

UCUENCA

Universidad de Cuenca

Facultad de Ingeniería

Carrera de Ingeniería de sistemas

Cosechador de metadatos en servidores de mapas para mejorar las búsquedas sobre la Infraestructura de Datos Espaciales de la Universidad de Cuenca

Trabajo de titulación previo a la obtención del título de Ingeniero de Sistemas


Autores:

Jhon Jairo Flores Pacheco

Jorge Enrique Encalada Ulloa

Director:

Carlos Villie Morocho Zurita

ORCID:  0000-0002-8196-2644

Cuenca, Ecuador

2024-02-22

Resumen

En los últimos años, el empleo de datos geoespaciales ha tenido un repentino crecimiento, convirtiéndose en un factor determinante para la toma de decisiones en una amplia gama de sectores. La información geoespacial, proveniente de diversas fuentes como sensores remotos, GPS y herramientas de información geográfica, ha experimentado un uso extenso en diversas áreas ambientales, urbanas, entre otras; siendo capaz de ofrecer una comprensión profunda del entorno que nos rodea. La calidad de la información geoespacial depende en gran medida de metadatos precisos y completos, fundamentales para su búsqueda y organización. Las Infraestructuras de Datos Espaciales (IDE) facilitan el acceso a la información geoespacial a través de entornos web; la IDE UCuenca ofrece servicios como mapas web, visualización de mapas interactivos y gestión de capas. Teniendo en cuenta que, la gestión inadecuada de metadatos puede llevar a la pérdida de datos relevantes y dificultar el intercambio de información, el presente proyecto de tesis propone un cosechador de metadatos para mejorar la IDE UCuenca, centralizando la información geográfica, mejorando sus servicios y buscando facilitar la toma de decisiones en la planificación urbana y otros sectores. Por último, con el fin de proporcionar orientación a futuros lectores y facilitar su implementación en otros contextos, se elaboró un documento de Buenas Prácticas para la implementación de un cosechador de metadatos geoespaciales con GeoNetwork.

Palabras clave: cosechador de metadatos, información geoespacial, datos geográficos, IDE UCUENCA



El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Cuenca ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por la propiedad intelectual y los derechos de autor.

Repositorio Institucional: <https://dspace.ucuenca.edu.ec/>

Abstract

In recent years, the use of geospatial data has experienced a sudden growth, becoming a determining factor in decision-making across a wide range of sectors. Geospatial information, originating from various sources such as remote sensors, GPS, and geographic information tools, has seen extensive use in various environmental, urban, and other areas, capable of providing a profound understanding of the surrounding environment. The quality of geospatial information largely depends on accurate and comprehensive metadata, which are essential for its retrieval and organization. Spatial Data Infrastructures (SDIs) facilitate access to geospatial information through web environments; the UCuenca SDI offers services such as web maps, interactive map visualization, and layer management. Considering that inadequate metadata management can lead to the loss of relevant data and hinder information exchange, this thesis project proposes a metadata harvester to enhance the UCuenca SDI, centralizing geographic information, improving its services, and seeking to facilitate decision-making in urban planning and other sectors. Finally, to provide guidance to future readers and facilitate its implementation in other contexts, a Best Practices document was developed for the implementation of a geospatial metadata harvester with GeoNetwork.

Keywords: metadata harvester, geospatial information, geographic data, IDE UCUENCA



The content of this work corresponds to the right of expression of the authors and does not compromise the institutional thinking of the University of Cuenca, nor does it release its responsibility before third parties. The authors assume responsibility for the intellectual property and copyrights.

Institutional Repository: <https://dspace.ucuenca.edu.ec/>

Índice de contenido

Resumen	2
Abstract	3
Índice de contenido	4
Índice de figuras	7
Índice de tablas	9
Agradecimientos	10
Agradecimientos	11
1. Introducción	12
1.1. Motivación y contexto.....	12
1.2. Planteamiento del problema	13
1.3. Solución propuesta	15
1.4. Objetivos.....	16
1.4.1. Objetivo general.....	16
1.4.2. Objetivos específicos	16
1.5. Metodología.....	16
1.6. Estructura del trabajo	18
2. Marco teórico y tecnológico	19
2.1. Metadatos geoespaciales	19
2.2. Sistema de Información Geográfica (SIG).....	19
2.2.1. Geoportales.....	20
2.3. Infraestructura de Datos Espaciales (IDE)	20
2.4. Gestor de metadatos.....	21
2.5. ArcGis.....	22
2.6. Estándares y normas para datos geoespaciales	22
2.6.1. Protocolo OAI-MPH	22
2.6.2. ISO 19115	23
2.6.3. ISO 19139.....	23
2.6.4. ISO 19112	23
2.6.5. FGDC CSDGM	23
2.6.6. Dublin Core	23
2.6.7. La Infraestructura de Información Espacial para Europa (INSPIRE)	24
2.6.8. Perfil Ecuatoriano de Metadatos (PEM).....	24
2.7. Cosechador o recolector de metadatos.....	25
3. Estado del arte	27

3.1.	Aplicación de la metodología PRISMA-ScR.....	27
3.2.	Definición de preguntas de investigación.....	27
3.3.	Búsqueda sistemática.....	28
3.3.1.	Criterios de inclusión.....	29
3.3.2.	Criterios de exclusión.....	29
3.3.3.	Fuentes de información.....	30
3.3.4.	Estrategia de búsqueda.....	30
3.3.5.	Selección de estudios relevantes.....	31
3.3.6.	Diagrama de flujo PRISMA.....	32
3.4.	Extraer datos de contexto.....	33
3.4.1.	Definición de variables a analizar.....	33
3.4.2.	Lectura completa de los documentos.....	34
3.4.3.	Resultados de documentos seleccionados.....	35
3.5.	Análisis y síntesis de los resultados.....	38
3.6.	Discusión y conclusiones de resultados.....	42
4.	Metodología.....	45
4.1.	Análisis de herramientas con enfoque basado en características.....	45
4.1.1.	Introducción de las herramientas a analizar.....	45
4.1.2.	CatMDEdit.....	47
4.1.3.	TKME.....	50
4.1.4.	Metavist.....	52
4.1.5.	EPA Metadata.....	53
4.1.6.	GeoNetwork.....	55
4.1.7.	GeoNode.....	57
4.1.8.	PyCSW.....	59
4.1.9.	ESRI ArcGIS.....	61
4.1.10.	PgMetadata.....	62
4.1.11.	Identificación de características claves.....	64
4.1.12.	Puntuación de las características.....	65
4.1.13.	Resultados de características de las herramientas.....	73
4.1.14.	Análisis de resultados y selección de herramienta.....	74
4.2.	Diseño y desarrollo del modelo conceptual.....	75
4.3.	Software requerido para la implementación del cosechador.....	78
4.3.1.	Jetty.....	78
4.3.2.	GeoNetwork.....	79
4.3.3.	PostGIS.....	80
4.3.4.	Elasticsearch.....	81
4.3.5.	Kibana.....	81

4.4.	Configuración y funcionamiento del cosechador de IG.....	82
4.4.1.	Identificación de fuente para cosechamiento.....	83
4.4.2.	Configuración de fuente por medio de catálogo de metadato.....	85
4.5.	Gestión de metadatos con información de la IDE UCuenca.....	87
5.	Resultados	92
5.1.	Cosechamiento de metadatos geoespaciales.....	92
5.1.1.	Análisis del conjunto de metadatos en los geoportales.....	92
5.1.2.	Cosechamiento de información	95
5.2.	Dificultades y limitaciones en el cosechamiento	97
5.3.	Análisis de usabilidad del cosechador de metadatos	101
6.	Conclusiones.....	114
6.1.	Conclusión del objetivo específico 1.....	116
6.2.	Conclusión del objetivo específico 2.....	117
6.3.	Conclusión del objetivo específico 3.....	117
6.4.	Conclusión del objetivo específico 4.....	118
6.5.	Recomendaciones	118
6.6.	Trabajos futuros	118
	Referencias.....	119
	Anexos	123

Índice de figuras

Figura 1.1. Estructura de la Plataforma IDE UCuenca.	14
Figura 2.1. Información representada de un GIS.....	20
Figura 3.1. Diagrama de las etapas del proceso de investigación PRISMA-ScR.....	27
Figura 3.2. Diagrama de flujo PRISMA.	33
Figura 3.3. Comparación de documentos seleccionados y resultados primarios.....	35
Figura 3.4. Esquema de IDE y nodos para la arquitectura.	43
Figura 4.1. Pantalla inicial de CatMDEdit.....	48
Figura 4.2. Importación de metadato en formato XML.	49
Figura 4.3. Detalles adicionales del metadato.	49
Figura 4.4. Pantalla de inicio de TKME.....	51
Figura 4.5. Plantilla para una edición de un archivo de metadato.....	51
Figura 4.6. Plantilla para una edición de un archivo de metadato.....	52
Figura 4.7. Edición sección Información de identificación.	53
Figura 4.8. Edición metadato en ArcGis Pro.....	54
Figura 4.9. Interfaz de EPA metadato.	55
Figura 4.10. Interfaz inicial con el buscador disponible.	56
Figura 4.11. Grafica de análisis de información que contiene la almacenada.....	57
Figura 4.12. Ventana inicial de Geonode con los datos de prueba.....	58
Figura 4.13. Opciones de búsqueda y filtros en Geonode.	59
Figura 4.14. Interfaz de inicio de herramienta PyCSW.....	60
Figura 4.15. Ventana de portal web con administración de capas con IG.	61
Figura 4.16. Captura de video ilustrativo de creación de metadato.	63
Figura 4.17. Captura de video ilustrativo de edición plantilla HTML.	64
Figura 4.18. Diagrama de barras de las herramientas y las características destacadas.	74
Figura 4.19. Modelo conceptual del cosechador de metadatos.	76
Figura 4.20. Docker file del proyecto GeoNetwork.	79
Figura 4.21. Imagen de GeoNetwork en archivo Docker-compose.yml.....	80
Figura 4.22. Imagen de Postgis en archivo Docker-compose.yml.	81
Figura 4.23. Imagen de Elasticsearch en archivo docker-compose.yml.....	81
Figura 4.24. Imagen de Kibana en archivo docker-compose.yml.	81
Figura 4.25. Despliegue de herramienta GeoNetwork.....	82
Figura 4.26. Búsqueda de fuentes en buscador de Google.	83
Figura 4.27. Servicios de Metadatos y Mapas del M. de Agricultura.	84
Figura 4.28. GeoNetwork del Ministerio de Agricultura.	84
Figura 4.29. Opciones de la herramienta GeoNetwork de la IDE UCuenca.	86
Figura 4.30. Opciones de cosechamiento de metadatos.	86
Figura 4.31. Ingreso de credenciales en GeoNetwork.....	88
Figura 4.32. Opción para creación y gestión de metadato en GeoNetwork.	88
Figura 4.33. Selección de plantilla ISO 19139 para creación de metadato.	89
Figura 4.34. Formulario de metadatos con los campos obligatorios y opcionales.	90
Figura 4.35. Metadato creado y buscarlo por nombre.....	90
Figura 4.36. Detalles del recurso y su metadato almacenado.	91
Figura 5.1. Resultado de cosechamiento de metadatos en GeoNetwork.....	92
Figura 5.2. GeoNetwork del SNI en enero del 2023.	94
Figura 5.3. GeoNetwork del SNI en julio del 2023.....	94
Figura 5.4. Servidor con sobrecarga de información local.	99

Figura 5.5. Plantilla de la norma ISO 19115 registrada en la configuración de GeoNetwork.

.....	100
Figura 5.6. Error de carga de la plantilla de la ISO 19115 para generar un nuevo registro.	101
Figura 5.7. Resultados Escenario 1, Tarea 1.....	102
Figura 5.8. Resultados Escenario 1, Tarea 2.....	103
Figura 5.9. Resultados Escenario 1, Tarea 3.....	103
Figura 5.10. Resultados Escenario 1, Tarea 4.....	104
Figura 5.11. Resultados Escenario 2, Tarea 1.....	105
Figura 5.12. Resultados Escenario 3, Tarea 1.....	105
Figura 5.13. Resultados Escenario 4, Tarea 2.....	106
Figura 5.14. Resultados Escenario 4, Tarea 2.....	107
Figura 5.15. Resultados Usabilidad, Pregunta 1.	107
Figura 5.16. Resultados Usabilidad, Pregunta 2.	108
Figura 5.17. Resultados Usabilidad, Pregunta 3.	108
Figura 5.18. Resultados Usabilidad, Pregunta 4.	108
Figura 5.19. Resultados Usabilidad, Pregunta 5.	109
Figura 5.20. Resultados Usabilidad, Pregunta 6.	109
Figura 5.21. Resultados Usabilidad, Pregunta 7.	109
Figura 5.22. Resultados Usabilidad, Pregunta 8.	110
Figura 5.23. Resultados Usabilidad, Pregunta 9.	110
Figura 5.24. Resultados Usabilidad, Pregunta 10.	110
Figura 5.25. Resultados Usabilidad, Pregunta 11.	111
Figura 5.26. Resultados Usabilidad, Pregunta 12.	111
Figura 5.27. Resultados Usabilidad, Pregunta 13.	111
Figura 5.28. Resultados Usabilidad, Pregunta 14.	112

Índice de tablas

Tabla 3.1. Cadenas de búsqueda.	28
Tabla 3.2. Fuentes de información y bases de datos.	30
Tabla 3.3. Resultados de búsqueda en bases de datos y fuentes externas.....	31
Tabla 3.4. Definición de variables.	33
Tabla 3.5. Resultado de herramientas para cosechamiento de metadatos.	36
Tabla 3.6. Resultado de componentes para una arquitectura de un cosechador.....	36
Tabla 3.7. Buenas prácticas para una arquitectura de un cosechador.	37
Tabla 3.8. Metodologías para evaluar la calidad de la cosecha.....	37
Tabla 3.9. Documentos técnicos y trabajos científicos de referencia para el control de calidad de datos geoespaciales.	38
Tabla 4.1. Escala de puntuación en multiplataforma.	66
Tabla 4.2. Escala de puntuación en uso de estándares en la herramienta.	67
Tabla 4.3. Puntuación de cosechamiento o recolección en la herramienta.	68
Tabla 4.4. Puntuación de gestión de bases de datos o archivos.	69
Tabla 4.5. Puntuación de búsqueda y opciones de búsqueda en la herramienta.	70
Tabla 4.6. Puntuación de la interfaz de usuario en la herramienta.	71
Tabla 4.7. Puntuación sobre la integración con otros sistemas que tiene la herramienta.....	72
Tabla 4.8. Puntuación de las características de las herramientas analizadas.....	73
Tabla 5.1. Instituciones con metadatos disponibles en los sitios web.	93
Tabla 5.2. Fuentes de sitios web de los catálogos.	95
Tabla 5.3. Cosechamiento de metadatos con las fuentes identificadas.	96
Tabla 5.4. Pruebas de cosechamiento de metadatos en servidor local.....	97
Tabla 5.5. Conteo final de respuestas en el análisis de usabilidad general.	112

Agradecimientos

Primeramente, quiero expresar mi agradecimiento a mis padres, mi familia, en especial a mi madre Sara. Su apoyo, cariño y amor incondicional que me ha transmitido a lo largo de toda mi vida han sido el pilar fundamental de mi formación.

Agradezco al Ingeniero Villie Morocho, por su orientación experta, sabiduría y paciencia durante todo el proceso de investigación. Sus valiosos consejos y retroalimentación han sido fundamentales para dar forma a este trabajo.

A mis amigos y compañeros de estudio, quienes me brindaron su apoyo moral y alentador en los momentos de mayor exigencia. Sus palabras de ánimo fueron el impulso necesario para continuar en los momentos de duda.

A cada uno de ustedes, gracias por su valiosa contribución y por formar parte de este importante logro académico en mi vida.

Jhon Flores

Agradecimientos

En este momento significativo de mi vida, me gustaría expresar mi profundo agradecimiento a todos aquellos que han aportado de alguna manera en mi trayectoria académica y en la culminación de esta tesis.

En primer lugar, a mis padres, les debo todo lo que soy. Su dedicación, valores y ejemplo han sido mi guía a lo largo de los años. Siempre han creído en mí, incluso cuando yo mismo dudaba, y su apoyo inquebrantable ha sido la fuerza impulsora detrás de mis logros.

A mi amada, Michelle Bravo, que ha sido una fuente constante de consejos, apoyo y sobre todo de amor incondicional. Su presencia, palabras de aliento y comprensión han sido cruciales para superar los momentos de estrés, desafío y sobre todo hicieron esta experiencia memorable. Gracias por estar a mi lado y por ser mi inspiración.

A mi mejor amigo y hermano, Christopher Hugo, que siempre ha estado incondicionalmente en las buenas y en las malas durante todo este camino. Regañándome y brindándome consejos oportunos. Agradezco su apoyo inquebrantable en esta etapa crucial de mi vida y en la tesis. Tu amistad es un tesoro invaluable. Gracias.

A mis compañeros de carrera y amigos, gracias por compartir este viaje conmigo. Juntos hemos superado desafíos, celebrado triunfos y pasado por varias anécdotas en este arduo camino.

A mi tutor, ingeniero Villie Morocho, mi más sincero agradecimiento por su orientación, paciencia y dedicación. Su experiencia y sabiduría fueron esenciales en esta tesis. Siempre estuvo dispuesto a escuchar, aconsejar y motivarnos para alcanzar esta meta académica.

En fin, mi eterna gratitud a cada uno de ustedes por el valioso granito de arena que han aportado a mi vida. A todos, ¡Gracias!

Jorge Encalada

1. Introducción

1.1. Motivación y contexto

Generalmente, las instituciones públicas y privadas, tanto en el ámbito científico y tecnológico, generan información con el fin de lograr una toma de decisiones acertada en diferentes ámbitos; desde la planificación urbana hasta la gestión de recursos naturales, dando como resultado, la información geográfica (IG) como uno de los activos más importantes de un estado (Choquehuanca, 2020).

Gracias al aumento en la demanda de IG, existe la necesidad de conocer sobre sus metadatos y cómo recuperarlos de forma eficiente. Los metadatos geospaciales (metageo) son los datos de los datos que describen características adicionales de un recurso convirtiéndose en un complemento clave para identificar dicha geoinformación o IG (Ideca, 2023). Por ello, las instituciones trabajan en la publicación de esta información en repositorios de alta calidad y acceso abierto, siguiendo diversos estándares. (Chaglla y Morocho, 2016) (Meneses et al., 2018). Por otro lado, no solo las instituciones, sino también las personas que generan y publican IG en medios digitales, como portales web o correos electrónicos, deben encargarse de la administración de estos metadatos (Morocho, 2014). Para la publicación y tratamiento de los metadatos, se han adoptado las normas ISO 19115-19115-2 y 19139, lo que permite aplicar un marco homogéneo en este procedimiento que es de obligatorio cumplimiento. Debido a la heterogeneidad de los datos, no se determinó un perfil único de metadatos que permita la publicación del mismo de forma estandarizada para todos los productos espaciales (Cid y Delgado, 2022), pero existen normas y reglamentos para cada región o país que se establecieron en base a la necesidad de las instituciones.

La motivación detrás de este proyecto de tesis está ligado a la necesidad que se ha observado al momento de la recuperación de IG. Esta búsqueda o recuperación de información tiene una relevancia especial y difiere significativamente de una búsqueda convencional en Google Maps. Si se hace una búsqueda usando el motor de Google con mapas, ésta se realiza sobre la posición (coordenadas) y las variables (por ejemplo, nombres). Sin embargo, cuando se desea conocer sobre la geoinformación de una institución privada o pública del país, la que es necesaria para proyectos académicos, científicos y sobre todo de planificación territorial; ya no es localizable a través de este medio (Julià et al., 2005) (Chaglla y Morocho, 2016). Debido a esto, algunas instituciones publican su IG en sus propios servidores, haciéndola accesible al público. Por este motivo, herramientas de búsqueda como Google Maps demuestran lo ineficientes que son para la búsqueda de este tipo de información y sus correspondientes metadatos.

Durante los últimos años, se ha observado un cambio de paradigma en el ámbito de la información geoespacial, pasando de los tradicionales Sistemas de Información Geográfica (SIG) hacia las modernas Infraestructuras de Datos Espaciales (IDE). Esta transformación se ha impulsado por la necesidad de compartir datos de manera eficiente en la comunidad científica, optimizando recursos y tiempo. Como resultado, ha surgido un nuevo enfoque que permite la accesibilidad y el intercambio de información geoespacial a través de la web, consolidando así el concepto de Infraestructuras de Datos Espaciales (IDE) (SENPLADES, 2016). Las IDE pueden ser definidas como sistemas de información geográfica (SIG) de código abierto, implementados en línea, que brindan acceso a un vasto volumen de datos geoespaciales, y ofrecen capacidad de análisis y consulta de datos adicionales para la información geográfica (Ortiz González, 2021) (Otero & De Lázaro Y Torres, 2019).

En la actualidad, contamos con diversas herramientas SIG que ofrecen servicios y recursos para desarrollar una IDE. Estas herramientas permiten mejorar la disponibilidad de información al integrar tecnología informática y cumplen una función esencial al capturar, analizar, almacenar, editar y representar datos georreferenciados (Del et al., 2018). Considerando la relevancia de los metadatos geográficos y la gestión de la información obtenida mediante herramientas SIG, surge una pregunta crucial: ¿Cómo podemos obtener estos metadatos de diversas fuentes y almacenarlos en un solo repositorio?, la respuesta radica en el uso de un cosechador de metadatos, una solución diseñada para recolectar, importar y almacenar localmente metadatos de información geoespacial provenientes de fuentes remotas y diversas instituciones (Ortiz González, 2021). Un cosechador de metageo tiene como objetivo gestionar y centralizar la información geoespacial junto con sus metadatos, facilitando el acceso a esta información mediante un motor de búsqueda y filtros basados en distintas características. Esta herramienta permite recolectar y consolidar datos geográficos provenientes de diversas fuentes, por ejemplo, catálogos de metadatos, asegurando el cumplimiento de estándares y protocolos de transferencia.

En este contexto, los cosechadores de metadatos se presentan como una herramienta muy útil para automatizar el proceso de búsqueda, recolección y gestión de información, lo que puede contribuir a mejorar la eficiencia y eficacia de diferentes procesos en diferentes ámbitos.

1.2. Planteamiento del problema

La Universidad de Cuenca cuenta con una Infraestructura de Datos Espaciales denominada IDE UCuenca, la cual funciona en la Dirección Zonal de Planificación 6. La visión de esta dirección zonal es fomentar y promover la consolidación de los territorios mediante la

provisión de servicios óptimos y el aprovechamiento eficiente del suelo. En este contexto, la IDE UCuenca juega un papel fundamental al ofrecer recursos geospaciales que respaldan la toma de decisiones y el desarrollo sostenible de la región (Secretaría Nacional de Planificación, 2023). La IDE UCuenca se enfoca en la administración de datos geospaciales de las parroquias de Cuenca, provincia de Azuay, y tiene como objetivo integrar información de la Universidad de Cuenca. La Figura 1.1 muestra claramente la estructura existente entre las diversas plataformas que conforman la IDE UCuenca.

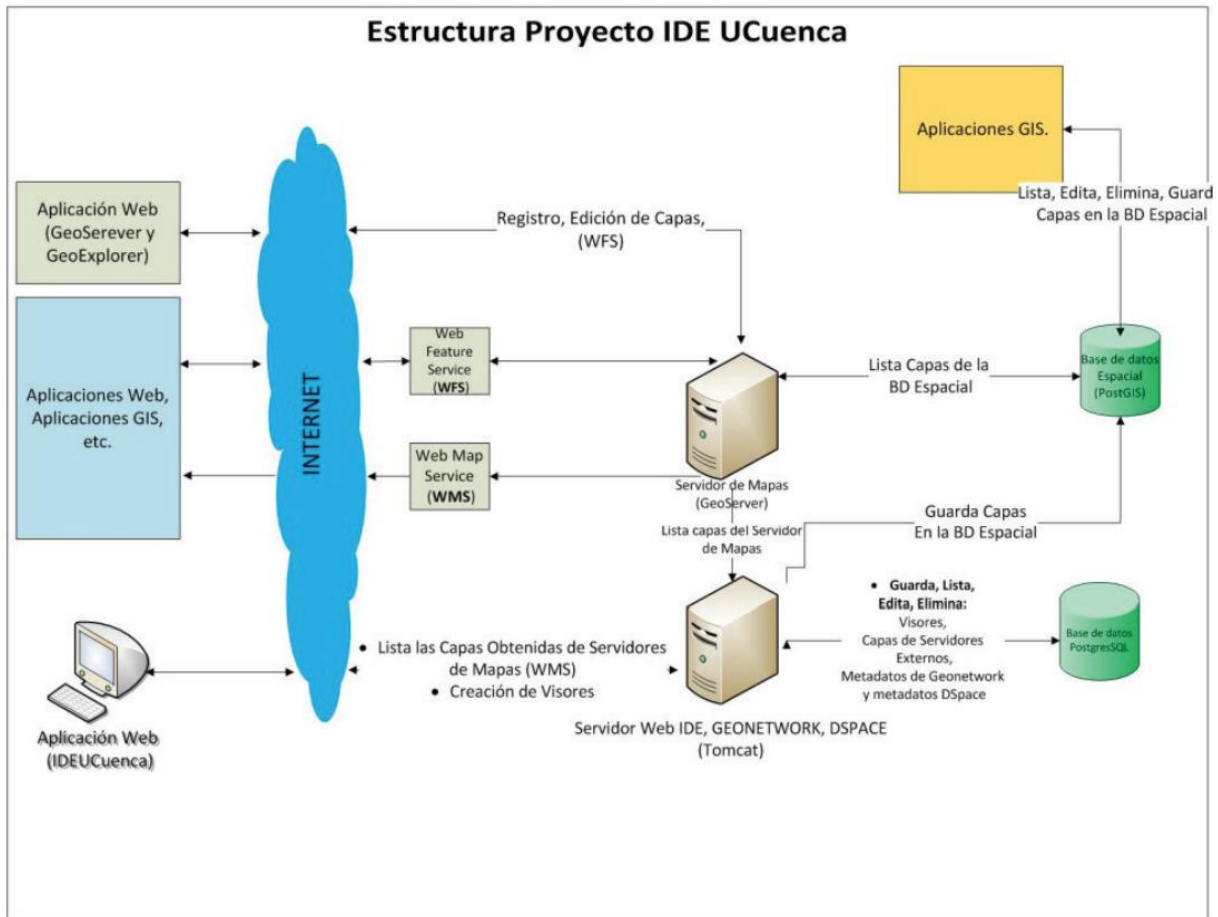


Figura 1.1. Estructura de la Plataforma IDE UCuenca. Fuente: (Morocho et al., 2016)

Actualmente, la IDE UCuenca carece de una herramienta de gestión de metadatos de IG, lo que dificulta el acceso, búsqueda y localización de los metadatos necesarios para la comunidad científica en el desarrollo de proyectos basados en esta información. La falta de esta herramienta también implica un consumo considerable de recursos y tiempo al tener que buscar los metadatos en diferentes instituciones.

La posibilidad de tener acceso a los datos y metadatos es crucial para la toma de decisiones en las instituciones y en el desarrollo de investigaciones futuras. En la sección anterior, se destacó una limitación importante presente en herramientas de búsqueda tradicionales, como

Google Maps, que se enfocan en proporcionar información general de lugares y fotos, sin ofrecer acceso a datos especializados, como catastros, uso de suelos y otros elementos georreferenciados esenciales para análisis detallados y aplicaciones específicas. En el contexto de obtener esta IG, el siguiente desafío radicaría en acceder a los metadatos asociados, los cuales suelen ser escasos o no se encuentran disponibles para el usuario en el sitio de consulta. Esta limitación ha demostrado ser una barrera significativa para la eficiencia en la localización de IG. A esto, se suma la cantidad de servidores de metadatos que brindan información como es el caso del Instituto Geográfico Militar (IGM), el Sistema Nacional de Información (SNI), Ministerios como de Agricultura, Medio ambiente, Desarrollo Urbano y Vivienda, Economía y Finanzas; además la Secretaría Nacional de Planificación, Gestión de Riesgos, entre otros.

Existen ciertos trabajos relacionados con respecto al cosechamiento de metadatos de la IG, pero éstos han sido solo un acercamiento al mismo; como es el trabajo de los autores Chaglla y Morocho; el cual, su estudio tiene como base el uso del sistema de gestión de metadatos y catalogador GeoNetwork, mismo que es configurado para cosechar metadatos de distintos nodos para una IDE subnacional y dando como resultado un sistema que soporta la integración de registros de metadatos con IG de fuentes heterogéneas (Chaglla & Morocho, 2016).

Tomando en cuenta esta problemática planteada, no se tiene un adecuado procedimiento para poner en marcha una herramienta que gestione eficientemente los metadatos y la implementación de la recolección o cosechador de los mismos. Estos datos estarían disponibles para ser compartidos y publicados, lo que permitiría que usuarios e instituciones lleven a cabo investigaciones y proyectos específicos, aprovechando al máximo la IG.

1.3. Solución propuesta

En este trabajo de tesis, se busca abordar el visible problema de la falta de accesibilidad a la información geoespacial en las Instituciones de Gobierno Autónomo Descentralizado (GADs). De este modo, incentivar a estas instituciones a generar sus propios metadatos mediante el uso de herramientas que permitan su publicación, fomentando así la colaboración con la IDE UCuenca para centralizar la información a través del cosechamiento de datos.

En este contexto, el objetivo principal de este trabajo es optimizar la búsqueda de IG en la IDE UCuenca, mediante la implementación de un cosechador de metadatos. Esta solución tiene como propósito agilizar la gestión de la información geoespacial, haciéndola aplicable de manera efectiva en diversas instituciones gubernamentales. En esta propuesta de tesis se considera trabajar en conjunto con el Servicio Nacional de Gestión de Riesgos y Emergencias

(SNGRE), el cual se ha iniciado una campaña a nivel nacional con los organismos encargados de la generación de información, en especial se ha reactivado la necesidad de regeneración y gestión de metadatos. De esta manera, se motiva y se da más relevancia a la propuesta de usar herramientas que puedan cosechar metadatos. Para esto, se despliega una herramienta de gestión de metadatos para mejorar la búsqueda y por ende el acceso a información en una IDE, acogiendo las normas y estándares establecidos a nivel nacional como es el Perfil Ecuatoriano de Metadatos (PEM), siendo un modelo para usos e implementaciones en otras entidades a nivel nacional. En adición, éste cosechador estará disponible para una posible integración con el Asistente Virtual (AV) desarrollado en la tesis titulada " Los AV como parte en los procesos de participación ciudadana de los planes de desarrollo y ordenamiento territorial" (Bustamante, 2023).

1.4. Objetivos

Para el desarrollo del presente trabajo de titulación, se han establecido los siguientes objetivos generales como específicos.

1.4.1. Objetivo general

Implementación de un cosechador de metadatos para mejorar la búsqueda de geoinformación en la IDE UCuenca, que además sea integrable a un AV.

1.4.2. Objetivos específicos

- Analizar y comparar diferentes gestores de metadatos existentes.
- Generar la arquitectura de software propuesta para el cosechador de metadatos de la IDE Ucuena con servicios para el uso del AV.
- Implementar el cosechador de metadatos.
- Evaluar la cosecha de metadatos con otros servidores nacionales.

1.5. Metodología

Para el desarrollo del presente trabajo de titulación se ha dividido en 4 etapas, que son las siguientes:

- 1. Planteamiento del problema:** en esta etapa, se aborda el problema de búsqueda y localización de metadatos geospaciales de la provincia del Azuay, en el cantón Cuenca. Además, se procede a verificar el estado de los servicios geospaciales disponibles en la

IDE Ucuenca y en diversas instituciones gubernamentales que cuentan con información del cantón.

2. **Revisión bibliográfica y estado del arte:** aquí se realiza una revisión teórica de los conceptos necesarios para el desarrollo del trabajo, así como una búsqueda exhaustiva y revisión de trabajos relacionados al tema de tesis. Para ello, se utiliza una metodología basada en una versión reducida de PRISMA-ScR (Prisma for Scoping Review) para elaborar el estado del arte.
3. **Desarrollo de la propuesta de un cosechador de metadatos:** en esta etapa se inicia con un análisis completo de las herramientas existentes para la gestión de metageo. Para esto se puso a prueba cada una de las herramientas y se identifican las características con mayor relevancia para seleccionar la herramienta más adecuada. Luego, se implementó un cosechador usando la herramienta seleccionada para la captura de información desde los servicios disponibles que tienen las instituciones.
 - a) **Análisis de herramientas de gestión de metadatos:** En esta etapa se realiza un análisis comparativo de las herramientas de gestión de metadatos de las cuales podemos destacar CatMDEdit, Metavist, GeoNetwork y GeoNode. En esta sección se analizan cada una de las herramientas, ventajas y desventajas, las características específicas, una consideración especial es verificar si es de código abierto. También se ejecutan pruebas de funcionamiento.
 - b) **Revisión de la disponibilidad de servidores con geoinformación o geoportales:** En esta etapa se realiza una revisión de la disponibilidad de servicios que existen en la web de instituciones públicas o privadas. Esto dará como resultado un cuadro de todos los servidores de metadatos y sus servicios. Además, de describir el proceso que se realiza para llegar a cada uno de los servidores disponibles y sus catálogos de metadatos.
 - c) **Arquitectura de software propuesta con los servicios del AV:** en esta etapa se diseña la Arquitectura del cosechador y la propuesta de la posible integración al proyecto principal “Los AV como parte en los procesos de participación ciudadana de los planes de desarrollo y ordenamiento territorial”.
 - d) **Caso de estudio sobre IDEUcuenca:** en esta etapa se inicia el despliegue de la herramienta seleccionada con el cosechador de metadatos.
4. **Evaluación del cosechador:** en esta fase se va a validar el cosechamiento realizado a través de cuestionarios realizados por especialistas en el área de planificación territorial,

el cual evaluará el cosechador como un producto que sea viable. Además, se evaluará la usabilidad del sistema implementado para el usuario final usando cuestionarios de usabilidad ya establecidos.

1.6. Estructura del trabajo

En el Capítulo 1, se introduce la motivación, el contexto y el planteamiento del problema que dio origen a esta investigación. Además, se presenta una descripción de la solución propuesta, así como los objetivos generales y específicos que guían el desarrollo del trabajo. En el Capítulo 2, se presenta el marco teórico y tecnológico que sustenta el desarrollo de la investigación. En esta sección, se abordan conceptos clave necesarios para comprender a fondo el trabajo. El Capítulo 3 se centra en el estado del arte y ofrece una revisión detallada de los trabajos relacionados con la gestión de metadatos geospaciales, especialmente aquellos que utilizan herramientas de gestión de metadatos. El Capítulo 4 abarca el marco metodológico, que incluye un análisis comparativo de herramientas de gestión de metadatos. Además, se describe el diseño y desarrollo del modelo conceptual. También se presenta el software seleccionado para el funcionamiento del cosechador de metadatos y se explica con más detalle el proceso de cosechamiento de dicha información. En el capítulo 5, se muestran los resultados que se obtienen del cosechamiento con las instituciones seleccionadas para este estudio. Además, se realiza un análisis de cada fuente de información, evaluando la disponibilidad de los servicios que ofrecen y su pertinencia para el proyecto. Finalmente, en el último capítulo, el Capítulo 6, se presentan las conclusiones, destacando los hallazgos más relevantes del proyecto de tesis, así como posibles trabajos futuros y recomendaciones para futuras investigaciones.

2. Marco teórico y tecnológico

En este capítulo, se introduce el concepto de metadatos y su relevancia en la era de la información, comprendiendo cómo estos “datos sobre datos” enriquecen la descripción y organización de contenidos. Asimismo, se examina las diversas tecnologías que permiten gestionar, capturar, procesar y presentar esta información esencial de manera coherente y eficaz. Finalmente, normas y estándares que definen como se deben describir y compartir los metadatos.

2.1. Metadatos geoespaciales

En lo referente al contenido que manejan los geoportales pueden ser datos o metadatos geoespaciales, donde se puede entender por datos geoespaciales a datos basados en el tiempo que están relacionados con una ubicación específica de la superficie de la Tierra (IBM, 2023). Mientras que se entiende por metadatos geográficos a los datos de los datos que describen características adicionales de un recurso convirtiéndose en un complemento clave para identificar adecuadamente dicha geoinformación (Ideca, 2023).

Dependiendo de la región a trabajar con metadatos geográficos, existen organizaciones que han ido desarrollando y definiendo recomendaciones para la creación de metadatos. Por ejemplo, la Organización Internacional de Normalización (ISO) define algunas normas, entre ellas las más populares son la ISO 19115 y 19139 para el intercambio de información geoespacial a nivel internacional (Federal Geographic Data Committee, 2023).

2.2. Sistema de Información Geográfica (SIG)

Un SIG o GIS por sus siglas en inglés, es una plataforma informática para la gestión de datos geoespaciales que permite mapear, integrar y analizar información geográfica para resolver problemas en investigación, planificación, ordenamiento y gestión geográfica. La plataforma genera bases de datos espaciales georreferenciadas que crean estructuras de datos que se vinculan, procesan, analizan y muestran los resultados en forma de mapas, tablas e informes (Narváez & Andrés, 2020). El término información geoespacial hace referencia al conjunto de datos que contiene una ubicación de forma explícita sobre la superficie de la tierra (Criollo Carlosama & Beatriz Alexandra, 2017).

