseWare (OCW) es una de las iniciativas emergentes más triunfantes en el camino hacia el Contenido Abierto. El gran impacto de la iniciativa ha llevado a muchas universidades internacionales y de prestigio a unirse al movimiento, creando sus propios proyectos en coordinación con OCW-Massachusetts Institute of Technology.
Solución	Iniciativa para aplicar los principios de Linked Data para optimizar la búsqueda y el descubrimiento de contenidos OpenCourseWare (OCW) creados y compartidos por las universidades.
Conclusiones	Los datos enlazados entregan una solución para filtrar y seleccionar semánticamente aquellos contenidos educativos abiertos, y que automáticamente están enlazados a la nube de datos abierta enlazada.
Tecnologías y Plataformas	Tecnologías: RDF, SPARQL. Vocabulario: LOCWD Recursos OER.

TABLA LXIII.

DESCRIPCIÓN RESULTADOS ES32.

Estudio	**ES32**

Problema	El marco de descripción de recursos (RDF) es una tecnología emergente para describir, publicar y vincular datos de ciencias de la vida. Como uno de los principales proveedores de datos y servicios de bioinformática, el Instituto Europeo de Bioinformática (EBI) se compromete a hacer que los datos sean fácilmente accesibles para la comunidad de manera que compensen la demanda existente.
Solución	La plataforma RDF de EBI para satisfacer la progresiva demanda de coordinar las actividades de RDF en todo el instituto y proporciona un nuevo punto de entrada para consultar y explorar recursos integrados disponibles en el EBI.
Conclusiones	Los datos que han sido anotados utilizando ontologías, como EFO y Gene Ontology, reconocen la integración de datos con otros conjuntos de datos de la comunidad y aportan la semántica para realizar consultas ricas.
Tecnologías y Plataformas	Tecnologías: RDF, SPARQL, URIs. RESTful web services, Experimental Factor Ontology (EFO). Vocabularios: Dublin Core, Data Catalog Vocabulary and Vocabulary of Interlinked Datasets.

TABLA LXIV.

DESCRIPCIÓN RESULTADOS ES33

Estudio	**ES33**
Problema	El conocimiento humano de la vida en la Tierra es vasto, no hay una manera fácil de consultar toda la información acumulada en cientos de años de investigación y documentación sobre biodiversidad. Incluso preguntas simples como "¿qué plantas tienen flores amarillas? o "¿qué comen los tiburones? Los biólogos han capturado y manejado información sobre morfología, comportamiento, historia de vida e interacciones ecológicas de muchas maneras diferentes, por lo que sería optimo una herramienta que afronte el reto.
Solución	Enciclopedia de la Vida (EOL) ha desarrollado TraitBank (http://eol.org/traitbank), un nuevo repositorio de datos de atributos (rasgos) del organismo. TraitBank agrega, gestiona y sirve datos de atributos para organismos a través del árbol de la vida, incluyendo características del ciclo vital, hábitats, distribuciones, relaciones ecológicas y otros tipos de datos. Expone el impacto de TraitBank en los usuarios y colaboradores y los desafíos y beneficios de su enfoque ligero y escalable para la integración de datos de biodiversidad.
Conclusiones	TraitBank proporciona una organización del conocimiento distribuido de las fuentes de datos en un marco ligero y escalable, facilita su recuperación y reutilización para una variedad de aplicaciones, que van desde análisis sintéticos a gran escala de la biodiversidad hasta productos de datos vinculados; proporciona a los administradores de muchas ontologías de dominio relevantes casos de uso que pueden ayudar a optimizar el desarrollo de sus recursos, en sus investigaciones.
Contribución	Sistema diseñado por la Enciclopedia de la Vida (EOL) para adquirir, organizar y servir datos de atributos de la biodiversidad a escala global a lo largo de todo el árbol de la vida.
Beneficios y limitaciones	Beneficios: vincular los registros de datos con los ontologías y vocabularios controlados pertinentes, gestión eficiente de un almacén de datos grande, búsquedas refinadas especificando un valor o rango de valores. Limitaciones: la adquisición de datos está guiada por las necesidades de sus audiencias, patrocinadores y socios.

Tecnologías y Plataformas	Bases de datos: GBIF, GIoBI, OBIS, Paleobiology. Formatos: Darwin Core, BIBO. Ontologías: PATO (Phenotypic Quality Ontology), ENVO (Environments Ontology), SIO (Semantic science Integrated Ontology).

TABLA LXV.

DESCRIPCIÓN RESULTADOS ES34

Estudio	**ES34**
Problema	La gestión de los datos de investigación demanda a menudo la creación o reutilización de esquemas de metadatos especializados para satisfacer los requisitos de metadatos de cada grupo de investigación. Las ontologías presentan varias ventajas sobre los esquemas de metadatos. En particular, pueden compartirse y mejorarse más fácilmente, proporcionando la flexibilidad necesaria para establecer relaciones entre conjuntos de datos y conceptos de distintos dominios.
Solución	Se presenta un experimento preliminar sobre el uso de ontologías para la descripción de conjuntos de datos de biodiversidad, con un fuerte enfoque en la dinámica de las especies individuales, la diversidad de especies, las comunidades biológicas y los ecosistemas, basada en la recomendación europea INSPIRE, para modelar la ontología de BIOME para el dominio de la biodiversidad.
Conclusiones	Utilizar ontologías más simples junto con una plataforma (Dendro) para gestionar los conjuntos de datos desde su creación, y la descripción mínima es una opción en la preservación y compartición de los conjuntos de datos.
Contribución	Prototipo de plataforma de gestión de datos de investigación colaborativa, para probar su aplicabilidad en un escenario de gestión de datos y creación de metadatos.
Beneficios y limitaciones	Beneficios: Facilidad para combinar diferentes enfoques y desarrollo incremental de modelos a diferentes niveles de granularidad, utilizar elementos adicionales si así lo requieren otras normas internacionales o las prácticas de trabajo de sus comunidades. Limitaciones: Cobertura incompleta de indicadores espaciales y taxonómicos. El acceso, utilización, y la organización responsable de los recursos.
Tecnologías y Plataformas	Formatos: RDF, OWL, JSON Lenguaje: SPARQL Ontología: Dendro, BIOME, Dublin Core, FOAF, CERIF

TABLA LXVI.

DESCRIPCIÓN RESULTADOS ES35

Estudio	**ES35**
Problema	El modelado agrícola ha sufrido durante mucho tiempo la fragmentación en la implementación de un modelo. Se están desarrollando muchos modelos, hay mucha redundancia, los modelos a menudo están mal acoplados, la reutilización de los componentes de los modelos es poco frecuente, y es muy difícil aplicar modelos para generar soluciones reales para el sector agrícola.
Solución	Evaluar las tecnologías relevantes para su madurez, desarrollo esperado y potencial para beneficiar a la comunidad de modelado agrícola. Las tecnologías consideradas abarcan métodos para el desarrollo colaborativo y para involucrar a las partes interesadas y a los usuarios en el desarrollo de manera transdisciplinaria. La modelización general de los sistemas agrícolas debe adoptar y absorber rápidamente los datos más recientes y las tecnologías de la información y las comunicaciones.

Conclusiones	El modelado en base a NextGen proporciona estructuras y funciones específicas de dominio, herramienta significativamente que contribuya con el objetivo de ayudar a conseguir una producción sostenible, beneficios positivos en la modelización agrícola.
Contribución	Evaluación de los desarrollos recientes de las TIC a través de cinco casos de uso que han sido formulados para apoyar la visión de una comunidad de modelado en el ámbito agrícola.
Beneficios y Limitaciones	Beneficios: comprender mejor los sistemas biológicos como los modos de operación manejados a través de sus prácticas agrícolas, oportunidades de los datos y productos en desarrollo de aplicaciones móviles, servicios web para uso directo por parte de los agricultores y las ONG. Limitaciones: limitaciones del rendimiento de cada región, la falta de mecanismos de financiación sostenibles para la recopilación y conservación a largo plazo de clases importantes de datos. Los costos de desarrollo y mantenimiento son significativos.
Tecnologías y Plataformas	Herramienta: NextGen tools. Aplicaciones colaborativas: Agricultural Model Intercomparison and Improvement Project, BIOMA

3. Discusión.

Realizada la selección y análisis de los estudios, damos respuesta a las preguntas de investigación que se han formulado.

P1. - ¿QUÉ TIPO DE CONTRIBUCIONES Y BENEFICIOS SE HAN REALIZADO A TRAVÉS DE LA WEB SEMÁNTICA Y LINKED DATA APLICADAS A BIODIVERSIDAD?

Analizados los diferentes documentos y la propuesta de solución que abarque de manera concreta problemas de biodiversidad con la proyección asociada a la Web Semántica y Linked Data se destaca el desarrollo de plataformas, herramientas y portales de biodiversidad mencionados en ES07, ES09, ES10, ES13, ES33, que proveen una satisfactoria manera de solucionar los vacíos de conocimientos, información e incluso el mostrar conjuntos de colecciones de datos biológicos. Explicado en ES19, ES21, ES23 se menciona la operatividad para grupos de personas que realizan investigación, educación u otros profesionales que les interese la conservación de la naturaleza y su medio ambiente.

Los estudios revisados en ES07, ES08, ES19, ES22, ES23, destacan a la plataforma GBIF, como el principal ente de extracción, publicación y generación de datos para compartir en el ámbito de Biodiversidad, la red de GBIF reúne diferentes fuentes mediante el uso del estándar Darwin Core, que forma la base del índice de GBIF.org de

cientos de millones de registros de ocurrencia de especies. Los editores brindan acceso abierto a sus conjuntos de datos mediante la designación de licencias de Creative Commons [65] legibles por máquina, lo que permite a los científicos, investigadores y otros aplicar los datos en cientos de publicaciones y documentos de políticas revisados por pares cada año. Muchos de estos análisis tratados en ES02, ES05, ES06, ES10, ES16, ES35 cubren temas que van desde los impactos del cambio climático, la propagación de plagas invasoras y exóticas, hasta las prioridades para la conservación y las áreas protegidas, la seguridad alimentaria y la salud humana. Las principales contribuciones destacadas muestran los proyectos más trascendentes, enfocados en datos de biodiversidad, guía de transformación de datos, la inclusión del estándar más actual desarrollado para investigadores y editores de datos.

Como beneficios se destaca: el tratamiento de los datos en el campo de la Biodiversidad de tal forma que sean manipulables por la máquina y refuerce el criterio de análisis realizado por una persona; la operatividad que generan el tratar con estándares comunes, mejor vistos en ES03, ES05, ES34 manejados en distintos sistemas, mayor reconocimiento de datos vinculados entre conjuntos de colecciones biológicos producto de la transformación de Linked Data manifestado en ES13, que en repositorios dispersos antiguamente son tratados de manera tradicional de forma manual.

Adicionalmente se notan limitaciones existentes, como es el vacío de conocimiento y el aislamiento de grupos de investigación en diferentes silos de información que no hacen posibles la total integración de los datos detallado en ES22, ES23, por lo que no se puede determinar cuan dispersos están los conjuntos de información acerca de biodiversidad propuesto en ES26, se puede mencionar también como limitante la carencia de incentivos para aplicar Linked Data tratado en ES02, así como el tiempo y financiamiento significativo que requiere ser accesible a este enfoque.

P2. - ¿CUÁLES SON LAS TECNOLOGÍAS DE WEB SEMÁNTICA Y LINKED DATA ACTUALMENTE APLICADAS?

Con los documentos analizados ES09, ES25, ES32 se determinan diferentes tecnologías utilizadas, como resultado para tomar en cuenta al aplicar los principios de Linked Data, resaltamos las más referenciadas en la TABLA LXVII. El trabajo ES24

estudio de información histórica con tecnologías RDF, OWL, SPARQL, vocabularios RDFa, Dublín core y herramientas en el desarrollo como IBMDB2, Virtuoso, OWLIM.

ES25 Utiliza vocabularios DCAT, Data Cube conjuntamente con herramientas RESTful y Apache Jena Fuseki, ES26 es un trabajo que además de las tecnologías mencionadas añade la tecnología de XML, herramientas de desarrollo como CubeViz emplean vocabularios destacados como SKOS.

ES28 se mantiene con las tecnologías ya mencionadas, sus herramientas Tripify, D2R, Drupal constituyen un nuevo marco organizacional y operativo para el manejo de datos del gobierno de Tailandia. El estudio ES29 combina la ontología BFO para su caso específico, en cambio ES30 expone métricas entre la calidad de sistemas diferentes como BLEU, NIST, chrF, MEANT, TER que contribuyan con la eficiencia de la traducción automática y exponer el uso de tecnologías de Web Semántica, el vocabulario LOCWD se resalta en ES31 al igual que los recursos OER utilizados en su mayoría por universidades; ES32 utiliza los servicios RESTful con ontologías EFO y Gene Ontology.

TABLA LXVII.

RESPUESTA RSL – PREGUNTA 2.

Portales Principales	Tecnologías Web Semántica y Linked Data	Vocabularios y Formatos
Se pone mayor énfasis en relucir a la plataforma GBIF[15] que forma la base del índice de GBIF.org de cientos de millones de registros de ocurrencia de especies, sin dejar de lado plataformas como EOL[16] con formatos TSV, SpeciesLink[17] con estándares como KML, y utilizando Api REST detallados en Chavan [9], Robertson [17], Baskauf [47].	Se utilizan diferentes tecnologías entre las más mencionadas tenemos: XML, RDF, SPARQL, OWL; apoyadas en las distintas herramientas como Protege, GraphDB, Jena, Virtuoso, Open Refine.	Destacamos entre los más mencionados: Dublín Core, XML Schema, FOAF, Darwin Core, y más extensiones de vocabularios RDF (ABCD, EOL), KML, TTL, TSV.

Elaborado por el Autor.

4. Conclusiones.

- Uno de los problemas más importantes encontrados en aplicar Linked Data hacia la Biodiversidad es la integración de datos, por el desafío permanente de colecciones biológicas muy dispersas.

[15] Global Biodiversity Information Facility, portal: http://gbif.org
[16] Encyclopedia of Life, portal: http://www.eol.org
[17] Portal: http://splink.cria.org.br

- El beneficio de la fusión desarrollado por marcos de trabajo hacia la interoperabilidad con Linked Data tiene un gran potencial en el monitoreo de ecosistemas.
- La generación de estándares y la reutilización de vocabularios por parte de las diferentes organizaciones que desean compartir información, causan un efecto positivo, y utilizadas sobre biodiversidad han promovido un gran avance en la distribución de información.
- Existe una carencia de terminología idónea que cubra el área de Biodiversidad en su totalidad.
- Las principales formas de unirse al enfoque de Linked Data tanto en la Nube como en Datasets particulares, poseen en su mayoría la estructura general de Modelar los datos en un vocabulario, contar con un almacenamiento triple-store y presentarla mediante una aplicación semántica o por una herramienta de visualización, la gran escogida Pubby utilizada por la DBPedia a nivel Mundial.

ANEXO 2: CRITERIOS DE SELECCIÓN.

TABLA LXVIII.

CRITERIOS DE SELECCIÓN REVISION SISTEMATICA DE LITERATURA.

Cadena de Búsqueda	Total	Incluido	No Incluido	Titulo	Año	Publicado en
(((((((((Semantic Web) OR Linked Data) OR Linked Open Data) AND applied*) OR model*) OR experiments*) OR conservation*) OR analysis*) AND Biodiversity).	59169	26	59143	EScience through the integration of data and models: A biodiversity scenario [49].	2015	Proceedings - 11th IEEE International Conference on eScience, eScience 2015
				The data paper: a mechanism to incentivize data publishing in biodiversity science [9].	2011	BMC bioinformatics
				Darwin core: An evolving community-developed biodiversity data standard [10].	2012	PLoS ONE
				Trends in access of plant biodiversity data revealed by Google Analytics [50].	2014	Biodiversity Data Journal
				Darwin-SW: Darwin Core-based terms for expressing biodiversity data as RDF [47].	2016	Semantic Web
				Elevating Natural History Museums' Cultural Collections to the Linked Data Cloud [51].	2013	CEUR Workshop Proceedings
				Biodiversity data should be published, cited, and peer reviewed [52].	2013	Trends in Ecology and Evolution.
				Biodiversity Observations Miner: A web application to unlock primary biodiversity data from published literature [53].	2018	Biodiversity Data Journal
				Biodiversity Databases and Tools [54].	2018	Reference Module in Life Sciences
				Global biodiversity monitoring: From data sources to Essential Biodiversity Variables [55].	2017	Biological Conservation
				The next generation of site-based long-term ecological monitoring: Linking essential biodiversity variables and ecosystem integrity [56].	2018	Science of the Total Environment
				Biodiversity Models: What If Unsaturation Is the Rule? [57].	2017	Trends in Ecology and Evolution.
				Transformando Datos de Biodiversidad en Linked Data [8].	2017	

	Hacia la Interoperabilidad Semántica para el Manejo Inteligente y Sostenible de Territorios de Alta Biodiversidad usando SmartLand-LD [11].	2018	RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação
	High Performance Methods for Linked Open Data Connectivity Analytics [58].	2018	
	Integration of Web APIs and Linked Data Using SPARQL Micro-Services — Application to Biodiversity Use Cases [59].	2018	Biodiversity Information Science and Standards.
	Theophrastus: On demand and real-time automatic annotation and exploration of (web) documents using open linked data [13].	2014	Journal of Web Semantics
	Terminology supported archiving and publication of environmental science data in PANGAEA [60].	2017	Journal of Biotechnology
	A Model of Provenance Applied to Biodiversity Datasets [34].	2016	Proceedings - 25th IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE 2016
	Data mining and linked open data – New perspectives for data analysis in environmental research [61].	2015	Ecological Modelling
	Understanding the link between population dynamics and biodiversity conservation through remote sensing and gridded population data integration [32].	2015	
	Improving Biodiversity Data Retrieval through Semantic Search and Ontologies [30]	2014	Proceedings - 2014 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology - Workshops, WI-IAT 2014
	A Gazetteer for Biodiversity Data as a Linked Open Data Solution [62].	2014	Proceedings of the Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE.

				TraitBank: Practical semantics for organism attribute data [63].	2016	Semantic Web
				Beyond INSPIRE: An Ontology for Biodiversity Metadata Records [35].	2014	On the Move to Meaningful Internet Systems: OTM 2014 Workshops.
				Towards a new generation of agricultural system data, models and knowledge products: Information and communication technology [64].	2017	Agricultural Systems
((“Semantic Web“) or (“Linked Data”)) and (“technologies”) and (“models”)*.	23890	9	23881	Semantic technologies for historical research: A survey [39].	2015	Jane's Defence Industry
				Linked Open Vocabularies (LOV): A gateway to reusable semantic vocabularies on the Web [46].	2017	Semantic Web
				Publishing a scorecard for evaluating the use of open-access journals using linked data technologies [40].	2015	IEEE CITS 2015 - 2015 International Conference on Computer, Information and Telecommunication Systems
				Supporting seamless learning with semantic technologies and situation awareness [41]	2014	Proceedings - 2014 International Conference on Intelligent Networking and Collaborative Systems, IEEE INCoS 2014
				A linked data approach to planning collaboration amongst local governments in Thailand [42]	2018	Proceeding of 2017 2nd International Conference on Information Technology, INCIT 2017
				Towards a health observatory conceptual model based on the semantic web [43]	2018	Procedia Computer Science
				Machine Translation using Semantic Web Technologies: A Survey [44].	2018	Journal of Web Semantics
				Consuming and producing linked open data: the case of OpenCourseWare [45].	2014	Program
				The EBI RDF platform: linked open data for the life sciences [38].	2014	Bioinformatics

Elaborado por el Autor.

ANEXO 3: RDF STORES

Teniendo en cuenta los diferentes sistemas que proporcionan la posibilidad de almacenar tripletas RDF, de manera que conste Sujeto - Predicado – Objeto, la Figura 37 amplia la gran variedad de alternativas.

19 systems in ranking, December 2018

Rank			DBMS	Database Model	Score		
Dec 2018	Nov 2018	Dec 2017			Dec 2018	Nov 2018	Dec 2017
1.	1.	1.	MarkLogic	Multi-model	14.28	+0.81	+3.13
2.	↑ 3.	↑ 3.	Virtuoso	Multi-model	2.48	+0.11	+0.67
3.	↓ 2.	↓ 2.	Apache Jena - TDB	RDF store	2.29	-0.17	0.00
4.	4.		Amazon Neptune	Multi-model	1.23	+0.17	
5.	↑ 6.	↓ 4.	AllegroGraph	Multi-model	0.91	+0.24	+0.32
6.	↓ 5.	↓ 5.	GraphDB	Multi-model	0.72	+0.03	+0.14
7.	7.	↓ 6.	Stardog	Multi-model	0.53	+0.00	+0.02
8.	8.	8.	Blazegraph	Multi-model	0.39	0.00	+0.18
9.	9.	↓ 7.	Algebraix	RDF store	0.39	0.00	+0.06
10.	10.	↑ 12.	RDF4J	RDF store	0.27	-0.01	+0.14
11.	↑ 12.	↓ 9.	Redland	RDF store	0.27	+0.01	+0.08
12.	↓ 11.	↑ 14.	4store	RDF store	0.26	-0.01	+0.17
13.	↑ 14.	↑ 16.	CubicWeb	RDF store	0.12	+0.02	+0.06
14.	↓ 13.	↓ 10.	RedStore	RDF store	0.11	-0.02	-0.04
15.	15.	↓ 13.	Mulgara	RDF store	0.08	0.00	-0.05
16.	16.	↓ 11.	BrightstarDB	RDF store	0.07	-0.01	-0.06
17.	↑ 18.	17.	Dydra	RDF store	0.05	+0.00	+0.01
18.	↓ 17.	↓ 15.	Strabon	RDF store	0.05	0.00	-0.03
19.	19.	↓ 18.	SparkleDB	RDF store	0.03	0.00	+0.03

Figura 37.Representación gráfica - vista RDF Stores [66].

Se mencionan en este proyecto para mayor criterio del lector otras alternativas de almacenamiento con más detalle.

MarkLogic.

El enfoque flexible y de múltiples modelos de MarkLogic le permite traer datos desde cualquier lugar (bases de datos relacionales, mainframes, servidores de archivos, Hadoop) o cualquier otra fuente. Una base de datos flexible y de múltiples modelos con almacenamiento nativo para archivos binarios JSON, XML, RDF, geoespaciales y grandes (por ejemplo, archivos PDF, imágenes, videos). Con este enfoque, es fácil obtener todos sus datos y hacer cambios más adelante. Es flexible Es rápido. Es iterativo. Obtenga los beneficios combinados de una tienda de documentos y una tienda

triple RDF, donde se almacenan los datos semánticos. La semántica es ideal para almacenar relaciones, lo cual es útil para mejorar la integración de datos, administrar metadatos y crear aplicaciones con datos altamente conectados, visto en Marklogic [67].

Amazon Neptune.

Es un servicio de base de datos de gráficos rápido, fiable y completamente administrado que le permite crear y ejecutar fácilmente aplicaciones que funcionen con conjuntos de datos fuertemente conectados. Su núcleo es un motor de bases de datos de gráficos de alto desempeño diseñado expresamente y optimizado para almacenar miles de millones de relaciones y consultar gráficos con una latencia de milisegundos. Amazon Neptune admite los modelos de gráficos populares Property Graph y RDF del W3C, así como sus lenguajes de consulta asociados Apache TinkerPop Gremlin y SPARQL, lo que le permite crear fácilmente consultas que navegan de manera eficiente por conjuntos de datos altamente conectados. Ofrece alta disponibilidad, con réplicas de lectura, recuperación a un momento dado, copias de seguridad continuas en Amazon S3 y replicación entre zonas de disponibilidad, detallado en Amazon [68].

Stardog.

Puede consultar, buscar y analizar datos, donde sea que esté, utilizando la tecnología escalable de Knowledge Graph. Dada la variedad de datos completa y cada velocidad de datos diferente unifique fuentes de datos dispares y externas y convierta los datos en información procesable. A través de Stardog studio le permite escribir consultas para el Knowledge Graph e incluye un editor de consultas SPARQL y una vista de resultados desplazable. Exportar los resultados de la consulta a su sistema de archivos ampliado en Stardog [69]–[71].

RDF4J.

Conocido como Sesame, es un potente marco de Java para procesar y manejar datos RDF. Esto incluye la creación, el análisis, el almacenamiento escalable, el razonamiento y la consulta con RDF y los datos vinculados. Ofrece una API fácil de usar que se puede conectar a todas las soluciones de base de datos RDF líderes. Le permite conectarse con puntos finales SPARQL y crear aplicaciones que aprovechan el poder de los datos vinculados y la Web Semántica. El marco ofrece una gran variedad de herramientas para que los desarrolladores aprovechen el poder de RDF y los estándares relacionados. es compatible con todos los formatos de archivo RDF convencionales, incluidos RDF / XML, Turtle, N-Triples, N-Quads, JSON-LD, TriG y TriX. ofrece dos bases de datos RDF listas para usar (la tienda en memoria y la tienda nativa) y, ofrece muchas soluciones de almacenamiento de terceros, Eclipse [72].

ANEXO 4: CUADRO METODOLOGÍAS PARA ONTOLOGÍAS.

TABLA LXIX.

METODOLOGÍAS PARA LA CREACION DE ONTOLOGÍAS PARTE 1.

Criterio	Cyc	Kactus	Grüninger y Fox	Methontology
1° Etapa	Codificación manual del conocimiento	Especificación de la aplicación, descripción del dominio de la ontología	Definición de escenarios, ver las aplicaciones de la ontología	Especificación para determinar alcance, objetivos, propósito, usuarios finales de la ontología.
2° Etapa	Utilización de Herramientas de Software para la codificación.	Diseño preliminar en base a categorías top level ontológicas	Determinar el ámbito de la ontología	Conceptualización en la creación de glosario de términos, taxonomía.
3° Etapa	Adquisición de nuevo conocimiento con herramientas de PLN o aprendizaje computacional.	Refinamiento y estructuración de la ontología.	Especificar la terminología	Formalización a través de un modelo conceptual en un modelo formal o semicomputable
4° Etapa			Formalización de las Interrogantes	Implementación de la ontología mediante su codificación en un lenguaje formal
5° Etapa			Especificación de axiomas formales	Actualización y corrección de la ontología.
6° Etapa			Verificación de la ontología	

Elaborado por el Autor

TABLA LXX.

METODOLOGÍAS PARA LA CREACION DE ONTOLOGÍAS PARTE 2.

Criterio	ONTOLOGY DEVELOPMENT 101	Ding y Foo.	On-to-Knowledge	Uschold y King
1° Etapa	Determinar el dominio y ámbito de la ontología.	Datos fuente: Vocabularios controlados, corpus de sentencias, extracción de texto libre, preguntas a usuarios.	Estudio de factibilidad, determinar factores que influyen en el éxito o fracaso.	Determinar el propósito de la Ontología
2° Etapa	Determinar la intención de uso de la ontología.	Métodos para la extracción de conceptos: las diferentes técnicas empleadas para la extracción de información (análisis sintáctico, procesamiento del lenguaje natural, implicación humana, etc.).	Especificación de requerimientos, desarrollo de una taxonomía informal.	Capturar los conceptos y relaciones entre términos.
3° Etapa	Reutilizar ontologías o vocabularios controlados existentes.	Métodos para la extracción de relaciones: puede ser de forma automática o basándose en algoritmos que en ocasiones se aplican de forma manual.	Refinamiento de la taxonomía acorde a los requerimientos y objetivos	Codificación de la ontología
4° Etapa	Enumerar los términos importantes del dominio.	Reutilización de ontologías: puede ser habitual utilizar otros instrumentos terminológicos.	Evaluación de la ontología , capacidad de respuesta y funcionamiento	
5° Etapa	Definir jerarquía de clases.	Representación de la ontología, que va desde la estructura jerárquica, pasando por la lógica de descripción hasta los grafos conceptuales y el XML.	Realizar cambios identificados y actualizar de acuerdo a las reglas especificadas	
6° Etapa	Crear las instancias.			