Entre sus características, dispone de la visualización de información geográfica en forma de mapas, en la Figura 2.1 se observa los diferentes tipos de información que se puede representar. Con respecto a un determinado territorio, los datos se organizan en diversas capas (layers), conteniendo aspectos o parte de la realidad; por ejemplo, tipo de suelos, asentamientos poblacionales, vías, entre otras.

Entre sus características, dispone de la visualización de información geográfica a través de mapas (ver Figura 2.1), que permiten representar diferentes tipos de información. Los datos se organizan en diversas capas (layers) para abarcar aspectos específicos de un territorio, como tipo de suelos, asentamientos poblacionales, vías, entre otros



Figura 2.1. Información representada de un GIS. Fuente: (Carlosama & Beatriz Alexandra, 2017)

Un SIG es un sistema que conecta personas, información geoespacial y tecnología informática. La calidad de la información geoespacial capturada juega un papel crucial en la eficiencia y el éxito de los proyectos desarrollados mediante SIG. (Astudillo Vargas, 2018)

2.2.1. Geoportales

Según Ramos (2017), un geoportal se define como un sitio web cuyo propósito es proporcionar al usuario acceso a recursos y servicios basados en información geográfica. Además, Mora (2016, citado en Godoy 2022) describe los geoportales como sitios web que ofrecen información, servicios y contenido público gubernamental que involucra información geoespacial (Ramos Chango, 2017) (Godoy, 2022).

Cardona et al. (2019, citado en Carlosama 2022) clasificó los geoportales en dos tipos: el primero se refiere a los tipos de servicios que ofrecen, mientras que el segundo hace referencia a las funciones que desempeñan (Criollo Carlosama & Beatriz Alexandra, 2017).

Según SENPLADES, la administración de la IG debe ser ágil y eficiente, lográndose a través del uso de catálogos de metadatos conectados a sus respectivas bases de datos y la prestación de servicios mediante la implementación de geoportales (Espinoza Martínez, 2017).

2.3. Infraestructura de Datos Espaciales (IDE)

Las Infraestructuras de Datos Espaciales (IDEs), según la definición de Bernabé & López (2012, citado en Otero 2019), son Sistemas de Información Geográfica (SIG) abiertos

implementados sobre la Red, lo que implica componentes distribuidos, interfaces estándares, interoperabilidad, coordinación, acceso a los datos y capacidad de análisis como objetivos principales. En contraste, Ortiz et al. (2021) las describen como portales que albergan un gran volumen de datos geográficos, caracterizados por sus metadatos que permiten consultar información adicional sobre los datos geográficos. Por otro lado, IDERA (2011, citado en Graciana 2019) afirma que las IDEs pretenden ser entornos solventes, permanentes y bien mantenidos para garantizar la accesibilidad a datos de calidad en todo momento (Otero & De Lázaro Y Torres, 2019) (Ortiz González, 2021) (Graciana Pombo, 2019).

Según Atapuma en su trabajo, menciona que la IDE es el conjunto de tecnologías, políticas, estándares y leyes que facilitan el acceso a datos espaciales. Posee una estructura virtual conformada por bases de datos y servicios de la información geográfica como son los metadatos, WMS y WFS, que siguen estándares nacionales e internacionales que regulan y certifican que la información disponible (Atapuma et al., 2020).

Existen trabajos relevantes y actuales que muestran el estado actual de las IDE y su relación con inteligencia artificial como es “Building a Gold Standard Dataset to Identify Articles About Geographic Information Science” (Lopez-Vazquez et al., 2022). Otro trabajo es “Virtual Assistants to bring geospatial information closer to a smart citizen”, que presenta una propuesta del uso de Asistentes Virtuales en una IDE (Morocho et al., 2022).

2.4. Gestor de metadatos

López (2016, citado en Santafe & Laiton 2021) nos dice que un gestor de metadatos es capaz de centralizar el control de los metadatos en un repositorio en el cual se puede manipular de manera eficiente y centralizada. Mientras que Romero (2016) indica que la falta de un gestor de metadatos con la información de cada recurso, genera desgaste y pérdida de tiempo, lo que se conlleva en una baja disponibilidad e integridad de los objetos de información a largo plazo, desorden en la información y poca eficiencia en la captura de datos y almacenamiento (Santafe & Laiton, 2021) (Romero. N, 2016).

Una de las herramientas más difundidas para la gestión de metadatos en el ámbito de información geoespacial, con un entorno estandarizado y descentralizado, es GeoNetwork. Entre sus características principales se encuentra el acceso a una base de datos georreferenciada, productos cartográficos y metadatos de diversas fuentes, lo que facilita el intercambio de información y la colaboración entre instituciones a través de internet (Meneses et al., 2018).

2.5. ArcGis

ArcGIS es un conjunto de software y herramientas desarrollado por Esri (Environmental Systems Research Institute) que se utiliza para la gestión, análisis y visualización de datos geoespaciales. ArcGIS ofrece una plataforma completa para trabajar con información geográfica, permitiendo a los usuarios crear, editar, analizar y compartir mapas, capas de datos y aplicaciones geoespaciales.

ArcGIS incluye diferentes componentes, como ArcGIS Desktop y Pro que son programas de escritorio dedicados a realizar análisis avanzados y creación de mapas interactivos. ArcGIS Online que brinda acceso, compartición y colaboración de mapas por medio de la web. Por otro lado, ArcGIS Server y Enterprise que proporcionan soluciones de implementación y administración de servicios en servidores propios (ESRI, n.d.).

Este software de ESRI, ampliamente utilizado en diversos sectores, abarca áreas como la planificación urbana, la gestión de recursos naturales, el análisis medioambiental, la infraestructura, la agricultura, la salud, la educación y muchos otros campos que aprovechan datos geoespaciales para la toma de decisiones en instituciones y organizaciones (Eafit, 2017).

2.6. Estándares y normas para datos geoespaciales

Se han establecido normas para la administración adecuada de la información en diversas instituciones. Estas normas garantizan la interoperabilidad de sistemas y herramientas heterogéneas desarrolladas bajo estándares comunes. Asimismo, permiten gestionar la estructura del recurso de información geoespacial, incluyendo sus campos o atributos.

2.6.1. Protocolo OAI-PMH

El protocolo Open Archive Initiative-Protocol for Metadata Harvesting (OAI-PMH) se utiliza para gestionar metadatos de diversas fuentes a través de Internet o en una red local. OAI-PMH establece directrices para listar, recuperar (cosechar) metadatos de un repositorio (OAI Service Providers) y exponer recursos (OAI Data Providers) para su cosecha por aplicaciones externas. Su principal característica es el uso del estándar XML, lo que permite una mejor estructuración y beneficia el transporte de recursos a través del protocolo HTTP (Aguilera et al., 2020).

2.6.2. ISO 19115

El estándar establece la definición de metadatos de propósito general, ofreciendo un modelo para describir información o recursos, incluyendo los elementos de metadatos adicionales necesarios para las imágenes geográficas. Además, la "Implementación de esquemas XML para conceptos fundamentales" establece los esquemas XML necesarios para asegurar la interoperabilidad y compatibilidad de manera independiente de las herramientas utilizadas. (Espinoza Martínez, 2017) (Geodetic, 2023)

2.6.3. ISO 19139

Esta norma "Información Geográfica – Metadatos –Implementación de esquemas XML", define el esquema XML para la ISO 19115 y 19119, que establece el lenguaje estándar para almacenar e intercambiar metadatos geoespaciales (Carrascosa Basterra, 2018).

2.6.4. ISO 19112

La siguiente norma, es de "Información geográfica - Referencia espacial por identificadores geográficos", define el esquema conceptual para las referencias espaciales basadas en identificadores geográficos. (Espinoza Martínez, 2017).

2.6.5. FGDC CSDGM

El "Content Standard for Digital Geospatial Metadata" desarrollado por el Comité Federal de Coordinación de Datos Geográficos (por sus siglas en inglés, FGDC) de los Estados Unidos, define un conjunto de elementos para la descripción de metadatos geográficos. Este estándar se utiliza principalmente en metadatos e información geoespacial generada en Estados Unidos, como es el caso de las imágenes Landsat (Wang et al., 2023).

2.6.6. Dublin Core

Dublin Core es la iniciativa del desarrollo de estándares para describir, identificar, encontrar los recursos de información y facilitar su recuperación, que resultó en 15 descriptores, metadatos que se transformarían en un estándar muy utilizado a nivel global (Espinoza Martínez, 2017). Los metadatos de este estándar, se pueden agrupar de la siguiente manera:

- Elementos relacionados al contenido: Título, tema, descripción, fuente, lenguaje, relación y cobertura.
- Elementos relacionados al recurso como propiedad intelectual: Autor, editor, colaborador y derechos.

- Elementos relacionados a la instancia del recurso: Fecha, tipo de recurso, formato e identificador.

Este estándar es principalmente usado en metadatos no espaciales (At These Coordinates, 2020), como es el caso de metadatos para repositorios digitales (ejemplo DSPACE).

2.6.7. La Infraestructura de Información Espacial para Europa (INSPIRE)

INSPIRE establece un marco para el intercambio de datos geoespaciales entre los estados miembros de la Unión Europea, promoviendo la cooperación y la armonización de metadatos geográficos. Esta infraestructura europea de datos espaciales, comparte información espacial medioambiental entre organizaciones del sector público, facilitando el acceso público a la información de toda Europa (Ortiz González, 2021).

INSPIRE comprende de 27 elementos principales, los cuales se dividen en 16 elementos obligatorios y 11 elementos condicionales. En el contexto de INSPIRE, se reconocen tres tipos diferentes de recursos geoespaciales: conjuntos de datos espaciales, series de conjuntos de datos espaciales y servicios (Specka et al., 2019).

2.6.8. Perfil Ecuatoriano de Metadatos (PEM)

En el territorio ecuatoriano, con la colaboración de varias organizaciones, en el año 2010 se desarrolló el Perfil Ecuatoriano de Metadatos para catalogar productos, el mismo que se define como un documento basado en las normas ISO 19115:2003, ISO 19115-2:2009 e ISO 19139:2005 que muestra, con base en un análisis exhaustivo y participativo, los acuerdos que se han logrado tomando como referencia la experiencia de varias instituciones en este tema (Secretaría Nacional de Planificación y Desarrollo, 2016) (Miyamoto, 2021).

El principal objetivo del PEM es instaurar todas las especificaciones técnicas necesarias que las instituciones públicas o privadas deben cumplir para la generación de IG; de este modo obtener mayor funcionalidad en la búsqueda, acceso, localización, utilidad y distribución de datos (Atapuma et al., 2020) (IEDG, 2010).

A continuación, se presenta las secciones que componen el PEM:

1. Información del metadato (obligatorio): Indica la entidad raíz que define el metadato.
2. Identificación (obligatorio): Información básica para identificar de modo único los datos o producto terminado.

3. Restricciones (opcional): Información concerniente a las restricciones existentes sobre los datos y puede ser especificada para información de restricciones legales y/o información de restricciones de seguridad.
4. Calidad de los datos (obligatorio): contiene una valoración general de la calidad del conjunto de datos; es decir, información sobre la calidad de los datos especificados o producto terminado.
5. Información de mantenimiento: Información sobre el alcance y la frecuencia de actualización de los datos.
6. Información de la representación espacial (opcional): Información sobre el mecanismo usado para representar información espacial en un conjunto de datos.
7. Información del sistema de referencia (opcional): Descripción del sistema de referencia espacial usado en el conjunto de datos o producto terminado. La información del sistema de referencia pretende identificar el tipo de coordenadas utilizadas, ya sean geográficas, locales o proyecciones cartográficas.
8. Información del contenido (opcional): Información sobre el catálogo de características de los datos y su descripción. La información del contenido pretende detallar las entidades y atributos de los datos descritos.
9. Identificación del catálogo de representación (opcional): Información que identifica el catálogo de representación utilizado. Tiene un elemento obligatorio que se emplea para especificar el catálogo de representación usado por el conjunto de datos.
10. Información de distribución (opcional): Información sobre el distribuidor y las opciones para obtener el producto terminado. La información de distribución tiene por objeto dar la información necesaria para que la solicitud de datos pueda llevarse a cabo.
11. Información de extensión de metadato (opcional): Información acerca de la especificación de extensiones definidas según las necesidades del usuario.
12. Información del modelo de aplicación (opcional): Este paquete contiene información sobre el modelo de aplicación utilizado para construir un conjunto de datos.

2.7. Cosechador o recolector de metadatos

Un cosechador de metadatos geoespacial es una herramienta que permite recolectar y consolidar información geográfica de diferentes catálogos de metadatos en un solo lugar a través de diferentes protocolos y estándares como OAI-PMH, OGC CSW (Open Geospatial

Consortium Catalogue Service for the Web.), entre otros (GeoNetwork Documentation OGC Services, n.d.; GeoNode Documentation GC Web Services, n.d.).

Los beneficios de integrar un cosechador de metadatos es la optimización de recursos como es ahorro de tiempo, mejora la precisión de los datos, garantiza la calidad de la información. Además, se puede identificar cada conjunto de datos sin ambigüedades al proporcionar información detallada y facilitan la distribución de información; de este modo, reducen la carga de trabajo del usuario final (Chaglla & Morocho, 2016).

3. Estado del arte

En esta sección se realizará una revisión a los diferentes documentos sobre la implementación de un cosechador de metadatos haciendo uso de herramientas de Sistemas de Información Geográfica (SIG), proveyéndonos diferentes ideas y perspectivas sobre el tema a trabajar. En esta parte se realizó una revisión sistemática de la literatura, de tal manera se pueda recopilar y sintetizar toda evidencia científica relacionado al tema.

3.1. Aplicación de la metodología PRISMA-ScR

La metodología usada para el proceso de investigación y revisión de la literatura es PRISMA-ScR (Prisma for Scoping Review) (Tricco et al., 2018). En la Figura 3.1 se puede observar este proceso conformado de las siguientes etapas o pasos metodológicos: 1) Definir preguntas de investigación, 2) La búsqueda sistemática, el cual describe el uso de términos de búsqueda, 3) Selección de fuentes relevantes, donde se toma en cuenta los criterios de inclusión y exclusión para los diferentes documentos o artículos resultantes de las bases de datos, 4) Extracción de datos, el cual se extrae detalles de cada estudio por medio de una plantilla de extracción, 5) Resumen de los resultados, 6) Discusión y para terminar las 7) Conclusiones.

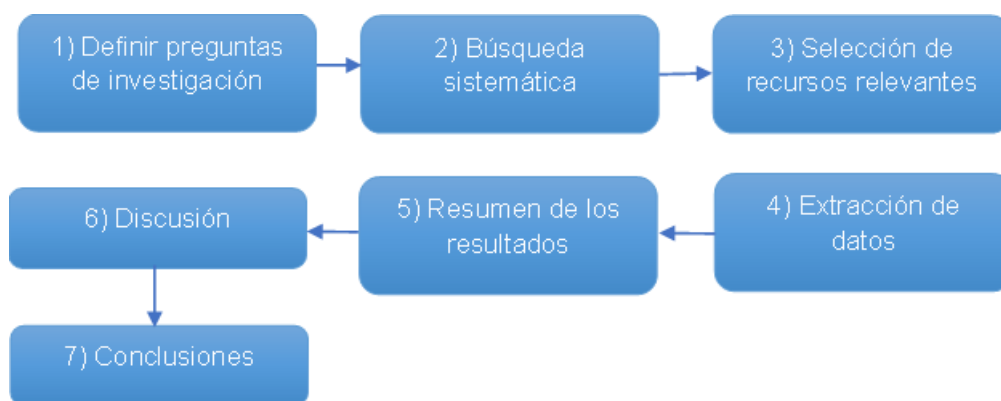


Figura 3.1. Diagrama de las etapas del proceso de investigación PRISMA-ScR. Fuente: Autoría propia

3.2. Definición de preguntas de investigación

La primera etapa consiste en definir las preguntas de investigación siguiendo el modelo PRISMA-ScR. Estas preguntas deben ser formuladas de manera precisa, enfocándose en una pregunta principal para la realización de una revisión sistemática. Estos interrogantes están diseñados para abordar los objetivos específicos del estudio de revisión y desempeñan un papel clave al guiar todo el proceso de investigación y revisión del material disponible.

A continuación, se define la pregunta principal de investigación.

P.I.: ¿Cómo el cosechamiento de metadatos puede mejorar la búsqueda de geoinformación?

En base a la pregunta principal, se definieron las siguientes preguntas secundarias que nos van a ayudar para dar una respuesta con más detalle a la pregunta principal, de tal forma conocer cómo puede mejorar o de qué manera mejora la búsqueda de los metadatos de información geográfica por medio de un cosechador de metadatos:

P.I.S. 1; ¿Qué características, ventajas y desventajas tienen los gestores de metadatos geoespaciales existentes?

P.I.S. 2; ¿Que se necesita para generar una arquitectura de software para un cosechador de metadatos geoespaciales?

P.I.S. 3; ¿Cuale son las buenas prácticas para la implementación de un cosechador de metadatos geoespaciales?

P.I.S. 4; ¿Cómo se puede evaluar la calidad y la eficacia de la cosecha de metadatos para garantizar la validez y la fiabilidad de los datos recolectados?

3.3. Búsqueda sistemática

Antes de llevar a cabo la búsqueda sistemática, es fundamental definir las palabras clave que se utilizarán en las distintas bases de datos, tales como Science Direct, Scopus, ACM Digital Library, IEEE Xplore, Google Scholar, entre otras..

Tabla 3.1. Cadenas de búsqueda. Fuente: Autoría propia

Tipo	Palabra clave	Palabra clave similar	Palabra clave español
Compuesta	geospatial AND metadata	geo spatial AND metadata	metadatos AND geoespaciales
Compuesta	geospatial AND harvesting	geo spatial AND harvesting	geoespacial AND recolección
Compuesta	metadata AND spatial AND harvesting	metadata AND spatial AND harvest	metadatos AND espacial AND recolección
Compuesta	geospatial AND metadata AND harvesting	geospatial AND metadata AND harvest	geoespacial AND metadatos AND recolección
Compuesta	geospatial AND ide	geo spatial AND ide	geoespacial AND ide
Compuesta	geospatial AND metadata AND ide	geo spatial AND metadata AND ide	geoespacial AND metadatos AND ide
Simple	Metadata	Metadata	Metadatos
Compuesta	tools AND managment AND metadata	tools AND managment AND metadata	herramientas AND gestión AND metadatos

La búsqueda se realiza utilizando las palabras clave descritas en la Tabla 3.1, empleando el operador AND para identificar documentos que contengan ambas palabras en los resultados. Esta búsqueda se efectúa en cada una de las bibliotecas digitales.

Posteriormente, se aplican criterios de inclusión y exclusión basados en el modelo PRISMA para seleccionar los estudios relevantes encontrados en la búsqueda sistemática.

3.3.1. Criterios de inclusión

Los criterios de inclusión definidos fueron los siguientes:

- Se ha definido un período de estudio entre 2018 y 2023, considerando la implementación de una tecnología que trabajará con metadatos. Para este propósito, se requiere utilizar información actualizada que no exceda los 5 años de antigüedad.
- Los resultados se basarán en las áreas de Ciencias de la Computación e Ingeniería, dado que estas disciplinas cuentan con una mayor cantidad de investigaciones relacionadas con el tema.
- Se considerarán tanto artículos como investigaciones escritas en inglés y español.
- Para la selección de trabajos, se dará relevancia a aquellos publicados en revistas especializadas en información geográfica y áreas afines a la geoinformación, así como a documentos de conferencias.
- Los trabajos que aborden temas sobre la calidad del cosechamiento de metadatos y el uso de estándares que contribuyan a la mejora del trabajo.
- Se incluirán trabajos relacionados con el uso e implementación de cosechadores de metadatos geoespaciales y herramientas que faciliten dicho proceso.
- Asimismo, se considerarán documentos que mencionen la mejora o actualización de una IDE para la gestión de datos espaciales.

3.3.2. Criterios de exclusión

Los criterios de exclusión definidos fueron los siguientes:

- Repetición de estudios en diferentes bases de datos.
- Resultados que carezcan de referencia a la fuente o no estén disponibles.
- Verificación de títulos de los documentos resultantes.

- Lectura de abstracts o resúmenes de los documentos.

3.3.3. Fuentes de información

En esta sección, se describen las bases de datos utilizadas y las fuentes de información empleadas para realizar la búsqueda de documentos relacionados con el tema de investigación. En la Tabla 3.2 se presentan las bases de datos seleccionadas, el período analizado, los tipos de documentos considerados para la revisión sistemática y los criterios de búsqueda descritos en las secciones 3.3.1 y 3.3.2.

Tabla 3.2. Fuentes de información y bases de datos. Fuente: Autoría propia

Base de datos	Período analizado	Tipo de documento	Criterio de búsqueda
Scopus	2018 - 2023	Artículos científicos e investigaciones	Fecha publicación, título, idioma, área investigación, abstract.
IEEE Xplore			
ACM Digital Library			
Science Direct		Tesis	
Fuentes externas: Google Scholar, Revista Cartográfica			

3.3.4. Estrategia de búsqueda

En esta sección, se presentan los primeros resultados obtenidos mediante la utilización de términos de búsqueda relevantes en diversas bases de datos mencionadas anteriormente. Se realizó una búsqueda en fuentes externas utilizando el motor de búsqueda Google Scholar, además de consultar bases de datos de universidades ecuatorianas como la Universidad San Francisco de Quito y revistas especializadas como La Revista Cartográfica, que contienen diversos documentos científicos provenientes de bases de datos como Scielo, Latindex y Dialnet, relacionados con la información geoespacial en general (Revista Cartográfica, n.d.).

Este trabajo de tesis se basa en la puesta en funcionamiento de tecnología, utilizando investigaciones realizadas desde 2018 hasta marzo de 2023. Para abordar conceptos y definiciones relacionadas con el tema, se ha empleado información que abarca el período desde 2010 hasta 2023.

En la Tabla 3.3 se presentan los resultados obtenidos mediante el criterio de búsqueda de las palabras clave previamente definidas en la sección 3.3. Estos resultados permiten identificar los estudios primarios relevantes y contribuyen a una búsqueda sistemática más eficiente de la literatura pertinente.

Tabla 3.3. Resultados de búsqueda en bases de datos y fuentes externas. Fuente: Autoría propia

Base de datos	Resultados
Scopus	45
IEEE Xplore	62
ACM Digital Library	173
Science Direct	121
Fuentes externas	40
TOTAL	441

Para obtener estos resultados, se utilizó una tabla que detalla los criterios de búsqueda aplicados en cada base de datos, incluyendo el rango de fecha de publicación, el área de investigación, el idioma y el tipo de documento. Al final de la tabla, se registró el número total de resultados obtenidos en cada base de datos. Con base en estos resultados iniciales, se llevó a cabo una selección de los documentos más relevantes para el presente trabajo, los cuales se encuentran detallados en el Anexo A.

3.3.5. Selección de estudios relevantes

En la selección de estudios relevantes, se consideraron principalmente los filtros disponibles en cada base de datos, como la fecha de publicación y el idioma (limitado a inglés y español debido a nuestra investigación). Además, se priorizó el área de investigación relacionada con Ingeniería y Ciencias de la Computación, ya que se observó que estos campos presentaban un mayor número de resultados en comparación con otras áreas, como Medio Ambiente, Literatura y Medicina.

Asimismo, la revisión de los títulos de los documentos fue esencial para identificar rápidamente aquellos trabajos que podrían tener relevancia para nuestro caso de estudio sobre cosechadores de metadatos con el propósito de mejorar una IDE.

Los resultados obtenidos de las fuentes externas se evaluaron utilizando los mismos criterios de búsqueda. Es importante destacar que la base de datos de Google Scholar arrojó una mayor cantidad de resultados en comparación con otras bases de datos. No obstante, se procedió a realizar una revisión de títulos y resúmenes, priorizando aquellos documentos

relacionados con el tema de investigación o provenientes de revistas pertinentes al área de estudio.

Para llevar a cabo este proceso, se utilizó una tabla de Excel en la que se registraron todos los documentos obtenidos, identificándolos con la cadena de búsqueda, el año y el título del documento. A continuación, se revisaron los posibles documentos duplicados, comenzando por filtrar los títulos para verificar su relevancia con nuestro caso de estudio. Para evitar duplicados debido al idioma del documento, se utilizó el identificador de objeto digital (DOI) para obtener los documentos únicos. Posteriormente, se procedió a la lectura de los abstracts o resúmenes de estos documentos.

3.3.6. Diagrama de flujo PRISMA

Para la documentación del proceso, nos hemos basado en el diagrama de flujo PRISMA. Este diagrama nos permite describir el proceso de identificación de documentos o artículos en diversas bases de datos.

El diagrama de flujo PRISMA está compuesto por diferentes secciones o etapas, cada una detallando las actividades realizadas y el flujo seguido para llegar a la selección de documentos relevantes. A continuación, procederemos a identificar y explicar cada una de estas secciones del diagrama de flujo.

- **Identificación:** esta sección inicia con la recopilación de todos los resultados obtenidos de las bases de datos, incluyendo aquellos provenientes de fuentes externas como revistas científicas.
- **Selección:** consiste en obtener un conjunto único de resultados tras eliminar duplicados. Para lograrlo, se empleó un identificador único digital presente en cada publicación, así como los títulos de las mismas.
- **Elegibilidad:** en esta fase del diagrama, se aplicaron criterios de inclusión y exclusión para obtener los resultados más significativos que se utilizarán en el desarrollo del trabajo.
- **Incluidos:** Estos son los documentos finales que presentan mayor relevancia y serán sometidos a una lectura exhaustiva.

En la Figura 3.2 se puede observar la estructura del diagrama de la búsqueda sistemática que se realiza y el total de resultados en cada sección.

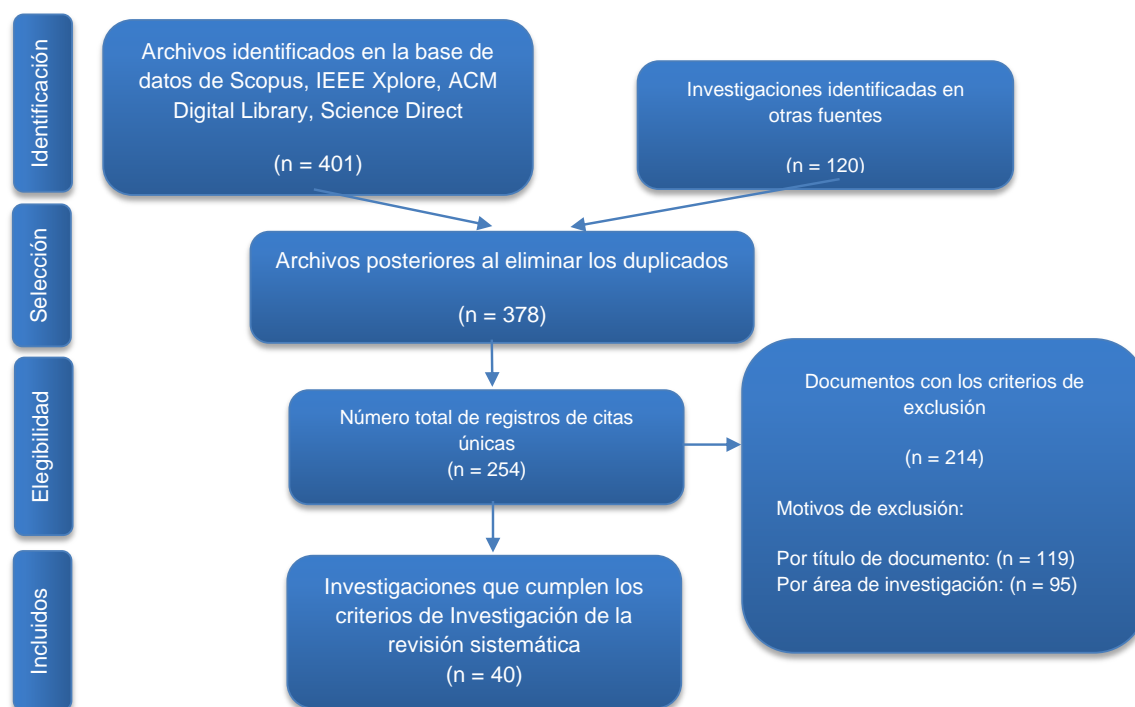


Figura 3.2. Diagrama de flujo PRISMA. Fuente: Autoría propia

3.4. Extraer datos de contexto

Tras la selección de los documentos de investigación pertinentes mediante criterios de inclusión, procedemos a la revisión de los abstracts o resúmenes de cada documento obtenido. De acuerdo con la metodología PRISMA, este proceso puede resultar en una recopilación de entre 20 y 60 artículos.

3.4.1. Definición de variables a analizar

Las variables relacionadas con las preguntas secundarias de investigación se detallan en la Tabla 3.2. Fuentes de información y bases de datos. Fuente: Autoría propia

Tabla 3.4. Definición de variables. Fuente: Autoría propia

Pregunta de investigación	Variables definidas
P.I.S. 1: ¿Qué características, ventajas y desventajas tienen los gestores de metadatos geoespaciales existentes?	Funcionalidad
	Flexibilidad
	Integración
	Interfaz de usuario

	Escalabilidad
	Soporte y mantenimiento
	Costo
P.I.S. 2: ¿Que se necesita para generar una arquitectura de software para un cosechador de metadatos geoespaciales?	Fuente de datos
	Protocolos de recolección de datos
	Estructura de los datos
	Tecnologías utilizadas
	Almacenamiento de datos
P.I.S. 3: ¿Cuales son las buenas prácticas para la implementación de un cosechador de metadatos geoespaciales?	Estándares de metadatos
	Definición de términos
	Formatos de datos
	Consistencia
P.I.S. 4: ¿Cómo se puede evaluar la calidad y la eficacia de la cosecha de metadatos para garantizar la validez y la fiabilidad de los datos recolectados?	Precisión de los metadatos
	Consistencia de los metadatos
	Actualidad de los metadatos
	Facilidad de uso de los metadatos
	Compatibilidad
	Actualización de metadatos

3.4.2. Lectura completa de los documentos

Realizando un análisis exhaustivo de los artículos seleccionados, se ha creado una tabla que resume cada artículo con su respectivo objetivo, variables definidas, estructura y resultados principales. La tabla incluye el título del documento, un enlace directo a la fuente, los nombres de los autores, el objetivo general del trabajo, la idea principal y los resultados obtenidos. Además, se incluyen las variables específicas a las cuales hace referencia el documento que se ha leído detalladamente. Se puede encontrar esta tabla como Anexo B adjunto al final del documento.

3.4.3. Resultados de documentos seleccionados

En esta sección se presentarán los resultados obtenidos a partir del análisis de los documentos mencionados en la sección 3.4.2. Estos resultados, obtenidos mediante un análisis cuantitativo, nos permitirán responder a las preguntas de investigación planteadas.

En la Figura 3.3, se muestra un diagrama de barras que compara el número total de documentos encontrados con el total de documentos seleccionados después de aplicar los criterios de búsqueda. Esta comparación permitirá evaluar la adecuación de los documentos seleccionados en relación con el conjunto inicial de resultados obtenidos en la búsqueda.

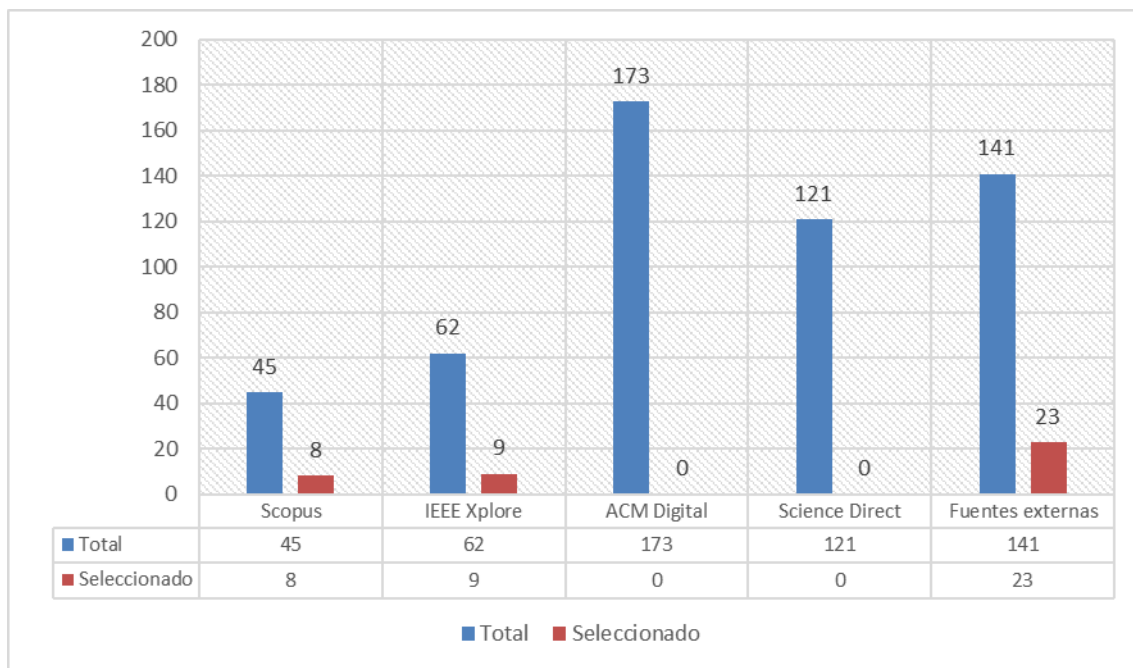


Figura 3.3. Comparación de documentos seleccionados y resultados primarios. Fuente: Autoría propia

A continuación, se presentan los resultados relacionados con las preguntas planteadas acerca de los gestores de metadatos y las implementaciones de cosechadores de metadatos geospaciales encontrados. Estos resultados se han resumido en una tabla para facilitar su visualización

- P.I.S. 1; ¿Qué características, ventajas y desventajas tienen los gestores de metadatos geospaciales existentes?

Tabla 3.5. Resultado de herramientas para cosechamiento de metadatos. Fuente: Autoría propia

Herramienta	Características	Ventajas	Desventajas	Documentos
GeoNetwork	<p>Soporta múltiples estándares de metadatos</p> <p>Permite la búsqueda y visualización de metadatos</p> <p>Herramienta personalizable</p>	<p>Facilita el descubrimiento y acceso a metadatos</p> <p>Comunidad activa y soporte técnico</p> <p>Interfaz intuitiva y fácil de usar</p>	<p>Necesita de recursos de hardware y capacidad de almacenamiento</p> <p>Necesita configuración y personalización</p>	6 documentos seleccionados en total
CatMEdit	<p>Hace uso de estándares ISO 19115 y FGDC</p> <p>Integración con sistemas de catálogo y repositorios</p> <p>Opciones para edición y validación de metadatos</p>	<p>Herramienta enfocada la edición y validación de metadatos</p> <p>Cumplimiento de estándares de calidad de forma sencilla</p> <p>Multiplataforma para sistemas operativos</p>	<p>Requiere conocimientos técnicos y capacitación para su uso</p> <p>Enfoque más orientado a la edición que a la recolección de metadatos</p>	1 documento seleccionado en total

Se observa que un número significativo de documentos ($n = 6$) hacen referencia al uso y utilidad de la herramienta GeoNetwork para la gestión de metadatos, así como la disponibilidad para su cosechamiento. Solo un estudio menciona la herramienta CatMEdit, mientras que los demás abordan el manejo de estándares y calidad de metadatos para instituciones públicas y privadas, pero no utilizan una herramienta específica para gestionar y recolectar metadatos.

- P.I.S. 2; ¿Que se necesita para generar una arquitectura de software para un cosechador de metadatos geoespaciales?

Tabla 3.6. Resultado de componentes para una arquitectura de un cosechador. Fuente: Autoría propia

Componente	Descripción	Total de documentos
Interfaz de usuario (UI)	Una interfaz para que los usuarios puedan interactuar con las diferentes opciones disponibles en la herramienta	5 documentos seleccionados en total
Almacenamiento de metadatos	Herramientas que hagan uso de bases de datos geoespaciales o manejo de archivos. De tal manera se pueda realizar una búsqueda y recuperación de metadatos para un uso posterior.	13 documentos seleccionados en total
Servidores de aplicaciones	Herramientas que necesitan de un servidor para desplegarlo en la web u otras que son instalables localmente en el sistema	3 documentos seleccionados en total
Conectividad en la red	El establecer y gestionar las conexiones con las fuentes de metadatos geoespaciales es primordial si realiza cosechamiento de diversas fuentes.	7 documentos seleccionados en total

Programación y planificación	Tenga opciones o funcionalidades para planificar tareas de cosechamiento de forma automática.	1 documentos seleccionados en total
-------------------------------------	---	-------------------------------------

- P.I.S. 3; ¿Cuales son las buenas prácticas para la implementación de un cosechador de metadatos geoespaciales?