Elaborado por el Autor

ANEXO 5. DOCUMENTACIÓN BLOG JARDÍN BOTÁNICO

En el siguiente anexo se muestra los documentos que se encuentran alojados en el blog principal del jardín botánico “Reinaldo Espinosa”.

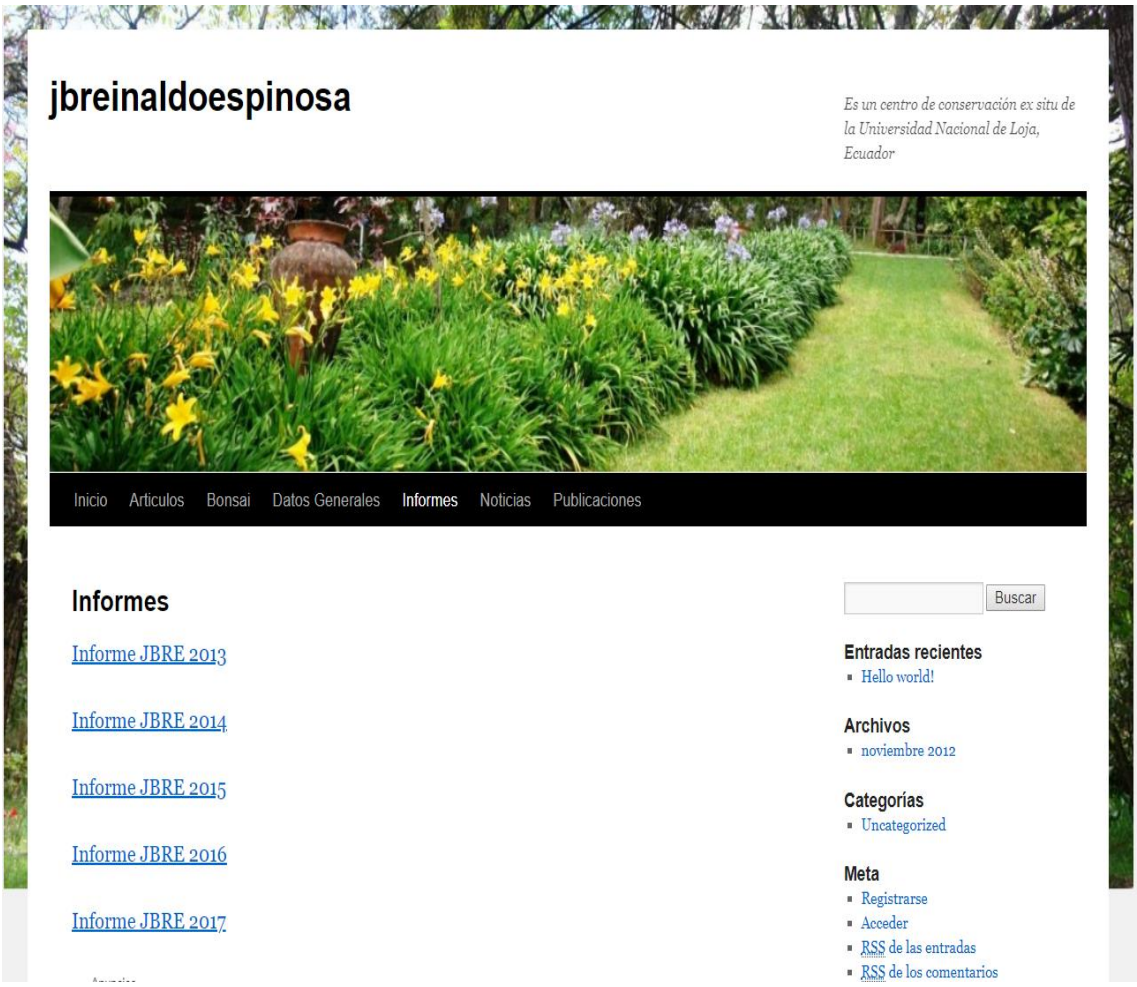


Figura 38. Representación gráfica Página de Informes JBRE.

El informe JBRE 2017 es un resumen de las actividades que se han desarrollado a lo largo del año 2017 y un porcentaje mensual de las visitas que se han atendido; así mismo como los porcentajes anuales y servicios puestos ante la comunidad universitaria, la ciudadanía en general (vista de informes en la Figura 38).

Figura 39.Representación gráfica de la vista de las Publicaciones JBRE.

UNIVERSIDAD NACIONAL DE LOJA
FACULTAD AGROPECUARIA Y DE RECURSOS NATURALES RENOVABLES
Jardín Botánico "Reinaldo Espinosa"
Casilla 11-01-249 o letra "B" Loja, Ecuador
Telefono 3027930
E-mail: herbario@unl.edu.ec
www.jbreinaldoespinosa.wordpress.com

INFORME DE LA GESTIÓN DEL JARDÍN BOTÁNICO "REINALDO ESPINOSA", 2017

Marcelo Gutiérrez Riofrío
TÉCNICO DEL JBRE

Zhofre Aguirre Mendoza
DIRECTOR EL JBRE

Loja, Ecuador
2017

Figura 40.Representación gráfica de la Portada Documento Informe JBRE 2017.

En la Figura 41 se puede apreciar la portada del documento floranativa.pdf, con la descripción del sendero de la flora nativa y la distribución breve de las diferentes secciones que conforman al jardín.

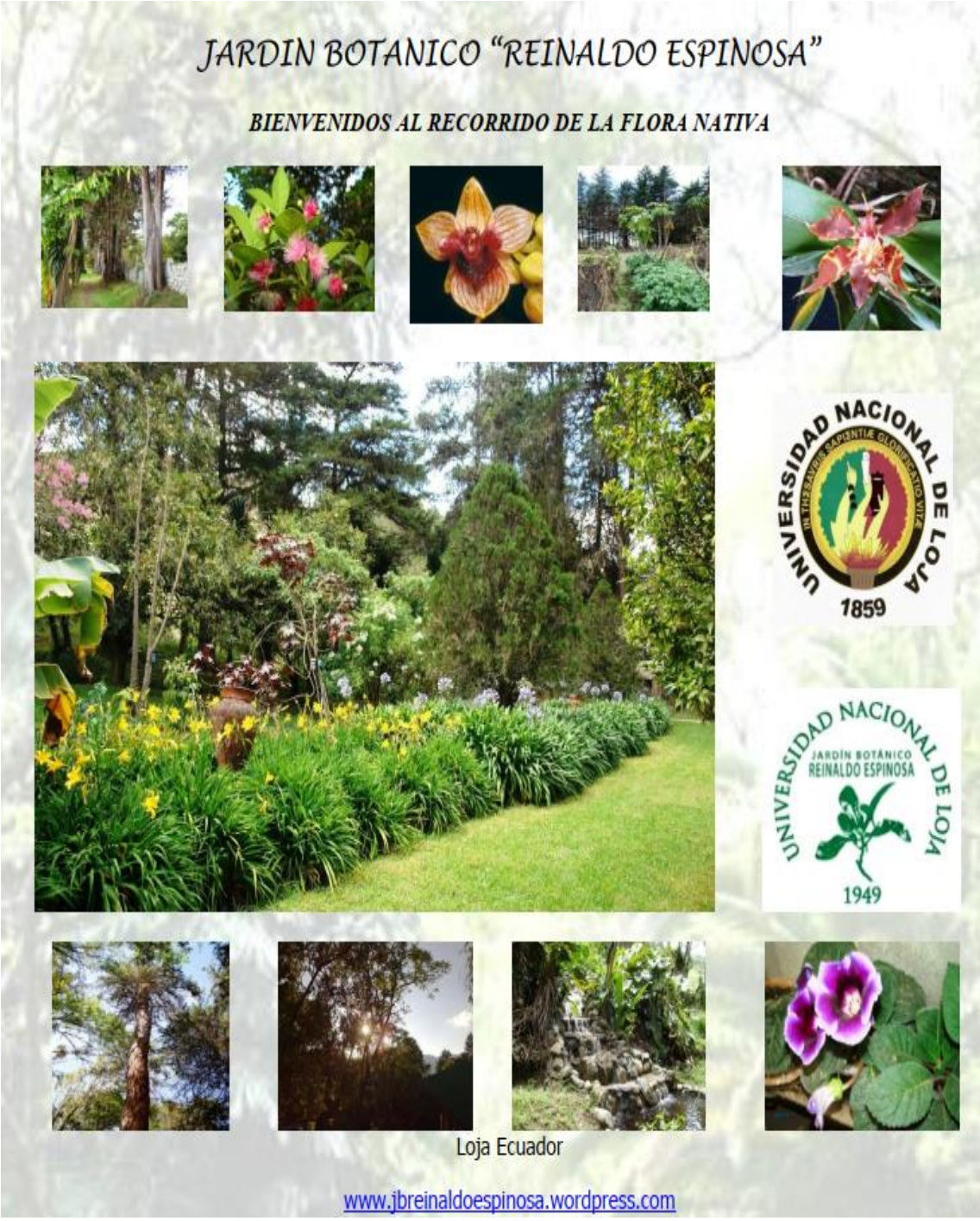


Figura 41.Representación gráfica del documento Flora Nativa JBRE.

En resumen, exponemos las diferentes publicaciones encontradas en el blog principal del jardín botánico y los formatos que en su mayoría son de solo lectura bajo el formato PDF.

TABLA LXXI.

DATOS PUBLICADOS POR EL BLOG PRINCIPAL DEL JARDÍN BOTÁNICO.

Nombre	**Sección**	**Archivo**
Artículos	https://jbreinaldoespinosa.wordpress.com/articulos/	Escenarios-para-la-enseñanza-de-la-Biodiversidad.pdf
Bonsai	https://jbreinaldoespinosa.wordpress.com/bonsai/	seccion-bonsai4.pdf
Informes (del 2013-2017)	https://jbreinaldoespinosa.wordpress.com/informes/	informe-jbre-2017.pdf
		informe-jbre-2016.pdf
		informe-jbre-2015.pdf
		informe-jbre-2014.pdf
		informe-jbre-2013.pdf
Publicaciones	https://jbreinaldoespinosa.wordpress.com/publicaciones-3/	
	Álbum de Orquídeas	Orquideaalbum.pdf
	“JBRE” Centro de conservación ex-situ de Loja.	Jardinbotanico.pdf
	Recorre la flora nativa del “JBRE”.	Floranativa.pdf
	Catalogo JBRE	catalogo-jbre.pdf
	Guía de Plantas útiles	guía-de-plantas-utiles1.pdf
	Pastos	Pastos.pdf

Elaborado por el Autor.

ANEXO 6. VOCABULARIO – ONTOLOGÍA LDJBRE


```
<?xml version="1.0"?>
<rdf:RDF xmlns="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#"
    xml:base="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre"
    xmlns:cc="http://creativecommons.org/ns#"
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:ns="http://www.w3.org/2006/vcard/ns#"
    xmlns:dbo="http://dbpedia.org/ontology/"
    xmlns:dwc="http://rs.tdwg.org/dwc/terms/"
    xmlns:org="http://www.w3.org/ns/org#"
    xmlns:owl="http://www.w3.org/2002/07/owl#"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:xml="http://www.w3.org/XML/1998/namespace"
    xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
    xmlns:foaf="http://xmlns.com/foaf/0.1/"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:skos="http://www.w3.org/2004/02/skos/core#"
    xmlns:terms="http://purl.org/dc/terms/"
    xmlns:ldjbre="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#">

   <owl:Ontology rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre">
      <owl:versionIRI
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre/1.0"/>
      <terms:creator rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Juan
Maldonado</terms:creator>
      <terms:license
rdf:datatype="http://www.w3.org/2001/XMLSchema#anyURI">https://creativecommons.
org/licenses/by-nc-sa/4.0/</terms:license>
      <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">Ontologia Jardin Botanico reinaldo Espinoza</rdfs:comment>
      <rdfs:label
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Vocabulario
ldjbre</rdfs:label>
   </owl:Ontology>
```

```
<!--
///////////////////////////////////////////////////////////////////////////////////////
//
// Object Properties
//
///////////////////////////////////////////////////////////////////////////////////////
 -->

<!-- http://dbpedia.org/ontology/location -->

<owl:ObjectProperty rdf:about="http://dbpedia.org/ontology/location">
    <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
    <rdfs:range rdf:resource="http://dbpedia.org/ontology/Place"/>
</owl:ObjectProperty>

<!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#belong_to -->

<owl:ObjectProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#belong_to">
    <owl:inverseOf
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#have"/>
    <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
    <rdfs:range
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Biological_collectio
n"/>

    <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">pertecece
a</rdfs:comment>
</owl:ObjectProperty>
```

```
<!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#contains -->

<owl:ObjectProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#contains">
    <owl:inverseOf
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#stored_in"/>
    <rdf:type
rdf:resource="http://www.w3.org/2002/07/owl#InverseFunctionalProperty"/>
    <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
    <rdfs:range
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Biological_collectio
n"/>
    <rdfs:comment>identifica el objeto que contiene los datos</rdfs:comment>
</owl:ObjectProperty>

<!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#corresponding_to -->

<owl:ObjectProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#corresponding_to">
    <owl:inverseOf
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#information_of"/>
    <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Bioclimate_data"/>
    <rdfs:range
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
    <rdfs:comment>relacion correspondiente al jardin botanico</rdfs:comment>
</owl:ObjectProperty>
<!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#has_a -->

<owl:ObjectProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#has_a">
    <owl:inverseOf rdf:resource="http://purl.org/dc/elements/1.1/isPartOf"/>
    <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University"/>
```

```
        <rdfs:range
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
    </owl:ObjectProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#has_category -->

    <owl:ObjectProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#has_category">
        <owl:inverseOf
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#isCategoryOf"/>
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#person"/>
        <rdfs:range
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#labor_role"/>
        <rdfs:comment>identifica la categoria laboral de la persona</rdfs:comment>
    </owl:ObjectProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#has_responsible -->

    <owl:ObjectProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#has_responsible">
        <owl:inverseOf
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#ischarge_of"/>
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
        <rdfs:range
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#person"/>
    </owl:ObjectProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#have -->

    <owl:ObjectProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#have">
```

```
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Biological_collectio
n"/>
        <rdfs:range
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
        <rdfs:comment>identifica la especie que tiene la coleccion</rdfs:comment>
    </owl:ObjectProperty>

        <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#information_of -->

    <owl:ObjectProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#information_of">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
        <rdfs:range
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Bioclimate_data"/>
        <rdfs:comment>datos bioclimaticos respecto al jardin</rdfs:comment>
    </owl:ObjectProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#isCategoryOf -->

    <owl:ObjectProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#isCategoryOf">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#labor_role"/>
        <rdfs:range
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#person"/>
    </owl:ObjectProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#ischarge_of -->

    <owl:ObjectProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#ischarge_of">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#person"/>
```

```
        <rdfs:range
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
        <rdfs:comment>relacion de la persona con el jardin</rdfs:comment>
    </owl:ObjectProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#stored_in -->

    <owl:ObjectProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#stored_in">
        <rdf:type
rdf:resource="http://www.w3.org/2002/07/owl#InverseFunctionalProperty"/>
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Biological_collectio
n"/>
        <rdfs:range
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
        <rdfs:comment>identifica el objeto donde se almacena los datos</rdfs:comment>
    </owl:ObjectProperty>

    <!-- http://purl.org/dc/elements/1.1/isPartOf -->

    <owl:ObjectProperty rdf:about="http://purl.org/dc/elements/1.1/isPartOf">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
        <rdfs:range
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University"/>
        <rdfs:comment>identifica a que universidad es parte</rdfs:comment>
    </owl:ObjectProperty>

    <!--
    ///////////////////////////////////////////////////////////////////////////////////////
    //
    // Data properties
    //
    ///////////////////////////////////////////////////////////////////////////////////////
```

```
    -->

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#bioclimate -->

    <owl:DatatypeProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#bioclimate">
        <rdfs:subPropertyOf
rdf:resource="http://www.w3.org/2002/07/owl#topDataProperty"/>
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Bioclimate_data"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">bioclima</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#category_use -->

    <owl:DatatypeProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#category_use">
        <rdfs:subPropertyOf
rdf:resource="http://www.w3.org/2002/07/owl#topDataProperty"/>
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">categoria del especie
segun su uso comun</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#height_data -->

    <owl:DatatypeProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#height_data">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Bioclimate_data"/>
```

```
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">altura
del Jardin botanico</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#id_bioclimate -->

    <owl:DatatypeProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#id_bioclimate">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Bioclimate_data"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">identificador de datos bioclimaticos</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#id_collection -->

    <owl:DatatypeProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#id_collection">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Biological_collectio
n"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">Identificador de coleccion</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#id_person -->

    <owl:DatatypeProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#id_person">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#person"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
```

```
        <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">Identificador de la persona</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#idbotanic -->

    <owl:DatatypeProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#idbotanic">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">Identificador del jardin botanico</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#idspecies -->

    <owl:DatatypeProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#idspecies">
        <rdfs:subPropertyOf
rdf:resource="http://www.w3.org/2002/07/owl#topDataProperty"/>
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">identificador de la especie
en la coleccion de datos</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#section -->
    <owl:DatatypeProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#section">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
```

```
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">seccion a la que pertenece
segun el jardin</rdfs:comment>
    </owl:DatatypeProperty>

        <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#surface -->

    <owl:DatatypeProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#surface">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Bioclimate_data"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Superficie</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#typeofclimate -->

    <owl:DatatypeProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#typeofclimate">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Bioclimate_data"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">tipo
de clima</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#ubicacion -->

    <owl:DatatypeProperty
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#ubicacion">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
```

```
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">ubicacion del Jardin
Botanico</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://purl.org/dc/elements/1.1/description -->

    <owl:DatatypeProperty rdf:about="http://purl.org/dc/elements/1.1/description">
        <rdfs:subPropertyOf
rdf:resource="http://www.w3.org/2002/07/owl#topDataProperty"/>
        <rdfs:domain>
            <owl:Class>
                <owl:unionOf rdf:parseType="Collection">
                    <rdf:Description
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Biological_collection"/>
                    <rdf:Description
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#labor_role"/>
                </owl:unionOf>
            </owl:Class>
        </rdfs:domain>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">descripcion del
objeto</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://rs.tdwg.org/dwc/terms/Family -->

    <owl:DatatypeProperty rdf:about="http://rs.tdwg.org/dwc/terms/Family">
        <rdfs:subPropertyOf
rdf:resource="http://www.w3.org/2002/07/owl#topDataProperty"/>
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment>familia a la que corresponde la especie</rdfs:comment>
```

```
    </owl:DatatypeProperty>

    <!-- http://rs.tdwg.org/dwc/terms/scientificName -->

    <owl:DatatypeProperty rdf:about="http://rs.tdwg.org/dwc/terms/scientificName">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Nombre
cientifico</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://rs.tdwg.org/dwc/terms/vernacularName -->

    <owl:DatatypeProperty rdf:about="http://rs.tdwg.org/dwc/terms/vernacularName">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Nombre
comun</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://www.w3.org/2006/vcard/ns#category -->

    <owl:DatatypeProperty rdf:about="http://www.w3.org/2006/vcard/ns#category">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">categoria de la
universidad</rdfs:comment>
    </owl:DatatypeProperty>
```

```
    <!-- http://www.w3.org/2006/vcard/ns#mbox -->

    <owl:DatatypeProperty rdf:about="http://www.w3.org/2006/vcard/ns#mbox">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">correo
de la entidad</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://www.w3.org/2006/vcard/ns#street-addres -->

    <owl:DatatypeProperty rdf:about="http://www.w3.org/2006/vcard/ns#street-addres">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">direccion de la
universidad</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://www.w3.org/ns/org#identifier -->

    <owl:DatatypeProperty rdf:about="http://www.w3.org/ns/org#identifier">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Identificador de la
organizacion</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://www.w3.org/ns/org#location -->

    <owl:DatatypeProperty rdf:about="http://www.w3.org/ns/org#location">
```

```
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
    </owl:DatatypeProperty>

    <!-- http://xmlns.com/foaf/0.1/familyName -->

    <owl:DatatypeProperty rdf:about="http://xmlns.com/foaf/0.1/familyName">
        <rdfs:domain
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#person"/>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">identifica los apellidos de
la persona</rdfs:comment>
    </owl:DatatypeProperty>

    <!-- http://xmlns.com/foaf/0.1/homepage -->
    <owl:DatatypeProperty rdf:about="http://xmlns.com/foaf/0.1/homepage">
        <rdfs:subPropertyOf
rdf:resource="http://www.w3.org/2002/07/owl#topDataProperty"/>
        <rdfs:domain>
            <owl:Class>
                <owl:unionOf rdf:parseType="Collection">
                    <rdf:Description
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
                    <rdf:Description
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University"/>
                </owl:unionOf>
            </owl:Class>
        </rdfs:domain>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">pagina web de
inicio</rdfs:comment>
    </owl:DatatypeProperty>
```

```
    <!-- http://xmlns.com/foaf/0.1/name -->
    <owl:DatatypeProperty rdf:about="http://xmlns.com/foaf/0.1/name">
        <rdfs:domain>
            <owl:Class>
                <owl:unionOf rdf:parseType="Collection">
                    <rdf:Description
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
                    <rdf:Description
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University"/>
                </owl:unionOf>
            </owl:Class>
        </rdfs:domain>
        <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <rdfs:comment
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">nombre del
objeto</rdfs:comment>
    </owl:DatatypeProperty>

<!--
    ///////////////////////////////////////////////////////////////////////////////////////
    //
    // Classes
    //
    ///////////////////////////////////////////////////////////////////////////////////////
-->
    <!-- http://dbpedia.org/ontology/Place -->

    <owl:Class rdf:about="http://dbpedia.org/ontology/Place">
        <owl:disjointWith
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Bioclimate_data"/>
        <owl:disjointWith
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
        <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Lugar
de Ubicacion</rdfs:comment>
```

```
    </owl:Class>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Bioclimate_data -->

    <owl:Class
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Bioclimate_data">
        <owl:disjointWith
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#labor_role"/>
        <owl:disjointWith
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#person"/>
        <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">Clase de los datos bioclimaticos del jardin Botanico</rdfs:comment>
    </owl:Class>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Biological_collection -->
    <owl:Class
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Biological_collection">
        <owl:disjointWith
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University"/>
        <owl:disjointWith
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#labor_role"/>
        <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">Coleccion de los datos</rdfs:comment>
    </owl:Class>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden -->

    <owl:Class
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden">
        <owl:disjointWith
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#labor_role"/>
        <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">Clase que representa al Jardin Botanico</rdfs:comment>
    </owl:Class>
```

```
  <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University -->
   <owl:Class
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University">
      <owl:disjointWith
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#person"/>
      <owl:disjointWith
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
      <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">Institucion a la que es parte el Jardin Botanico</rdfs:comment>
   </owl:Class>
   <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#labor_role -->
   <owl:Class
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#labor_role">
      <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">Clase de categoria laboral</rdfs:comment>
   </owl:Class>

   <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#person -->

   <owl:Class rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#person">
      <owl:equivalentClass rdf:resource="http://xmlns.com/foaf/0.1/person"/>
      <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">Persona encargada del Jardin</rdfs:comment>
   </owl:Class>

   <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species -->
   <owl:Class
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species">
      <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">Elementos encontrados en la coleccion de datos</rdfs:comment>
   </owl:Class>
      <!-- http://xmlns.com/foaf/0.1/person -->

   <owl:Class rdf:about="http://xmlns.com/foaf/0.1/person"/>
```

```
<!--
    ///////////////////////////////////////////////////////////////////////////////////
    //
    // Individuals
    //
    ///////////////////////////////////////////////////////////////////////////////////
     -->

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#BREv1 -->
    <owl:NamedIndividual
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#BREv1">
        <rdf:type
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_Garden"/>
        <contains
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#datasetv1"/>
        <information_of
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Biov1"/>
        <dc:isPartOf
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#UNL"/>
        <idbotanic
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">JBREv1</idbotanic>
        <org:location rdf:datatype="http://www.w3.org/2001/XMLSchema#string">E682-
Loja</org:location>
        <foaf:homepage rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">www.jbreinaldoespinosa.wordpress.com</foaf:homepage>
        <foaf:name rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Reinaldo
Espinosa</foaf:name>
    </owl:NamedIndividual>
    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Biov1 -->
    <owl:NamedIndividual
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Biov1">
        <rdf:type
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Bioclimate_data"/>
        <corresponding_to
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#BREv1"/>
```

```
        <bioclimate rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Sub-
húmedo templado</bioclimate>
        <height_data rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Sobre el
nivel del mar 2140 m</height_data>
        <id_bioclimate
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Bioclimate1</id_bioclimate
>
        <surface rdf:datatype="http://www.w3.org/2001/XMLSchema#string">8
Hectareas</surface>
        <typeofclimate
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Templado lluvioso,
mesotérmico, frío e isotermál.</typeofclimate>
        <rdfs:comment rdf:datatype="http://www.w3.org/2000/01/rdf-
schema#Literal">Datos Bioclimaticos jardin Botanico Reinaldo
Espinosa</rdfs:comment>
    </owl:NamedIndividual>
    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Esp1 -->
    <owl:NamedIndividual
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Esp1">
        <rdf:type
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
        <category_use
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Cereales</category_use>
        <idspecies
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Esp1</idspecies>
        <section rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Plantas
Andinas</section>
        <dwc:Family
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Poaceae</dwc:Family>
        <dwc:scientificName
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Oryza sativa
L.</dwc:scientificName>
        <dwc:vernacularName
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Arroz</dwc:vernacularNa
me>
```

```
    </owl:NamedIndividual>
    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Esp10 -->
    <owl:NamedIndividual
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Esp10">
        <rdf:type
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
        <category_use
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Cereales</category_use>
        <idspecies
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Esp10</idspecies>
        <section rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Plantas
Andinas</section>
        <dwc:Family
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Poaceae</dwc:Family>
        <dwc:scientificName
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Panicum miliaceum
L.</dwc:scientificName>
        <dwc:vernacularName
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Mijo</dwc:vernacularNam
e>
    </owl:NamedIndividual>
    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Esp100 -->
    <owl:NamedIndividual
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Esp100">
        <rdf:type
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
        <category_use
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Pastos</category_use>
        <idspecies
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Esp100</idspecies>
        <section rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Medicinales
y Ornamentales</section>
        <dwc:Family
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Poaceae</dwc:Family>
```

```
        <dwc:scientificName
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Avena fatua
L.</dwc:scientificName>
        <dwc:vernacularName
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Avena
forrajera</dwc:vernacularName>
    </owl:NamedIndividual>

    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Esp101 -->

    <owl:NamedIndividual
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Esp101">
        <rdf:type
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
        <category_use
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Pastos</category_use>
        <idspecies
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Esp101</idspecies>
        <section rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Medicinales
y Ornamentales</section>
        <dwc:Family
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Poaceae</dwc:Family>
        <dwc:scientificName
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Urochloa brizantha
(Hochst. ex A.Rich) R. Webster</dwc:scientificName>
        <dwc:vernacularName
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Braquiaria</dwc:vernacula
rName>
    </owl:NamedIndividual>
    <!-- http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Esp102 -->
    <owl:NamedIndividual
rdf:about="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Esp102">
        <rdf:type
rdf:resource="http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species"/>
```

```
        <category_use
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Pastos</category_use>
        <idspecies
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Esp102</idspecies>
        <section rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Medicinales
y Ornamentales</section>
        <dwc:Family
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Poaceae</dwc:Family>
        <dwc:scientificName
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Bromus catharticus M.
Vahl.</dwc:scientificName>
        <dwc:vernacularName
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Cebadilla</dwc:vernacular
Name>
    </owl:NamedIndividual>
```

ANEXO 7. DESCRIPCIÓN LDJBRE

La realización de la ontología LDJBRE, otorga la identificación de los recursos que se modelan alrededor del entorno asociado al jardín botánico “Reinaldo Espinosa”, de tal manera que describimos las diferentes clases, propiedades y relaciones que están contenidas en el vocabulario.