Tabla 3.7. Buenas prácticas para una arquitectura de un cosechador. Fuente: Autoría propia

Buena practica	Descripción	Total de documentos
Validación de metadatos	Hacer uso de procesos u opciones que validen y garanticen la calidad de los metadatos que se han recolectado	6 documentos seleccionados en total
Uso de estándares	Uso de estándares validados y establecidos, como ISO 19115/119, Dublin Core, para la consistencia y compatibilidad de los metadatos recolectados.	20 documentos seleccionados en total

- P.I.S. 4; ¿Cómo se puede evaluar la calidad y la eficacia de la cosecha de metadatos para garantizar la validez y la fiabilidad de los datos recolectados?

Tabla 3.8. Metodologías para evaluar la calidad de la cosecha. Fuente: Autoría propia

Método	Descripción	Total de documentos
Análisis de la frecuencia de actualización	Verificar la frecuencia que actualizan los metadatos, revisando si los cambios se pueden visualizar correctamente	1 documentos seleccionados en total
Revisión manual de metadatos	Hacer una revisión manual de los metadatos que se recolectaron para identificar inconsistencias, errores, o posibles datos que pueden faltar.	1 documentos seleccionados en total

Los resultados obtenidos muestran un valioso aporte (n= 20) en el uso de estándares para la creación, gestión y almacenamiento de metadatos, obtenidos durante la captura de IG. Es esencial almacenar estos metadatos en bases de datos espaciales o sistemas de archivos compatibles con los sistemas operativos disponibles y las herramientas utilizadas.

A pesar de que la búsqueda sistemática se inició en 2018, aproximadamente el 60% de los artículos y documentos seleccionados se publicaron en los años 2018 y 2019. Entre los demás años, se observa un mayor énfasis en el tratamiento de los metadatos, siguiendo estándares y reglas para mantener una información de alta calidad, tal como se expone en el trabajo de Vedor De Paula sobre el control de calidad de los datos geoespaciales. Este estudio se fundamentó en recomendaciones técnicas y trabajos científicos que describen procedimientos para el control de calidad en los metadatos, exactitud posicional, geometría,

entre otros aspectos (Vedor De Paula et al., 2019). Los detalles sobre los documentos técnicos referentes al control de calidad se presentan en la Tabla 3.8.

Tabla 3.9. Documentos técnicos y trabajos científicos de referencia para el control de calidad de datos geoespaciales. Fuente: (Vedor De Paula et al., 2019)

Fuente	Descripción
(ISO, 2003a)	Determina los criterios para la evaluación de la calidad de los datos geoespaciales
(ISO, 2003b)	Principios de metadatos geoespaciales
(ISO, 2013)	Establece los principios para describir la calidad de los datos geoespaciales (actualización de la ISO 19114)
(CONCAR, 2009)	Describe procedimientos para rellenar metadatos de datos geoespaciales
(DSG, 2011)	Especificación técnica para la adquisición de datos geoespaciales vectoriales
(DEVILLERS; JEANSOULIN, 2006)	Presenta conceptos de calidad de datos geoespaciales, con el objetivo de minimizar los riesgos de utilización indebida de datos en la toma de decisión

En el análisis de los documentos de control de calidad, se destaca una atención detallada en la especificación y criterios para el llenado de metadatos, asegurando el cumplimiento de estándares internacionales y nacionales.

3.5. Análisis y síntesis de los resultados

En los últimos años, ha habido un enfoque constante en la mejora de las IDEs. Se han realizado diversos estudios que destacan el uso de herramientas GIS para la administración integral de la IG, desde su adquisición mediante captura de sensores, drones y otros dispositivos, hasta su procesamiento y generación de servicios y archivos para su posterior publicación en servidores, con el propósito de utilizarlos en investigaciones.

En la revisión de la literatura, se han identificado 40 documentos que abordan la gestión de metadatos mediante el uso de herramientas GIS. Estos estudios se centran en definir y consolidar la información geoespacial junto con sus respectivos metadatos, siguiendo los estándares internacionales ISO 19115 y 19115-2, especializados en la creación de metadatos y servicios geoespaciales. Además, se ha considerado la norma ISO 19139 para la creación del esquema XML basado en la ISO 19115 y 19119 (Carrascosa Basterra, 2018).

Entre las características más relevantes identificadas en los documentos seleccionados se encuentran las siguientes:

- Estándares y normas para la gestión de IG

Estudios mencionaban que una forma de mantener una estructuración adecuada de la información geoespacial y los metadatos, es basándose en las normal y estándares nacionales e internacionales que se hayan establecido. Esto se debe a que la información que generan puede ser muy diversa y no tener una estructura adecuada y ordenada para presentar la información; por tal motivo, se hace uso de estándares ampliamente aceptados, como el Protocolo de Descubrimiento y Acceso a Metadatos Geoespaciales (OGC CSW).

En el trabajo titulado "Desarrollo e implementación de un marco de generación automática de metadatos para SDI utilizando software de código abierto: un estudio de caso de la Infraestructura de Datos Espaciales de la India (NSDI)", se ha desarrollado un framework y un estándar de metadatos para gestionar automáticamente elementos de metadatos mediante el uso de software de código abierto (OSS) para la Infraestructura de Datos Espaciales del Gobierno de la India (NSDI). Utilizando estos estándares, se ha aplicado un prototipo de IDE de salud pública en la ciudad de Prayagraj, lo que beneficiará la búsqueda, descubrimiento, acceso y visualización de datos y metadatos geoespaciales (Agrawal et al., 2022).

En Latinoamérica, la tesis titulada "Elaboración de Bases de Datos Geográficos para la Planificación y Gestión de Áreas Protegidas" emplea herramientas SIG para la creación de una base de datos espacial. El trabajo destaca el uso del Perfil de Metadatos Geoespaciales de Brasil (Perfil MGB) y proporciona una detallada explicación de los campos de metadatos requeridos y opcionales para la estructura de la información (Vedor De Paula et al., 2019).

Por otro lado, en nuestro país, un trabajo menciona que en la ciudad de Santo Domingo existen datos geográficos o IG dispersos en diversas instituciones públicas y privadas, como municipios, cámaras de turismo, centros comerciales, entre otros. Estos datos se encuentran en formatos no estandarizados tanto en su generación como en su visualización (Meneses et al., 2018).

- Herramientas usadas

En los estudios seleccionados, se ha identificado el uso de diversas herramientas SIG, como GeoNetwork y ArcGIS, para la creación, estandarización y almacenamiento de datos geoespaciales, así como para la elaboración de productos cartográficos. Además, la herramienta CatMDEdit para la edición de metadatos y sus respectivas secciones. Estos trabajos también hicieron uso de estándares reconocidos, tales como ISO 19115, ISO 19139, ISO 19112, el estándar Dublin Core y la Normativa Ecuatoriana de desarrollo de IDES (Espinoza Martínez, 2017).

Una parte de los documentos resultantes ($n = 6$) se han referido a la herramienta GeoNetwork como un gran apoyo para proyectos en la gestión de metadatos. Además, de gobiernos o instituciones gubernamentales que hacen uso de la misma para la publicación y compartición de metadatos. Por otro lado, la herramienta enfocada para cosechamiento de metadatos y administración de los mismos de forma centralizada no se ha mencionado en estos mismos estudios.

Un estudio relevante realizado por Chaglla y Morocho en 2016, describe la implementación de un cosechador de metadatos en GeoNetwork con el propósito de mejorar el acceso a información en una IDE subnacional. El enfoque consiste en la integración de nodos, cada uno de ellos con su propio servidor GeoNetwork, permitiendo la recolección de Información Geográfica (IG) y su almacenamiento en un repositorio centralizado (Chaglla & Morocho, 2016).

Con esto se puede evidenciar que aún no se ha explorado a fondo la capacidad de la herramienta, ni la forma en que administra de manera centralizada los metadatos provenientes de diversas fuentes.

En el documento titulado "REPOSITORIO DE RIESGOS ASOCIADOS A FENÓMENOS NATURALES Y ANTRÓPICOS EN MÉXICO", se emplea la herramienta GeoNetwork para proporcionar una plataforma flexible y de acceso gratuito a través de Internet. La interfaz amigable de esta herramienta la hace especialmente útil para la toma de decisiones tanto en instituciones públicas como privadas. Su objetivo es ayudar a entender y evaluar los diversos riesgos generados por fenómenos naturales y antrópicos, así como para conocer las medidas preventivas establecidas para hacerles frente (Fernández Eguiarte et al., 2019). En el trabajo titulado 'Implementación de una plataforma de Infraestructura de Datos Espaciales', se destaca la importancia de utilizar el 'Catalog Service of Web' a través de la herramienta GeoNetwork. Esta aplicación resulta indispensable para establecer las características de la información vectorial y ráster según las normas ISO 19115:2003 y ISO 19115-2:2009 (Astudillo Vargas, 2018).

En un trabajo aborda la recolección de metadatos a partir de la selección de protocolos de la fuente principal, el catálogo GeoNetwork de la Infraestructura de Datos Espaciales de la República Argentina (IDERA). El objetivo principal de este trabajo es proporcionar datos geográficos (metadatos) de la IDE para fines educativos (Freddo et al., 2021).

Por otro lado, en el trabajo de Aníbal, se menciona la herramienta CatMDEdit, la cual propone un método para generar metadatos que cumple con los estándares ISO y facilita la creación de metadatos para archivos vectoriales y ráster. El uso de CatMDEdit mejora

significativamente la gestión de metadatos y optimiza los recursos involucrados en el proceso (Aníbal et al., 2018).

- Calidad en la creación y gestión de los metadatos

Existen trabajos que hacen referencia a la importancia de la calidad de los metadatos al obtener información geoespacial a partir de ortomosaicos, fotografías satelitales o datos capturados por sensores, y cómo se aplican en su estudio. Sin embargo, la mayoría de estos trabajos mencionan de manera general que la calidad de los metadatos es fundamental en el levantamiento de información geoespacial. Algunas instituciones no cumplen con las normas establecidas por el gobierno, como INSPIRE en España o PEM en Ecuador, que definen los parámetros o atributos necesarios para obtener metadatos con información geográfica o georreferenciada adecuada (Vedor De Paula et al., 2019).

En cuanto a la calidad del cosechamiento, ninguno de los estudios seleccionados la menciona, excepto un documento que aborda la importancia de tener metadatos de calidad y cómo completarlos adecuadamente para generar IG de forma adecuada.

- Planificación territorial

En el campo de la planificación territorial, la aplicación de la información geoespacial ha demostrado ser especialmente valiosa. Se le ha dado un enfoque significativo a los metadatos, que proporcionan geoposicionamiento en tiempo real o coordenadas, generando datos precisos y exactos mediante sistemas de captura de información satelital. Por ejemplo, uno de los trabajos menciona que, en el ámbito gubernamental, la Secretaría Nacional de Planificación y Desarrollo (SENPLADES) establece en sus documentos para el Desarrollo de Infraestructuras de Datos Espaciales la necesidad de construir una arquitectura basada en Internet, con el uso de protocolos y servicios web (Espinoza Martínez, 2017).

El trabajo "Diagnóstico para definir una propuesta de perfil ecuatoriano de catastro mediante la aplicación de la norma ISO 19152 (LADM): Caso de estudio en Barrios Miranda y La Católica", examina las metodologías y modelos para el catastro y ordenamiento territorial en las áreas seleccionadas. Además, se generaron metadatos para cada capa publicada en GeoServer mediante la herramienta GeoNetwork, siguiendo la plantilla del PEM correspondiente a cada capa (Atapuma et al., 2020).

Como se puede observar, estos estudios demuestran la relevancia de dar mantenimiento y gestionar adecuadamente la información geoespacial y sus metadatos. Sin embargo, persisten desafíos y oportunidades de investigación para optimizar y otorgar un mayor grado

de importancia a los cosechadores, especialmente abordando aspectos como la heterogeneidad de los datos y la continua actualización de los metadatos en una IDE.

El enfoque de ciertas instituciones u organizaciones en defender la gestión privada de los metadatos y la captura personalizada de información según las necesidades específicas puede ser comprensible. No obstante, es relevante resaltar que esta opción podría presentar ciertos inconvenientes significativos. Entre ellos, se encuentran los costos elevados asociados, así como las posibles restricciones en los servicios disponibles para determinados usuarios. Además, este enfoque podría requerir asignar una considerable cantidad de recursos para el desarrollo e implementación adecuada de la solución. Por lo tanto, es fundamental examinar cuidadosamente estos aspectos antes de decidir la estrategia de gestión de metadatos más adecuada para una IDE o sistema geoespacial.

3.6. Discusión y conclusiones de resultados

En base a los resultados que se obtuvieron en la sección 3.4.3; ahora se responderá las respuestas de las preguntas secundarias planteadas para el desarrollo del estado del arte.

P.I.S. 1; ¿Qué características, ventajas y desventajas tienen los gestores de metadatos geoespaciales existentes?

Entre sus características que podemos destacar, en base a la revisión literaria realizada, proporcionan un conjunto estructurado de campos para describir los datos geoespaciales, los cuales son como el tema, la escala, el sistema de coordenadas, así como el formato de información. Esta estructura de campos, está basada en un conjunto de directrices y estándares ya aprobadas a nivel internacional y nacional, como es el caso de PEM creado por el IGM para la descripción de metadatos geoespaciales en el país.

Una ventaja clave del uso de gestores de metadatos geoespaciales es la aplicación de estándares y normas reconocidas, como ISO 19115/19139, FGDC y Dublin Core. Estos estándares garantizan coherencia y compatibilidad entre diversas fuentes de datos y usuarios, aportando mayor valor y validez a los resultados obtenidos. Además, fomentan la colaboración entre diferentes usuarios y organizaciones que utilizan esta información. Por otro lado, los gestores de metadatos agilizan la búsqueda y acceso a la información geoespacial, optimizando la eficiencia operativa y reduciendo el tiempo dedicado a dichas tareas.

En cuanto a las desventajas, la implementación y configuración inicial de un gestor de metadatos puede resultar compleja y requerir recursos técnicos y de personal. Además, es importante considerar la necesidad de un mantenimiento constante para asegurar la

actualización de la información. Por otro lado, la herramienta puede tener una curva de aprendizaje para comprender la herramienta, lo que implica la necesidad de personal con conocimientos en la misma para lograr un funcionamiento óptimo.

Además, actualmente algunas instituciones buscan mejorar la disponibilidad y accesibilidad de su información geoespacial para el público en general, implementando servidores de mapas y servicios correspondientes. Esto enfatiza la importancia de compartir información geoespacial para que las personas sin conocimientos técnicos puedan acceder, buscar y filtrar datos según sus necesidades.

P.I.S. 2; ¿Qué se necesita para generar una arquitectura de software para un cosechador de metadatos geoespaciales?

En base al análisis se encontró que gran parte de los documentos seleccionados no hacen referencia a una arquitectura de software para un cosechador, pero según el único artículo realizado por Chaglla y Morocho, hace mención de una propuesta de arquitectura de cosechamiento en el cual proporciona un detalle de un esquema de la estructura en el cual está confirmado por un Nodo Central que es el servidor de CEDIA y los nodos del cual se realizara la recolección de los metadatos. En la Figura 3.4 se muestra el esquema que proponen los autores.

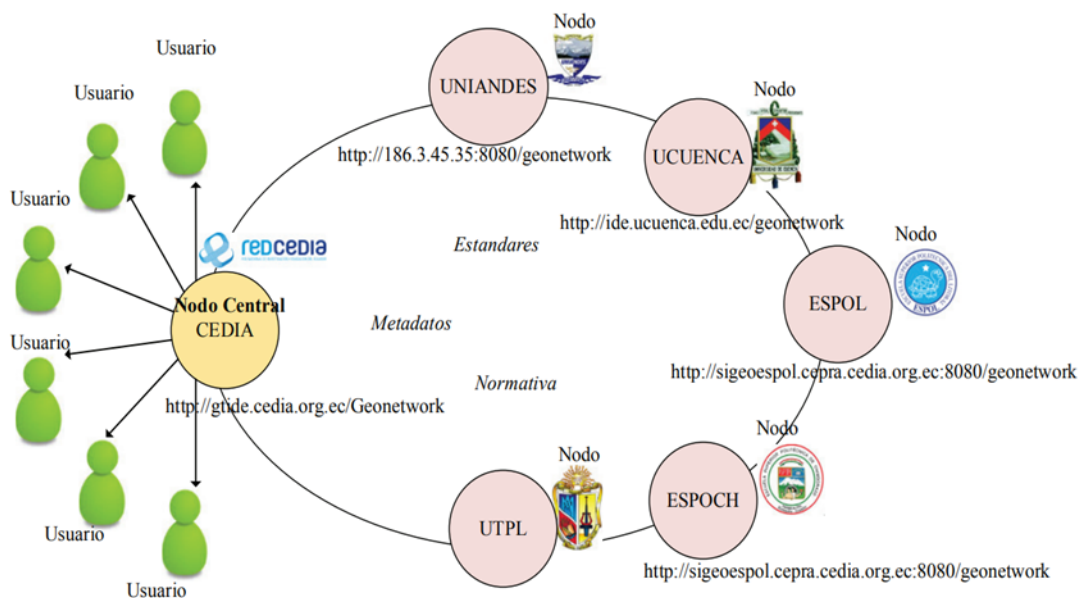


Figura 3.4. Esquema de IDE y nodos para la arquitectura. Fuente: (Morocho y Chaglla)

P.I.S. 3; ¿Cuáles son las buenas prácticas para la implementación de un cosechador de metadatos geoespaciales?

En los estudios analizados, se ha observado que una buena práctica es la verificación de la calidad de los metadatos antes de su publicación. Es fundamental seguir estándares y protocolos reconocidos en la industria para su generación y creación, como el estándar ISO 19139 para la codificación de metadatos geoespaciales y el protocolo OGC CSW (Catalogue Service for the Web) para su acceso y búsqueda. Asimismo, se debe considerar la escalabilidad del cosechador de metadatos para gestionar un volumen creciente de información a medida que la infraestructura mejore.

Por otro lado, es fundamental abordar la seguridad del cosechador, garantizando que cuente con los permisos y credenciales adecuados para acceder de manera segura a las fuentes de datos y proteger la privacidad de la información sensible. Esto implica configurar adecuadamente la red en la que se encuentra el cosechador y asegurar la robustez del sistema que actuará como servidor.

P.I.S. 4; ¿Cómo se puede evaluar la calidad y la eficacia de la cosecha de metadatos para garantizar la validez y la fiabilidad de los datos recolectados?

Para la evaluación de calidad y eficacia del cosechamiento, es fundamental asegurar que los metadatos cumplan con los estándares internacionales y nacionales, que definen los elementos esenciales para su correcta generación. Además, realizar una revisión minuciosa de los metadatos para garantizar su claridad, relevancia y adecuación a los datos geoespaciales asociados. De esta forma, se fomenta la reutilización de metadatos de calidad por parte de otros usuarios e instituciones.

Los resultados obtenidos demuestran que, a pesar de los avances significativos en la gestión de metadatos mediante herramientas y bases de datos geográficas, existe una escasez de información sobre el uso de cosechadores de metadatos geoespaciales para mejorar una IDE. Los hallazgos resaltan desafíos persistentes relacionados con la heterogeneidad de los datos, disponibilidad de servicios y recursos de los servidores, actualización de metadatos e interoperabilidad entre diferentes plataformas y sistemas. Estos hallazgos proporcionan una sólida base para comprender el panorama actual de los cosechadores de metadatos geoespaciales y ofrecen información valiosa para futuras investigaciones y desarrollos en este campo.

4. Metodología

En este capítulo, se presenta la implementación de un cosechador que mejora las búsquedas de IG en la IDE UCuenca. Se describe en detalle el diseño de la arquitectura del cosechador y su funcionamiento. Para este trabajo, se utiliza una metodología cuantitativa que incluye un análisis de las herramientas de gestión de metadatos con un enfoque basado en características. Luego, se selecciona una herramienta específica para implementar el diseño de la arquitectura del cosechador. Finalmente, se despliega la herramienta junto con el módulo del cosechador, considerando los recursos necesarios y la configuración adecuada para llevar a cabo la recolección de la información.

4.1. Análisis de herramientas con enfoque basado en características

El objetivo de este análisis comparativo es identificar una herramienta que administre los metadatos de forma sencilla y cumpla con los requisitos necesarios para la recolección o cosecha de metadatos.

Se propone realizar un análisis de características para identificar las más relevantes y evaluar la idoneidad de la herramienta según nuestras necesidades. Para ello, se lleva a cabo una revisión detallada del sistema o herramienta, describiendo sus características y funcionalidades. A continuación, se seleccionan las características más destacadas para la comparación y análisis de la herramienta elegida. Posteriormente, se asigna una puntuación a cada característica identificada, lo que permitirá determinar la herramienta más adecuada para nuestro caso de estudio.

4.1.1. Introducción de las herramientas a analizar

A continuación, examinaremos de manera concisa las herramientas que serán evaluadas y describiremos las impresiones iniciales y las descripciones que se pueden extraer directamente de sus sitios web, el grupo de herramientas que se consideran son los catálogos, editores y gestores de metadatos. De este modo, se obtuvo un primer vistazo de la herramienta.

- **CatMDEdit:** herramienta enfocada a la edición de metadatos y descripción de recurso de información geográfica, facilitando la documentación del mismo. Uso de estándares de edición y gestión de metadatos ISO 19115 y Dublín Core (ISO 15836). Presentación de metadatos usando HTML y archivos en Excel.
- **TKME:** editor de metadatos formales, se encarga de documentar de forma estructurada conforme al estándar de contenido para metadatos geoespaciales digitales. Fue

desarrollado por el Comité Federal de Datos Geográficos (FGDC) y puede ejecutarse en sistemas Unix o Linux, además de sistemas Windows.

- **Metavist:** software de edición de metadatos para el estándar de metadatos espaciales FGDC y el Perfil de Datos Biológicos. El software está disponible desde finales de 2004 y desarrollado en Visual Basic.Net. Actualmente tiene dos versiones la 1.0 que es estable y la 2.0 que está en versión Beta.
- **EPA Metadata:** esta herramienta se conoce también como el Editor de metadatos de la EPA (EME). La Agencia de Protección del medio ambiente de los Estados Unidos, sus siglas en inglés EPA, es quien se encarga de la toma de decisiones con respecto a las áreas de mayor riesgo ambiental y que a través del uso de IG pueden llevar a cabo la misión de proteger la salud humana y medio ambiente (About EPA's Geospatial Program | US EPA, n.d.). EPA metadata permite crear y editar registros que cumplen con los requisitos de metadatos de la EPA.
- **GeoNetwork:** aplicación de catalogo que gestiona recursos espaciales referenciados. Sus funciones a destacar son la edición y búsqueda de metadatos, así como el uso de un visor de mapas web.
- **Geonode:** es una aplicación y plataforma basada en la web para desarrollar sistemas de información geoespacial (GIS) y para implementar IDE.
- **PyCSW:** la herramienta habilita al usuario para desplegar catálogos de metadatos en línea, haciendo uso de formatos fundamentados en estándares reconocidos. Se configura de manera sencilla en un servidor web Apache, lo que posibilita el acceso a metadatos almacenados en forma de archivos XML dentro de una base de datos.
- **ESRI ArcGIS:** es un completo sistema que permite recopilar, organizar, administrar, analizar, compartir y distribuir información geográfica. Una de las plataformas reconocidas a nivel mundial para crear y utilizar SIG.
- **PgMetadata:** es un plugin para personas que hacen uso de QGIS como su aplicación GIS principal y PostgreSQL como su principal sistema de almacenamiento de datos vectoriales.

4.1.2. CatMDEdit

Esta herramienta tiene como objetivo principal la edición de metadatos para facilitar la documentación de recursos. Además, la gestión de diferentes repositorios de metadatos, la selección y filtrado de registros de metadatos almacenados.

4.1.2.1. *Características generales*

Uso del estándar "ISO 19115. Información Geográfica - Metadatos" (ISO 19115:2003/Cor 1 2006, ISO/TS 19139:2007 Codificación XML). Interfaces de edición adaptadas a diferentes perfiles de metadatos como NEM (Núcleo Español de Metadatos) definida por el Consejo Superior Geográfico Nacional de España, las Normas de implementación de INSPIRE para metadatos y el Perfil de metadatos WISE y sus siglas son "Sistema de Información del Agua para Europa" (WISE).

Generación automática de metadatos para algunos formatos de archivo de datos: Shapefile, DGN, ECW, FICC, GeoTiff, GIF/GFW, JPG/JGW, PNG/PGW (CatMDEdit, n.d.).

La licencia para la utilización de esta herramienta es de software libre, puede redistribuirlo y/o modificarlo según los términos de la Licencia pública general reducida de GNU publicada por la Free Software Foundation. Es útil pero no tiene garantía implícita de comerciabilidad o idoneidad para un propósito particular (CatMDEdit, n.d.). A continuación, se va a mostrar las opciones y funcionalidades que se pueden observar en la aplicación:

En la Figura 4.1 se puede apreciar la pantalla de inicio y la interfaz hasta el momento se puede observar que tiene dos ejemplos de metadatos con los estándares Dublín Core, ISO 19115/19119 e INSPIRE.

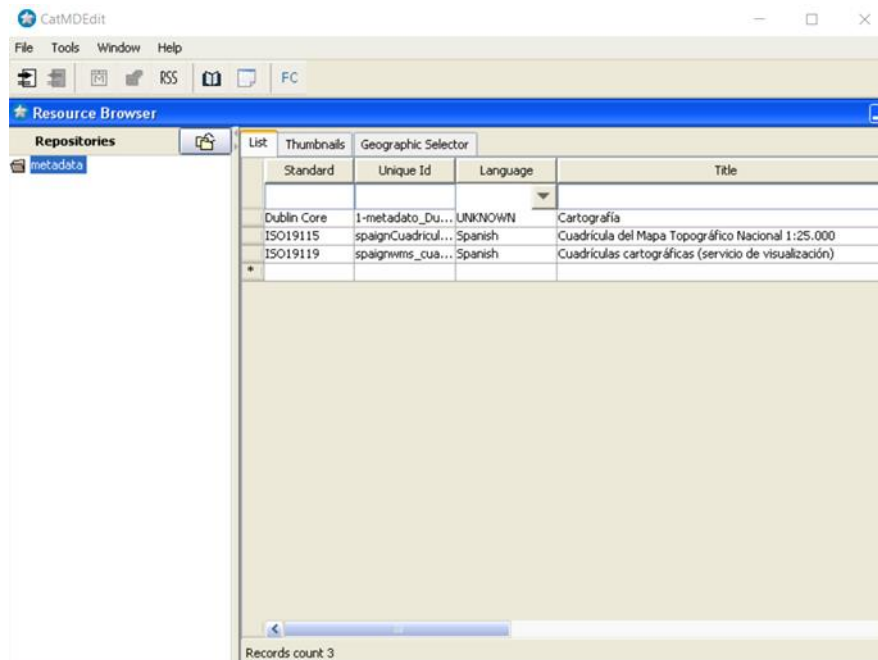


Figura 4.1. Pantalla inicial de CatMDEdit. Fuente: Autoría propia

La interfaz se divide en tres secciones distintas: en la parte superior, se ubica la barra de tareas que ofrece funcionalidades como la importación de metadatos desde archivos XML y la generación de GeoRSS, una extensión del formato Really Simple Syndication (RSS) utilizada por otros sistemas consumidores de dichos archivos. A la izquierda, se encuentra la sección de gestión de repositorios, que permite administrar los metadatos importados organizándolos en directorios. En la parte derecha, se presentan en forma de lista los metadatos pertenecientes al repositorio seleccionado. La Figura 4.2 muestra el proceso de importación de un metadato, en el cual se puede elegir entre el estándar predeterminado o seleccionar uno en particular.

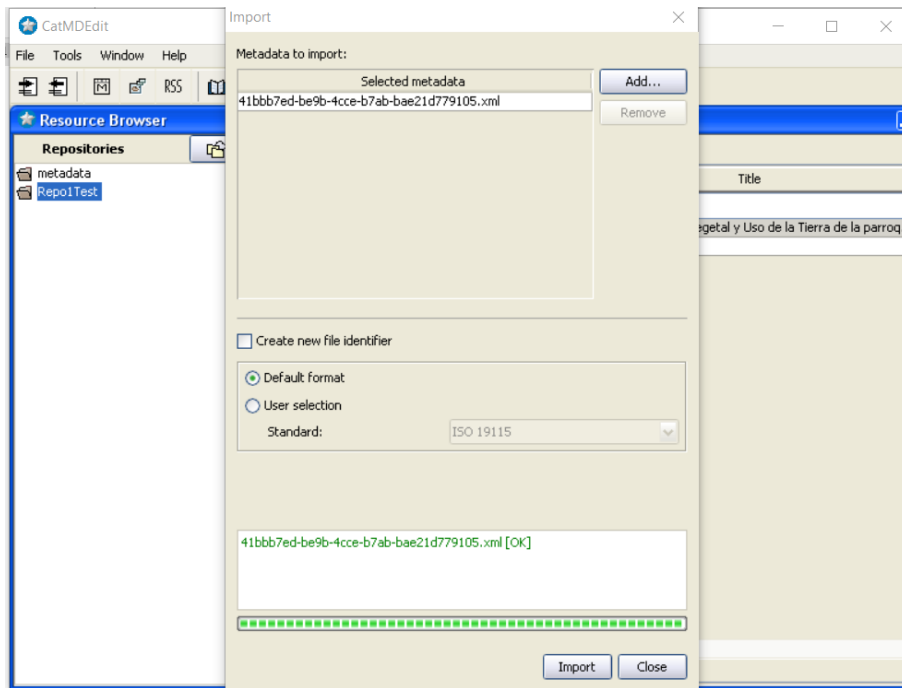


Figura 4.2. Importación de metadato en formato XML. Fuente: Autoría propia

Algo interesante que destacar es el detalle que se puede obtener al mostrar toda la información del metadato; para eso, debe seleccionar el metadato que se desea y con click derecho se tienen las opciones para gestionar el metadato como es actualizar, abrir metadato, abrir recurso. En la Figura 4.3 se muestra las ventanas que se abren al momento de buscar más información.

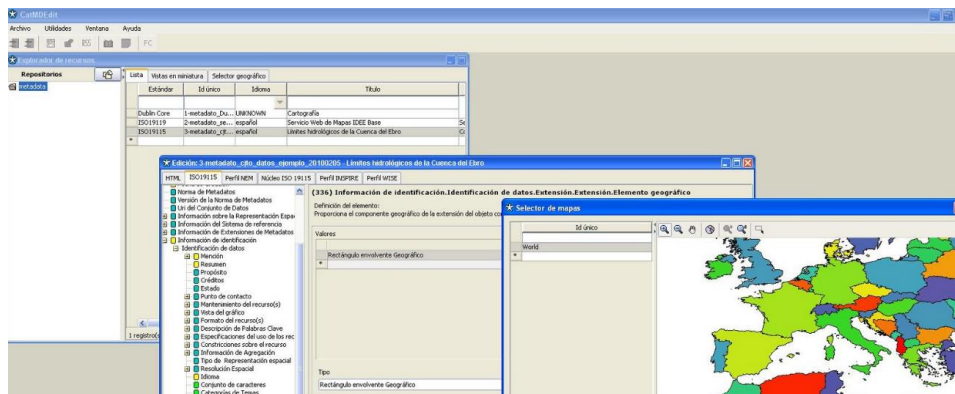


Figura 4.3. Detalles adicionales del metadato. Fuente: Blog IDEE

Algo que se puede mencionar es la forma en que busca o filtra los metadatos, en cada columna que se tiene en el inicio se pantalla ya seleccionado el repositorio de los metadatos, solo utiliza un filtro por cadena de texto en la parte superior de cada una de las columnas.

4.1.2.2. *Análisis de la herramienta*

CatMDEdit es una herramienta sencilla de escritorio de código abierto enfocada en la edición de metadatos geoespaciales. A pesar de no ofrecer un módulo de cosecha de metadatos integradas, posee una interfaz intuitiva para gestionar metadatos en diferentes formatos. Además, provee de la generación automática de metadatos a partir de la operación "getCapabilities" soportada por un servicio que cumple con las Especificaciones OGC. La instalación es sencilla para sistemas Windows y Linux, pero da problemas de lentitud en mostrar ciertas ventanas.

4.1.3. TKME

Principalmente, se trata de un editor de metadatos que comparte una cantidad considerable de su código con su producto complementario, llamado "mp", que funciona como un compilador de metadatos formales. La función principal de "mp" es verificar la estructura sintáctica de documentos que contienen metadatos formales, asegurando que cumplan con el estándar FGDC (FGDC Geospatial Standards — Federal Geographic Data Committee, s. f.). Además, este compilador tiene la capacidad de representar esta información en varios formatos útiles. El propósito del editor de metadatos es simplificar significativamente el proceso de creación de metadatos conformes al estándar establecido.

4.1.3.1. *Características generales*

Entre las características iniciales del sistema, se destaca su capacidad para llevar a cabo la modelación y organización de información de forma semántica avanzada. Adicionalmente, posibilita la creación de esquemas de metadatos personalizados mediante la utilización de diversos vocabularios y ontologías, lo que otorga una notable flexibilidad para ajustarse a una variedad de dominios y requerimientos específicos (Tk Metadata Editor, n.d.).

A continuación, en la Figura 4.4 se presenta la interfaz de inicio de la herramienta al ejecutar el programa.

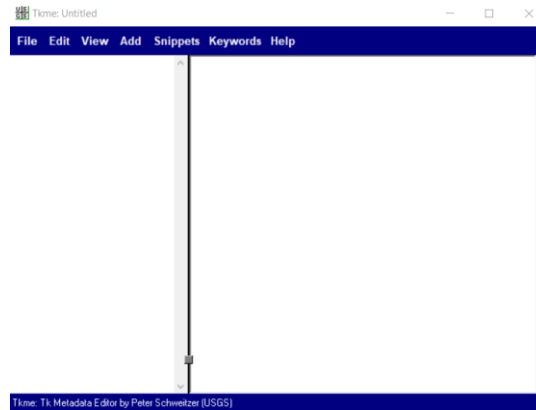


Figura 4.4. Pantalla de inicio de TKME. Fuente: Autoría propia

El programa actualmente se ejecuta exclusivamente en entornos de escritorio. La interfaz de inicio consta de dos secciones principales: en la sección izquierda, se muestra el archivo de metadatos al cargarlo; aquí también se visualiza la estructura completa del método. Por otro lado, en la sección derecha, se lleva a cabo la edición del contenido de cada atributo del metadato en proceso. La Figura 4.5 demuestra cómo se carga una plantilla para completar la información de los metadatos.

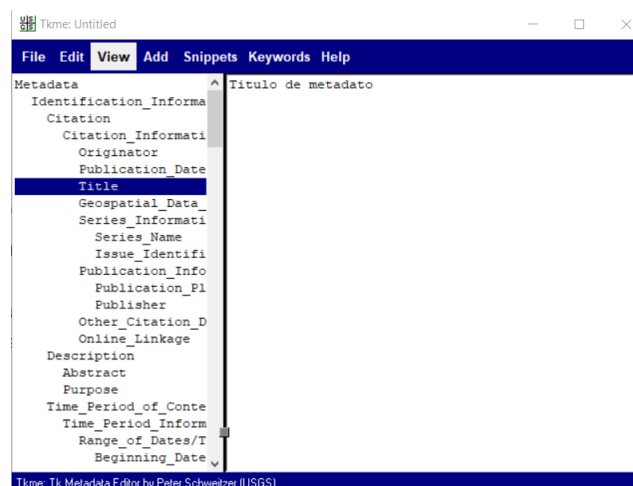


Figura 4.5. Plantilla para una edición de un archivo de metadato. Fuente: Autoría propia

4.1.3.2. Análisis de la herramienta

TKME es una aplicación editora de metadatos formales que ofrece un enfoque eficiente para la creación, administración y representación de información estructurada y semántica en el ámbito geoespacial. Esta herramienta permite la importación de metadatos en formatos XML y JSON, siendo su propósito principal la adecuada representación de la estructura de los metadatos. Por lo tanto, para lograr una gestión óptima de los metadatos, se requiere un cierto nivel de familiaridad con los conceptos de metadatos.

4.1.4. Metavist

Esta herramienta esta destinada para la creación de metadatos que se basan en archivos. Al igual que la herramienta anterior se basa en los estándares de FGDC y adicionalmente del Perfil de Datos Biológicos. Para la edición de los metadatos se hace uso solo de archivos que se encuentren almacenados en formato XML y de igual manera da como resultados en el mismo formato (Metavist - Editor for the Federal Geographic, n.d.).

4.1.4.1. Características generales

La herramienta ha sido desarrollada por el Servicio Forestal del USDA - North Central Research y está diseñada para su ejecución en entornos Windows. En la Figura 4.6 se presenta la interfaz inicial al iniciar el programa, donde se brinda la opción de crear un nuevo registro de metadatos.