Clases:

rdf:Class	**ldjbre:Bioclimate_data**
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Bioclimate_data
rdf:subClassOf	owl:Thing
rdfs:comment	Clase de los datos bioclimaticos del jardin Botanico

rdf:Class	**ldjbre:Biological_collection**
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Biological_collection
rdf:subClassOf	owl:Thing
rdfs:comment	Colección de los datos

rdf:Class	**ldjbre:Botanic_garden**
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Botanic_garden
rdf:subClassOf	owl:Thing
rdfs:comment	Clase que representa al jardín botánico

rdf:Class	**ldjbre:person**
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#person
rdf:subClassOf	owl:Thing
rdf:EquivalentTo	foaf:person
rdfs:comment	Persona a cargo del jardín botánico

rdf:Class	**ldjbre:Labor_role**
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#Labor_role
rdf:subClassOf	owl:Thing
rdfs:comment	Clase de categoría laboral

rdf:Class	**ldjbre:species**
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species
rdf:subClassOf	owl:Thing
rdfs:comment	Elementos encontrados en la colección de datos

rdf:Class	ldjbre:University
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#University
rdf:subClassOf	owl:Thing
rdfs:comment	Institución a la que es parte el jardín botánico

Propiedades:

ldjbre:bioclimate	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#bioclimate
rdf:subPropetyOf	owl:TopDataProperty
rdfs:comment	bioclima
rdfs:domain	ldjbre:Bioclimate_data
rdfs:range	rdfs:Literal

ldjbre:category_use	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#category_use
rdf:subPropetyOf	owl:TopDataProperty
rdfs:comment	Categoría del especie según su uso común
rdfs:domain	ldjbre:specie
rdfs:range	rdfs:Literal

ldjbre:height_data	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#height_data
rdf:subPropetyOf	owl:TopDataProperty
rdfs:comment	Altura del jardín botanico
rdfs:domain	ldjbre:Bioclimate_data
rdfs:range	rdfs:Literal

ldjbre:id_bioclimate	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#id_bioclimate
rdf:subPropetyOf	owl:TopDataProperty
rdfs:comment	Identificador de datos bioclimaticos
rdfs:domain	ldjbre:Bioclimate_data
rdfs:range	rdfs:Literal

ldjbre:id_collection	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#id_collection
rdf:subPropetyOf	owl:TopDataProperty
rdfs:comment	Identificador de coleccion
rdfs:domain	ldjbre:Biological_collection
rdfs:range	rdfs:Literal

ldjbre:id_person	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#bioclimate
rdf:subPropetyOf	owl:TopDataProperty
rdfs:comment	Identificador de la persona
rdfs:domain	ldjbre:person
rdfs:range	rdfs:Literal

ldjbre:idbotanic	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#idbotanic
rdf:subPropetyOf	owl:TopDataProperty
rdfs:comment	Identificador del jardín Botanico
rdfs:domain	ldjbre:Botanic_garden
rdfs:range	rdfs:Literal

ldjbre:idspecies	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#bioclimate
rdf:subPropetyOf	owl:TopDataProperty
rdfs:comment	Identificador de la especie en la colección de datos
rdfs:domain	ldjbre:species
rdfs:range	rdfs:Literal

ldjbre:section	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#section
rdf:subPropetyOf	owl:TopDataProperty
rdfs:comment	Sección perteneciente según el jardín botánico
rdfs:domain	ldjbre:speciez
rdfs:range	rdfs:Literal

ldjbre:surface	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#surface
rdf:subPropetyOf	owl:TopDataProperty
rdfs:comment	Superficie
rdfs:domain	ldjbre:Bioclimate_data
rdfs:range	rdfs:Literal

ldjbre:typeofclimate	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#bioclimate
rdf:subPropetyOf	owl:TopDataProperty
rdfs:comment	Tipo de clima
rdfs:domain	ldjbre:Bioclimate_data
rdfs:range	rdfs:Literal

ldjbre:ubicacion	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#ubicacion
rdf:subPropetyOf	owl:TopDataProperty
rdfs:comment	Ubicación del jardín botanico
rdfs:domain	ldjbre:Botanic_garden
rdfs:range	rdfs:Literal

ldjbre:belong_to	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#belong_to
rdfs:comment	Pertenece a
rdfs:domain	ldjbre:species
rdfs:range	ldjbre:Biological_collection

ldjbre:contains	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#contains
rdfs:comment	Identifica al objeto que contiene los datos
rdfs:domain	ldjbre:Botanic_garden
rdfs:range	ldjbre:Biological_collection

ldjbre:correspondin_to	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#correspondin_to
rdfs:comment	Relación correspondiente al jardín botánico
rdfs:domain	ldjbre:Bioclimate_data
rdfs:range	ldjbre:Botanic_garden

ldjbre:has_a	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#has_a
rdfs:comment	Relación de la universidad con el jardín botánico
rdfs:domain	ldjbre:University
rdfs:range	ldjbre:Botanic_garden

ldjbre:has_category	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#has_category
rdfs:comment	Identifica la categoría laboral de la persona
rdfs:domain	ldjbre:person
rdfs:range	ldjbre:Labor_role

ldjbre:has_responsible	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#has_responsible
rdfs:comment	Relación del jardín botánico con la persona a cargo
rdfs:domain	ldjbre:Botanic_garden
rdfs:range	ldjbre:person

ldjbre:information_of	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#information_of
rdfs:comment	Datos bioclimáticos respecto al jardín.
rdfs:domain	ldjbre:Botanic_garden
rdfs:range	ldjbre:Bioclimate_data

ldjbre:isCategoryOf	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#isCategoryOf
rdfs:comment	Relaciona a la persona y su función laboral
rdfs:domain	ldjbre:Labor_role
rdfs:range	ldjbre:person

ldjbre:ischarge_of	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#ischarge_of
rdfs:comment	Relación de la persona con el jardin
rdfs:domain	ldjbre:person
rdfs:range	ldjbre:Botanic_garden

ldjbre:stored_in	
rdf:Type	rdf:Property
Uri	http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#stored_in
rdfs:comment	Identifica al objeto donde se almacenan los datos
rdfs:domain	ldjbre: Biological_collection
rdfs:range	ldjbre: Botanic_garden

ANEXO 8. VERIFICACIÓN RDF

Validation Results Your RDF document validated successfully. **Triples of the Data Model**

Subject
http://localhost:8080/jardinbotanico/resource/Esp1
http://localhost:8080/jardinbotanico/resource/Esp1
http://localhost:8080/jardinbotanico/resource/Esp1
http://localhost:8080/jardinbotanico/resource/Esp1

Predicate
http://www.w3.org/1999/02/22-rdf-syntax-ns#type
http://www.w3.org/1999/02/22-rdf-syntax-ns#type
http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#idspecies
http://rs.tdwg.org/dwc/terms/scientificName
http://rs.tdwg.org/dwc/terms/Family
http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#category_use

Object
http://www.w3.org/2000/01/rdf-schema#Class
http://jbrespinosa.unl.edu.ec/ontologies/JBRE/ldjbre#species
"Esp1"
"Oryza sativa L."
"Poaceae"
"Cereales"

Figura 42. Representación gráfica Extracto Validación W3C.

ANEXO 9. ENTREVISTA.

Objetivo Principal: Investigar el entorno tecnológico actual del jardín botánico “Reinaldo Espinosa”, proyectos con enfoque de Linked Data, documentación en plataformas digitales.

Se pretende indagar sobre cuáles son los medios digitales que posee el jardín botánico, cuáles y como están expuestos los datos, en qué formato, de qué manera está estructurada, su actualización y la existencia o carencia de tecnologías u proyectos con enfoque de Web Semántica y Linked Data; por lo que se presenta la documentación de la entrevista realizada en base a las preguntas:

M.G.: Representa al Ing. Marcelo Gutiérrez técnico del Jardín Botánico

¿Cómo es la exposición de información del Jardín Botánico en medios digitales?
M.G.: dentro de nuestro blog www.jbreinaldoespinosa.wordpress.com exponemos la información, los requerimientos que pueden solicitarse en el público con informes y eventos que pueden surgir, en las redes sociales se mantenía algo de información como Facebook, pero actualmente muy poco; se está trabajando en la base de datos con FileMaker con 150 ejemplares ya registrados.

¿Cuál es la Plataforma digital con la que cuenta el Jardín Botánico?
M.G.: La plataforma principal que tenemos es el blog que más nos ayuda, exponemos allí la información del jardín botánico.

¿Dónde están almacenados los datos?, ¿En qué formato?, ¿quién investiga o consulta los datos?; ¿Cuál es el tiempo en el que los datos cambian o actualizan?
M.G.: Las publicaciones que actualmente se encuentran, que son como alrededor de 8 se componen de la gestión de cada año del jardín botánico, si algún ejemplar ingresa, al igual que alguna noticia considerada relevante, toda la elaboración de los documentos son en formato PDF, puesta a disposición para cualquier persona que esté interesada, el tiempo de actualizan de los informes es anual, y la modificación de los inventarios u noticias relevantes dependen del ingreso de nuevas especies, como se susciten los eventos es de forma inmediata.

¿Existen recursos digitales actualmente disponibles para los investigadores, docentes y estudiantes bajo el enfoque de Web Semántica o Linked Data?

M.G.: No poseemos tal información, solo lo que se encuentra en los documentos, como le comenté anteriormente, el jardín esta presto para cualquier investigación relacionada con la conservación, las plantas.

¿Existen proyectos propuestos o en desarrollo en beneficio del Jardín Botánico con la utilización de servicios u herramientas tecnológicas? (si es el, caso con enfoque de Web Semántica o Linked Data).

M.G.: Claro que se trabajó en propuestas interesantes como el desarrollo de recorridos virtuales, realidad aumentada con códigos QR, pero lastimosamente la gran mayoría queda solo en papel, ya que su ejecución no es realizada; pero se trabaja en temas de conservación, realmente la actualidad de esos trabajos tanto del departamento de Investigación de la Universidad es algo que se está trabajando.

ANEXO 10. EVALUACIÓN DE LA ONTOLOGÍA.

La evaluación de la ontología desarrollada se basó en la aplicación del Método triangular donde se constituyen buenas prácticas de otros métodos existentes y de la Ingeniería de Software. Su realización abarca tres flancos detallados en López [29]:

1) La verificación de los casos de prueba que permite evaluar el cumplimiento de los requerimientos de las ontologías expresados en las preguntas de competencia.

2) Los razonadores evalúan las propiedades lógico-formales de las ontologías durante todo el ciclo de vida.

3) La lista de chequeo permite evaluar el diseño de la ontología derivando en su evolución y continuo mejoramiento.

"Las entradas del método son: el modelo ontológico, la lista de chequeo con los errores más comunes y los documentos de casos de prueba diseñados a partir de las preguntas de competencia. La salida del método es el resumen de la evaluación", López [29].

Verificación del diseño a partir de una lista de chequeo.

Se propone una lista de errores comunes detectados al evaluar las ontologías que describimos en la tabla.

TABLA LXXII.
LISTA DE CHEQUEO ERRORES COMUNES EN ONTOLOGÍAS.

Tipo de Error	Elemento a Verificar
I	Una misma clase es definida como subclase y superclase al mismo tiempo en diferentes niveles de la taxonomía.
II	Uso excesivo de la relación es-un.
III	Existencia de más de un concepto principal.
IV	Existencia de clases incompletas que provocan ambigüedad por no estar correctamente documentadas
V	Falta de conocimiento disjunto. No se declara la disyunción entre conceptos provocando una incorrecta formalización del conocimiento.
VI	Falta de exhaustividad. Se declaran subclases sin tener en cuenta la división completa de los conceptos en partes.
VII	Existencia de términos repetidos.
VIII	Poca especificación o delimitación de las propiedades que provoca un pobre razonamiento.
IX	No se corresponden los elementos del dominio con los conceptos declarados o no se corresponde el conocimiento del dominio con los conceptos, relaciones y axiomas declarados.
X	Existencia de redundancia entre las extensiones disjuntas de un concepto.
XI	No se tiene en cuenta la traducción de los conceptos de la taxonomía a otros idiomas.
XII	Falta de estandarización. Los nombres de los términos no siguen un estándar.

El diseño de la ontología fue revisado de forma que se detectaron errores de tipo III, VII y XI; este último refleja que la ontología no posee el factor de internacionalización, porque no se creyó necesario (la terminología en inglés es generalmente utilizada y

aplicada), se evaluó como correcto la elaboración final del diseño de la ontología.

Verificación de los requisitos de la ontología.

Para demostrar que las ontologías pueden ser utilizadas según lo previsto se verificaron 5 casos de pruebas que engloban preguntas de competencia a través de la ejecución del razonador integrado en Protégé HermiT(v1.3.8.).

TABLA LXXIII. EVALUACIÓN ONTOLOGÍA CASO DE PRUEBA 1.

Caso de Prueba Pregunta de Competencia 1 (CPQ1)	
Pregunta de Competencia	¿Cuantas especies se encuentran representadas en la ontología LDJBRE?
Escenario de prueba	Mediante protege se encuentran almacenadas las instancias de la clase species con un total de 416.
Resultado esperado	Al aplicar el razonador deberá desplegar la lista de 416 instancias que sean de la clase species.
Resultado obtenido	Satisfactorio. El resultado generado muestra un listado de 416 Instancias de clase species.

TABLA LXXIV. EVALUACIÓN ONTOLOGÍA CASO DE PRUEBA 2.

Caso de Prueba Pregunta de Competencia 2 (CPQ2)	
Pregunta de Competencia	¿Cuál es la información de la especie 307 identificada en la ontología?
Escenario de prueba	A través protégé se debe filtrar la información con el identificador de especie equivalente a 307
Resultado esperado	Se debe desplegar los data properties de la instancia de especie con idspecie igual a 307
Resultado obtenido	Satisfactorio. Se obtuvieron los diferentes usos donde el filtro se aplicó a la especie con el id igual a 307 mostrando la relación belong_to , los data properties de category_use, Family, scientificName y demás características con su valor.

TABLA LXXV. EVALUACIÓN DE ONTOLOGÍA CASO DE PRUEBA 3.

Caso de Prueba Pregunta de Competencia 3 (CPQ3)	
Pregunta de Competencia	¿Cuál es la persona encargada del jardín botánico, rol de desempeño que ocupa?
Escenario de prueba	La persona posee una propiedad denominada ischarge_of que indica que está a cargo de algo.
Resultado esperado	Dentro de la instancia p1 representativo de persona se debe observar el valor de ischarge_of la instancia BREv1 acordes al jardín botánico.
Resultado obtenido	Satisfactorio. Ejecutado el razonador se obtuvo las instancias de p1, p2 representativas hacia la clase persona, sin embargo, dentro de la navegación al observar su uso se evidencio las relaciones “has_responsible” e “ischarge_of” que especifican a p1 como el responsable de “BREv1” instancia de la clase Botanic_Garden, así como la persona a cargo “role1” del mismo jardín botánico cuyo valor es “Director”.

TABLA LXXVI. EVALUACIÓN DE ONTOLOGÍA CASO DE PRUEBA 4.

Caso de Prueba Pregunta de Competencia 4 (CPQ4)	
Pregunta de Competencia	¿Qué relación posee el jardín botánico y La Universidad Nacional de Loja?
Escenario de prueba	La ontología posee la instancia BREv1 y UNL para poder buscar la relación existente.
Resultado esperado	La ontología posee un object property de dc:isPartOf que está orientado a la instancia de la clase University.
Resultado obtenido	Satisfactorio. Se obtiene al obtejo BREv1 que posee explícitamente la definición "dc:isPartOf" cuyo valor es UNL instancia de la clase University que refleja al individuo del jardín como parte de la Universidad, así como la Universidad tiene un jardín botánico mostrado por medio de "has_a" obtenido.

TABLA LXXVII. EVALUACIÓN DE ONTOLOGÍA CASO DE PRUEBA 5.

Caso de Prueba Pregunta de Competencia 5 (CPQ5)	
Pregunta de Competencia	¿Cuáles son las especies en la ontología que representan a la sección de Bonsáis?
Escenario de prueba	Protégé posee la clase species con los dataproperty section que representan las diferentes secciones de las especies del jardín botánico.
Resultado esperado	La ontología posee un total de 8 especies representativas con la sección de bonsáis
Resultado obtenido	Satisfactorio. Se obtuvieron 8 instancias de la clase species cuyo data property "section" sea igual al valor filtrado "Bonsai".

Verificación de propiedades lógico-formales

Los resultados de evaluar las preguntas de competencia son transparentes al poder verificar las inferencias del razonador Hermit con los detalles de las instancias almacenadas en Protégé.

- Verificación Pregunta de Competencia 1 (CPQ1).

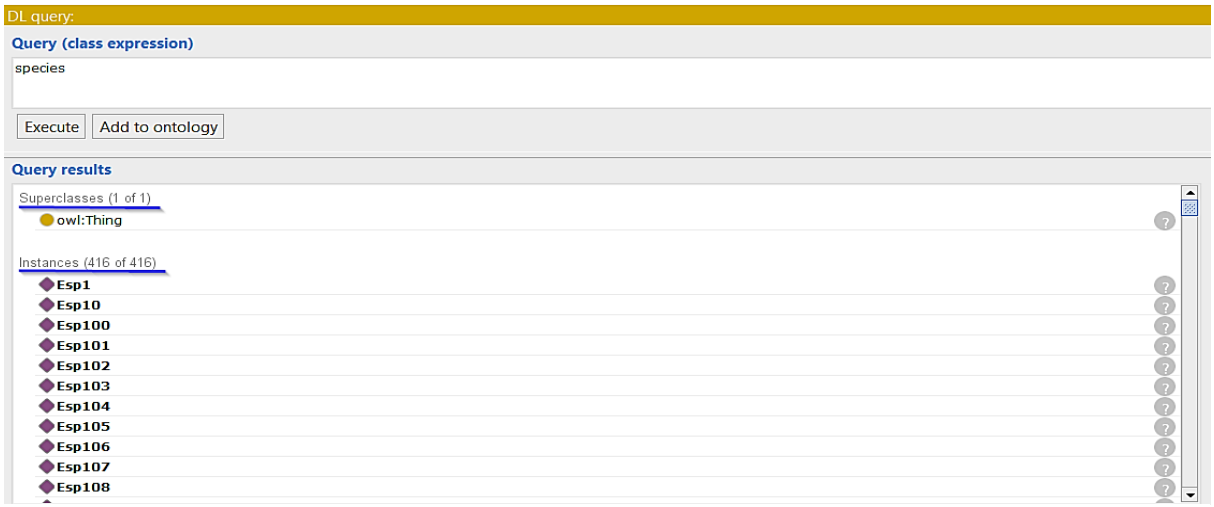


Figura 43. Representación gráfica verificación CPQ1 - Protégé

La ontología apunta a la representación de los datos del jardín botánico “Reinaldo Espinosa”, razón por la cual se plantea la pregunta: ¿Cuantas especies se encuentran representadas en la ontología LDJBRE?, como resultado (ver Figura 43) el razonador nos muestra la superclase correspondiente, seguido de 416 instancias almacenadas representativas a la clase “species”, obteniendo un desenlace muy satisfactorio.

- Verificación Pregunta de Competencia 2 (CPQ2).

  Planteando la pregunta: ¿Cuál es la información de la especie 307 identificada en la ontología?, se utilizó el filtro para obtener la instancia con los datos referentes al id del elemento 307, donde en la Figura 44 se observa al elemento y el razonador encuentra las propiedades y relaciones, así como, el tipo de clase a la que representa de manera satisfactoria.

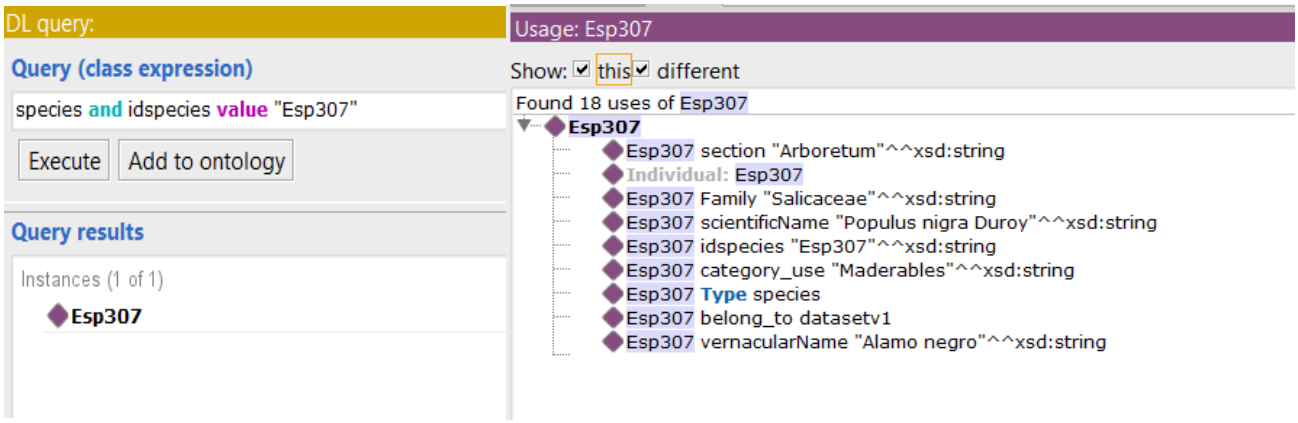


Figura 44. Representación gráfica verificación CPQ2 – Protégé.

- Verificación Pregunta de Competencia 3 (CPQ3).

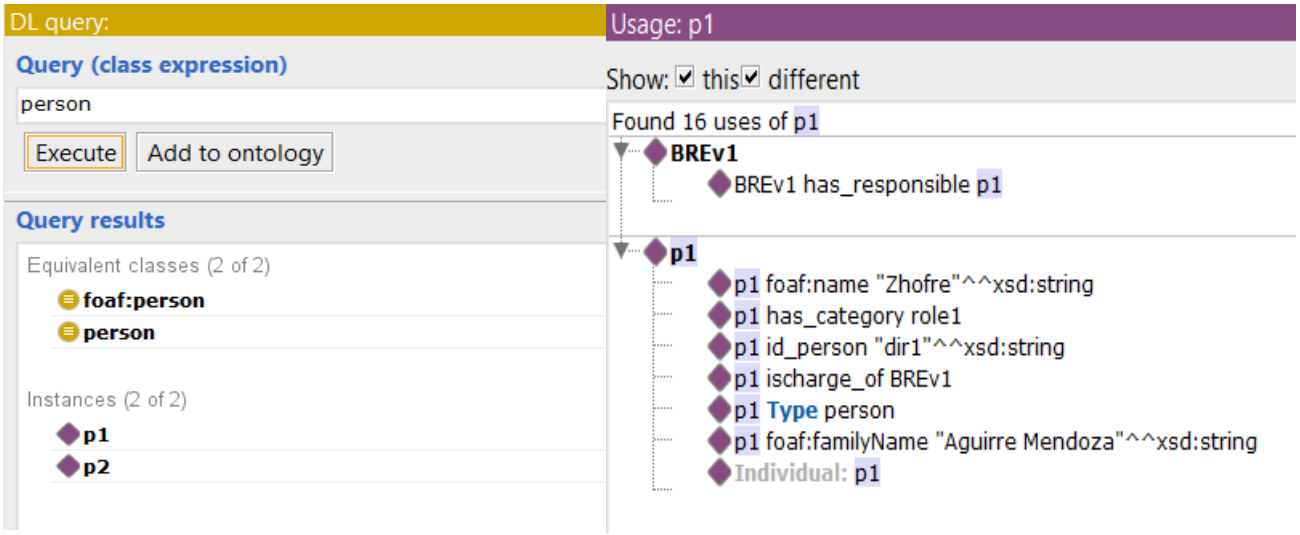


Figura 45. Representación gráfica verificación CPQ3 – Protégé.

Para resolver la pregunta: ¿Cuál es la persona encargada del jardín botánico, rol de desempeño que ocupa?, se obtuvo (Figura 45) dos instancias de personas con sus clases equivalentes, donde se tuvo que indagar los usos que poseen en la ontología y se generó de manera satisfactorio el resultado que plasma al individuo “p1” con la categoría “role1” cuyo valor es “Director” como la persona encargada del jardín definida en la relación “ischarge_of” del “BREv1”.

- Verificación Pregunta de Competencia 4 (CPQ4).

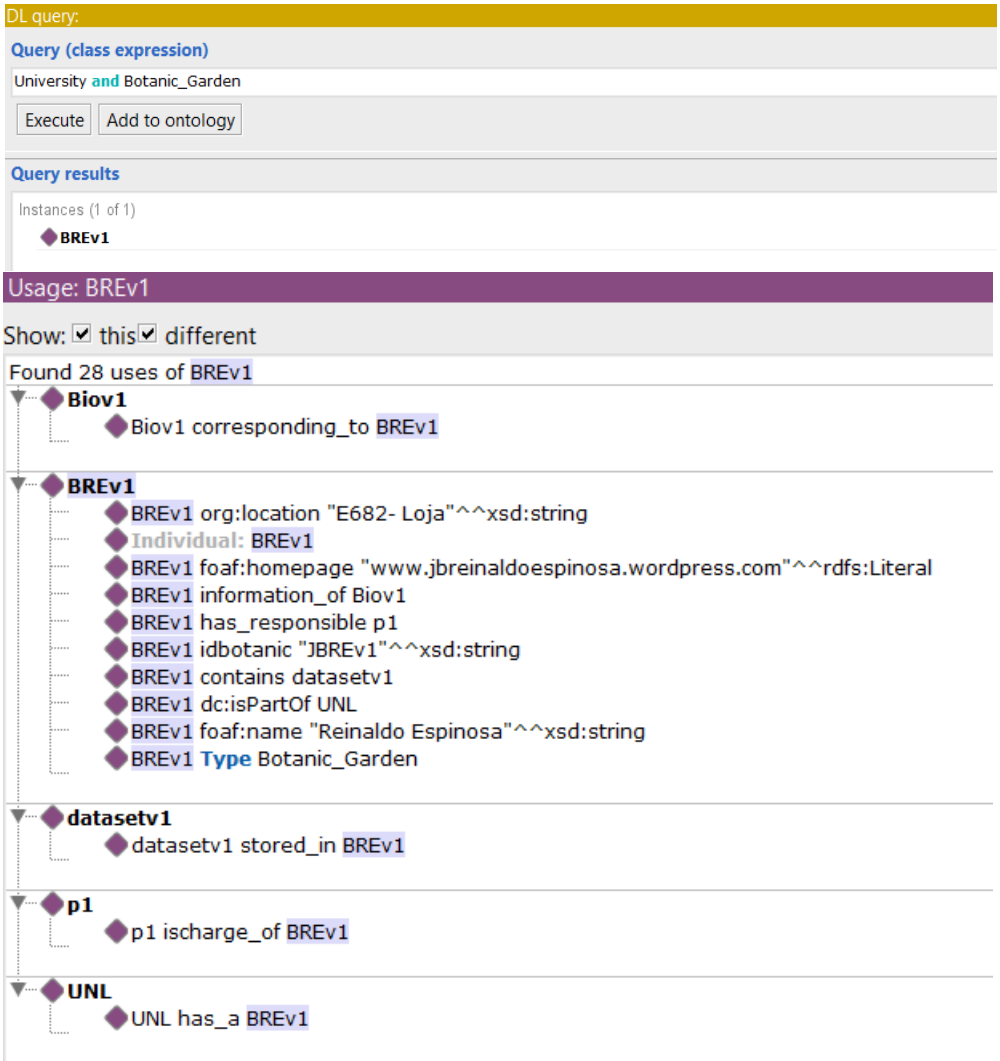


Figura 46. Representación gráfica verificación CPQ4 – Protégé.