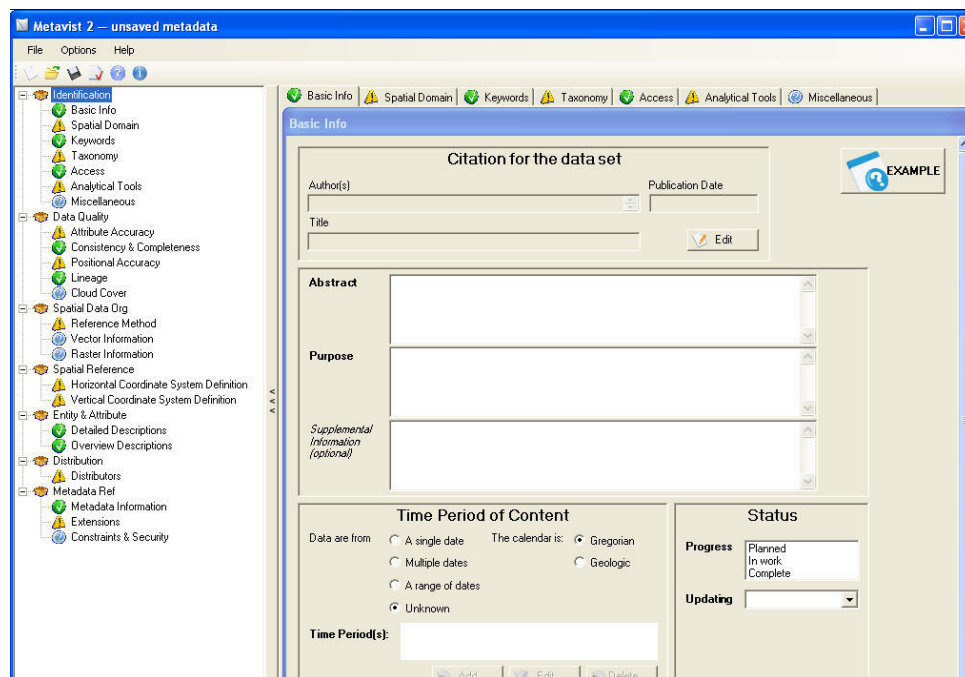


Figura 4.6. Plantilla para una edición de un archivo de metadato. Fuente: Autoría propia

Al acceder a la ventana de inicio y al crear un nuevo registro de metadatos, se evidencia la presencia de las siete secciones principales del estándar de metadatos en la parte superior. Estas secciones son: identificación, calidad de datos, organización de los datos espaciales, referencias espaciales, entidades y atributos, distribución, y referencia de metadatos.

Al comenzar la edición de cada una de estas secciones, se despliega una ventana que contiene un formulario con campos que requieren ser completados. Es importante destacar que tanto la herramienta en sí como las secciones en las que se divide se adhieren a estándares específicos. Los campos obligatorios dentro de estos formularios deben ser

llenados para proporcionar los metadatos necesarios, a menos que se opte por no completar cierta información.

En la Figura 4.7 se presenta un ejemplo concreto de la ventana que se activa al editar la sección de información de identificación.

Figura 4.7. Edición sección Información de identificación. Fuente: Creating Metadata using Metavist Software

4.1.4.2. Análisis de herramienta

Metavist es un programa editor de metadatos que permite la edición de archivos locales en formato XML, seguido de la generación de un resultado en el mismo formato. Esta herramienta ha sido desarrollada exclusivamente para el sistema operativo Windows. Debido a esta limitación, su utilidad en diferentes sistemas operativos es limitada. La adhesión a estándares juega un papel fundamental, ya que la interfaz visualiza los campos y atributos obligatorios y opcionales correspondientes al estándar de metadatos con el que se está trabajando.

4.1.5. EPA Metadata

El EPA Metadata Editor Pro es la última versión disponible, lanzada en el año 2020, de una herramienta de software personalizada diseñada para facilitar la documentación de conjuntos de datos geoespaciales. Esta herramienta opera como un plugin o software complementario

que enriquece la experiencia de edición de metadatos dentro del entorno del software ArcGIS Pro.

4.1.5.1. Características generales

La meta principal de esta herramienta es permitir a profesionales especializados en el campo generar, modificar y validar metadatos que estén en conformidad con la especificación técnica establecida por la organización EPA, así como con el estándar ISO 19115 respaldado por el Comité Federal de Datos Geográficos (FGDC).

A lo largo de los años, la EPA ha desarrollado varias versiones significativas de la herramienta. Uno de los aspectos más destacables de las versiones lanzadas es que, a partir de la versión 5.0, se incorporó una característica adicional: la capacidad de editar metadatos geospaciales. Esta funcionalidad no estaba disponible en las versiones anteriores. Para aquellos interesados, tanto los archivos de instalación como el código fuente de todas las versiones de EME están disponibles en el sitio oficial de GitHub.

Para llevar a cabo la edición de los metadatos, es necesario contar con una instalación previa de ArcGIS Pro. Para ello, es recomendable poseer un entendimiento previo de las diversas opciones disponibles para la gestión de metadatos. Afortunadamente, la interfaz resulta intuitiva y guarda similitudes con productos familiares de Microsoft, tales como Word y Excel. En las Figura 4.8 y Figura 4.9 se puede apreciar la ventana con la opción para la edición del recurso, así como la ventana para la edición de metadatos.

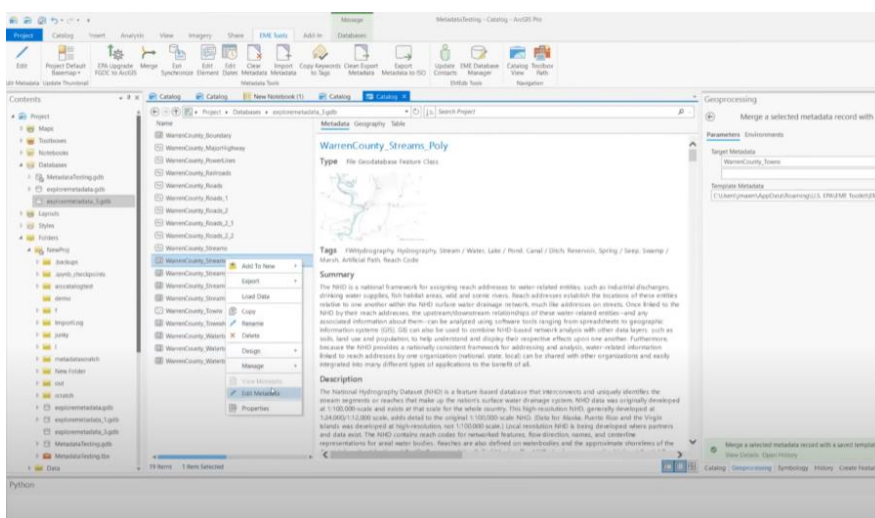


Figura 4.8. Edición metadato en ArcGis Pro. Fuente: Customizing the metadata experience in ArcGIS Pro

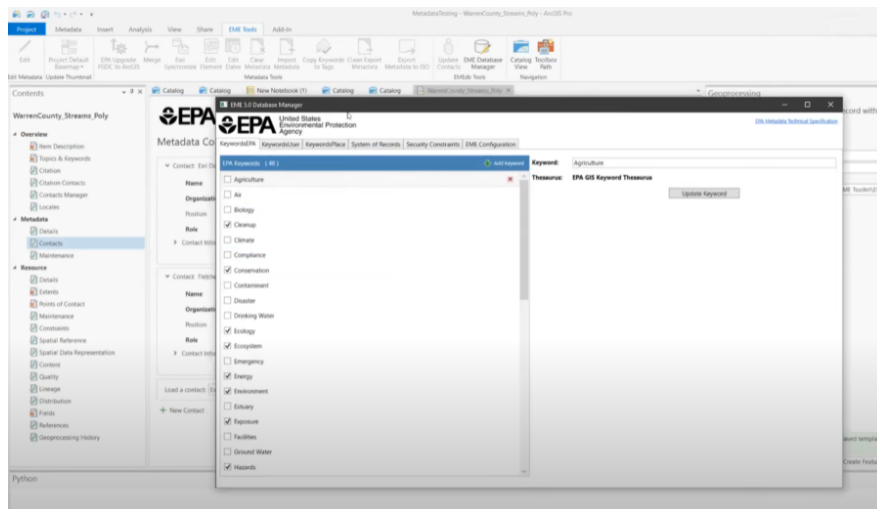


Figura 4.9. Interfaz de EPA metadata. Fuente: Customizing the metadata experience in ArcGIS Pro

4.1.5.2. Análisis de herramienta

EPA Metadata Editor Pro ha sido específicamente diseñado para adherirse a los rigurosos estándares de metadatos establecidos por la EPA, con el propósito de simplificar la descripción de datos ambientales. Esta iniciativa garantiza que los metadatos generados cumplan con todos los requisitos necesarios, fomentando así la interoperabilidad y el intercambio de datos con otras instituciones y sistemas que también siguen los estándares de la EPA. Además, destaca por su capacidad para manejar datos en formato XML, lo que agiliza la generación y manipulación de información. Esta herramienta está focalizada en la creación y edición de metadatos dentro del entorno del software ArcGIS Pro, lo que permite una integración fluida en el sistema.

4.1.6. GeoNetwork

Es una aplicación de catálogo para administrar recursos referenciados espacialmente. Proporciona funciones avanzadas de edición y búsqueda de metadatos, incorpora un visor web de mapas interactivo y se basa en estándares abiertos. Proporciona acceso a servicios OGC (WMS, WMTS) y estándares (KML, OWS). Soporte nativo de los formatos de metadatos ISO19115, ISO19119, ISO19139, ISO19110, FGDC y Dublin Core que se utiliza habitualmente para los portales de datos abiertos. Basado en principios de del software libre y de código abierto.

4.1.6.1. Características generales

Entre las características clave de esta herramienta se encuentra la implementación de un buscador integrado que, en sus versiones más recientes, utiliza Elastic Search para potenciar su funcionalidad. La herramienta ofrece una interfaz web intuitiva que permite la búsqueda

de datos geoespaciales en diversos catálogos de manera eficiente. Esta búsqueda admite tanto la búsqueda de texto completo como la búsqueda basada en facetas, como palabras clave, tipos de recursos u organizaciones. Estas opciones de búsqueda avanzadas permiten a los usuarios refinar sus consultas y acceder de manera ágil a los registros de su interés. En la Figura 4.10 se presenta la captura de pantalla inicial al abrir la aplicación, ilustrando su interfaz.

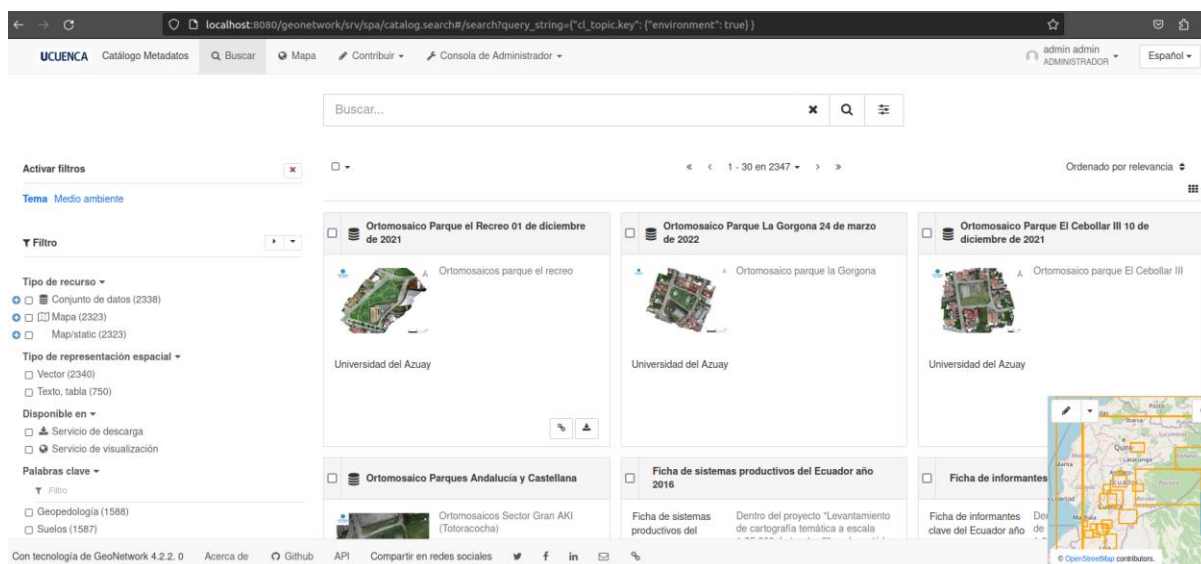


Figura 4.10. Interfaz inicial con el buscador disponible. Fuente: Autoría propia

En su propio sitio, se destaca la característica de configuración sencilla. Una consola de administración brinda un acceso rápido a la configuración del sistema, permitiendo la gestión eficiente de cuentas de usuarios y grupos, así como la configuración de sistemas de clasificación directamente desde la interfaz web. Además, se puede programar la recopilación de metadatos procedentes de múltiples fuentes, entre las que se incluyen: Perfil ISO OGC-CSW 2.0.2, OAI-PMH y un nodo GeoNetwork (GeoNetwork, n.d).

La configuración distintiva del "Harvester" o recolector de catálogos radica en su capacidad para importar metadatos desde fuentes remotas y almacenarlos en el catálogo local. Además, permite la automatización de procesos de recolección (harvesting) y sincronización de metadatos en catálogos distribuidos. Esta herramienta es compatible con diversas plataformas, incluyendo GeoNetwork, CSW, OGC WxS GetCapabilities, Z39.50, WebDav, ArcSDE, Thredds, OGC WFS Features y OAI-PMH.

En adición, una herramienta de análisis de información que brinda un seguimiento y generación de informes proporcionan información resumida sobre el contenido del catálogo y estadísticas sobre la búsqueda. En la Figura 4.11 se puede mostrar un ejemplo de la el análisis que realiza la herramienta y lo presenta gráficamente.

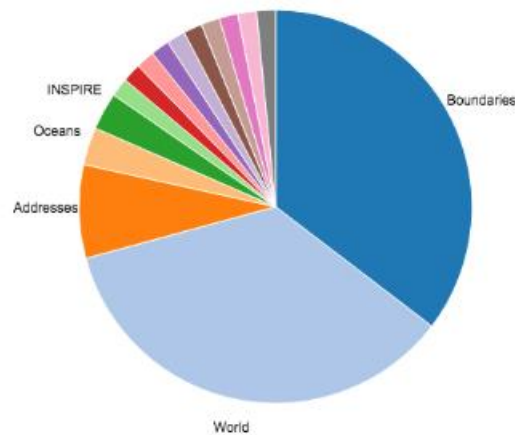


Figura 4.11. Grafica de análisis de información que contiene la almacenada. Fuente: (Geonode, n.d)

4.1.6.2. Análisis de la herramienta

GeoNetwork es una herramienta valiosa para las organizaciones y comunidades que manejan grandes volúmenes de datos geoespaciales. Proporciona una manera eficiente de catalogar, buscar y compartir recursos geoespaciales, mejorando la colaboración y la toma de decisiones basadas en la ubicación. Su enfoque en estándares abiertos y personalización lo convierte en una opción sólida para instituciones gubernamentales, instituciones académicas, organizaciones sin fines de lucro y cualquier entidad que necesite administrar información geoespacial de manera efectiva. Ya que GeoNetwork es una aplicación basada en Java, significa que es compatible con varios sistemas operativos como son Windows, Linux y MacOS.

4.1.7. GeoNode

GeoNode es una plataforma para la gestión y publicación de IG. Reúne proyectos de software de código abierto maduros y estables bajo una interfaz fácil de usar que permite a los usuarios no especializados compartir datos y crear mapas interactivos (GeoNode Documentation, n.d.).

4.1.7.1. Características generales

Las herramientas integradas en GeoNode permiten la creación integrada de datos, metadatos y visualización de mapas. Permite la gestión de IG para compartirlo públicamente o restringirse para permitir el acceso solo a usuarios específicos. También está diseñado para ser una plataforma flexible que los desarrolladores de software pueden ampliar, modificar o integrar para cumplir con los requisitos de sus propias aplicaciones.

Al ejecutar el proyecto, la herramienta despliega una pantalla inicial que presenta un buscador y un menú con diversas opciones como datos, mapas, gestores y un dashboard. Esta estructura mejora la experiencia de navegación para el usuario al facilitar la identificación de recursos específicos. La Figura 4.12 ilustra la ubicación de la opción “Datos” en la interfaz de la herramienta.

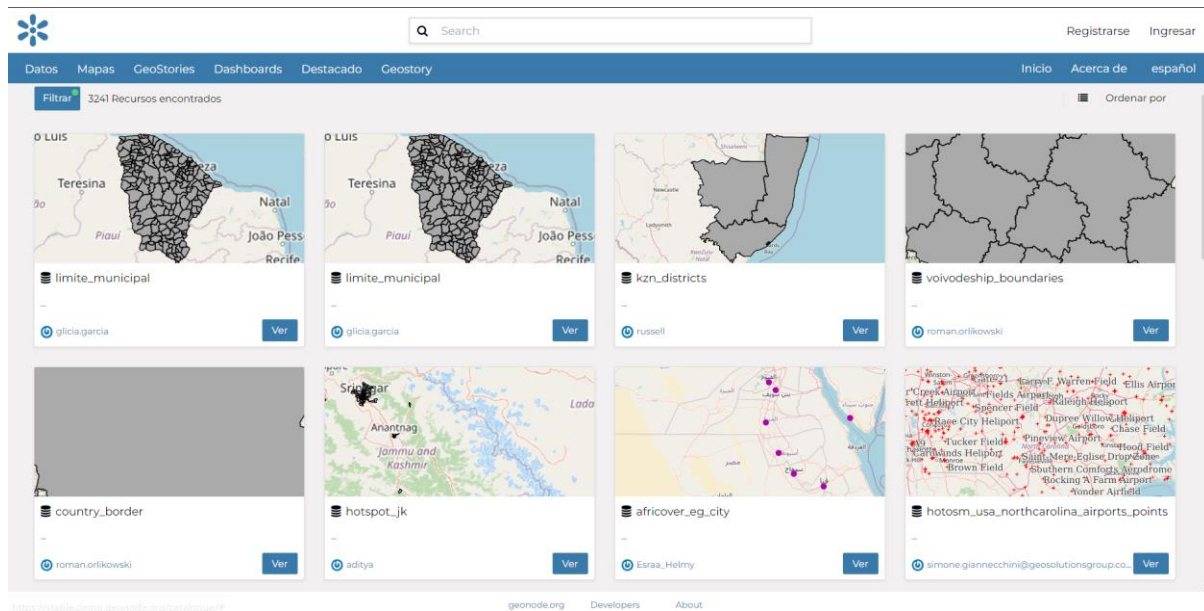


Figura 4.12. Ventana inicial de Geonode con los datos de prueba. Fuente: (Geonode, n.d)

En relación con la búsqueda de información dentro de la herramienta, tiene diversos filtros que permiten la selección y localización precisa de recursos específicos. Por ejemplo, es posible aplicar filtros basados en el tipo de conjuntos de datos, como Vectores, Raster o series de tiempo. Estas opciones de búsqueda están ilustradas en la Figura 4.13 para una comprensión visual más clara.

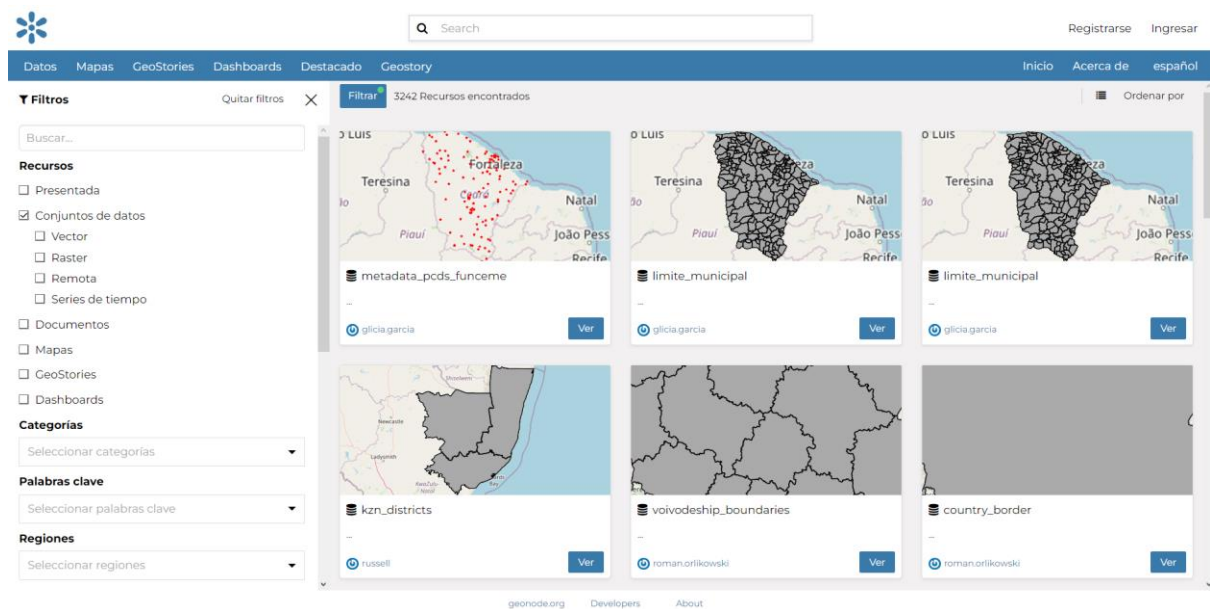


Figura 4.13. Opciones de búsqueda y filtros en Geonode. Fuente: (Geonode, n.d)

4.1.7.2. Análisis de herramienta

Geonode es una herramienta sumamente valiosa dirigida a organizaciones y comunidades involucradas con la IG. La plataforma simplifica el proceso de difusión e intercambio de datos geográficos, lo cual, a su vez, puede potenciar la colaboración y facilitar la toma de decisiones fundamentadas en datos geográficos. Es importante destacar que la adopción de Geonode dependerá de las necesidades concretas de cada entidad y de su capacidad para personalizar y administrar la plataforma de acuerdo con sus requerimientos específicos. Esto se debe a que esta herramienta está orientada principalmente a la visualización efectiva de la información geoespacial y, en particular, sobresale por su capacidad para crear IDE mediante las sólidas herramientas que se encuentran integradas en la plataforma.

Además, es factible llevar a cabo la instalación de Geonode en servidores que utilizan sistemas operativos Windows. No obstante, es esencial considerar que gran parte de las herramientas y componentes vinculados a Geonode están diseñados para una integración más sólida con sistemas Linux.

4.1.8. PyCSW

PyCSW (Python Catalog Service for the Web) es un software de código abierto que permite la implementación de un servidor de catálogo de metadatos compatible con los estándares de OGC para proporcionar servicios de descubrimiento de datos geoespaciales. El estándar CSW define interfaces comunes para encontrar, explorar y consultar metadatos sobre datos espaciales, servicios web y recursos relacionados (PyCSW Documentation, n.d.).

4.1.8.1. Características generales

Una de sus características principales consiste en proporcionar a los usuarios la capacidad de realizar búsquedas de metageo mediante consultas avanzadas. Además, permite filtrar los resultados obtenidos para así acceder a información detallada acerca de los recursos disponibles. Este sistema es compatible con múltiples formatos de metadatos, tales como ISO 19115, ISO 19139, FGDC CSDGM, Dublin Core, así como otros formatos personalizados.

PyCSW, siendo un software de código abierto distribuido bajo la licencia MIT, es versátil en términos de ejecución ya que es compatible con diversas plataformas importantes como Windows, Linux y Mac OS X (PyCSW Documentation, n.d.).

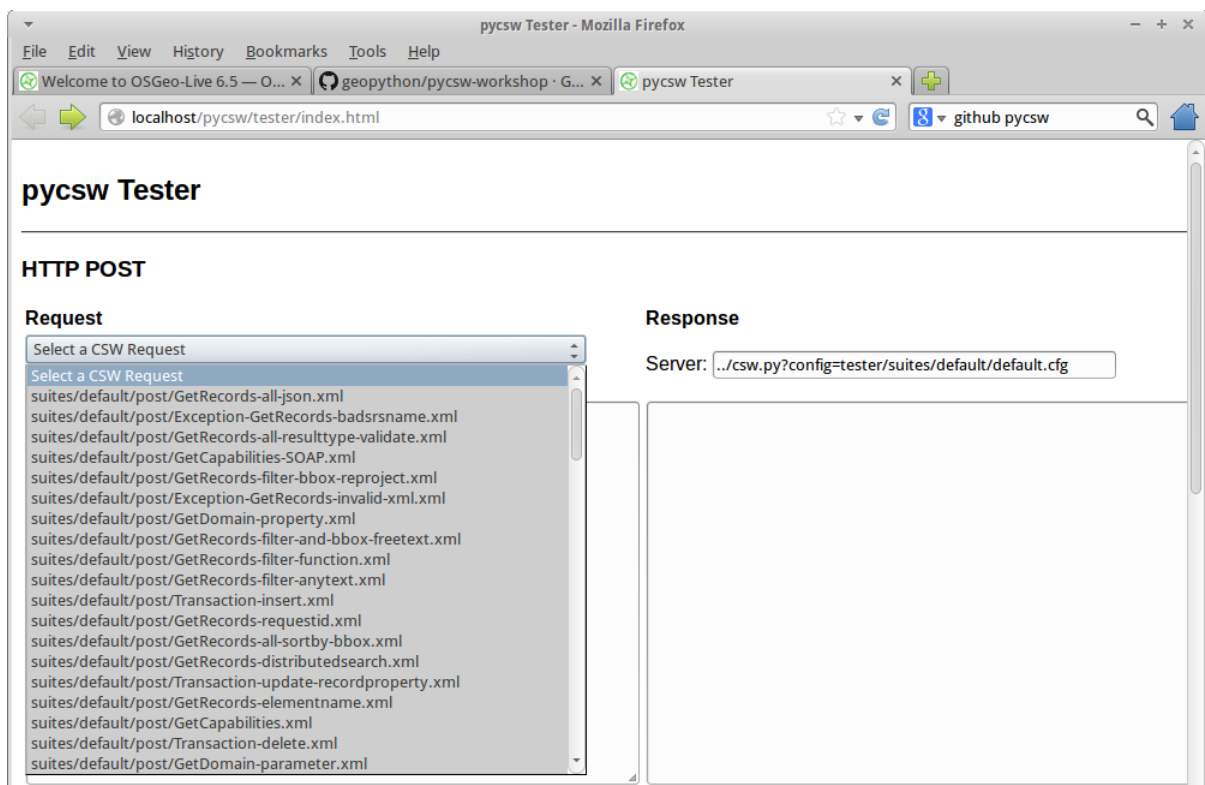


Figura 4.14. Interfaz de inicio de herramienta PyCSW. Fuente: (PyCSW Documentation, n.d.)

En la Figura 4.14, se presenta la interfaz inicial de la herramienta, exhibiendo información de prueba disponible. Asimismo, se aprecia la implementación en un entorno operativo Linux junto con la estructura de la interfaz. En la columna izquierda, se encuentra una lista desplegable que permite al usuario acceder a diversas solicitudes POST predefinidas, las cuales están codificadas en formato XML y pueden ser enviadas mediante pycsw.

4.1.8.2. Análisis de herramienta

PyCSW es una herramienta útil para aquellos que necesitan gestionar y proporcionar acceso a datos geospaciales a través de catálogos de metadatos. Su capacidad para cumplir con los estándares de OGC y su flexibilidad para la personalización y extensión lo hacen adecuado para una variedad de proyectos y aplicaciones en el campo de la geoinformación. Sin embargo, como con cualquier software, es importante considerar sus ventajas y limitaciones en función de los requisitos que se van a considerar para tomar la decisión de utilizarlo.

4.1.9. ESRI ArcGIS

ESRI (con sus siglas en inglés Environmental Systems Research Institute) es una compañía líder en tecnología geoespacial y SIG. ESRI es conocida por su plataforma ArcGIS, que proporciona un conjunto completo de herramientas para crear, administrar, analizar y visualizar datos geospaciales (ArcGIS Resource Center, n.d.).

4.1.9.1. Características generales

a característica principal que destaca de la herramienta es su habilidad para recopilar datos provenientes de una amplia gama de fuentes, que incluyen GPS, imágenes satelitales, encuestas y bases de datos, entre otras. Estos datos son presentados de manera interactiva mediante mapas personalizables. En la Figura 4.15 se exhibe el visor de mapas de muestra en el sitio oficial, a través del cual es posible visualizar información detallada sobre las capas disponibles en la plataforma web.

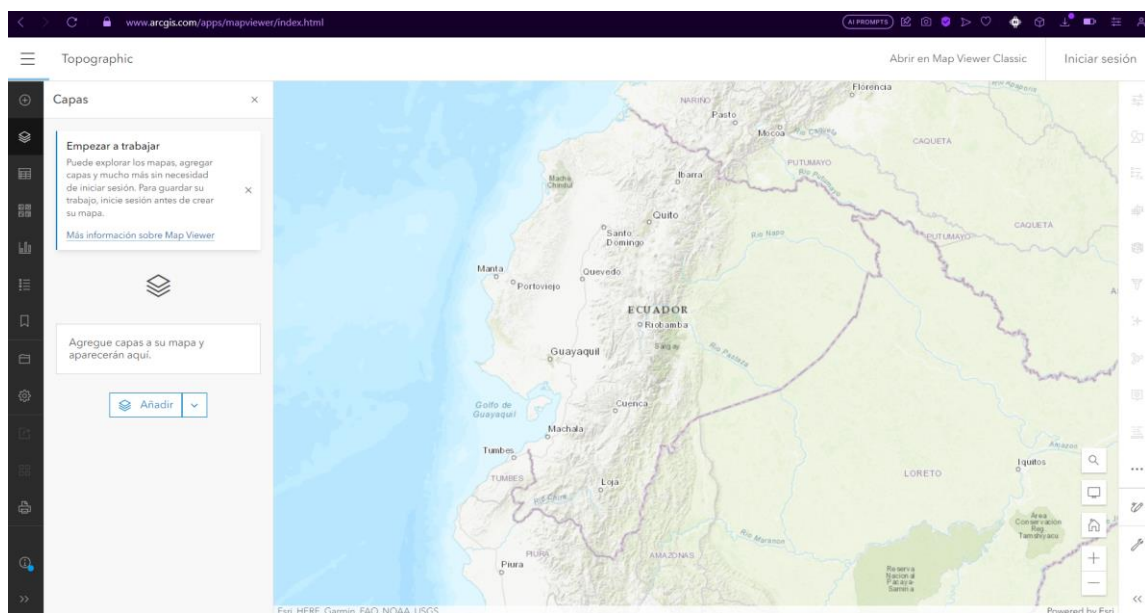


Figura 4.15. Ventana de portal web con administración de capas con IG. Fuente: (Arcgis, n.d)

La plataforma ArcGIS desempeña un papel fundamental al posibilitar la publicación de información geográfica, garantizando su accesibilidad para cualquier usuario interesado. Esta versátil herramienta se encuentra disponible en múltiples plataformas, incluyendo navegadores web, dispositivos móviles como smartphones y sistemas de escritorio. Asimismo, simplifica la gestión de datos geoespaciales al brindar facilidades para la organización, almacenamiento y administración eficiente de grandes volúmenes de información geográfica.

Este sistema abarca una variedad de componentes, que incluyen software especializado, una infraestructura en línea alojada en la nube, herramientas de nivel profesional y recursos configurables, como plantillas de aplicaciones, además de mapas base predefinidos y contenido exclusivo compartido por una comunidad activa de usuarios. La versatilidad de este sistema se ve respaldada por su compatibilidad con diversas plataformas de servidores y servicios en la nube, lo que facilita la colaboración y el intercambio de recursos, asegurando así la disponibilidad inmediata de información crucial para procesos de planificación y toma de decisiones (ArcGIS Resource Center, n.d.).

4.1.9.2. *Análisis de herramienta*

ArcGIS es una herramienta poderosa y versátil en el ámbito de los SIG, que ha sido utilizada en diversos sectores como la planificación urbana, la gestión ambiental, la agricultura, la salud, la gestión de recursos naturales y muchas otras áreas. Su capacidad para trabajar con datos geoespaciales, realizar análisis sofisticados y crear visualizaciones efectivas lo convierte en una herramienta esencial para tomar decisiones informadas basadas en la ubicación. Sin embargo, debido a su enfoque profesional y la curva de aprendizaje asociada, puede requerir tiempo y capacitación para aprovechar al máximo todo el potencial que brinda la herramienta.

Aunque ESRI ofrece una amplia gama de herramientas y soluciones geoespaciales, su software no está disponible bajo licencia de código abierto, este caso es una excepción a las herramientas previamente analizadas como son GeoNetwork, Geonode, PyCSW que son potentes herramientas para gestión de IG como lo es igualmente ArcGIS.

4.1.10. **PgMetadata**

Los metadatos de la IG de las capas se almacenan dentro de su base de datos PostgreSQL, en un esquema dedicado. Se admiten campos clásicos, como el título, la descripción, las categorías, los temas, los enlaces y las propiedades espaciales de sus datos: extensión, proyección, etc.

4.1.10.1. Características generales

La característica principal de PgMetadata no está concebida como una aplicación de catálogo que permita la búsqueda y descarga de conjuntos de datos. Su propósito radica en simplificar el uso y la gestión de metadatos dentro de QGIS. Esto habilita la búsqueda de datos para acceder a las capas correspondientes o examinar los metadatos de las capas alojadas en PostgreSQL. A través de la integración con QGIS, es posible acceder a los metadatos con facilidad utilizando la conexión a PostgreSQL asociada a la capa (PgMetadata, n.d.).

Adicionalmente, se destacan otras funcionalidades de PgMetadata, entre las que se incluyen:

- Personalización de la plantilla HTML para la adaptación de los metadatos visualizados.
- Generación de una hoja de metadatos en formatos como PDF, HTML o DCAT.
- Distribución de los metadatos como un catálogo DCAT, que es la especificación europea para la descripción de catálogos de datos abiertos (DCAT-AP 2.0.1, n.d.). Esta información puede ser adquirida posteriormente por aplicaciones externas, por ejemplo, mediante el uso de GeoNetwork.

En la Figura 4.17 se muestra el uso del complemento en QGIS, el cual nos indica una personalización de la plantilla HTML de los metadatos que se crean previamente como se observa en la Figura 4.16.

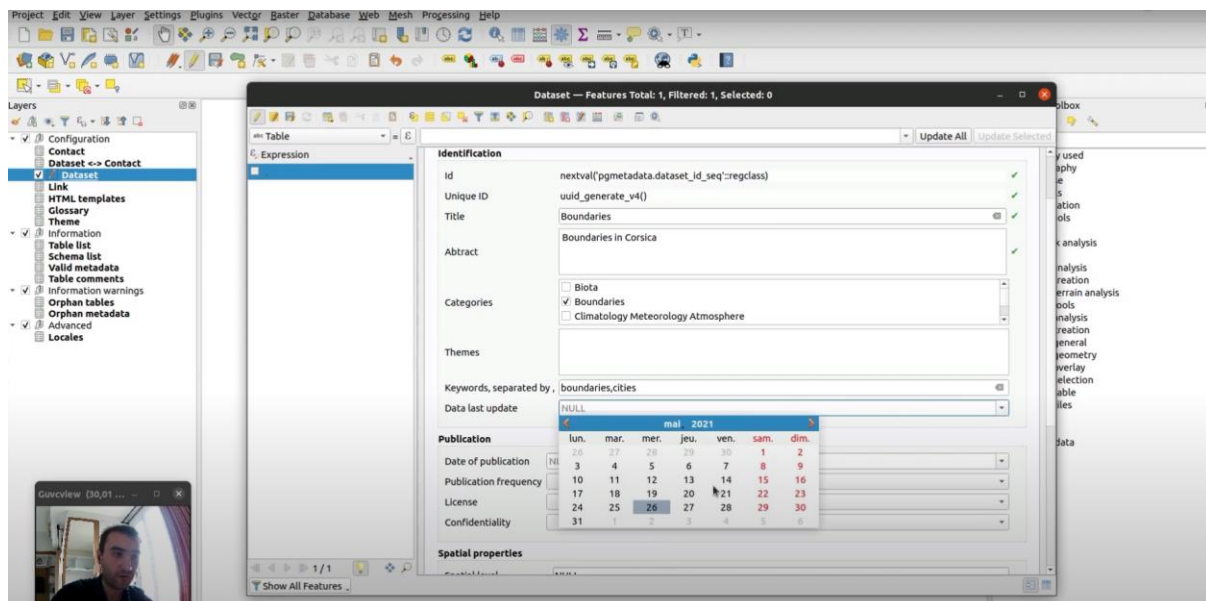


Figura 4.16. Captura de video ilustrativo de creación de metadato. Fuente: (PgMetadata, n.d.)