Si la ontología pretende modelar los datos reales del jardín botánico, así como sus relaciones, se pregunta: ¿Qué relación posee el jardín botánico y La Universidad Nacional de Loja?; cuyo elemento obtenido fué la instancia “BREv1” de la clase Botanic_Garden que aplicando el razonador muestra dentro de su uso (Figura 46) la relación “isPartOf” valor referente a “UNL” instancia de la clase “University” reflejando que es parte de la Universidad; resultado satisfactorio.

- Verificación Pregunta de Competencia 5 (CPQ5).

  La ontología abarca la realidad de los datos del jardín botánico “Reinaldo Espinosa”, tal que, las secciones inmersas en él están descritas de manera idónea, para su verificación se pregunta: ¿Cuáles son las especies en la ontología que representan a la sección de Bonsáis?; obteniendo de manera satisfactoria las 8 instancias representativas de la sección “Bonsai” de la clase “species” del jardín botánico.

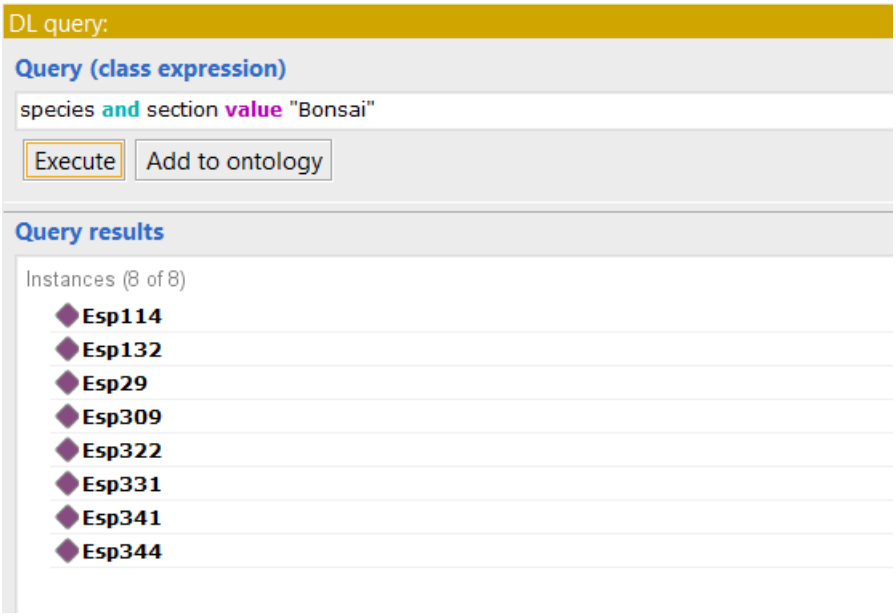


Figura 47. Representación gráfica verificación CPQ5 – Protégé.

ANEXO 11. MÉTODO EXPERIMENTO.

La experimentación consistió en modelar Linked Data en un área seleccionada, enfocada a la biodiversidad en la Universidad Nacional de Loja (jardín botánico "Reinaldo Espinosa"). El proceso experimental expuesto en Bocco [26] contempla las actividades como:

a. Definición del alcance.

   El objetivo es elaborar un modelo de Web Semántica y Linked Data enfocado en el campo de Biodiversidad, lo que se pretende obtener con el experimento es analizar si su nivel de impacto repercute en la conservación de la biodiversidad, exponiendo un punto de vista hacia los investigadores que realizan el despliegue de búsqueda de información en el dominio antes mencionado en un contexto limitado a las especies del JBRE de la Universidad Nacional de Loja.

TABLA LXXVIII.

DEFINICIÓN DE OBJETIVOS DE EXPERIMENTACIÓN.

Objeto de Estudio	Propósito	Aspecto de calidad	Perspectiva	Contexto
Proceso de generar Linked Data	Analizar	Efectividad	Investigador, editores de datos.	Sujetos
Modelo	Evaluar	Impacto	Comunidad científica	Objetos

Elaborado por el Autor.

b. Planificación

   La planificación realizada producto de generar los objetivos, se dividen en seis tareas que esclarezcan la ejecución del experimento.

   - Selección del Contexto: El experimento del presente TT pretende analizar el impacto de las tecnologías de web Semántica y Linked Data en la Biodiversidad, caso específico del JBRE; por lo tanto, las características del experimento se rigen en un entorno académico ejecutado por el estudiante investigador que permitan aseverar una operatividad más eficaz hacia los investigadores, y demás personas conocedoras de biodiversidad que la actualmente vigente.
   - Formulación de la Hipótesis: generado el contexto experimental la hipótesis que constituye este experimento se aplica para obtener la respuesta a la pregunta de investigación tratada como "¿Cuál será el impacto tecnológico de utilizar Linked Data en la conservación de la biodiversidad?", factor preponderante para el cumplimiento del objetivo principal en la definición del modelo de Web

Semántica y Linked Data enfocado en el campo de Biodiversidad con los datos del jardín botánico “Reinaldo Espinosa”.

H1,0: No existe una diferencia significativa en el impacto de la conservación de la biodiversidad, utilizando el modelo con tecnologías de Web Semántica y Linked Data realizando consultas en lenguaje Sparql con los recursos del jardín botánico “Reinaldo Espinosa”. H1,1: ¬H1,0.

El objetivo del contraste de hipótesis será rechazar las hipótesis nulas y posiblemente aceptar las alternativas.

- Selección de Variables: definidos los objetivos según la plantilla GQM[18] los factores del experimento se encuentran en la TABLA LXXIX.

TABLA LXXIX.

VARIABLES DEL EXPERIMENTO.

Variables Independientes	Variables Dependientes	Escala de Medición	Rango de Valores
P Linked Data	Efectividad	Disponibilidad de los datos	Correcta, Incorrecta.
Modelo	Impacto	Eficiencia del Modelo	Alto, Medio, Bajo.

Elaborado por el Autor.

La escala de medición se obtiene de la siguiente manera Coronel, Alarcón, Torres [73]:

Calidad de los datos: se engloba la Disponibilidad, es decir, la definición del grado al que los datos (o una porción de ellos) están presentes, alcanzables y listos para su uso. Para poder medir el grado de disponibilidad se guardan las respuestas obtenidas cada vez que se realiza una consulta al Endpoint a lo largo de la evaluación, Coronel, Alarcón, Torres [73].

$$D = 1 - \frac{Solicitudes\ erroneas}{solicitudes\ totales} \quad (1)$$

Eficiencia del Modelo: Los procesos y tecnologías involucradas para garantizar la conformidad de los valores de los datos. Los atributos con los cuales se mide la calidad del documento son:

- Validez sintáctica: La validez sintáctica es definida como el grado en que un documento RDF se ajusta a la especificación del formato de serialización. La serialización es el proceso de trasladar una estructura de datos u objeto a un formato que pueda ser almacenado y luego reconstruido. En esta métrica se consideran errores de sintaxis en RDF/XML, tipos de datos de literales mal

[18] GQM: acrónimo de Goal – Question – Metric, plantilla para definir objetivos.

formados y literales incompatibles con los valores posibles para el tipo de dato, Coronel, Alarcón, Torres [73].

$$VS = \frac{datos\ incorrectos}{cantidad\ de\ propiedades} \quad (2)$$

- Precisión semántica: La precisión semántica se define como el grado en que los valores de los datos correctamente representan hechos del mundo real, Coronel, Alarcón, Torres [73].

$$PS = \frac{valores\ incorrectos}{cantidad\ de\ propiedades} \quad (3)$$

- Correctitud: Este criterio determina si los enlaces a sitios Web externos contienen información relacionada con el recurso en cuestión, Coronel, Alarcón, Torres [73].

$$Co = \frac{enlaces\ incorrectos}{enlaces\ presentes} \quad (4)$$

- Impacto: Su criterio radica en la obtención del valor promedio de las medidas de VS, PS, CO, al cual se engloba de manera general como el valor de las respuestas del modelo en aspectos de sintaxis, representación e información relacionada a los recursos en cuestión, Coronel, Alarcón, Torres [73].

$$I = 1 - \frac{(VS+PS+Co)}{3} \quad (5)$$

La Figura 48 muestra la escala de calificación final, el valor del impacto que corresponde al rango de valores correspondientes al resultado de las variables de la experimentación.

VALORES	ESCALA
0.8 - 1	Alto
0.5 - 0.79	Medio
0 - 0.49	Bajo

Figura 48. Escala de Medición con resultados finales.
Elaborado por el Autor

- Selección de Sujetos: Para la aplicación de los tratamientos del experimento, el sujeto determinado es el investigador del TT, usando el muestreo no

probabilístico por conveniencia se pretenden contrastar los resultados con el director del TT, dentro del experimento se cuentan con los registros de 417 especies recolectadas del portal web principal correspondiente al JBRE.

- Elección del Diseño: Dado las restricciones de tiempo del TT el experimento a ejecutar cuenta con un diseño simple de inter-sujetos, es decir, se ejecutó un tratamiento para el conjunto de datos obtenidos.
- Instrumentación: Con el fin de contrastar la hipótesis formulada se preparó el objeto experimental que constaba de los datos.CSV con los elementos del JBRE, herramientas para revisar las tareas como Open Refine, Protege y el validador de la W3C. Las preguntas que se plantearon se muestran en la TABLA XX para explotar la información, y base fundamental para poder sacar conclusiones posteriormente.

c. Operación

Con la planificación ya realizada la operación del experimento se detalla en tres actividades que involucran como los datos se han de analizado.

- Preparación: para la realización del experimento se tuvo en cuenta las herramientas y su disponibilidad, si bien es cierto que son Open Source, la preparación radicó en el punto fundamental de llegar a detectar errores a corregir, como el tiempo oportuno a tomar en cuenta, ya que la herramienta Virtuoso permite la licencia de prueba de 30 días (gratis), el costo de la licencia a futuro varía en función del tiempo para su utilización.
- Ejecución: el experimento se realizó en un entorno local, ejecutado por el investigador, donde previamente eran conocidas las interrogantes y señalados los datos almacenados, su posterior consulta trata como objetivo de solventar las preguntas, a la par de contrastar que los resultados sean coherentes, estén completos; según el conjunto de datos que se construyó a través de la extracción de información del portal web principal del jardín botánico “Reinaldo Espinosa”.
- Validación de los datos: Después de la ejecución del experimento, la investigación recolectó los datos necesarios que solventen las dudas de las preguntas, así como, la redacción de las correspondientes respuestas se muestran en la TABLA XX; las consultas y su resultado se muestran en las figuras de Resultados sección 3. Representación de resultados y evaluación de la experimentación.

d. Análisis e interpretación.

Para poder analizar los datos que se emplearon en el modelo producto del proceso de generación y publicación de Linked Data se realizó el siguiente procedimiento.

- Realizar un estudio de todas las especies encontradas en la muestra y agruparlas de acuerdo a sus diferentes secciones.
- Contrastar la hipótesis de la experimentación a través de los resultados obtenidos en el paso anterior para decidir si proporcionan una contribución oportuna, realizando consultas del modelo.
- Determinar si los datos a través de los procesos previamente formulados, con el fin de explotar más información sean correctos, completos y coherentes.

Análisis Dataset de las especies del Jardín Botánico "Reinaldo Espinosa".

Las especies albergadas en el jardín datan de 417 respecto a la muestra obtenida del modelo basado en el archivo datos.CSV, las cuales según su consumo cotidiano poseen un uso diferente divido en muchas categorías, puestas a consideración en la TABLA LXXX y en la Figura 49.

TABLA LXXX.

CATEGORIAS DE USO DEL JARDÍN BOTÁNICO "REINALDO ESPINOSA".

Categoría de uso	Número de especies	Porcentaje (%)
Maderables	54	12.949
Cereales	11	2.637
Leguminosas	11	2.637
Oleaginosas	10	2.398
Hortalizas	33	7.913
Tubérculos	14	3.357
Azucareras	2	0.479
Especias	13	3.117
Pastos	30	7.194
Ornamentales	106	25.419
Aromáticas y Medicinales	44	10.551
Venenosas	6	1.438
Narcóticas	7	1.678
Textiles	8	1.918
Frutales	60	14.388
Bebibles	4	0.959
Otros	4	0.959
Total	417	100

Elaborado por el Autor.

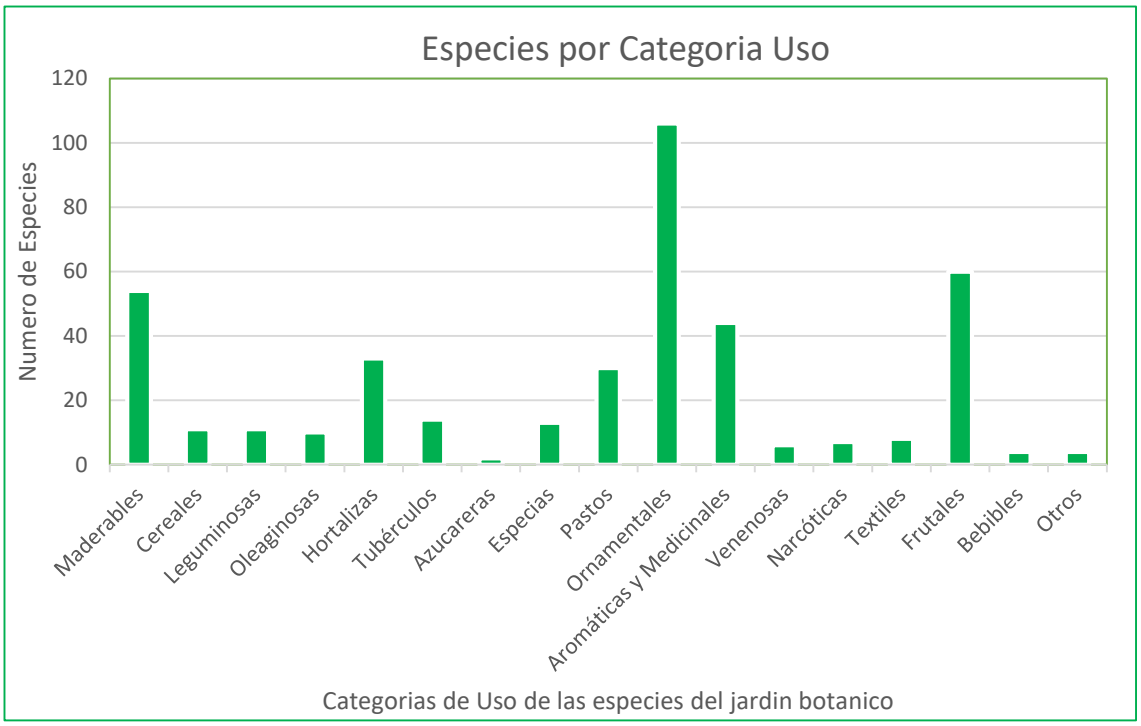


Figura 49.Representación gráfica de ejemplares categorizados por uso del JBRE.

Elaborado por el Autor.

De acuerdo al uso que se pueden dar se encuentran 17 categorías distintas; pero se destacan 4 categorías que sobreponen más del 10% del conjunto de 417 especies, por lo tanto, la primordial categoría a resaltar es “Ornamentales” con 25.42%, centradas en la estimación de la belleza y el perfume de las especies dando soporte generalmente al arte de jardinería, la segunda categoría “Frutales” con 14.39%, donde además de exponer el sabor exquisito de las frutas el interés de esta sección se enfoca en el apoyo de la investigación relacionada a vitaminas, minerales, aminoácidos; la tercera “Maderables” correspondiente al 12.95% referente a las plantas maderables, donde la madera es el principal material para la construcción, fabricación de muebles, tallado artesanal y otros usos; finalmente la cuarta categoría “Aromáticas y Medicinales” con 10.55% tiene su uso notable en la industria farmacéutica y lo relacionado a la utilización de las plantas en su estado natural en el desarrollo de tratamientos de curación de muchas enfermedades.

A continuación, se establecen las diferentes secciones que posee el jardín botánico analizando el porcentaje que posee cada una, relacionado a la muestra inicial de especies.

TABLA LXXXI.

SECCIONES DEL JARDÍN BOTÁNICO “REINALDO ESPINOSA”.

Sección	Número de especies	Porcentaje (%)
Arboretum	50	11.99
Plantas Andinas	91	21.82
Medicinales y Ornamentales	240	57.55
Orquideario	11	2.64
Plantas Xerofíticas	17	4.08
Bonsái	8	1.92
Total	417	100

Elaborado por el Autor.

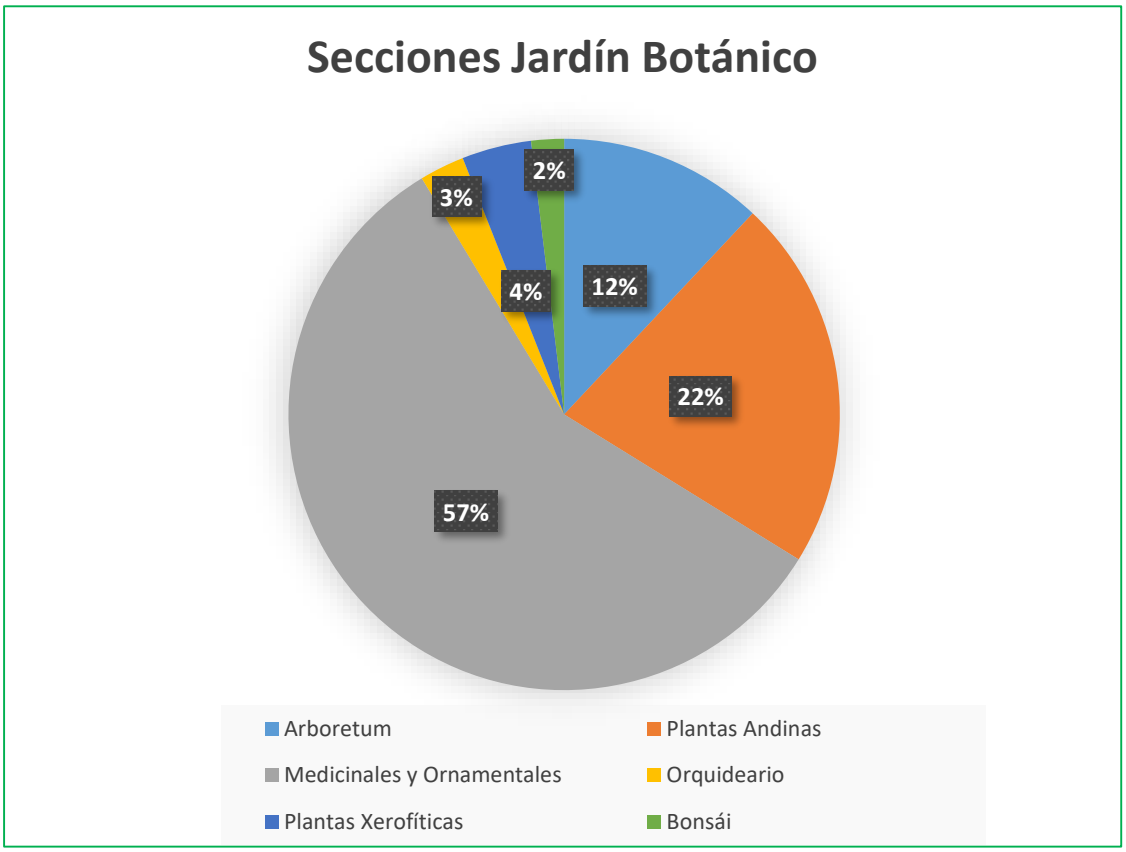


Figura 50.Representación gráfica de las secciones jardín botánico “Reinaldo Espinosa”.

Elaborado por el Autor.

Se puede observar con gran claridad que de acuerdo al conjunto de datos de 417 especies la sección de “Medicinales y Ornamentales” cuenta con 57.55%, es el principal subconjunto de elementos con mayor cantidad dentro de la muestra modelada; seguida de “Plantas Andinas” con 21.82% y “Arboretum” con 11.99% agrupaciones que muestran supremacía con respecto a las otras tres secciones restantes.

Contraste de Hipótesis.

Realizado el estudio de las secciones del jardín botánico se puede mencionar que, el análisis elaborado se resume en la supremacía dentro del conjunto de datos de muestra de las especies referentes a la sección de Ornamentales y Medicinales, por lo tanto, se expone la hipótesis propuesta H1,0 y su tratamiento en la TABLA LXXXII.

TABLA LXXXII.
CARACTERÍSTICAS DEL EXPERIMENTO APLICACIÓN LINKED DATA.

Hipótesis nula	No existe una diferencia significativa en el impacto de la conservación de la biodiversidad, utilizando el proceso de generación de Linked data y el modelo con los recursos del jardín botánico "Reinaldo Espinosa". H1,1: ¬H1,0
Lugar de realización	Ciudad de Loja (Ecuador).
Variables dependientes	Efectividad del proceso de generación y publicación de Linked Data medida a través de la *disponibilidad* (D) de los datos. Eficiencia del modelo medido a través del *impacto (I)* de las variables concernientes a la validez sintáctica, precisión semántica y correctitud.
Variables Independientes	Proceso Linked Data. Modelo.

Elaborado por el Autor.

La referencia de la publicación de los datos en formato PDF, de solo lectura, mostrado en la TABLA LXXI, muestra otra perspectiva en la calidad de los datos cuando se encaminan por los procesos de Linked Data hacia el formato RDF, al realizar el estudio de las secciones del jardín botánico con este enfoque, existe más eficiencia al trabajar con datos mejor estructurados, organizados, aportando con la información que se solicite al consultarla, difícil de lograr en tiempo y coherencia de validación con documentos de solo lectura con un alto porcentaje de error al ser manipulados con la percepción humana.

Los resultados de aplicar las fórmulas con el objetivo de evaluar los aspectos de Efectividad e Impacto planteados en el experimento se muestran en la TABLA LXXXIII, donde se contemplan individualmente el resumen de las métricas asociadas a las variables de efectividad, impacto, se identifican las columnas referentes a las variables del experimento, el valor final obtenido, las columnas siguientes correspondientes al rango y escala de calificación planificada con anterioridad, contemplando en la columna final la representación de si se puede o no rechazar la hipótesis nula.

TABLA LXXXIII.
RESULTADOS FINALES DE LA APLICACIÓN DEL EXPERIMENTO.

Variable	Valor	Escala Calificación	Rango correspondiente	Respuesta Hipótesis
Efectividad	0.9	Alto	0.8 – 1	NO
Impacto	0.8828	Alto	0.8 – 1	SI

Elaborado por el Autor.

Teniendo en cuenta los valores de la experimentación correspondiente a la variable de efectividad de los procesos de Linked Data no se puede rechazar la hipótesis, a pesar que el valor de la disponibilidad de los datos es de un 90% lo que demuestra su alto porcentaje de utilización; sin embargo, no se puede considerar con bastante firmeza que el proceso de Linked Data sea preponderante en la conservación de biodiversidad (posible alternativa), se debe para ello realizar un contraste con el flujo de procesos tratado en el jardín botánico “Reinaldo Espinosa”. Referente a la variable de Impacto considerando las métricas y su resultado final de 88% se puede rechazar la hipótesis nula, en vista que, se demuestra la eficiencia del modelo mediante y su impacto, sus valores de validez sintáctica, precisión semántica y correctitud son significativos (solicitudes con valores, datos y enlaces incorrectos son bajos) al momento de exponer la información de los recursos del jardín botánico “Reinaldo Espinosa”.

Finalmente se puede determinar que ejecutada la experimentación de los datos, el modelo representativo de los recursos del jardín botánico “Reinaldo Espinosa” posee datos correctos, coherentes y completos, verificados de manera automática a través de las herramientas utilizadas en la generación y publicación de Linked Data, así como, la validación manual efectuada por el investigador donde se comparó si los recursos pertenecían adecuadamente a una explícita descripción de lo que cada elemento representaba, con coherencia y abstracción completa de su información, mostrado en la TABLA XXVIII.

e. Presentación.

El Material educativo que representa el conocimiento técnico y empírico generado durante la realización del presente experimento se encuentra explícitamente descrito en la vigente memoria.

ANEXO 12. LICENCIA CREATIVE COMMONS.

Juan Maldonado 2019. Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional.

ANEXO 13. CERTIFICADO TRADUCCIÓN FINE TUNED ENGLISH

FINE-TUNED ENGLISH
LANGUAGE INSTITUTE
Líderes en la Enseñanza del Inglés

Ing. María Belén Novillo
DOCENTE DE FINE-TUNED ENGLISH CÍA LTDA.

C E R T I F I C A:

Que el documento aquí compuesto es fiel traducción del idioma español al idioma inglés del resumen de tesis titulada "ANÁLISIS Y EXPERIMETNACIÓN CON LINKED DATA APLICADO A BIODIVERSIDAD" autoría de Juan Andrés Maldonado Ajila, egresado de la carrera de Ingeniería en Sistemas de la Universidad Nacional de Loja.

Lo certifica en honor a la verdad y autorizo al interesado hacer uso del presente en lo que a sus intereses convenga.

Loja, 11 de septiembre de 2019

FINE - TUNED ENGLISH - CÍA. LTDA. - LOJA - ECUADOR

Ing. María Belén Novillo
DOCENTE DE FINE-TUNED ENGLISH CÍA LTDA.

Fine-Tuned English Cía. Ltda. | Teléfono 2578899 | Email venalfine@finetunedenglish.edu.ec | www.finetunedenglish.edu.ec

Loja: Macará entre Miguel Riofrío y Rocafuerte
Catamayo: Av. 24 de Mayo 08 - 21 y Juan Montalvo Telfs. 2678442
Zamora: García Moreno y Pasaje 12 de Febrero Telfs. 2608169
Yantzaza: Jorge Mosquera y Luis Bastidas Edificio "Coop. Sindicato de Choferes"

ANEXO 14. LISTA DE ACRÓNIMOS

Se expone la tabla de acrónimos descritos en el desarrollo del TT.

TABLA LXXXIV. TABLA DE ACRÓNIMOS EMPLEADOS.

ACRÓNIMO	DESCRIPCIÓN
OWL	Referente a los términos en inglés de Web Ontology Language
URI	Refleja los términos Unified Resource Identified
XML	Expresión inglesa eXtensible Markup Language
RDF	Términos en ingles de Resource Description Framework
SPARQL	Expresión inglesa Search Protocol and RDF Query Language
LOD	Referente a los términos en ingles de Linked Open Data
LODUM	Engloba los términos Linked Open Data University of Munster
W3C	Se refiere a World Wide Web Consortium
GBIF	Términos en inglés de Global Biodiversity Information Facility
BK	Se refiere a los términos de Base Knowledge
FOAF	Refleja los términos Friend Of A Friend
SLF4J	Términos en inglés de Simple Logging Façade for Java
XOL	Se refiere a XML-based Ontology Exchange Language
OIL	Términos en inglés de Ontology Inference Layer
RSL	Referente a los términos de Revisión Sistemática de Literatura
JBRE	Corresponde a Jardín Botánico Reinaldo Epinosa
LOV	Términos en inglés de Linked Open Vocabularies
LDJBRE	Corresponde a Linked Data Jardín Botánico Reinaldo Espinosa
CRUD	Expresión inglesa Create, Read, Update and Delete
TDWG	Refleja los términos Taxonomic Data Working Group
ABCD	Términos en inglés de Access to biological collections data
CoL	Expresión inglesa de Catalogue of life
EBV	Términos en ingles de Environments Biodiversity Variables
DCAT	Expresión inglesa de Data Catalog Vocabulary
OBOE	Corresponde a Extensible Observation Ontology
BFO	Expresión inglesa de Basic Formal Ontology
OCW	Términos en inglés de OpenCourseWare
EFO	Referente a los términos Experimental Factor Ontology
EOL	Corresponde a la Enciclopedia de la Vida
ENVO	Expresión inglesa de Environments Ontology
PATO	Términos en inglés de Phenotypic Quality Ontology
SIO	Términos en ingles de Semantic Science Integrated Ontology
GQM	Corresponde a los términos Goal – Question – Metric