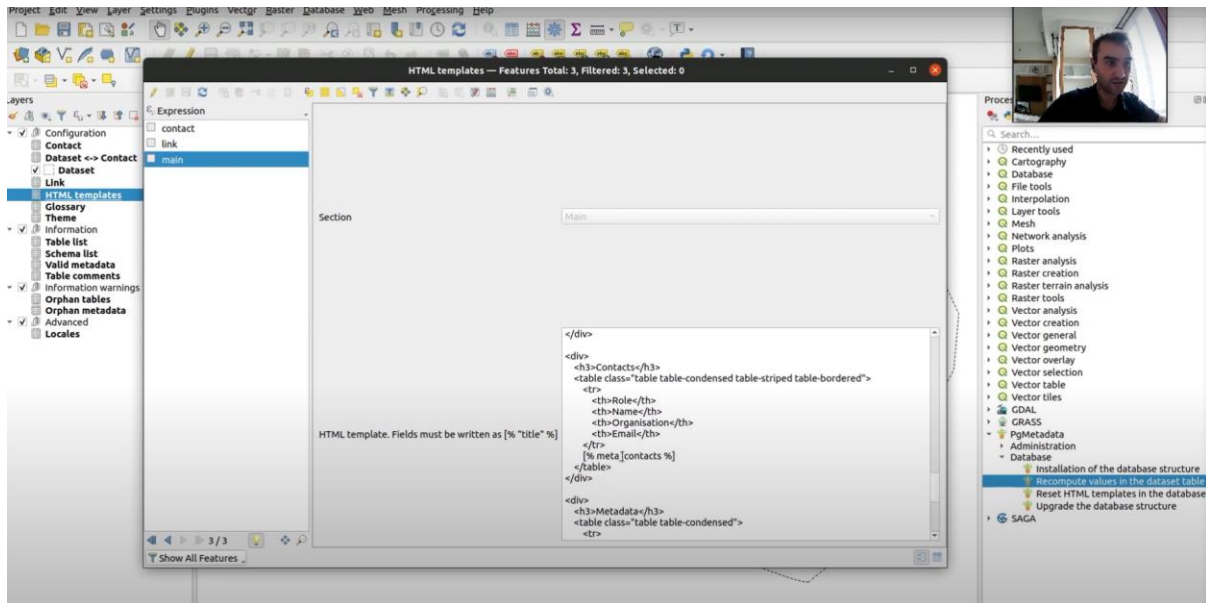


Figura 4.17. Captura de video ilustrativo de edición plantilla HTML. Fuente: (PgMetadata, n.d.)

4.1.10.2. Análisis de la herramienta

PgMetadata es un complemento diseñado para usuarios de QGIS, que facilita la gestión de metadatos dentro de la plataforma. Aunque no habilita una catalogación de información per se, brinda la posibilidad de generar resultados que pueden ser recopilados a través de la herramienta GeoNetwork, tal como se detalló en las características del propio complemento. En consecuencia, la relevancia de este plugin radica en su capacidad para satisfacer las necesidades específicas de cada institución en la gestión de metadatos asociados a las capas manipuladas en el entorno de QGIS.

4.1.11. Identificación de características claves

En esta sección, se identificarán las características clave que serán evaluadas de las herramientas de gestión de metadatos. Las características seleccionadas se consideran fundamentales para el cumplimiento del objetivo del trabajo de investigación. A continuación, se describen las características clave identificadas:

- **Multiplataforma:** esta característica trata de ver la capacidad de la herramienta para ser ejecutado en diferentes plataformas y sistemas operativos. Se evalúa la compatibilidad con plataformas populares como Windows, macOS y iOS, Linux y sus distribuciones y Android.
- **Compatibilidad con estándares:** En esta característica se evalúa la capacidad del gestor de metadatos para soportar y cumplir con los estándares de metadatos ampliamente aceptados, como Dublin Core, FGDC, INSPIRE, entre otros. Se considerará

la capacidad de los gestores de metadatos para almacenar, organizar y visualizar los metadatos de acuerdo con estos estándares.

- **Cosechamiento (Harvesting) o recolección de metadatos:** en este caso se refiere a la capacidad para recolectar automáticamente metadatos de una variedad de fuentes y repositorios como pueden ser institucionales, catálogos bibliográficos. Se evaluará la capacidad al realizar la extracción y recolección automatizada de metadatos de sitios web, bases de datos, sistemas de archivos, servicios, entre otros. Además, si para el cosechamiento tiene alguna importancia el tipo de formato o estándares que se va a obtener de las fuentes.
- **Gestión de base de datos o archivos:** hace referencia a la capacidad de la herramienta para administrar eficientemente bases de datos o archivos que contienen metadatos. Se evaluará la capacidad para almacenar y organizar los metadatos de manera estructurada, permitiendo una gestión eficaz y accesible de la información.
- **Búsquedas y consultas avanzadas:** la característica hace referencia a las capacidades de búsqueda que ofrece la herramienta. Se tomará en cuenta la capacidad para realizar búsquedas basadas en metadatos específicos, filtros por palabras clave o etiquetados, además de la recuperación de información de manera eficiente.
- **Interfaz de usuario intuitiva:** esta característica se refiere a la facilidad de uso y comprensión de la interfaz de usuario de la herramienta. Se evaluará la claridad de la disposición de los elementos de la interfaz, la facilidad de navegación y la capacidad de los usuarios para realizar tareas sin dificultades, evitando la necesidad de una curva de aprendizaje prolongada.
- **Integración con otros sistemas:** en este caso se evaluará la capacidad que tiene la herramienta para integrarse con otros sistemas y aplicaciones relevantes. Se considerará la capacidad de importar y exportar metadatos desde y hacia otras herramientas; entre estas pueden ser: sistemas de gestión de bases de datos, repositorios digitales u otros gestores de metadatos.

4.1.12. Puntuación de las características

A continuación, puntuaremos las características clave que se seleccionaron en la sección anterior. La puntuación que se asigna a cada característica nos permitirá comparar y analizar las fortalezas y debilidades de cada gestor en relación con las necesidades específicas para seleccionar un gestor de metadatos para cumplir con el objetivo del trabajo.

La asignación de puntuación será de escala numérica el cual utilizamos una escala del 1 al 5, donde 1 indica un bajo rendimiento o ausencia de la característica, y 5 representa un rendimiento óptimo o presencia completa de la característica. Cabe destacar que estas puntuaciones son subjetivas y se basan en observaciones y análisis de cada herramienta que se ha realizado. Ahora se va a describir un resumen de cada característica con su respectiva puntuación.

- **Multiplataforma**

Tabla 4.1. Escala de puntuación en multiplataforma. Fuente: Autoría propia

Grado	Descripción
1	Baja compatibilidad multiplataforma
2	Compatibilidad limitada
3	Compatibilidad media
4	Buena compatibilidad multiplataforma
5	Alta compatibilidad multiplataforma

Baja compatibilidad multiplataforma: tiene una compatibilidad limitada o nula con diferentes plataformas. Es posible que solo funcione en un sistema operativo específico o esté restringida a un conjunto limitado de dispositivos.

Compatibilidad limitada: ofrece cierta compatibilidad con algunas plataformas adicionales, pero todavía presenta limitaciones significativas en términos de su capacidad para ejecutarse en diferentes sistemas operativos.

Compatibilidad media: tiene una compatibilidad aceptable con varias plataformas. Puede ejecutarse en diferentes sistemas operativos o dispositivos, pero aún pueden existir algunas limitaciones en ciertos casos.

Buena compatibilidad multiplataforma: la herramienta es compatible con múltiples plataformas y puede ejecutarse en una amplia gama de sistemas operativos y dispositivos sin problemas significativos. Es capaz de adaptarse y funcionar correctamente en diferentes entornos.

Alta compatibilidad multiplataforma: herramienta con una compatibilidad excelente con diversas plataformas. Puede ejecutarse sin problemas en diferentes sistemas operativos y

dispositivos, ofreciendo una experiencia consistente y completa en cada uno de ellos. La aplicación está diseñada para ser altamente interoperable y adaptable.

- **Compatibilidad con estándares**

En esta característica se ha considerado principalmente la compatibilidad con estándares internacionales como por ejemplo ISO y Dublin Core. Además, se considera también si las herramientas tienen la posibilidad de crear sus propias plantillas de estándares que rigen en la región en la que se genera la IG, esto asegura la compatibilidad con perfiles nacionales como es el caso de PEM en Ecuador.

Tabla 4.2. Escala de puntuación en uso de estándares en la herramienta. Fuente: Autoría propia

Grado	Descripción
1	Baja uso de estándares
2	Incorporación limitada de estándares
3	Incorporación adecuada
4	Buena incorporación
5	Excelente incorporación de estándares

Baja uso de estándares: la herramienta tiene una baja ausencia de estándares en su funcionalidad. Puede carecer de compatibilidad con formatos y protocolos estándar, lo que limita su interoperabilidad y capacidad de integración con otros sistemas.

Incorporación limitada de estándares: muestra una compatibilidad limitada con algunos estándares de metadatos, pero su implementación es parcial o incompleta. La herramienta trabaja con algunos formatos o protocolos estándar, pero no ofrece soporte completo para todas las especificaciones relevantes.

Incorporación adecuada: aquí demuestra un uso adecuado a los estándares de metadatos esenciales. Puede trabajar con los formatos y protocolos comunes utilizados en la gestión de metadatos, aunque puede haber limitaciones en algunos aspectos.

Buena incorporación: la herramienta muestra una sólida adhesión a los estándares de metadatos más ampliamente aceptados y utilizados en la industria. Ofrece soporte completo para una amplia gama de formatos y protocolos estándar. Cumple con los requisitos y regulaciones relacionados con los estándares de metadatos establecidos en su dominio de aplicación.

Excelente incorporación de estándares: dispone de un nivel excepcional de aceptación a los estándares de metadatos. Cumple con los estándares más rigurosos y avanzados de la industria y proporciona soporte completo para una amplia gama de formatos, protocolos y especificaciones.

- **Cosechamiento (Harvesting) o recolección de metadatos:**

Tabla 4.3. Puntuación de cosechamiento o recolección en la herramienta. Fuente: Autoría propia

Grado	Descripción
1	Cosechamiento bajo o nulo
2	Cosechamiento regular
3	Cosechamiento medio
4	Alto cosechamiento
5	Excelente cosechamiento

Cosechamiento bajo o nulo: aquí presenta deficiencias significativas en el cosechamiento de metadatos o no disponibilidad del mismo. Puede tener problemas de compatibilidad con diferentes formatos de datos o fuentes de información. Es posible que la herramienta no pueda extraer correctamente o la calidad de los metadatos es muy baja. Además, puede requerir una intervención manual constante y no ofrece automatización en el proceso.

Cosechamiento regular: cumple con los requisitos básicos de cosechamiento o recolección de metadatos, pero no destaca en términos de funcionalidad o rendimiento. Tiene ciertas limitaciones en la obtención de algunos tipos de archivos o fuentes de información determinadas.

Cosechamiento medio: en la herramienta presenta un rendimiento aceptable en el cosechamiento de metadatos. Es capaz de extraer metadatos de la mayoría de los tipos de archivos y fuentes de información relevantes. Proporciona resultados consistentes y de calidad en la extracción de metadatos. Ofrece una automatización adecuada y puede manejar tareas básicas de cosechamiento de manera eficiente.

Alto cosechamiento: La herramienta demuestra un buen rendimiento en el cosechamiento o recolección de metadatos. Es capaz de manejar una amplia variedad de formatos de archivos y fuentes de información con precisión. Proporciona metadatos de alta calidad y

confiabilidad. Ofrece una automatización avanzada y eficiente, reduciendo significativamente la intervención manual requerida.

Excelente cosechamiento: La herramienta es altamente efectiva y eficiente en el cosechamiento o recolección de metadatos. Es compatible con prácticamente todos los formatos de archivos y fuentes de información relevantes. Proporciona metadatos precisos, completos y de alta calidad en todas las situaciones. Ofrece una automatización completa y robusta, minimizando la necesidad de intervención manual y maximizando la productividad.

- **Gestión de base de datos o archivos**

Tabla 4.4. Puntuación de gestión de bases de datos o archivos. Fuente: Autoría propia

Grado	Descripción
1	Gestión base de datos baja
2	Gestión regular de base de datos
3	Gestión media de base de datos
4	Alta gestión de base de datos
5	Excelente gestión con base de datos

Gestión base de datos baja: indica un rendimiento deficiente o una funcionalidad limitada en esta área. Es posible que la herramienta carezca de características básicas de gestión de base de datos o archivos, capacidad de indexar y buscar eficientemente en grandes conjuntos de datos.

Gestión regular de base de datos: la herramienta cuenta con algunas funcionalidades básicas, pero aún tiene limitaciones en comparación con otras herramientas. Carece de ciertas características avanzadas de gestión de base de datos o archivos, como la capacidad de optimizar consultas complejas.

Gestión media de base de datos: la herramienta muestra un rendimiento promedio o estándar en esta área. La herramienta cumple con las funcionalidades básicas esperadas para la gestión de base de datos o archivos y soporte para una variedad de formatos de archivos. Sin embargo, puede carecer de características más avanzadas o especializadas que se encuentran en herramientas de puntuación más alta.

Alta gestión de base de datos: el rendimiento de la herramienta es superior y una funcionalidad sólida en esta gestión. La herramienta ofrece funcionalidades avanzadas y

completas para la gestión de base de dato, como la capacidad de realizar consultas complejas, así como soporte para una amplia gama de formatos de archivos que puede almacenarse y gestionarse. Proporciona una experiencia de gestión de base de datos o archivos robusta y confiable.

Excelente gestión con base de datos: la herramienta tiene un rendimiento excepcional y una funcionalidad de vanguardia en esta área. Ofrece características como indexación eficiente, escalabilidad para grandes volúmenes de datos. Por tanto, la herramienta es altamente confiable y potente para la gestión de base de datos.

Tabla 4.5. Puntuación de búsqueda y opciones de búsqueda en la herramienta. Fuente: Autoría propia

Grado	Descripción
1	Baja capacidad de búsqueda y consultas avanzadas
2	Funcionalidad limitada de búsqueda y consultas avanzadas
3	Funcionalidad moderada de búsqueda y consultas avanzadas.
4	Buena funcionalidad de búsqueda y consultas avanzadas.
5	Excelente funcionalidad de búsqueda y consultas avanzadas.

Baja capacidad de búsqueda y consultas avanzadas: se refiere a que tiene una funcionalidad limitada para realizar búsquedas y consultas avanzadas. No ofrecen filtros complejos, operadores booleanos o búsquedas con expresiones regulares. La búsqueda de metadatos se basa principalmente en criterios básicos y simples.

Funcionalidad limitada de búsqueda y consultas avanzadas: la herramienta proporciona algunas opciones para realizar búsquedas y consultas avanzadas, pero con limitaciones significativas. La limitación de la funcionalidad puede ser en términos de flexibilidad y capacidad de personalización.

Funcionalidad moderada de búsqueda y consultas avanzadas: se refiere a que ofrece algunas capacidades para realizar búsquedas y consultas, aunque no están completamente desarrolladas. Es posible utilizar filtros y operadores para refinar las búsquedas, pero pueden faltar opciones más avanzadas.

Buena funcionalidad de búsqueda y consultas avanzadas: en este caso la herramienta brinda opciones sólidas para realizar búsquedas y consultas. La búsqueda es efectiva y permite una recuperación con más detalle de los metadatos.

Excelente funcionalidad de búsqueda y consultas avanzadas: la herramienta ofrece una amplia gama de capacidades para realizar búsquedas y consultas. La funcionalidad de la búsqueda es eficiente, lo que permite una recuperación adecuada de los metadatos.

- Interfaz de usuario

Tabla 4.6. Puntuación de la interfaz de usuario en la herramienta. Fuente: Autoría propia

Grado	Descripción
1	Baja, poca intuitiva.
2	Media-baja, interfaz aceptable.
3	Media, aceptable y con expectativas básicas.
4	Media alta, interfaz bien diseñada y agradable visualmente
5	Interfaz de usuario excepcional.

Baja: hace referencia a que presenta una interfaz de usuario poco intuitiva y confusa. La navegación puede ser difícil y poco clara, lo que dificulta el uso eficiente de la herramienta. La disposición de elementos o la organización de los menús no sean coherentes.

Media-baja: hace referencia a que ofrece una interfaz de usuario aceptable, pero aún presenta algunas deficiencias. Puede existir cierta desorganización en los elementos de la interfaz. Es posible que ciertas opciones no sean tan accesibles o que requieran de un tiempo de aprendizaje adicional para entenderlos.

Media: la interfaz es adecuada y cumple con las expectativas básicas. Los elementos de la interfaz están organizados de manera coherente y siguen convenciones de diseño estándar. La navegación es intuitiva y los usuarios pueden encontrar fácilmente las funciones que necesitan. Sin embargo, pueden existir mejoras en términos de diseño visual.

Media-Alta: la herramienta presenta una interfaz de usuario bien diseñada y agradable visualmente. La navegación es intuitiva y los elementos de la interfaz están ubicados de manera lógica y fácilmente accesible. Los usuarios que usan la herramienta pueden realizar sus tareas de forma eficiente.

Alta: la herramienta tiene una interfaz de usuario excepcional. La atención al diseño y la usabilidad es evidente en cada aspecto de la interfaz. El diseño visual es atractivo, y la disposición de los elementos, la consistencia y la retroalimentación visual están

cuidadosamente implementados. La herramienta tiene la capacidad de proporcionar una experiencia de usuario de alta calidad.

- **Integración con otros sistemas**

Tabla 4.7. Puntuación sobre la integración con otros sistemas que tiene la herramienta. Fuente: Autoría propia

Grado	Descripción
1	Integración mínima o nula
2	Integración básica con algunos sistemas
3	Integración aceptable con sistemas
4	Buena integración con sistemas
5	Excelente integración con sistemas

Integración mínima o nula: la interoperabilidad de la herramienta con otras aplicaciones o plataformas es mínima o inexistente. Requiere esfuerzos significativos de desarrollo o personalización adicional para lograr la integración con otros sistemas.

Integración básica con algunos sistemas: la herramienta presenta limitaciones significativas con otros sistemas. La interoperabilidad entre herramientas es parcial o restringida a un conjunto específico de sistemas.

Integración aceptable con sistemas: hace referencia a una integración aceptable con una variedad de sistemas comunes. La comunicación entre otros sistemas es suficiente para cubrir las necesidades básicas. Posiblemente, sea necesario configuraciones adicionales para una integración personalizada.

Buena integración con sistemas: dispone de una buena integración con una amplia gama de sistemas y tecnologías. Proporciona mecanismos y API que facilitan la integración y la interacción con otros sistemas de manera eficiente.

Excelente integración con sistemas: hace referencia a que la interoperabilidad de la herramienta con otras aplicaciones o plataformas es completa y abarca prácticamente todos los escenarios. Dispone de una amplia variedad de mecanismos y protocolos para una buena integración.

4.1.13. Resultados de características de las herramientas

Ahora se exponen los resultados del análisis de cada herramienta en función de las características previamente seleccionadas y su correspondiente puntuación. En la Tabla 4.8, se asigna a cada gestor de metadatos según su desempeño en las características evaluadas. Estas puntuaciones se fundamentan en nuestra evaluación subjetiva de cada gestor, reflejando así nuestras prioridades y requisitos específicos.

Tabla 4.8. Puntuación de las características de las herramientas analizadas. Fuente: Autoría propia

Herramienta	Multiplataforma	Estándares	Harvesting	Base de datos	Búsqueda	Interfaz de usuario	Integración
CatMDEdit	2 (W, L)	3	1	1	1	1	1
TKME	2 (W, L)	3	1	1	1	1	1
Metavist	1 (W)	3	1	1	1	3	1
EPA Metadata Editor	1 (W)	4	1	4	3	4	3
GeoNetwork	5 (W, L, M)	5	5	5	5	5	5
GeoNode	4 (W, L)	4	4	5	5	5	4
PyCSW	5 (W, L, M)	5	3	5	5	5	3
ESRI ArcGIS	5 (W, L, M, A, I)	5	4	5	5	5	4
PgMetadata	5 (W, L, M)	1	1	4	4	3	3

Para la característica de compatibilidad multiplataforma, se han identificado los sistemas operativos con los que cada una de las herramientas es compatible: Windows (W), Linux (L), MacOS (M), Android (A) e iOS (I). Es importante resaltar que las puntuaciones de compatibilidad pueden variar según el contexto y las necesidades de cada usuario. Por ejemplo, algunos podrían utilizar la herramienta para crear metadatos sin la necesidad de compartirlos, sino solo para visualizarlos en el contexto de investigaciones dentro de la institución. Otro caso podría involucrar la gestión centralizada de metadatos, lo que permitiría el manejo en diferentes servidores con diversas fuentes de metadatos.

Es relevante tener en cuenta que estas puntuaciones representan solo una faceta del proceso de evaluación. Un análisis detallado debe complementar estas puntuaciones antes de tomar

una decisión definitiva respecto al gestor de metadatos más apropiado para nuestra aplicación.

4.1.14. Análisis de resultados y selección de herramienta

En la siguiente sección, procederemos a analizar y discutir los resultados derivados de la evaluación de las características. En este análisis, se destacarán tanto las fortalezas como las debilidades inherentes a cada gestor, resaltando su pertinencia en el contexto de este trabajo de tesis.

A continuación, en la Figura 4.18 se presenta un cuadro comparativo de las características prominentes en todas las herramientas. Esta representación gráfica proporciona una visión sucinta de los servicios disponibles junto con su relevancia en el logro del objetivo central de este proyecto de investigación.

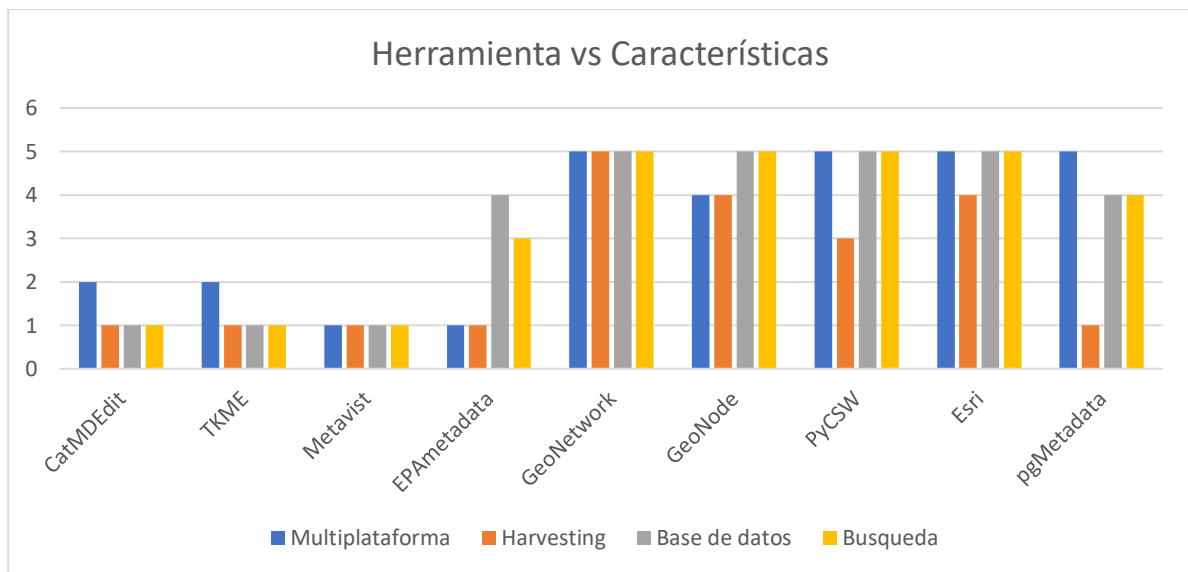


Figura 4.18. Diagrama de barras de las herramientas y las características destacadas. Fuente: Autoría propia

Dada la puntuación obtenida, es claro que existe una ventaja significativa en el uso de las herramientas GeoNetwork y Geonode en el proceso de recolección de Información IG. Estas herramientas, además, exhiben un desempeño destacado en la gestión de bases de datos, aspecto crucial, especialmente al tratar con la administración de volúmenes considerables de datos geoespaciales.

Por otro lado, se puede apreciar que los catalogadores de metadatos como GeoNetwork, Geonode, Pycsw y ESRI ArcGIS (ArcCatalog) muestran un rendimiento superior y destacan en aspectos como la utilización de bases de datos, los métodos de búsqueda y la recolección de información. Esta tendencia podría atribuirse a su naturaleza como herramientas más integrales y a su constante actualización, lo que les permite mantenerse a la vanguardia.

Si bien los editores de metadatos están diseñados específicamente para la edición de metadatos y su estructura, incluida la validación de los datos según estándares establecidos, es importante destacar que estas herramientas suelen ser más simples y presentar limitaciones en comparación con otras características cruciales para alcanzar los objetivos de esta investigación, por ejemplo, la función de cosechamiento de metadatos. Una de las limitaciones notables es su enfoque en la gestión de archivos locales, lo que representa una desventaja evidente en relación con las herramientas catalogadoras. Esto se debe en parte a que los catalogadores cuentan, por defecto, con una base de datos integrada o configurables para la integración y almacenamiento de información.

La parte más destacada del análisis realizado radica en la impresionante capacidad de extracción que exhiben ciertas herramientas. Estas herramientas no solo se limitan a recopilar información de diversas fuentes, sino que también se centran en la gestión a través de una base de datos interna o un sistema de archivos propio. Entre estas opciones, resalta GeoNetwork por múltiples razones, incluida su habilidad para recolectar metadatos y su eficacia en la búsqueda de información, respaldada por un motor de búsqueda incorporado en las últimas versiones disponibles. Además, la implementación mediante el uso de Docker proporciona una valiosa alternativa, especialmente en términos de la capacidad multiplataforma que ofrece la herramienta.

Considerando todo el análisis previo, podemos concluir que la herramienta GeoNetwork exhibe un sólido rendimiento en la mayoría de las características evaluadas, lo que la posiciona como una alternativa idónea para entornos que demandan una búsqueda avanzada y una integración efectiva con otros sistemas a través del empleo de estándares. Además, es importante destacar que GeoNetwork se utiliza ampliamente en instituciones y organizaciones con el propósito de establecer catálogos, facilitando el acceso a la Información Geoespacial (IG) y operando como un servidor para la recopilación de metadatos.

En contraste, GeoNode no se presenta como una elección pertinente en nuestro caso de estudio, ya que esta herramienta está enfocada en la creación de plataformas participativas para datos geoespaciales, en lugar de desempeñar el rol de servidor dedicado a la gestión de grandes volúmenes de metadatos, como es el caso de GeoNetwork.

4.2. Diseño y desarrollo del modelo conceptual

En la siguiente sección, se presenta la arquitectura conceptual del modelo del cosechador, junto con su implementación prevista en un servidor web. Esta elección de arquitectura se basa en un análisis exhaustivo, lo que ha llevado a la selección de la herramienta más

adecuada. En este contexto, se opta por la utilización de contenedores Docker para la implementación y despliegue de la aplicación.

La herramienta seleccionada para este propósito es GeoNetwork. Esta elección está respaldada no solo por los resultados positivos obtenidos en el análisis previo, sino también por referencias en documentos seleccionados que destacan su eficacia en la gestión de metadatos y su módulo de cosecha. Uno de los aspectos destacables es su capacidad para simplificar la configuración del proceso de cosecha, eliminando la necesidad de conocimientos técnicos altamente especializados para su configuración y ejecución.

A continuación, en la Figura 4.19 se expone el diseño del modelo conceptual en el cual se identifican las entidades y relaciones que componen un cosechador de metadatos, utilizando la herramienta GeoNetwork.

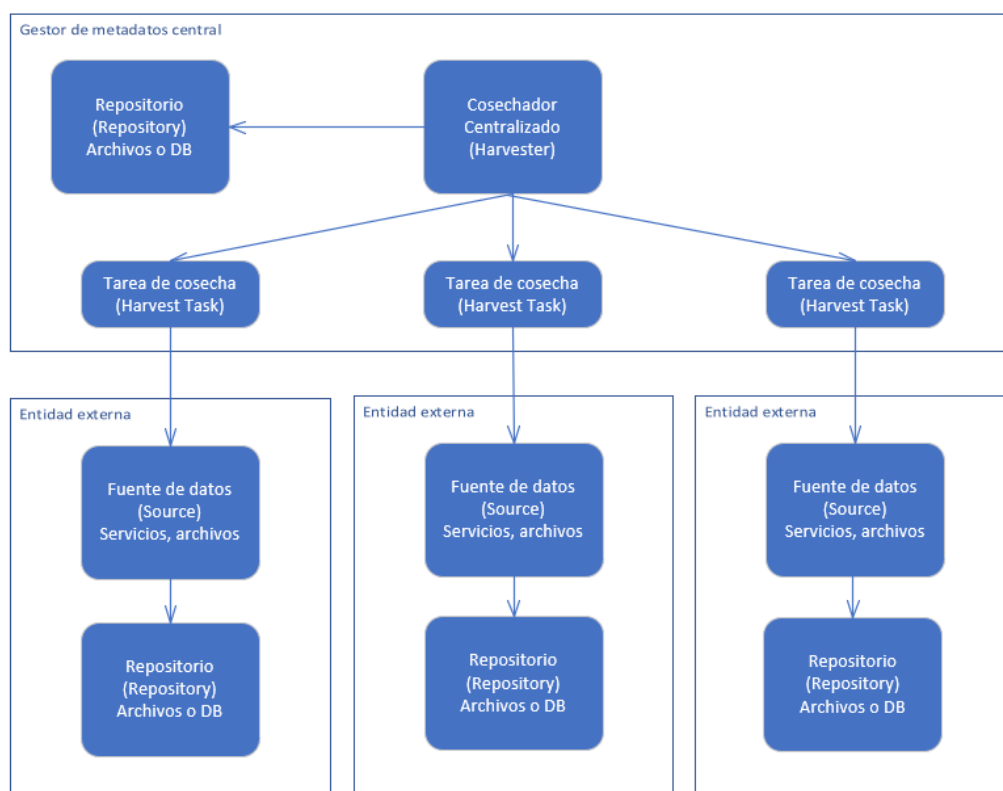


Figura 4.19. Modelo conceptual del cosechador de metadatos. Fuente: Autoría propia

El modelo conceptual que se plantea para el diseño del cosechador de metadatos está compuesto por las siguientes entidades y relaciones:

Cosechador o recolector: se la entidad encargada de recopilar los metadatos de diversas fuentes. En este contexto, estas fuentes son instituciones ya sean públicas o privadas que tienen sus datos disponibles en servidores, que pueden ser catalogadores, servicios web, bases de datos, entre otros. Si la herramienta de recolección dispone de un historial o

configuración, es posible realizar un seguimiento de los metadatos recolectados exitosamente, así como identificar posibles errores ocurridos durante el proceso. Esto permite a un usuario, incluso sin conocimientos técnicos especializados en el proceso de recolección, verificar el estado y progreso del proceso de recolección de metadatos.

Fuente de datos o información: hace referencia a la proveniencia de los datos a partir de los cuales se extraerán los metadatos procedentes de diversas instituciones que cuentan con la información disponible.

Repositorio: representa el espacio designado para el almacenamiento de los metadatos recopilados, esto puede ser bases de datos geográficas. La elección del formato para gestionar la información almacenada varía según las prácticas de cada institución, pudiendo incluir metadatos almacenados en formatos como XML y JSON.

Tarea de cosecha: se refiere en la recolección de metadatos en un determinado momento. Esto incluye la configuración de la fuente de la cual se recolectarán los datos, la programación temporal de la ejecución de la tarea, así como la elección del protocolo o estándar que se utilizará para llevar a cabo la cosecha.

Por otro lado, en cuanto a las relaciones entre las entidades, se presentan de la siguiente manera:

- Un cosechador puede estar vinculado con múltiples fuentes, las cuales pueden ser asociadas a varios cosechadores a su vez.
- Un cosechador puede asociarse a un repositorio donde se almacenará toda la información recolectada de las fuentes.
- Un cosechador tiene la capacidad de gestionar varias tareas de cosecha, en las cuales se encargará de llevar a cabo la acción de recolección correspondiente para una fuente específica. Es importante destacar que una tarea de cosecha está directamente relacionada con un único cosechador.
- Cada tarea de cosecha está enlazada con una única fuente, desde la cual se extraerán los metadatos pertinentes.
- Finalmente, cada tarea de cosecha está asociada con un único repositorio, en el cual se almacenan los metadatos que han sido recopilados

Este modelo conceptual proporciona una visión global que permite entender la estructura para la arquitectura del cosechador de metadatos. Dicha arquitectura se materializará a través de una herramienta especializada en la gestión de metadatos geoespaciales.

4.3. Software requerido para la implementación del cosechador

Se sabe de antemano que la instalación o implementación de una aplicación o software requiere de programas base para garantizar su funcionamiento óptimo. En el caso específico de poner en marcha y configurar el cosechador, se ha optado por utilizar la herramienta de código abierto previamente mencionada: GeoNetwork. Además, requiere el uso de una base de datos, en este caso hemos seleccionado PostgreSQL. A continuación, procederemos a detallar el proceso seguido para lograr un despliegue adecuado y funcional.

Dado que se utilizarán contenedores (Docker) para implementar la aplicación, se ha elegido el sistema operativo Linux. En el contexto de la gestión de contenedores, es necesario generar y establecer la estructura de un archivo conocido como `docker-compose.yml`. Este archivo contiene información sobre los servicios, imágenes de contenedor, configuraciones de red y otros elementos necesarios para crear un entorno coherente y funcional. Se puede acceder a esta información en el proyecto alojado en el siguiente enlace en GitHub: <https://github.com/geonetwork/docker-geonetwork>. Este repositorio está disponible para su clonación o descarga sin restricciones.

4.3.1. Jetty

La versión del proyecto GeoNetwork que se está utilizando es la 4.2.2. Esta configuración se encuentra especificada en el archivo `Dockerfile`. Además, la configuración del servidor de aplicaciones Jetty como la base para ejecutar la herramienta GeoNetwork. Jetty es un contenedor Java para servlets y servidores web ampliamente utilizado en la implementación de aplicaciones web (Eclipse Jetty, n.d.). Los detalles sobre la configuración del archivo `Dockerfile` para el proyecto GeoNetwork se encuentran ilustrados en la Figura 4.20.

```

1 | RUN jetty:9-jdk8
2
3 ENV DATA_DIR /catalogue-data
4 ENV JAVA_OPTS -Dorg.eclipse.jetty.annotations.AnnotationParser.LEVEL=OFF \
5 -Djava.security.egd=file:/dev/./urandom -Djava.awt.headless=true \
6 -Xms2048M -Xss2048M -Xmx3072M -XX:+UseConcMarkSweepGC \
7 -Dgeonetwork.resources.dir=${DATA_DIR}/resources \
8 -Dgeonetwork.data.dir=${DATA_DIR} \
9 -Dgeonetwork.codeList.dir=/var/lib/jetty/webapps/geonetwork/WEB-INF/data/config/codelist \
10 -Dgeonetwork.schema.dir=/var/lib/jetty/webapps/geonetwork/WEB-INF/data/config/schema_plugins
11
12 USER root
13 RUN apt-get -y update && \
14 apt-get -y install \
15 curl \
16 unzip && \
17 rm -rf /var/lib/apt/lists/* && \
18 mkdir -p /${DATA_DIR} && \
19 chown -R jetty:jetty ${DATA_DIR} && \
20 mkdir -p /var/lib/jetty/webapps/geonetwork && \
21 chown -R jetty:jetty /var/lib/jetty/webapps/geonetwork
22
23 USER jetty
24 ENV GN_FILE geonetwork.war
25 ENV GN_VERSION 4.0.6
26 ENV GN_DOWNLOAD_MD5 793732cb9c723e73857a4da73b78451b
27
28 RUN cd /var/lib/jetty/webapps/geonetwork/ && \
29 curl -fSL -o geonetwork.war \
30 https://sourceforge.net/projects/geonetwork/files/GeoNetwork_opensource/v${GN_VERSION}/${GN_FILE}/download && \
31 echo "${GN_DOWNLOAD_MD5} *geonetwork.war" | md5sum -c && \
32 unzip -q geonetwork.war && \
33 rm geonetwork.war
34
35 COPY ./docker-entrypoint.sh /geonetwork-entrypoint.sh
36 ENTRYPOINT ["/geonetwork-entrypoint.sh"]
37 CMD ["java","-jar","/usr/local/jetty/start.jar"]
38
39 VOLUME [ "${DATA_DIR}" ]

```

Figura 4.20. Docker file del proyecto GeoNetwork. Fuente: (GitHub - Official Docker Image GeoNetwork, n.d.).

El producto de este proceso consiste en obtener la imagen del proyecto GeoNetwork, para luego proceder a su configuración en el archivo docker-compose.yml.

4.3.2. GeoNetwork

Dentro del archivo docker-compose.yml, se procede a definir las imágenes necesarias para desplegar el proyecto. Específicamente, para la imagen de GeoNetwork, se adopta la convención de nomenclatura que consiste en utilizar el nombre de la herramienta seguido de su versión, en este caso, “geonetwork:4.2.2”. En cuanto a la asignación de puertos, se opta por utilizar el puerto 8080, tanto a nivel del contenedor como en el servidor de acceso. Esta configuración detallada se encuentra visualmente representada en la Figura 4.21.

```
geonetwork:
  image: geonetwork:4.2.2
  restart: always
  ports:
    - 8080:8080
  environment:
    DATA_DIR: /catalogue-data

    JAVA_OPTS: >
    -Dorg.eclipse.jetty.annotations.AnnotationParser.LEVEL=OFF
    -Djava.security.egd=file:/dev/./urandom -Djava.awt.headless=true
    -Xms512M -Xss512M -Xmx2G -XX:+UseConcMarkSweepGC
    -Dgeonetwork.resources.dir=/catalogue-data/resources
    -Dgeonetwork.data.dir=/catalogue-data
    -Dgeonetwork.codeList.dir=/var/lib/jetty/webapps/geonetwork/WEB-INF/data/config/codelist
    -Dgeonetwork.schema.dir=/var/lib/jetty/webapps/geonetwork/WEB-INF/data/config/schema_plugins

  ES_HOST: elasticsearch
  ES_PROTOCOL: http
  ES_PORT: 9200
  KB_URL: http://kibana:5601

  # Note: defining the following variable will activate the postgresql
  # configuration and will require an external database (see the database
  # service below)
  #POSTGRES_DB_HOST: database
  #POSTGRES_DB_NAME: geonetwork
  #POSTGRES_DB_USERNAME: geonetwork
  #POSTGRES_DB_PASSWORD: geonetwork

  GEONETWORK_DB_TYPE: postgres
  GEONETWORK_DB_HOST: database
  GEONETWORK_DB_PORT: 5432
  GEONETWORK_DB_NAME: geonetwork
  GEONETWORK_DB_USERNAME: geonetwork
  GEONETWORK_DB_PASSWORD: geonetwork

  volumes:
    - geonetwork:/catalogue-data
  depends_on:
    - kibana
    - elasticsearch
    - database
```

Figura 4.21. Imagen de GeoNetwork en archivo Docker-compose.yml. Fuente: (GitHub - Official Docker Image GeoNetwork, n.d.).

Además, es necesario configurar dos herramientas interdependientes en los parámetros de la imagen de GeoNetwork, las cuales facilitarán la búsqueda de metadatos en la plataforma. Estas herramientas son Elasticsearch, que actuará como motor de búsqueda, y Kibana. Asimismo, se requiere incluir en la configuración la imagen de la base de datos PostgreSQL, utilizando la extensión de PostGIS. Esta configuración permitirá contar con una base de datos geoespacial adecuada para el almacenamiento de la información.

4.3.3. PostGIS

Se ha optado por utilizar la versión 14-3.2 de PostGIS como imagen para la base de datos geoespacial. Esta elección se fundamenta en la compatibilidad con otros programas y software alojados en el mismo servidor de implementación que ya operan con esta versión específica. Al mantener la coherencia entre las versiones, evitamos la necesidad de emplear múltiples versiones para herramientas individuales, como GeoNetwork. Para visualizar la configuración detallada del archivo docker-composer.yml, se hace referencia a la Figura 4.22.


```
database:
  image: postgres/postgis:14-3.2
  ports:
    - "5432:5432"
  environment:
    POSTGRES_USER: geonetwork
    POSTGRES_PASSWORD: geonetwork
    POSTGRES_DB: geonetwork
  volumes:
    - pgdata:/var/lib/postgresql/data
```

Figura 4.22. Imagen de Postgis en archivo Docker-compose.yml. Fuente: Autoría propia

4.3.4. Elasticsearch

La versión del motor de búsqueda Elasticsearch utilizada en la configuración es la 7.6.2. Esta elección se basó en consideraciones específicas, ya que se encontraron desafíos para lograr un arranque adecuado. Tras evaluar diferentes opciones, se determinó que la versión 7.6.2 era la más adecuada para la herramienta en cuestión. Para más detalles verifique la Figura 4.23Figura 4.23, donde se presenta la configuración realizada.

```
elasticsearch:
  image: elasticsearch:7.6.2
  environment:
    discovery.type: single-node
  volumes:
    - esdata:/usr/share/elasticsearch/data
```

Figura 4.23. Imagen de Elasticsearch en archivo docker-compose.yml. Fuente: Autoría propia

4.3.5. Kibana

La herramienta Kibana es un componente adicional que se emplea junto al motor de búsqueda. Su función principal consiste en llevar a cabo un análisis de datos a gran velocidad y en una escala considerable, lo que permite obtener observabilidad, fortalecer la seguridad y facilitar las labores de búsqueda (Kibana: Explora, Visualiza y Descubre Datos, n.d.). La configuración necesaria para el archivo docker-compose se presenta en la Figura 4.24.

```
kibana:
  image: kibana:7.6.2
  volumes:
    - ./kibana/kibana.yml:/usr/share/kibana/config/kibana.yml
```

Figura 4.24. Imagen de Kibana en archivo docker-compose.yml. Fuente: Autoría propia

Con respecto a la elección de la versión de Kibana, es aconsejable optar por la misma versión que se utiliza para el motor de búsqueda, dado que Kibana depende de la compatibilidad con dicho motor para un funcionamiento óptimo.

Un aspecto que no ha sido abordado en las secciones anteriores referentes a cada imagen se relaciona con la configuración de los volúmenes en el archivo docker-compose. Estos volúmenes desempeñan un papel crucial, ya que almacenan los datos generados por cada imagen una vez que han sido implementadas o están en funcionamiento. En este contexto, se han definido tres volúmenes específicos: “geonetwork”, “esdata” y “pgdata”. Cada uno de estos volúmenes almacena datos fundamentales para el proyecto de GeoNetwork, incluidas las tareas de cosechamiento, así como la información en la base de datos asociada a esta herramienta.

4.4. Configuración y funcionamiento del cosechador de IG

A continuación, se detalla el proceso de recopilación de datos de diversas instituciones con información geoespacial. Para llevar a cabo este proceso, es necesario tener desplegada la herramienta de gestión de metadatos. En este caso, la herramienta se encuentra disponible en la siguiente dirección URL: <https://ide.ucuenca.edu.ec/geonetwork/>. Al acceder a esta URL, se redirige al catálogo que se ilustra en la Figura 4.25.

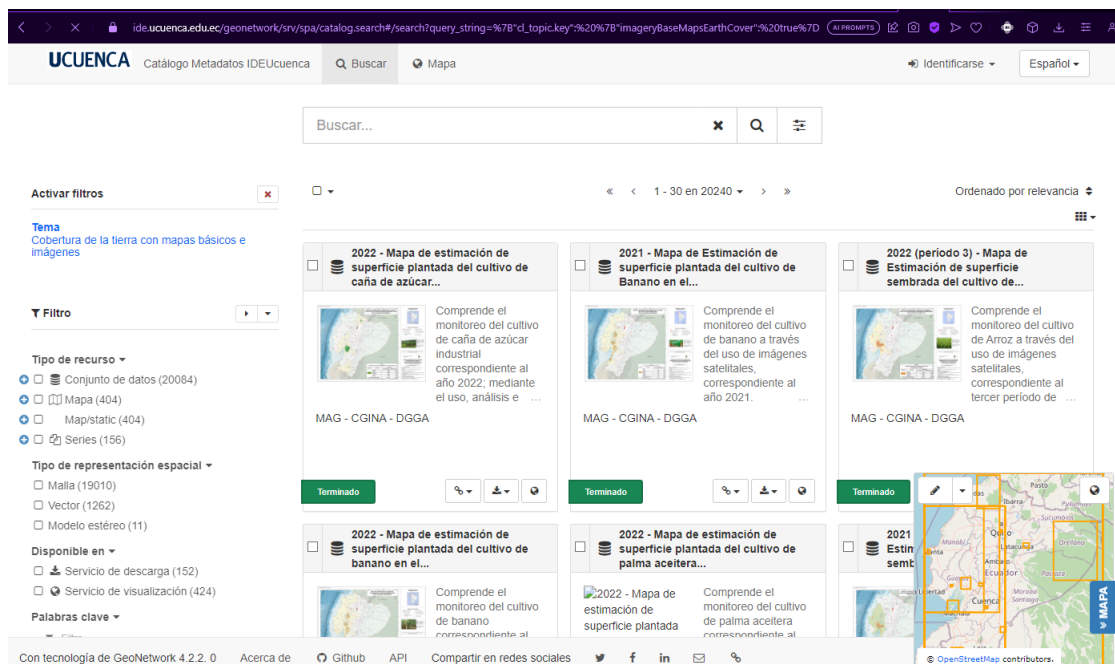


Figura 4.25. Despliegue de herramienta GeoNetwork. Fuente: Autoría propia

Para iniciar el proceso de recolección de metadatos, es fundamental disponer de fuentes de datos que serán sometidas a este procedimiento. En las siguientes secciones, procedemos a

la identificación de estas fuentes y a la configuración específica de la tarea de recolección de datos correspondiente a cada una de ellas.

4.4.1. Identificación de fuente para cosechamiento

Aquí se detalla el proceso de identificación y localización de las fuentes individuales dentro de cada institución, con el propósito de obtener los metadatos respectivos. Para determinar la gama de servicios disponibles, se inicia con la búsqueda en los sitios web oficiales de las instituciones. Para esto, se emplea el motor de búsqueda de Google ingresando el nombre de la institución junto a términos como "geoportal" o "geonetwork" y adicionalmente el nombre de la institución. La Figura 4.26 visualiza el proceso de búsqueda y muestra los resultados obtenidos de manera gráfica.



Figura 4.26. Búsqueda de fuentes en buscador de Google. Fuente: Autoría propia

Después de localizar el portal oficial de la institución, es posible determinar si ofrece servicios tales como Metadatos, Geoservicios o Geoportal. Una vez que uno de estos servicios es identificado, es factible acceder para explorar las opciones disponibles. En el contexto de nuestra investigación, nos enfocaremos en analizar el catálogo de metadatos GeoNetwork y los Servicios CSW. Esto se encuentra ejemplificado en la Figura 4.27, donde se detallan los servicios presentes en el sitio web oficial del Ministerio de Agricultura.

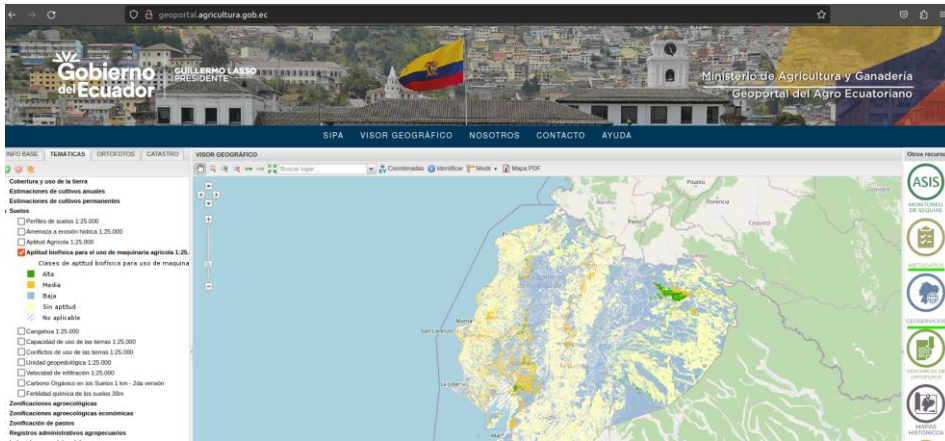


Figura 4.27. Servicios de Metadatos y Mapas del M. de Agricultura. Fuente: Autoría propia

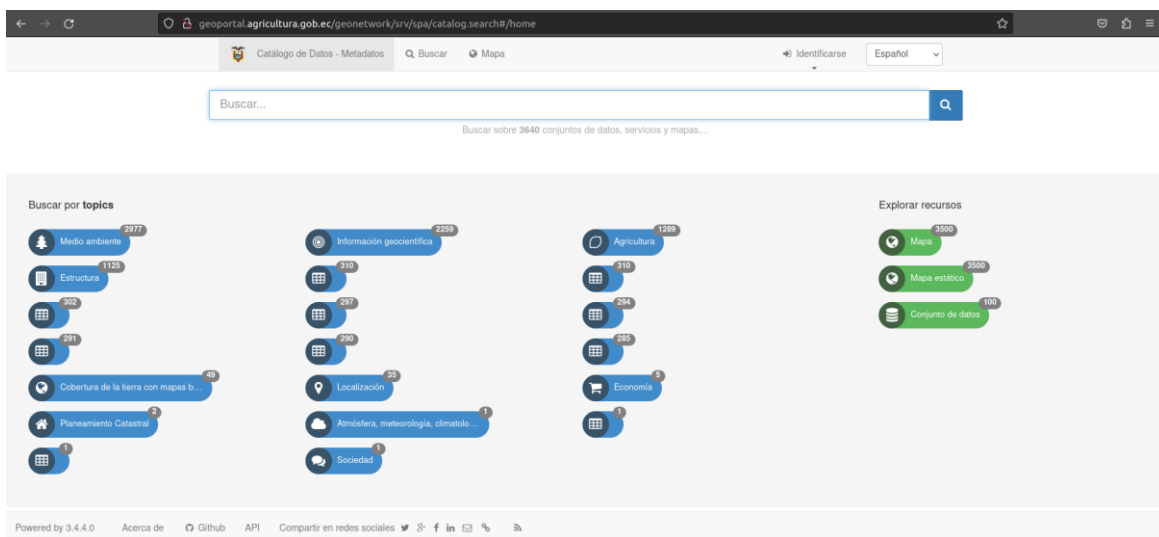


Figura 4.28. GeoNetwork del Ministerio de Agricultura. Fuente: Autoría propia

Como se puede apreciar en la Figura 4.28, la plataforma de gestión de metadatos se encuentra desplegada en un servidor de la institución. A través de este portal, es posible visualizar todos los metadatos generados y verificar su disponibilidad.

Hasta el momento, se han identificado las siguientes fuentes que contienen metadatos geospaciales:

- Sistema Nacional de Información de Tierras (SIGTIERRAS)
- Ministerio del Ambiente
- Universidad del Azuay
- IGM Instituto Geográfico Militar
- MAG Ministerio de Agricultura y Ganadería

- Sistema Nacional de Información (SNI)

Un aspecto destacado para mencionar es que varios gobiernos e instituciones de relevancia a nivel internacional han implementado IDEs. Estas IDEs tienen como propósito brindar una diversidad de datos y recursos geográficos para ser compartidos con el público en general. Un ejemplo notable es el geoportal desarrollado por el Instituto Geográfico Militar (IGM), el cual ofrece una gama de servicios, incluyendo un catálogo de datos basado en software libre, un visor de mapas y la capacidad de descargar información geográfica en diversos formatos (Espinoza Martínez, 2017). Otro ejemplo es la IDESC de Santiago de Cali que ofrece servicios de descargas de mapas, catálogo de metadatos y un geovisor (Ortiz González, 2021). Además, en la IDE denominada como NSDI de la India (Agrawal et al., 2022).

4.4.2. Configuración de fuente por medio de catálogo de metadato

Ahora que hemos definido la fuente de metadatos que se va a recolectar, es necesario proceder con la configuración en nuestro gestor de metadatos previamente implementado en el servidor. En primer lugar, utilizaremos las credenciales de administrador para acceder a las opciones avanzadas que ofrece GeoNetwork. Estas opciones incluyen la gestión de metadatos y plantillas, la recolección, las estadísticas, los estados y la configuración, entre otros.

En particular, nos enfocaremos en la opción de "Recolección". Al seleccionar esta opción, seremos redirigidos a una ventana que presenta los diferentes tipos de servicios y opciones disponibles para el proceso de recolección de metadatos. Esta información está ilustrada en detalle en las Figura 4.29 y Figura 4.30.

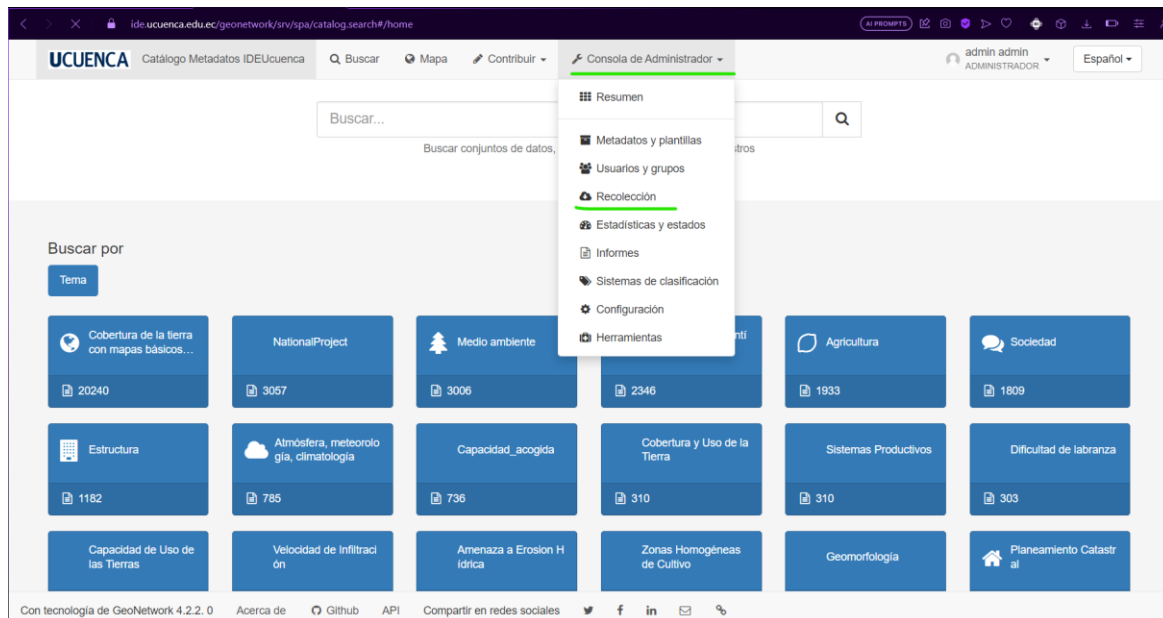


Figura 4.29. Opciones de la herramienta GeoNetwork de la IDE UCuenca. Fuente: Autoría propia

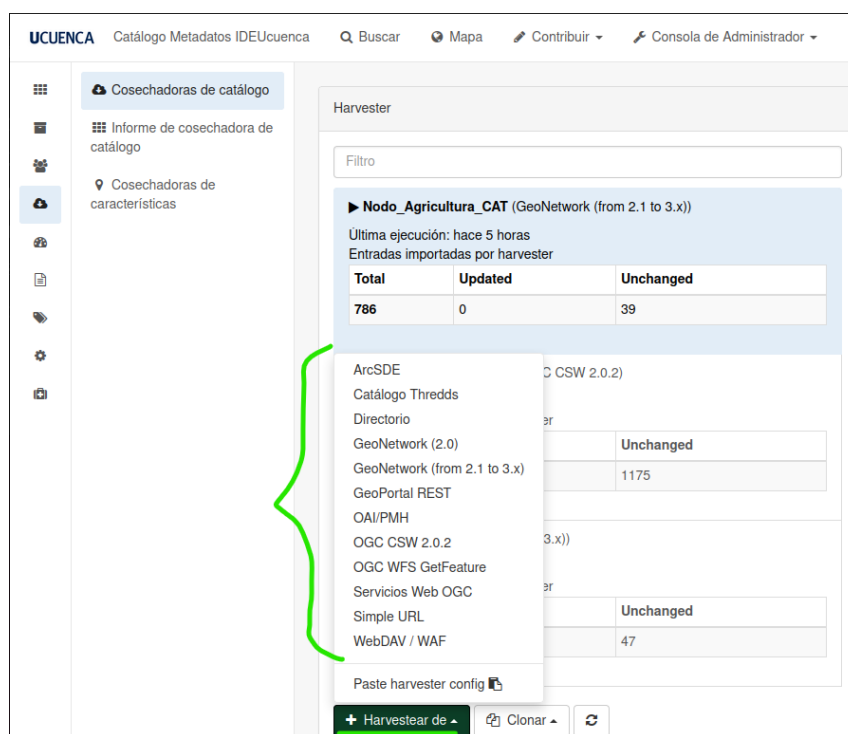


Figura 4.30. Opciones de cosechamiento de metadatos. Fuente: Autoría propia

Se puede observar que en los diversos métodos de recolección de datos se emplean Servicios OGC, los cuales posibilitan la cosecha de servicios CSW y catálogos de metadatos GeoNetwork en sus versiones 2.1 y 3.x, siendo estas las opciones disponibles en las instituciones. Un caso ejemplar mencionado previamente es el Catálogo de Metadatos del Ministerio de Agricultura y Ganadería.

Para iniciar el proceso de recolección de metadatos, será necesario obtener el enlace de despliegue del proyecto GeoNetwork de la institución, que en este ejemplo corresponde a <http://geoportal.agricultura.gob.ec/geonetwork/>. Esta información es crucial para identificar el catálogo con los metadatos que se van a recolectar. En el caso de este tipo de recolección, se seleccionará la opción GeoNetwork (desde 2.1 hasta 3.x). Posteriormente, aparecerá un formulario que deberá completarse con los datos requeridos para la configuración del proceso de recolección. Los campos que deben llenarse en este formulario son:

- Nombre del nodo: un nombre que se puede asignar para identificar la fuente.
- Grupo: esto se puede colocar para identificar el grupo de investigación al que se asigna al configurar el recolector.
- Usuario que se asignó para realizar la recolección.
- Configuración de conexión de GeoNetwork (from 2.1 to 3.x).
- Filtro de búsqueda: puede ser opcional si desea solo una cierta IG que se desea recolectar.
- Privilegios para las entradas importadas con harvester.

Después de completar todos estos datos, el usuario procede a hacer clic en el botón de guardar y, posteriormente, ejecutar el proceso de cosechamiento. En caso de que se presente algún error, la herramienta proporciona mensajes específicos para identificar posibles causas, como la indisponibilidad del servicio de GeoNetwork o la caída del enlace, por mencionar algunas situaciones.

4.5. Gestión de metadatos con información de la IDE UCuenca

En esta sección, se proporcionará una descripción detallada del proceso paso a paso para crear metadatos en la herramienta GeoNetwork. Estos metadatos son esenciales para mejorar la IDE UCuenca. A continuación, se presenta una guía completa sobre cómo realizar este proceso:

1. Acceso a GeoNetwork e inicio de sesión

Para creación y gestión de los metadatos, puede hacer a través del enlace donde esta desplegado la herramienta: <https://ide.ucuenca.edu.ec/geonetwork/> y se encuentra disponible en la IDE. En GeoNetwork, se inicia sesión. (ver Figura 4.31).

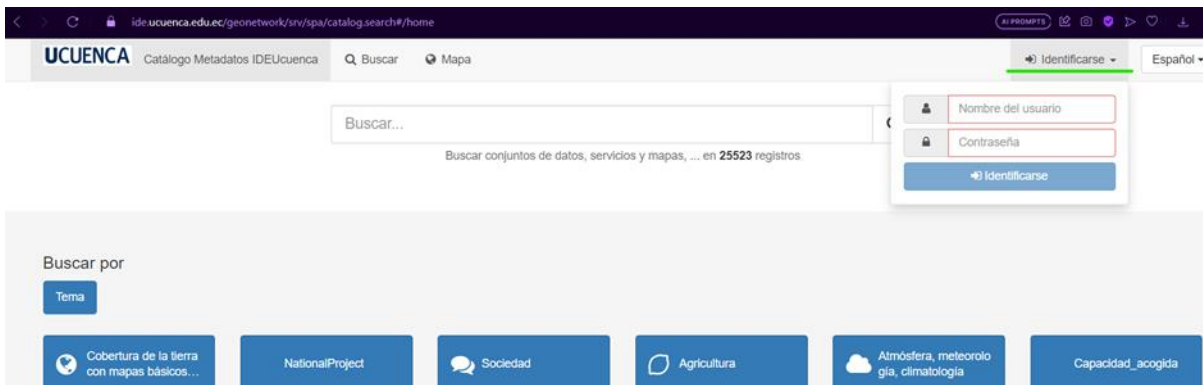


Figura 4.31. Ingreso de credenciales en GeoNetwork. Fuente: Autoría propia

2. Creación del metadato

Para crear un nuevo metadato, en la opción "Contribuir" se despliega un menú con la opción "Añadir una nueva entrada". Esto llevará al formulario de creación de metadatos (ver Figura 4.32).

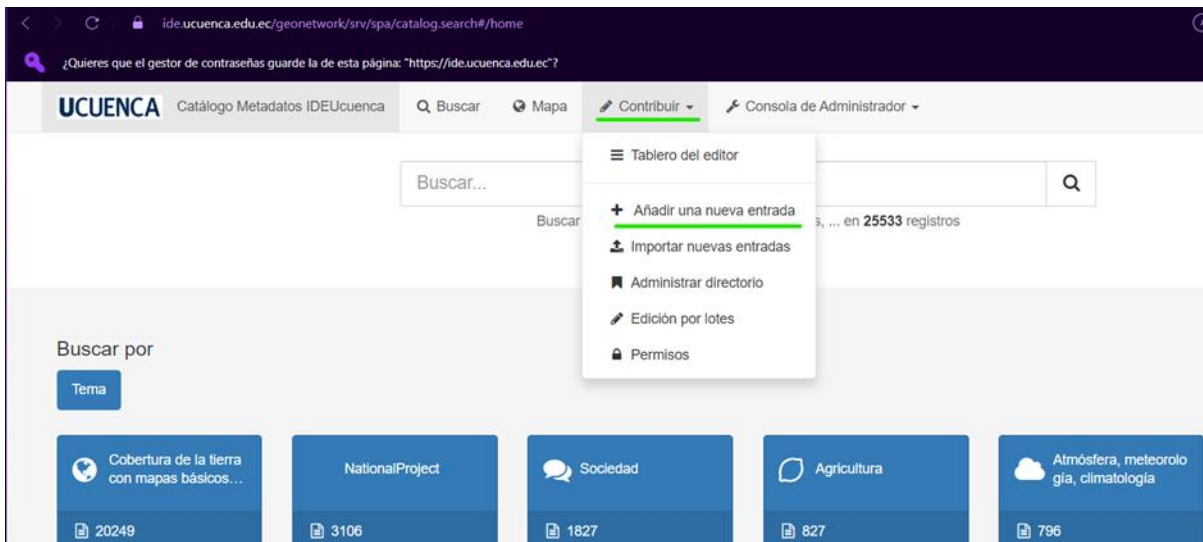


Figura 4.32. Opción para creación y gestión de metadato en GeoNetwork. Fuente: Autoría propia

3. Creación y gestión del metadato

En el proceso de creación del metadato dentro de GeoNetwork, es necesario elegir una plantilla que cumpla con las especificaciones de la norma ISO 19139 o ISO 19115, que rige la estructura y los estándares de metadatos geoespaciales determinados en el PEM (IEDG, 2010). Como ejemplo se ha procedido a crear el metadato correspondiente a la Capa de Red

Sísmica de la Universidad de Cuenca almacenada en el servidor GeoServer de la IDE UCuenca (ver Figura 4.33)

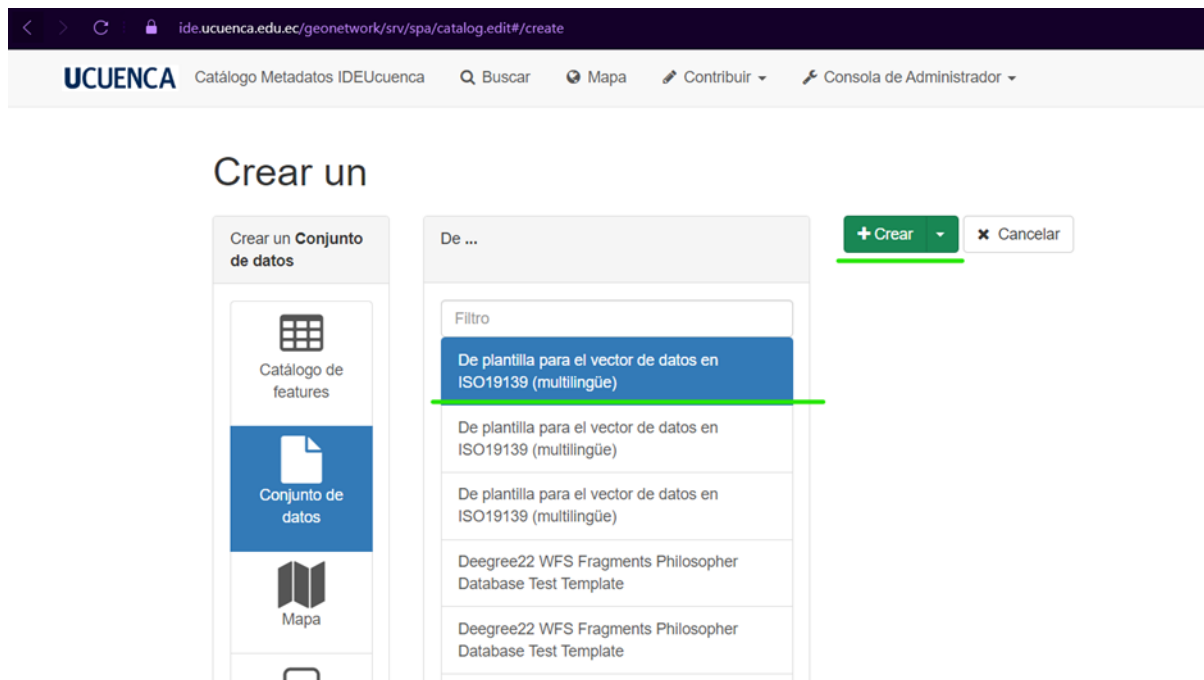


Figura 4.33. Selección de plantilla ISO 19139 para creación de metadato. Fuente: Autoría propia

Se debe completar los campos obligatorios y/u opcionales que dispone en la plantilla en el formulario de metadatos. Los campos obligatorios en la plantilla ISO 19139 son los siguientes:

- Título: El título del recurso geoespacial.
- Resumen: Una descripción breve del recurso.
- Organización: Nombre de la entidad que genera el recurso.
- Fecha de publicación: Fecha en la que se va a publicar el recurso geoespacial.
- Palabras clave: Términos que describan el contenido.
- Categoría: La categoría del recurso (por ejemplo, mapa, imagen, conjunto de datos).
- Formato: El formato del recurso (por ejemplo, shapefile, imagen JPEG).

Por otro lado, si en el caso de que se disponga archivos adicionales relacionados al metadato, como documentos, imágenes o enlaces a recursos, se pueden cargar en las secciones de “Descargar” y “Recursos asociados” (ver Figura 4.34).

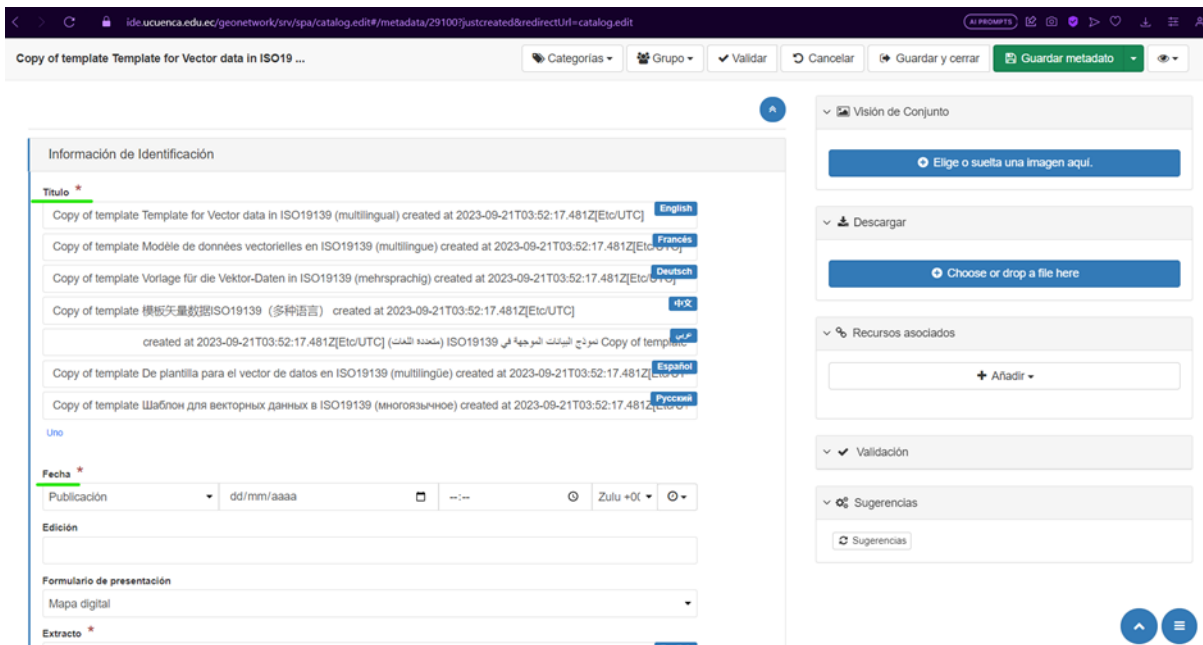


Figura 4.34. Formulario de metadatos con los campos obligatorios y opcionales. Fuente: Autoría propia

Una vez almacenado el metadato, será posible visualizar como se muestra en las Figura 4.35 y Figura 4.36.

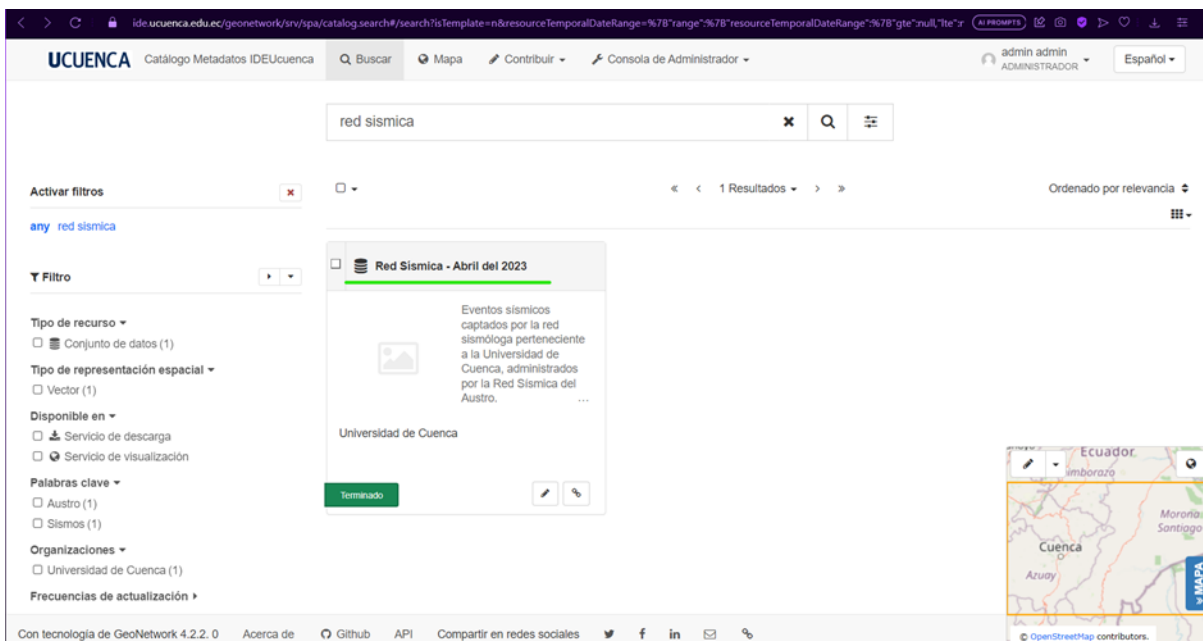


Figura 4.35. Metadato creado y buscarlo por nombre. Fuente: Autoría propia

The screenshot shows a web application interface for a metadata catalog. At the top, there is a navigation bar with the UCUENCA logo and various utility links like 'Buscar', 'Mapa', and 'Contribuir'. Below this, a search bar and action buttons (Editar, Borrar, Gestionar registro, Descargar, Modo de visualización) are visible. The main content area features a title 'Red Sísmica - Abril del 2023' with a green 'Terminado' status badge. A descriptive paragraph follows: 'Eventos sísmicos captados por la red sísmológica perteneciente a la Universidad de Cuenca, administrados por la Red Sísmica del Austro.' Below the text is a 'Discover data' section containing a map of the region around Cuenca, Ecuador, with labels for various locations like Loja, Guayaquil, Cuenca, and Machala. At the bottom of the visible content is an 'Enlaces' section.

Figura 4.36. Detalles del recurso y su metadato almacenado. Fuente: Autoría propia

5. Resultados

5.1. Cosechamiento de metadatos geoespaciales

A continuación, en la Figura 5.1 se muestran los resultados obtenidos del cosechamiento realizado a cada uno de los servicios de catálogos que tienen las diferentes instituciones públicas identificadas, el mismo que se encuentra en el IDE de la Ucuena en el siguiente enlace <https://ide.ucuenca.edu.ec/geonetwork> . Además, se consideró los recursos de software que se tenían a disposición, y pruebas necesarias, para trabajar de la mejor forma en el proceso de recolección.

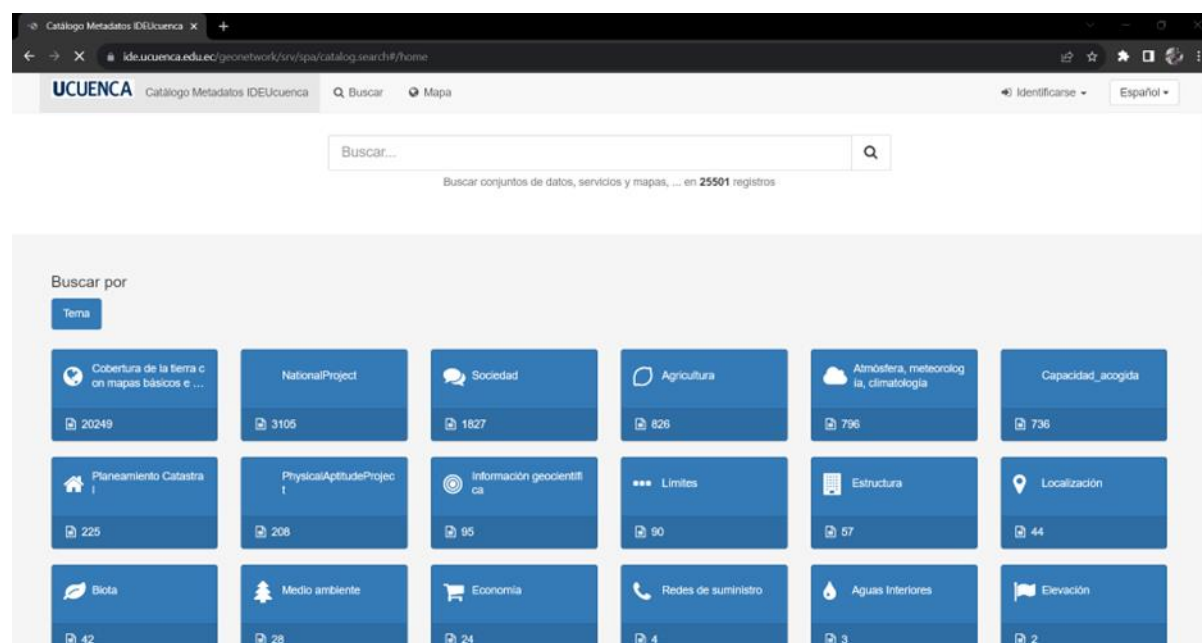


Figura 5.1. Resultado de cosechamiento de metadatos en GeoNetwork. Fuente: Autoría propia

5.1.1. Análisis del conjunto de metadatos en los geoportales

En esta sección, se llevó a cabo un exhaustivo análisis de los servicios de catálogos de metadatos en diversas instituciones del país. El objetivo principal era identificar y cuantificar el total de metadatos disponibles en cada una de estas plataformas, con el propósito de comprender la riqueza y la diversidad de la información que se encuentra accesible a través de estos sistemas.

Para lograr este objetivo, se identificaron instituciones representativas del país que cuentan con servicios de catálogos de metadatos en sus sitios web, como son:

Tabla 5.1. Instituciones con metadatos disponibles en los sitios web. Fuente: Autoría propia

Institución	Total metadatos (Enero-2023)	Total metadatos (Julio-2023)
Sistema Nacional de Información de Tierras (SIGTIERRAS)	4264	4264
Ministerio del Ambiente	425	351
Universidad del Azuay	19	47
Instituto Geográfico Militar (IGM)	6616	13079
Ministerio de Agricultura y Ganadería (MAG)	3635	3640
Sistema Nacional de Información (SNI)	48671	12579
Total	63630	33960

Luego de identificar el total de metadatos a cosechar, se procedió a realizar un seguimiento al total de metadatos existentes. Esto se debe a que, desde un inicio, ciertas instituciones optaron por actualizar las herramientas a una versión más actual o movían toda la información a un portal con un dominio diferente. Un ejemplo claro fue el IGM, que desde el principio que se realizó la búsqueda sobre los servicios que tenían se logró obtener un catálogo de metadatos con un aproximado de 480 en total, pero por información que nos brindaron las personas encargadas de la plataforma nos ayudaron con un catálogo que tiene mucha más información, pero no está enlazado directamente con el portal del IGM. Y este catálogo que nos brindaron tenían un aproximado de 12.599 metadatos, el cual fue mucho mayor al que se encuentra directamente en el sitio web. Por otro lado, el SIN en su catálogo desde el inicio de la investigación ha tenido un total aproximado de 48617 metadatos y trabajaba con la versión 3.0.2.0, pero en la actualidad se encuentra con 12579 en una versión superior de GeoNetwork 4.2.3. En la Figura 5.2 se puede apreciar una captura del geoportal en la fecha de enero del 2023 y en la Figura 5.3 se tiene una captura del sitio web en el mes de julio hasta la actualidad.

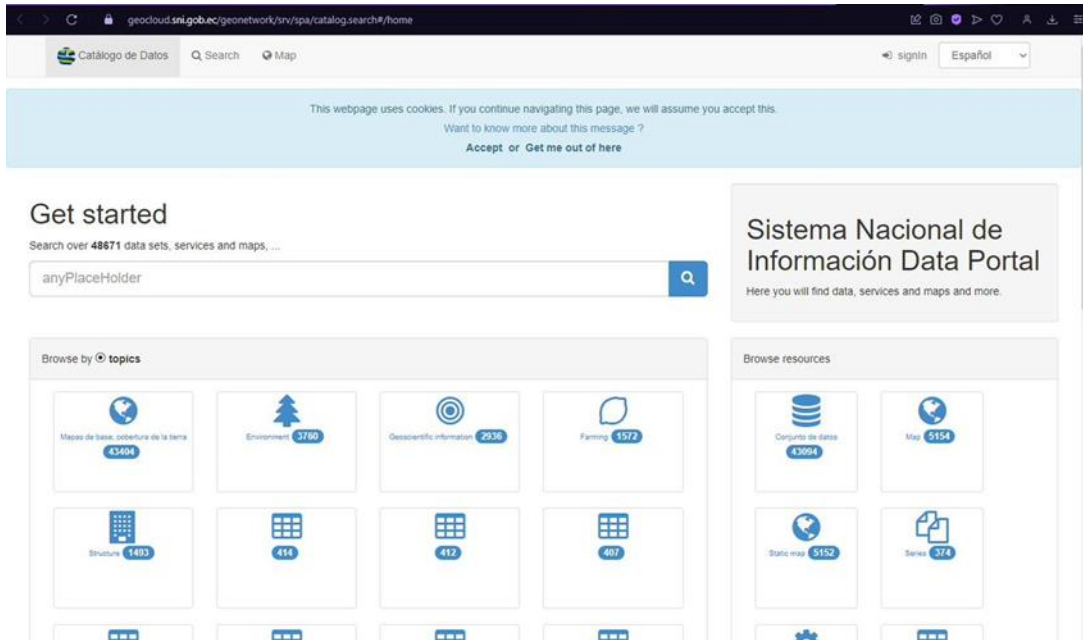


Figura 5.2. GeoNetwork del SNI en enero del 2023. Fuente: Autoría propia

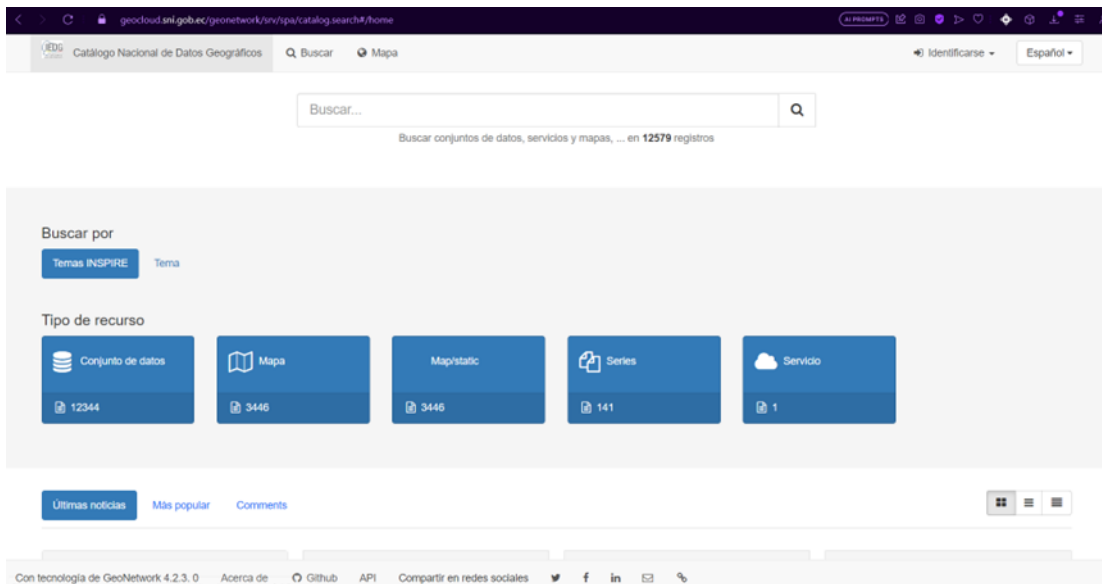


Figura 5.3. GeoNetwork del SNI en julio del 2023. Fuente: Autoría propia

Como se puede apreciar, el total de metadatos que contiene el catálogo del SNI puede variar por razones de mantenimiento en el servidor de aplicaciones, que puede ser por cambio de versión en la herramienta, o el dominio en el que se encuentra, y para cumplir con esa meta es posible realizar una migración de la información de un portal a otro. Este tipo de variaciones en los metadatos pueden ocurrir en la mayoría de instituciones ya que están a la expectativa de que la herramienta se mantenga actualizada y accesible para los usuarios.

Los resultados revelaron una gran variabilidad en la cantidad de metadatos disponibles en los servicios de catálogos. Algunas de ellas brindan una amplia gama de metadatos que

permitían una descripción detallada de los recursos, mientras que otras se limitaban a proporcionar información básica.

5.1.2. Cosechamiento de información

Para lograr el objetivo de este trabajo, se puso en funcionamiento el proceso del cosechamiento o recolección de metadatos en la herramienta seleccionada GeoNetwork, haciendo uso de los servicios de catálogos en las instituciones previamente mencionadas en la sección 5.1.1. Cada una de estas instituciones contiene diversos tipos de recursos con IG, como son mapas, documentos e imágenes.

Las fuentes de cada uno de los catálogos describen en la Tabla 5.2 Se debe considerar que estos enlaces actualmente están disponibles, pero puede darse que éstos no se encuentren disponibles por motivos de mantenimiento a la plataforma y será necesario buscar en el sitio web de la institución.

Tabla 5.2. Fuentes de sitios web de los catálogos. Fuente: Autoría propia

Institución	GeoNetwork
Sistema Nacional de Información de Tierras (SIGTIERRAS)	http://metadatos.sigtierras.gob.ec/geonetwork/srv/spa/catalog.search#/home
Ministerio del Ambiente	http://ide.ambiente.gob.ec/geonetwork/srv/spa/catalog.search;jsessionid=317E30241AFAF2FD77AD9C5480ACF4A6#/home
Universidad del Azuay	http://gis.uazuay.edu.ec:8080/geonetwork/
Instituto Geográfico Militar (IGM)	https://www.geoportaligm.gob.ec/geonetwork/srv/spa/catalog.search#/home https://www.geoportaligm.gob.ec/catalogo/srv/spa/catalog.search#/home
Ministerio de Agricultura y Ganadería (MAG)	http://geoportal.agricultura.gob.ec/geonetwork/srv/spa/catalog.search#/home
Sistema Nacional de Información (SNI)	https://geocloud.sni.gob.ec/geonetwork/srv/spa/catalog.search#/home

Para iniciar con el proceso de recolección, se configuró cada uno de los catálogos y el resultado se describe en la Tabla 5.3.

Tabla 5.3. Cosechamiento de metadatos con las fuentes identificadas. Fuente: Autoría propia

Institución	Version GeoNetwork de fuente	Total de fuente	Total Cosechado
Sistema Nacional de Información de Tierras (SIGTIERRAS)	3.0.3.0	4264	0
Ministerio del Ambiente	3.4.4.0	351	0
Universidad del Azuay	3.4.2.0	47	47
Instituto Geográfico Militar (IGM)	4.0.6.0	6117	0
	3.0.3.0	6962	5133
Ministerio de Agricultura y Ganadería (MAG)	3.4.4.0	264	251
Sistema Nacional del Información (SNI)	3.0.2.0	48671	20070

El resultado obtenido con cada uno de los catálogos es distinto con respecto al tiempo transcurrido; esto significa que, al momento de recolectar la información, ciertos catálogos terminan el proceso mucho antes que otros. Esto no se debe a la gran cantidad que puede tener el catálogo, sino a la configuración del servidor y la red en la que se encuentre la herramienta desplegada.

El periodo de cosechamiento que se tomó en cuenta fue alrededor de 90 a 120 días. La programación del tiempo para iniciar el cosechamiento en cada institución fue de 7 o 15 días, 7 días para catálogos con un total aproximado de 5000 metadatos y 15 días a catálogos mayores a 5000 metadatos. La razón del tiempo para la programación de cosechamiento, fue porque la actualización de información en los catálogos se realizaba cada semana o dos semanas aproximadamente.

El tiempo de 90 a 120 días que se estimó, fue porque en pruebas realizadas en un servidor local, con un intervalo de tiempo de 30 minutos, se descubrió que la recolección en ciertos catálogos era más lenta y en otros más rápida. En la Tabla 5.4 se describe de manera detallada los recursos utilizados y el tiempo de cosechamiento de prueba.

Tabla 5.4. Pruebas de cosechamiento de metadatos en servidor local. Fuente: Autoría propia

Institución	Versión GeoNetwork de fuente	Total de fuentes analizado	Total Cosechados	Tiempo
Sistema Nacional de Información de Tierras (SIGTIERRAS)	3.0.3.0	4264	1832	30 min
Ministerio del Ambiente	3.4.4.0	351	351	25 min
Universidad del Azuay	3.4.2.0	47	47	10 min
Instituto Geográfico Militar (IGM)	4.0.6.0	6117	0	0 min
	3.0.3.0	6962	291	30 min
Ministerio de Agricultura y Ganadería (MAG)	3.4.4.0	3640	1832	20 min
Sistema Nacional de Información (SNI)	3.0.2.0	48671	805	30 min

Los resultados finales revelaron que el proceso de cosechamiento fue exitoso en un 45.27 %, 25506 de un total de 63630. Esto se debe a que el catálogo del SNI, tarda más de 12 días en cosechar solamente un aproximado de 7000 metadatos, por tanto, el proceso de cosechamiento llega a cancelarse o terminar de forma brusca.

Por otro lado, se logró extraer una cantidad significativa de IG sobre los recursos disponibles en los catálogos, con la excepción del SNI, lo que permitió obtener una visión panorámica de la oferta de información en cada institución. Además, se observó que algunas instituciones proporcionaban metadatos más completos y estructurados, lo que facilitaba la identificación y recuperación de recursos específicos. Por otro lado, no fue posible obtener información de portales como SIGTIERRAS y el Ministerio del Ambiente, ya que estos portales se encontraban inactivos o dejaban de funcionar con frecuencia

5.2. Dificultades y limitaciones en el cosechamiento

El proceso de cosechamiento en la herramienta de GeoNetwork, se puede considerar que es de fácil configuración, pero la falta de información de cómo se debe configurar de manera correcta y adecuada es muy escasa. Se identificaron ciertos desafíos y limitaciones que se encontraron al momento de realizar el cosechamiento, entre los que más se destacaron son las siguientes:

- **Configuración en panel de cosechamiento o recolección (Harvesting).**

El proceso del cosechamiento en GeoNetwork a primera vista es relativamente sencillo, ya que es completar un formulario y guardarlo para posteriormente ejecutarlo, pero al momento de realizar la primera prueba de cosechamiento salían errores, los cuales no se podían identificar si es un problema de la fuente del catálogo o de la configuración en el cosechador.

Luego de una revisión en diferentes foros y fuentes en la web, se pudo encontrar información sobre como completar la información para el cosechamiento. Además, con ayuda de personas especialistas encargadas del catálogo de metadatos del IGM, nos recomendaron la configuración que utilizan para realizar la recolección.

- **Cosechamiento con asignación de memoria RAM de 512 MB a 1024 MB.**

Un error que se identificó desde un inicio, fue el despliegue y el proceso del cosechamiento con cualquier servicio de catálogo de metadatos o servicio que contenga metadatos con IG. Si en el caso de que un servidor en el cual se despliega la herramienta no tiene los suficientes recursos asignados para el cosechamiento, es posible que no permita realizar el proceso para ninguna fuente de información. Se obtuvieron mejores resultados al trabajar asignando memoria RAM de 2GB y superior.

Por otro lado, un servidor que tenga en total 6GB en RAM, y se asigna 2GB para el cosechamiento, existe la posibilidad de que llegue a detenerse el proyecto de GeoNetwork por completo conjuntamente con el cosechamiento. Esto se debe a que además de asignar una cantidad de memoria al cosechamiento (2GB), se debe asignar al proyecto (2GB) para un adecuado arranque del sistema; esto llevaría a un total de 4 GB en uso en total y con un servidor que tenga 6GB puede llevar a tener problemas en el cosechamiento. Además, el almacenamiento de información es indispensable, por tanto, el espacio en disco debe ser considerable y dependiente de la cantidad se está cosechando, ya que no solo almacena registros en una base de datos geográfica (BDG) sino documentos e imágenes.

- **Configuración de red y firewall en servidor de instituciones.**

Luego de ya revisar una configuración que sea adecuada para el cosechamiento, se tuvieron inconvenientes con los servicios de los catálogos al momento de acceder a los mismo en la configuración del cosechador. Esto se debía a que, las instituciones no configuraban la red o permisos en el servidor, para proporcionar acceso al momento de enviar solicitudes de parte del cosechador.

Por otro lado, si ya se podía visualizar el cosechamiento en un historial, se cosechaba muy poco en un gran periodo de tiempo o el servicio del catálogo llegaba a caerse. En el transcurso de tiempo, las personas que estaba encargado de los catálogos verificaban la configuración

en su red y su disponibilidad, que al notificarnos se podía ver un cosechamiento más fluido de la información.

- **Integración del cosechador con el IDE UCuenca.**

En el contexto del despliegue del recolector junto con el IDE UCuenca, nos enfrentamos a varios desafíos y complicaciones. Uno de los principales obstáculos fue la necesidad específica de utilizar Docker para poner en funcionamiento el servidor de la IDE UCuenca. Esto, a su vez, planteó la dificultad de encontrar la configuración más adecuada, que incluía la elección de la versión de GeoNetwork, la configuración del almacenamiento local y de la base de datos, así como la gestión de los recursos disponibles para el proyecto.

Inicialmente, se optó por la versión 3.0.1 de GeoNetwork debido a la disponibilidad de una mayor documentación para dicha versión. Sin embargo, esta elección generó problemas al intentar recolectar datos de otros GeoNetworks con versiones más antiguas o más recientes. Por lo tanto, se decidió migrar a la versión 4.2.2, versión más reciente y estable de hasta la fecha de GeoNetwork, lo que abordó en gran medida los problemas mencionados anteriormente. Aun así, esta elección requirió un esfuerzo adicional en la búsqueda de documentación y guías de configuración, ya que eran realmente escasas para esta versión.

Otra dificultad que surgió en este contexto fue la gestión de los recursos disponibles en el servidor para el proyecto. Si no se configuraba adecuadamente el almacenamiento de datos, los metadatos, documentos e imágenes se almacenaban localmente en el servidor, lo que resultaba en una pérdida de espacio para cualquier otra gestión, provocando en algún punto la caída del servidor como se puede observar en la Figura 5.4. Este punto, en gran medida, implicó un proceso de prueba y error para configurar correctamente la base de datos PostgreSQL junto con su extensión PostGIS, con el objetivo de garantizar que los datos se almacenaran y recuperaran de manera efectiva en la IDE.

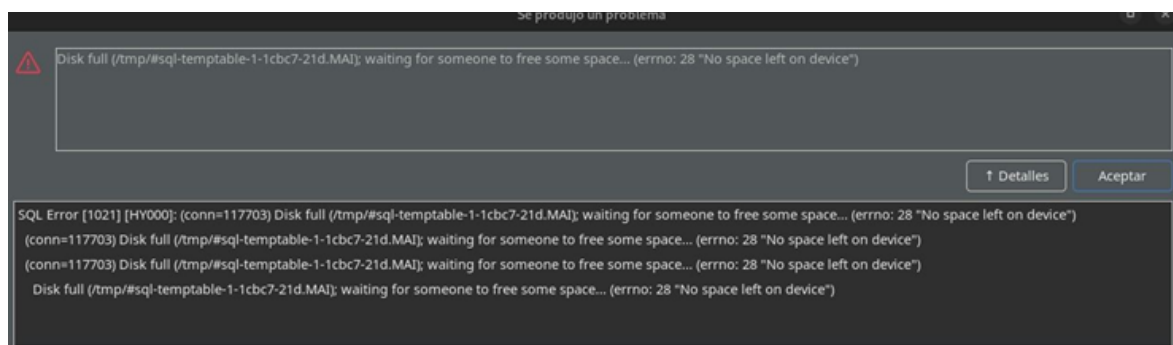
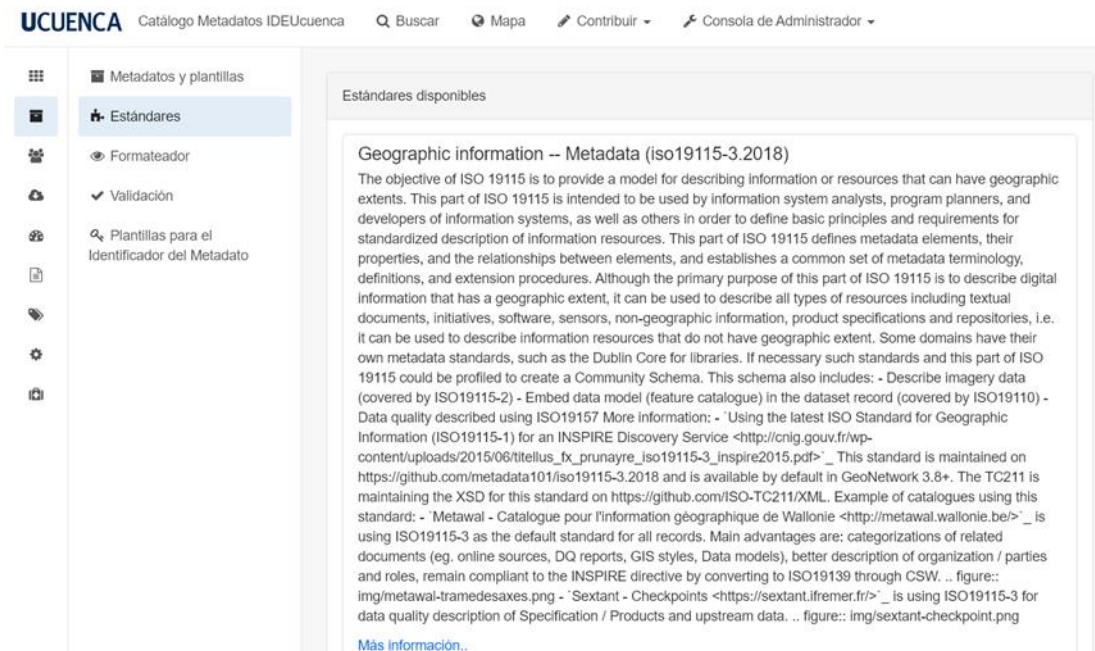


Figura 5.4. Servidor con sobrecarga de información local. Fuente: Autoría propia

- **Carga de Plantillas para el ingreso de metadatos.**

Se detectó un inconveniente con el cosechador de GeoNetwork integrado en el IDE al intentar crear un nuevo registro bajo las normas ISO configuradas. En la Figura 5.5, se observa que la ISO 19115 está cargada; no obstante, en la Figura 5.6, se aprecia que no estaba disponible como plantilla para la creación de un conjunto de datos. Esta limitación impidió realizar el registro de metadatos utilizando la ISO 19115, obligando a utilizar exclusivamente la norma ISO 19139.



The screenshot shows the UCUENCA metadata catalog interface. The top navigation bar includes the UCUENCA logo, the text 'Catálogo Metadatos IDEUcuenca', and search and navigation icons. A sidebar on the left contains a menu with options: 'Metadatos y plantillas', 'Estándares' (highlighted), 'Formateador', 'Validación', 'Plantillas para el Identificador del Metadato', and several other icons. The main content area is titled 'Estándares disponibles' and displays the details for the 'Geographic information -- Metadata (iso19115-3.2018)' standard. The text describes the objective of ISO 19115, its intended use by information system analysts and developers, and its role in defining metadata terminology and requirements. It also lists various applications and standards that utilize ISO 19115-3, such as the Dublin Core for libraries and the INSPIRE Discovery Service. A link for 'Más información..' is provided at the bottom of the standard's description.

Figura 5.5. Plantilla de la norma ISO 19115 registrada en la configuración de GeoNetwork. Fuente: Autoría propia

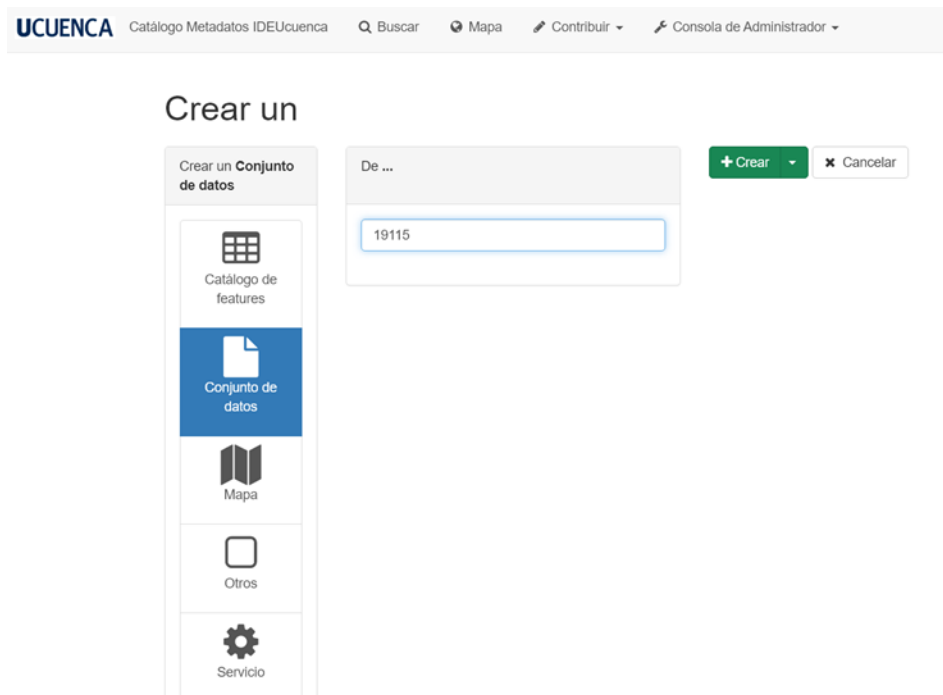


Figura 5.6. Error de carga de la plantilla de la ISO 19115 para generar un nuevo registro. Fuente: Autoría propia

5.3. Análisis de usabilidad del cosechador de metadatos

Para llevar a cabo el análisis de usabilidad, se optó por utilizar el cuestionario PSSQU (Post-Study System Usability Questionnaire) debido a su destacado índice de aceptación en el ámbito de evaluaciones de usabilidad. Este cuestionario fue administrado a estudiantes y egresados de la carrera de ingeniería de sistemas de la Universidad de Cuenca, con conocimientos en Geoinformación, logrando un total de 30 participantes a través de un formulario en Google Forms disponible en el siguiente enlace: <https://forms.gle/PhrKtSAn5FjMxWry7>.

La aplicación del cuestionario se centró en cuatro escenarios distintos:

- Búsqueda y acceso a datos Geoespaciales.
- Integración con otras herramientas.
- Soporte y Documentación.
- Visualización de datos Geoespaciales.

Adicionalmente, se incorporaron preguntas generales sobre el uso de GeoNetwork en este cuestionario.

A continuación, se presenta una descripción detallada de cada escenario, sus tareas correspondientes, la evaluación general y los resultados específicos obtenidos:

Escenario 1: Búsqueda y acceso a datos Geoespaciales

- **Tarea 1: Búsqueda de información Geoespacial por Criterios Específicos.** Utilice GeoNetwork para buscar información geoespacial sobre inundaciones en la provincia del Azuay. (1. Poco Complicado – 7. Muy Complicado)

Objetivo: Evaluar la facilidad percibida por los usuarios al realizar la búsqueda de información geoespacial sobre inundaciones en la provincia del Azuay utilizando GeoNetwork.

Resultados: En la Figura 5.7, se muestran los resultados obtenidos con una media total obtenida fue de 2,53. Este valor sugiere que la percepción de los participantes fue de manera favorable al momento de la realización de la tarea, indicando una facilidad de uso del sistema.

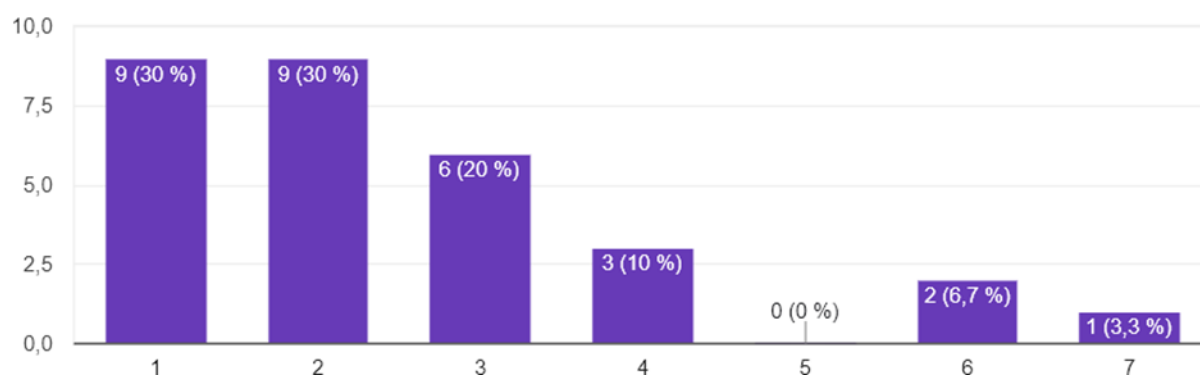


Figura 5.7. Resultados Escenario 1, Tarea 1. Fuente: Autoría propia

- **Tarea 2: Filtrado y Refinamiento de Resultados.** Después de realizar la búsqueda en GeoNetwork, utilice las opciones de filtro y refinamiento para reducir los resultados a un conjunto específico de datos geoespaciales. (1. Poco Complicado – 7. Muy Complicado)

Objetivo: Evaluar la facilidad de uso al aplicar opciones de filtro y refinamiento en GeoNetwork para reducir los resultados de búsqueda a un conjunto específico de datos geoespaciales.

Resultados: En la Figura 5.8 se muestran los resultados obtenidos con una media total obtenida fue de 2,57, indicando una percepción favorable de la facilidad de filtrado y refinamiento de resultados en GeoNetwork.

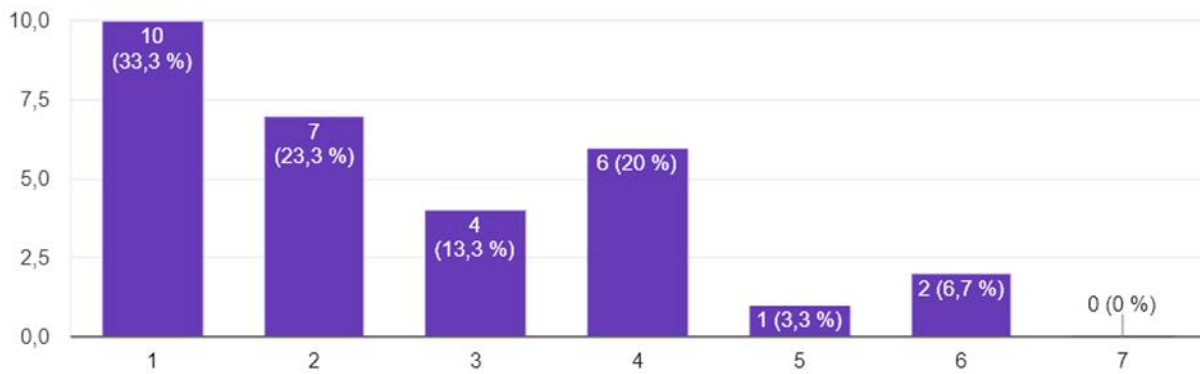


Figura 5.8. Resultados Escenario 1, Tarea 2. Fuente: Autoría propia

- **Tarea 3: Descarga de Datos Seleccionados.** Seleccione un conjunto de datos geospaciales sobre inundaciones en el Azuay y descárgalo desde GeoNetwork en el formato que prefiera. (1. Poco Complicado – 7. Muy Complicado)

Objetivo: Evaluar la facilidad de descargar un conjunto de datos geospaciales específicos sobre inundaciones en la provincia del Azuay desde GeoNetwork en el formato preferido.

Resultados: En la Figura 5.9, se muestran los resultados con una media total obtenida fue de 2,37, reflejando una percepción positiva en cuanto a la facilidad de descarga de datos seleccionados.

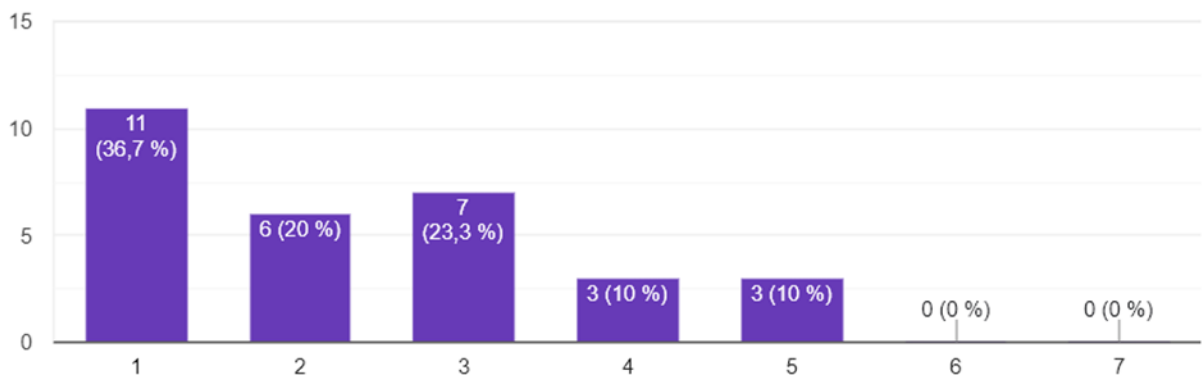


Figura 5.9. Resultados Escenario 1, Tarea 3. Fuente: Autoría propia

Media Total = 2,37

- **Tarea 4: Evaluación de la Documentación de los Datos Descargado.** Después de descargar los datos, consulte la información incluida en el mismo (metadatos) para

comprender mejor su contenido y contexto. Evalúe la utilidad de la documentación y si le ayudó a comprender los datos de manera efectiva. (1. Muy Útil – 7. Poco Útil)

Objetivo: Evaluar la utilidad de la documentación (metadatos) asociada con los datos descargados para comprender su contenido y contexto.

Resultados: En la Figura 5.10 se muestran los resultados obtenidos con una media total obtenida fue de 2,7, indicando una percepción favorable en cuanto a la utilidad de la documentación para comprender de manera efectiva los datos descargados.

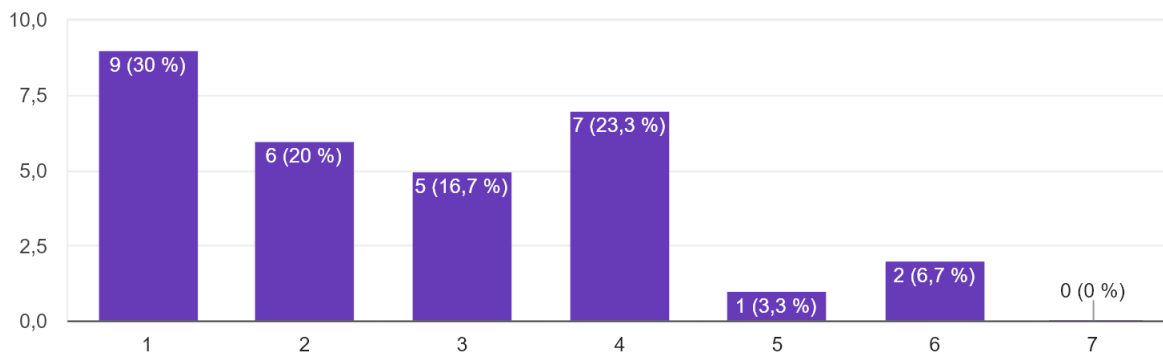


Figura 5.10. Resultados Escenario 1, Tarea 4. Fuente: Autoría propia

Escenario 2: Integración con Otras Herramientas

- **Tarea 1: Exportación de Datos.** Existen diversas herramientas que hacen uso de información Geoespacial y dependiendo de la misma, los tipos de archivos que soportan son diferentes. Geonetwork tiene la posibilidad de almacenar y descargar los metadatos en varios tipos de archivos. Realiza diferentes búsquedas y señala la cantidad de tipos diferentes de archivos que lograste identificar.

Objetivo: Evaluar la versatilidad de GeoNetwork para la exportación de metadatos en diferentes tipos de archivos, considerando la diversidad de herramientas que hacen uso de información geoespacial.

Resultados: En la Figura 5.11 presenta los resultados, en el cual el resultado más votado fue la identificación de 3 tipos de archivos diferentes en los cuales se pueden almacenar y descargar los metadatos utilizando GeoNetwork.

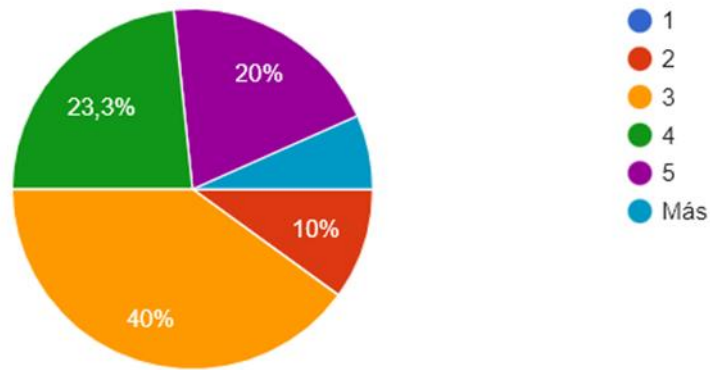


Figura 5.11. Resultados Escenario 2, Tarea 1. Fuente: Autoría propia

Escenario 3: Soporte y Documentación

- **Tarea 1: Utilización de la Documentación.** Encuentra un conjunto de datos geoespaciales en GeoNetwork que te interese. Luego, busca la documentación o metadatos asociados con ese conjunto de datos para obtener información detallada sobre su contenido y origen. Evalúa la utilidad de la documentación y si te ayudó a comprender mejor el conjunto de datos. (1. Poco Complicado – 7. Muy Complicado)

Objetivo: Evaluar la efectividad de la documentación o metadatos asociados con un conjunto de datos geoespaciales encontrado en GeoNetwork, con el propósito de obtener información detallada sobre su contenido y origen.

Resultados: En la Figura 5.12 se puede observar los resultados para esta tarea, en el cual la media total fue de 2,33, indicando que los participantes percibieron en promedio una dificultad moderada al utilizar la documentación para comprender el conjunto de datos seleccionado.

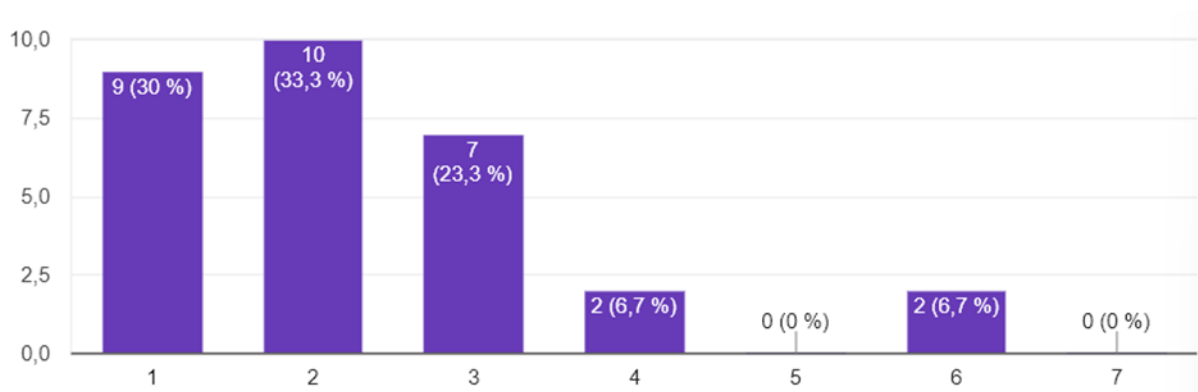


Figura 5.12. Resultados Escenario 3, Tarea 1. Fuente: Autoría propia

Escenario 4: Visualización de Datos Geoespaciales

- **Tarea 1: Interacción con Mapas Interactivos.** Utiliza GeoNetwork para abrir un mapa interactivo que te permita realizar acciones como acercar, alejar, hacer clic en objetos geográficos para obtener información adicional y cambiar la simbología de los datos. Evalúa la facilidad de interacción con el mapa interactivo y la utilidad de estas características. (1. Poco Complicado – 7. Muy Complicado)

Objetivo: Evaluar la facilidad de interacción con mapas interactivos en GeoNetwork, incluyendo acciones como acercar, alejar, hacer clic en objetos geográficos para obtener información adicional, y cambiar la simbología de los datos.

Resultados: La Figura 5.13 refleja los resultados obtenidos y la media total fue de 2,37, indicando una percepción moderada en cuanto a la facilidad de interacción con mapas interactivos y la utilidad de sus características.

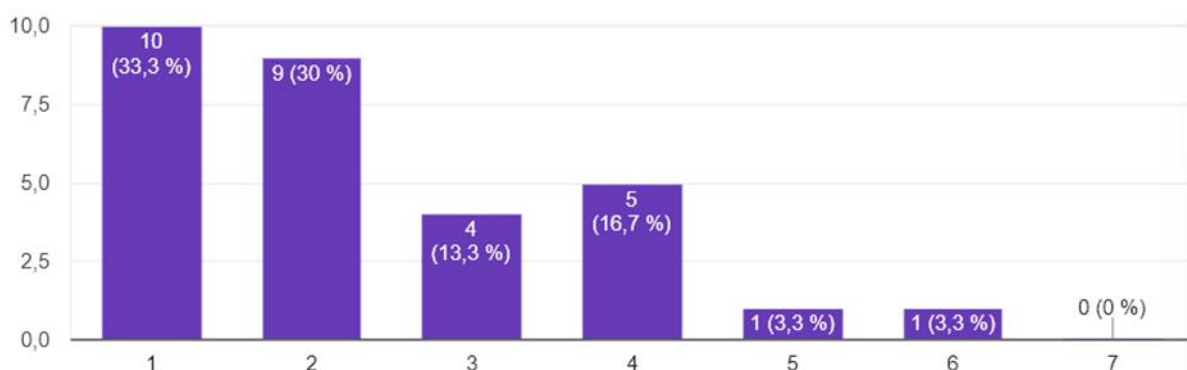


Figura 5.13. Resultados Escenario 4, Tarea 2. Fuente: Autoría propia

- **Tarea 2: Exploración de Capas de Datos.** Abre un mapa en GeoNetwork que contenga múltiples capas de datos geoespaciales (por ejemplo, capas de relieve, hidrología y límites administrativos). Evalúa la facilidad de exploración de estas capas, la capacidad de activar/desactivar capas y si pudiste encontrar la información que necesitabas. (1. Poco Complicado – 7. Muy Complicado)

Objetivo: Evaluar la facilidad de exploración de capas de datos geoespaciales en GeoNetwork, incluyendo la activación/desactivación de capas y la capacidad de encontrar información necesaria.

Resultados: En la Figura 5.14 se muestra los resultados obtenidos en el cual la media total fue de 3,00, indicando una percepción moderada en cuanto a la facilidad de explorar capas de datos y la capacidad de encontrar información necesaria.

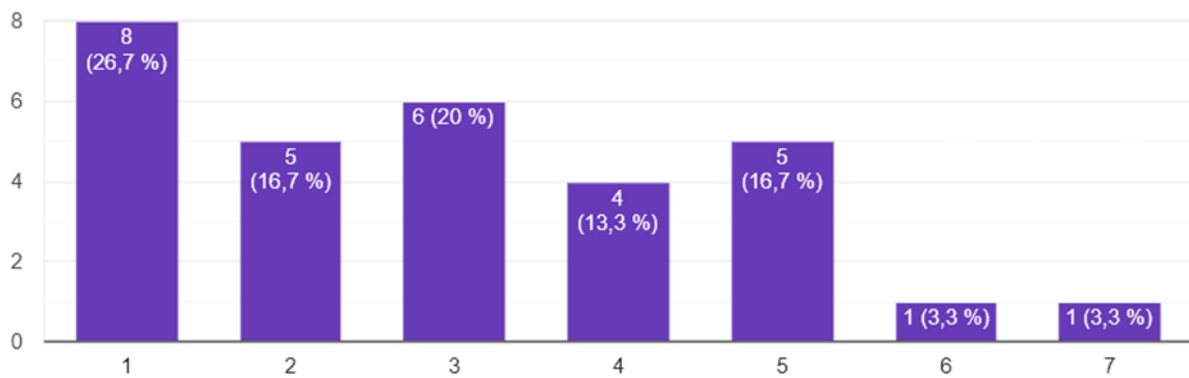


Figura 5.14. Resultados Escenario 4, Tarea 2. Fuente: Autoría propia

Evaluación general del Sistema

En la evaluación general del sistema se especifican solo 14 preguntas de las 16 propuestas que tiene un cuestionario PSSUQ debido a que 2 items hacen referencia al manejo de errores de la aplicación, los mismos que no se pueden modificar en geonetwork. También, cabe mencionar que para cada pregunta se utilizó una sola métrica. (1. Totalmente en Acuerdo – 7. Totalmente en Desacuerdo):

1. En general, estoy satisfecho con lo fácil que es utilizar este sistema. Ver Figura 5.15

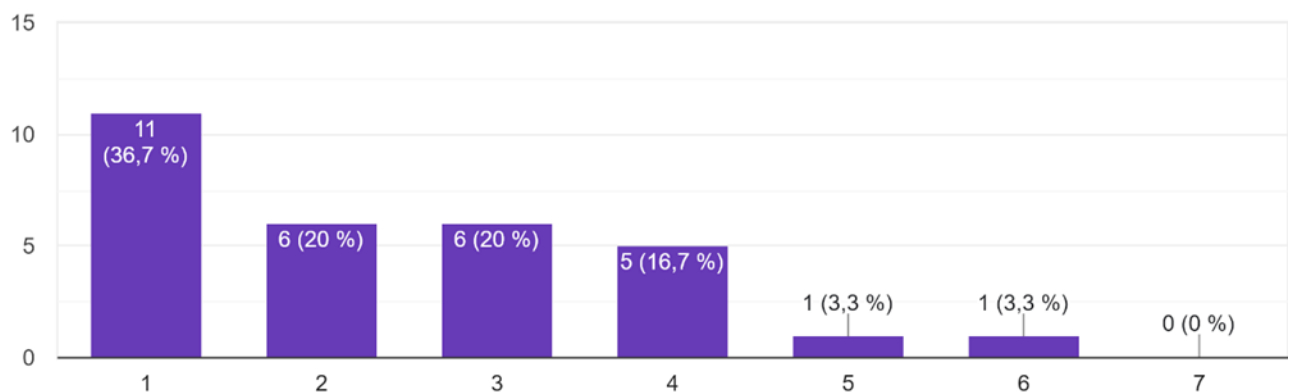


Figura 5.15. Resultados Usabilidad, Pregunta 1. Fuente: Autoría propia

2. Fue sencillo utilizar este sistema. Ver Figura 5.16

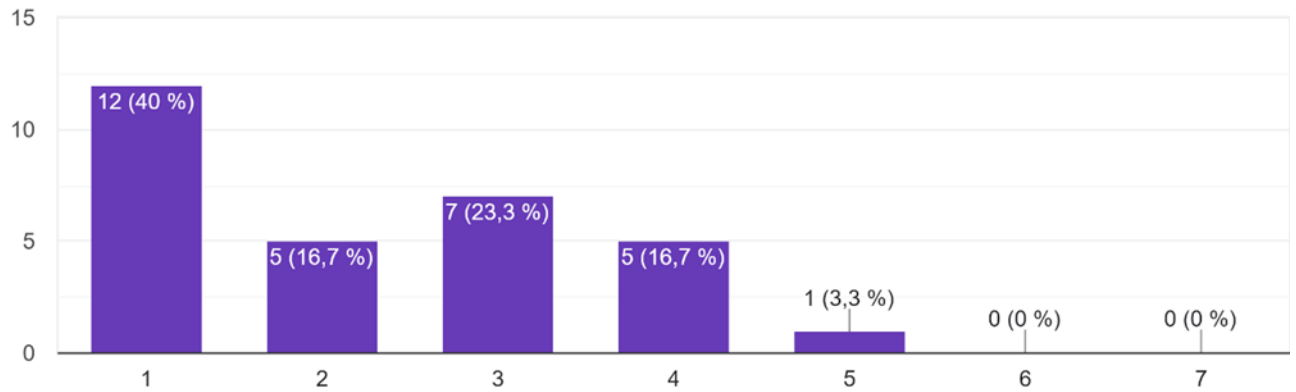


Figura 5.16. Resultados Usabilidad, Pregunta 2. Fuente: Autoría propia

3. Pude completar las tareas propuestas rápidamente. Ver Figura 5.17

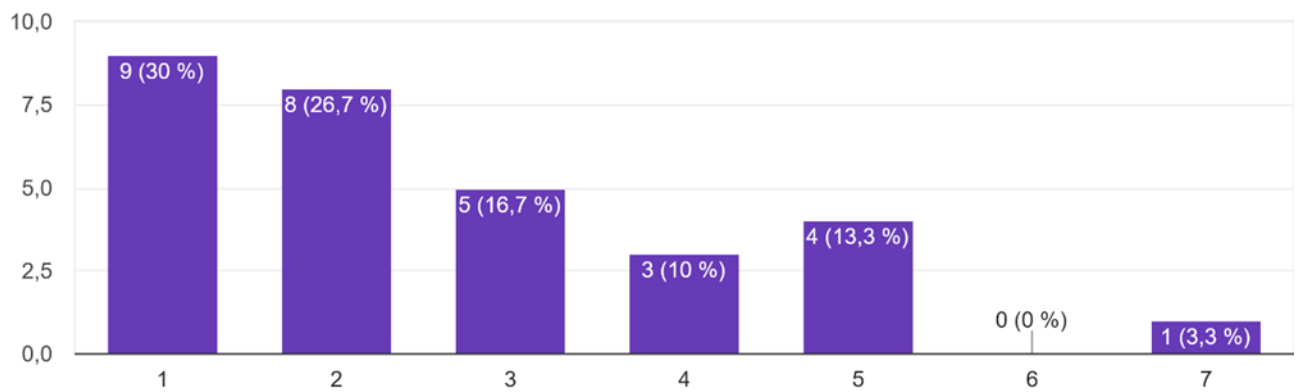


Figura 5.17. Resultados Usabilidad, Pregunta 3. Fuente: Autoría propia

4. Me sentí cómodo usando este sistema. Ver Figura 5.18

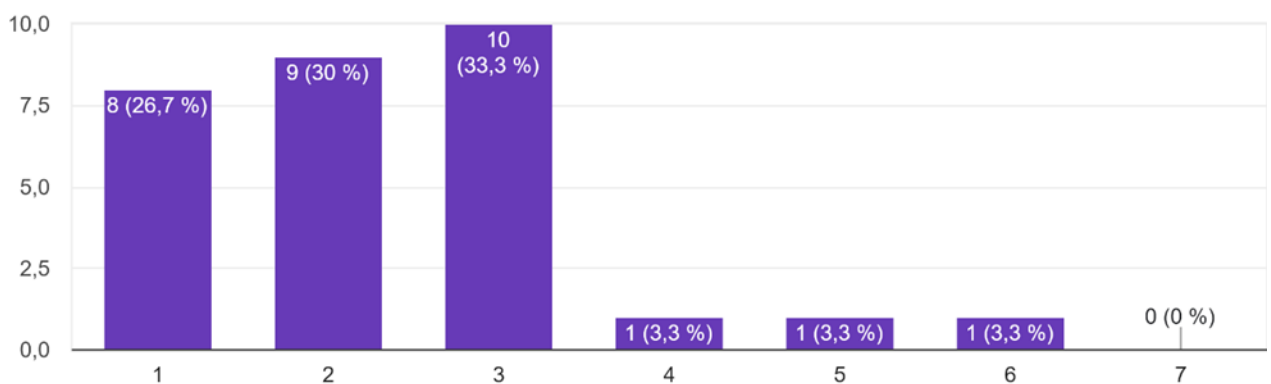


Figura 5.18. Resultados Usabilidad, Pregunta 4. Fuente: Autoría propia

5. Fue fácil aprender a utilizar este sistema. Ver Figura 5.19

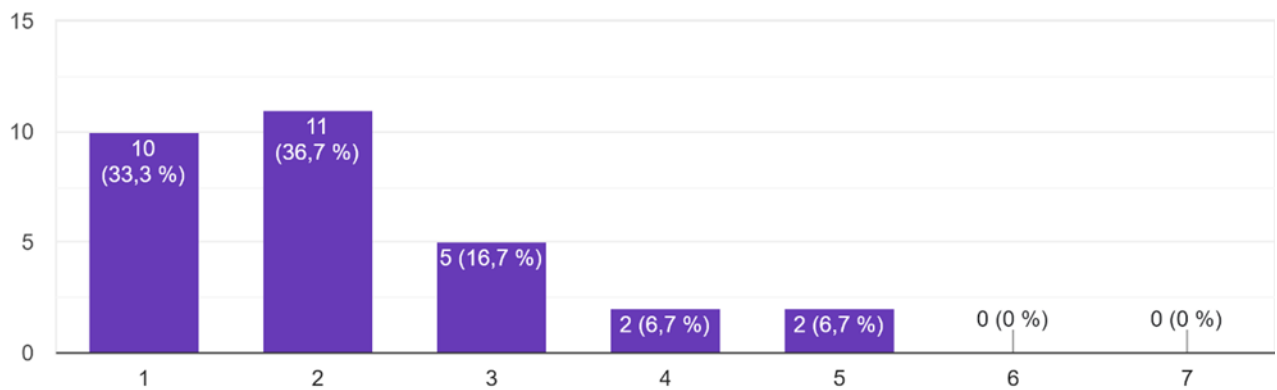


Figura 5.19. Resultados Usabilidad, Pregunta 5. Fuente: Autoría propia

6. Creo que podría buscar Geoinformación de manera más efectiva con Geonetwork. Ver Figura 5.20

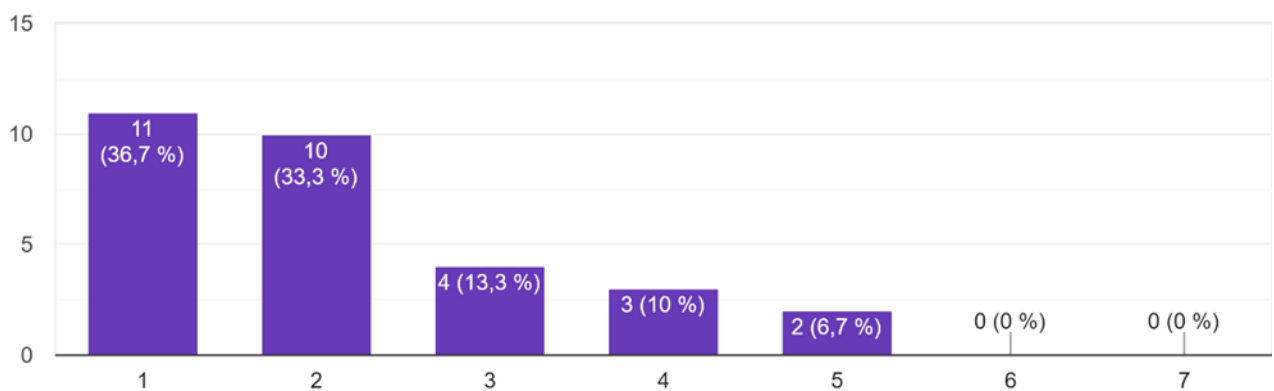


Figura 5.20. Resultados Usabilidad, Pregunta 6. Fuente: Autoría propia

7. Cada vez que cometía un error al usar el sistema, podía recuperarlo fácil y rápidamente. Ver Figura 5.21

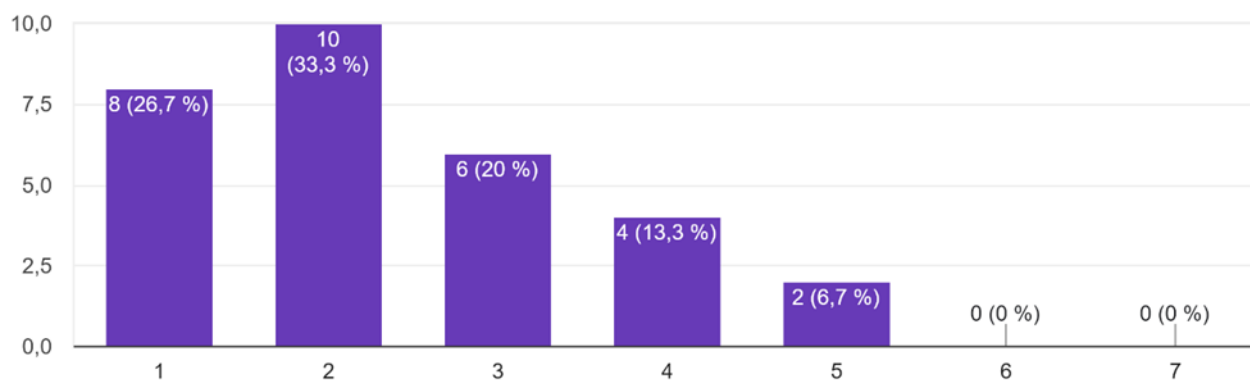


Figura 5.21. Resultados Usabilidad, Pregunta 7. Fuente: Autoría propia

8. Fue fácil encontrar la información que necesitaba. Ver Figura 5.22

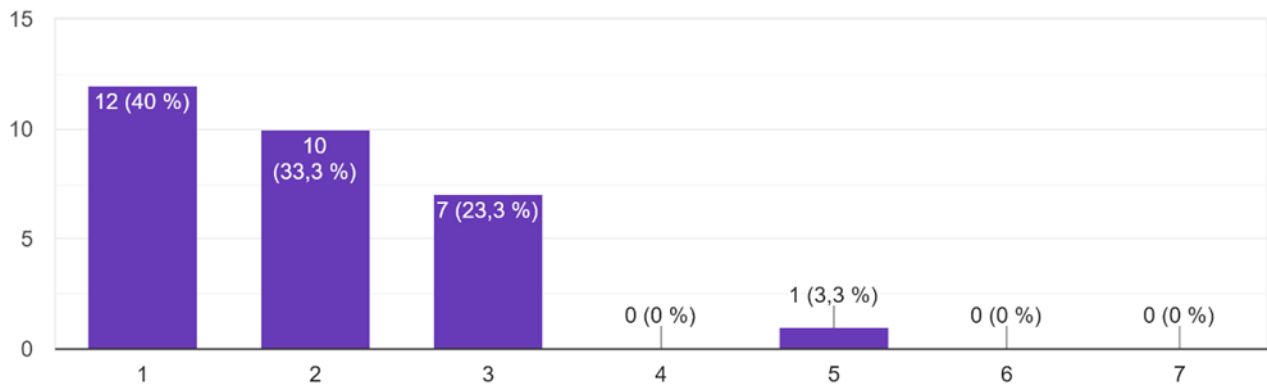


Figura 5.22. Resultados Usabilidad, Pregunta 8. Fuente: Autoría propia

9. La información fue efectiva para ayudarme a completar las tareas. Ver Figura 5.23

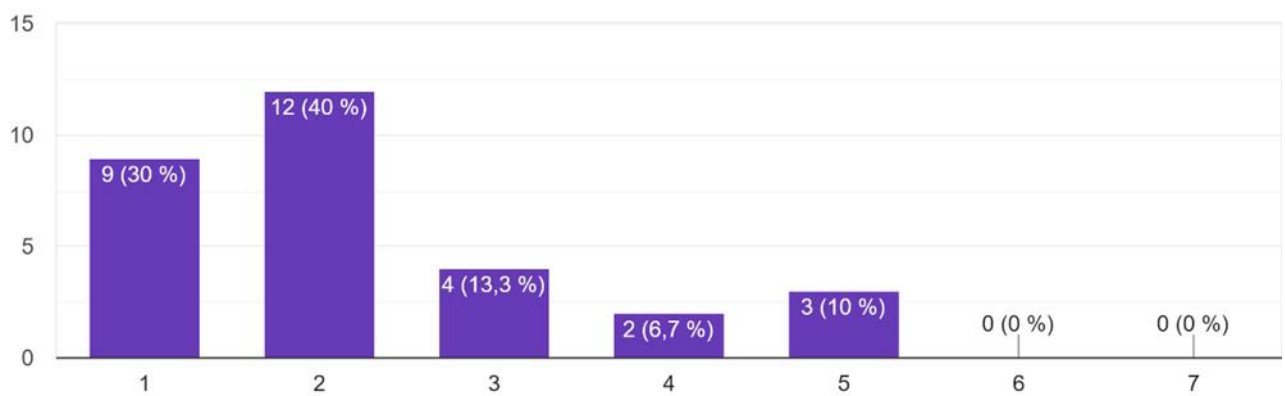


Figura 5.23. Resultados Usabilidad, Pregunta 9. Fuente: Autoría propia

10. La organización de la información en las pantallas del sistema fue claro. Ver Figura 5.24

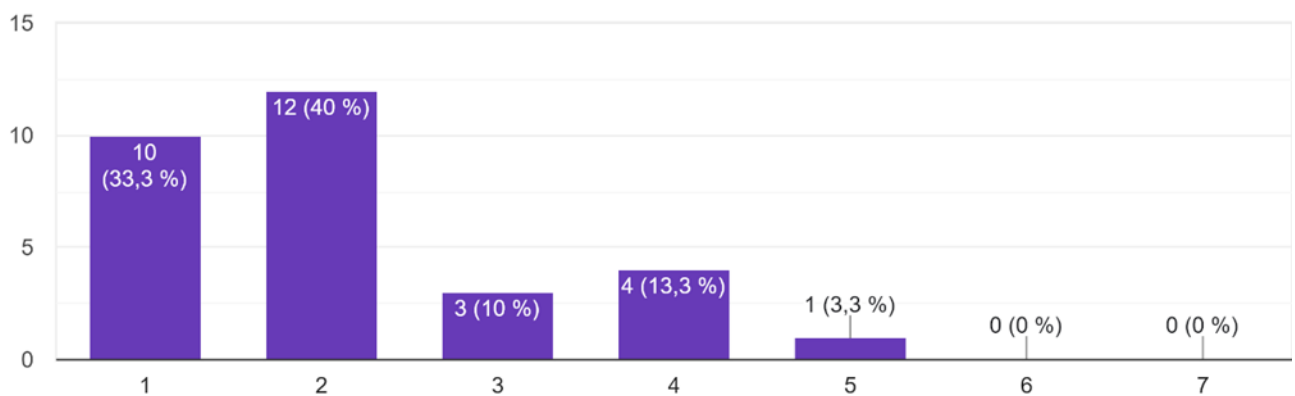


Figura 5.24. Resultados Usabilidad, Pregunta 10. Fuente: Autoría propia

11. La interfaz de Geonetwork fue agradable. Ver Figura 5.25

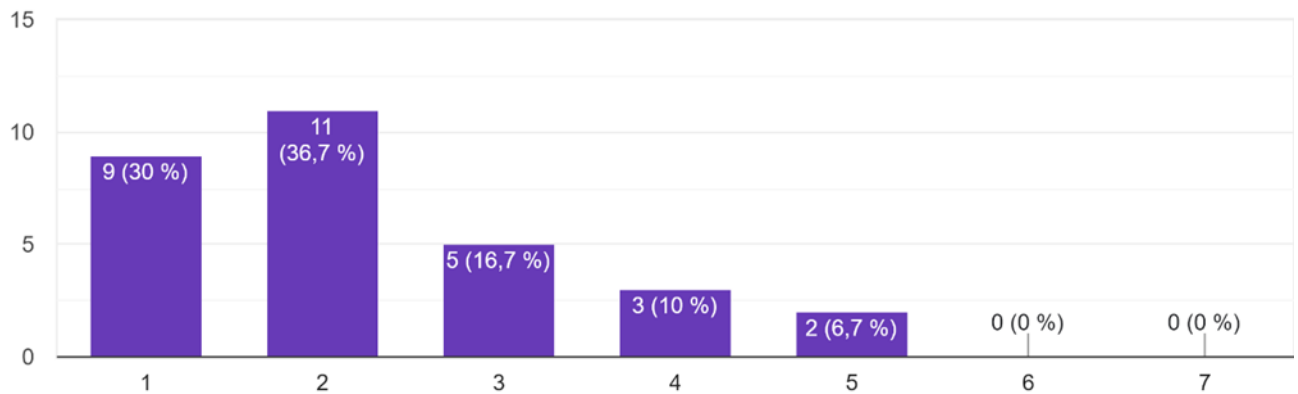


Figura 5.25. Resultados Usabilidad, Pregunta 11. Fuente: Autoría propia

12. Me gustó usar la interfaz que posee Geonetwork. Ver Figura 5.26

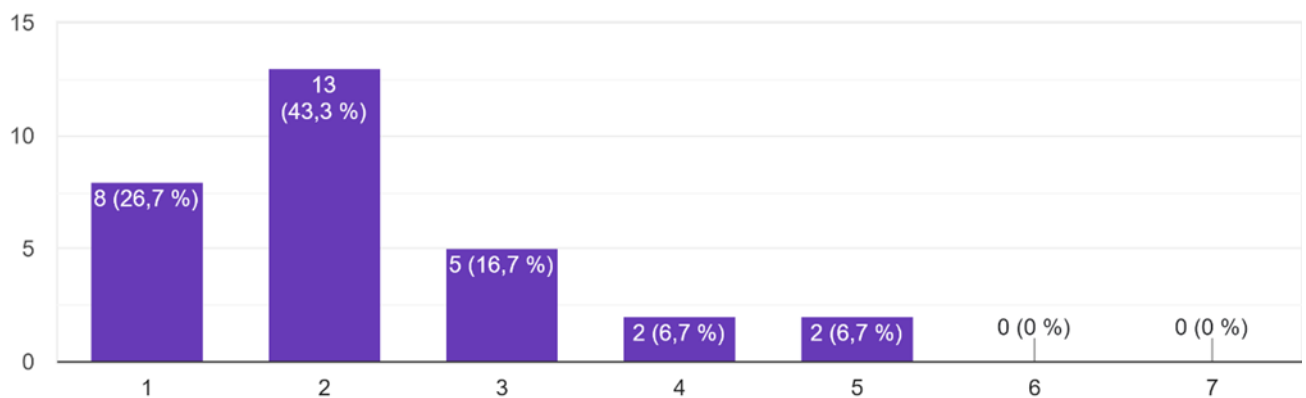


Figura 5.26. Resultados Usabilidad, Pregunta 12. Fuente: Autoría propia

13. Este sistema tiene todas las funciones y capacidades que podría esperar al buscar geoinformación. Ver Figura 5.27

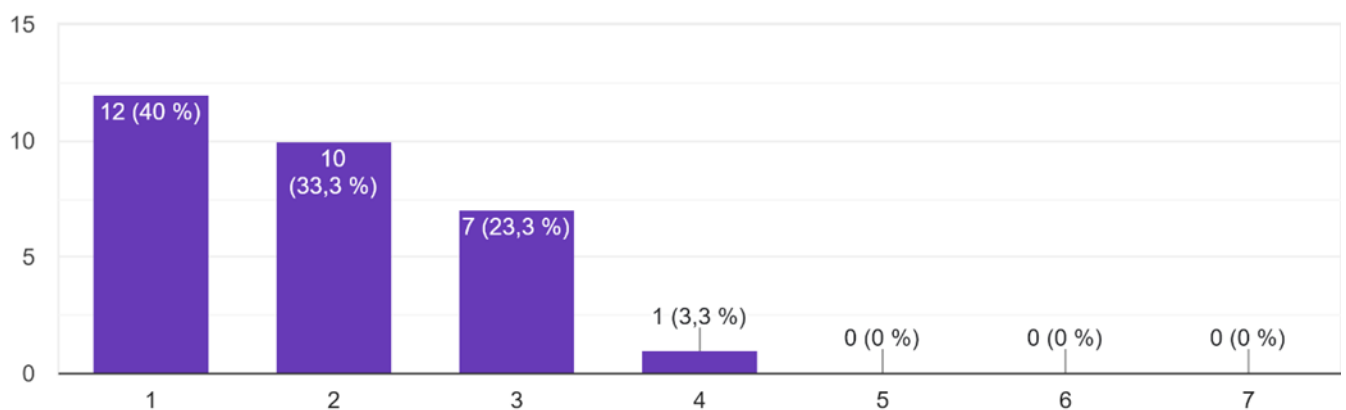


Figura 5.27. Resultados Usabilidad, Pregunta 13. Fuente: Autoría propia

14. En general, estoy satisfecho con este Geonetwork. Ver Figura 5.28

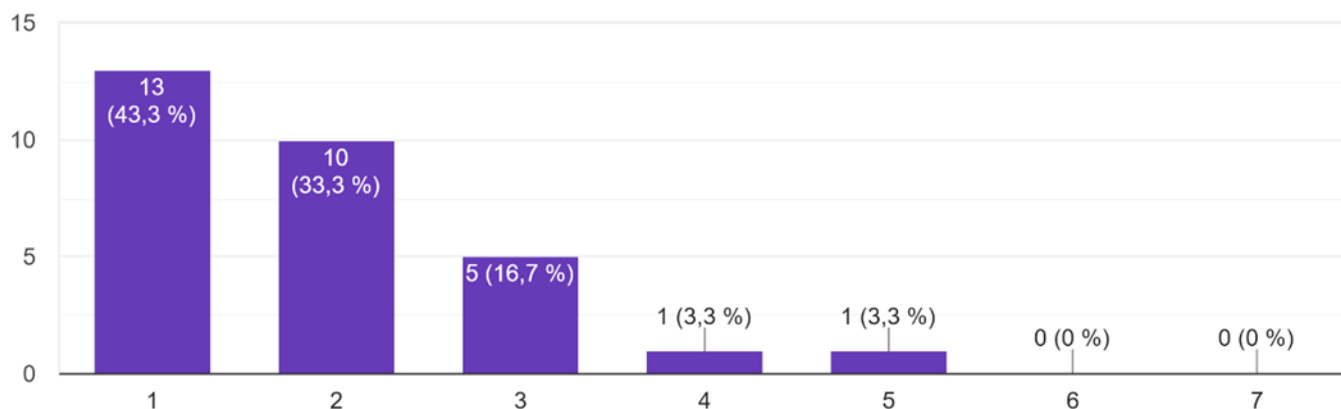


Figura 5.28. Resultados Usabilidad, Pregunta 14. Fuente: Autoría propia

Para este apartado, también se lo puede resumir mediante la siguiente tabla:

Tabla 5.5. Conteo final de respuestas en el análisis de usabilidad general. Fuente: Autoría propia

	1	2	3	4	5	6	7	m	s
1. En general, estoy satisfecho con lo fácil que es utilizar este sistema.	11	6	6	5	1	1	0	2,40	1,40
2. Fue sencillo utilizar este sistema.	12	5	7	5	1	0	0	2,27	1,26
3. Pude completar las tareas propuestas rápidamente.	9	8	5	3	4	0	1	2,63	1,61
4. Me sentí cómodo usando este sistema.	8	9	10	1	1	1	0	2,37	1,22
5. Fue fácil aprender a utilizar este sistema.	10	11	5	2	2	0	0	2,17	1,18
6. Creo que podría buscar Geoinformación de manera más efectiva con Geonetwork.	11	10	4	3	2	0	0	2,17	1,23
7. Cada vez que cometía un error al usar el sistema, podía recuperarlo fácil y rápidamente.	8	10	6	4	2	0	0	2,40	1,22
8. Fue fácil encontrar la información que necesitaba.	12	10	7	0	1	0	0	1,93	0,98
9. La información fue efectiva para ayudarme a completar las tareas	9	12	4	2	3	0	0	2,27	1,26

10. La organización de la información en las pantallas del sistema fue claro.	10	12	3	4	1	0	0	2,13	1,14
11. La interfaz de Geonetwork fue agradable.	9	11	5	3	2	0	0	2,27	1,20
12. Me gustó usar la interfaz que posee Geonetwork.	8	13	5	2	2	0	0	2,23	1,14
13. Este sistema tiene todas las funciones y capacidades que podría esperar al buscar geoinformación.	12	10	7	1	0	0	0	1,90	0,88
14. En general, estoy satisfecho con este Geonetwork	13	10	5	1	1	0	0	1,90	1,03

Media Total = 2,22

Utilidad del sistema (1 - 6): **Media = 2,33**

Calidad de la Información (7 - 10): **Media = 2,18**

Calidad de la interfaz (11 - 14): **Media = 2,07**

En resumen, luego del análisis de usabilidad con los usuarios se ha llegado a los siguientes resultados: En el Escenario 1, podemos afirmar que el proceso de búsqueda se caracteriza por su simplicidad, dado que el promedio de las tareas 1, 2 y 3, que cubren todos los aspectos de la búsqueda es de 2.49. Además, en la tarea 4, que evalúa la utilidad de los metadatos extraídos la media total de 2.7, esto sugiere que los metadatos obtenidos son valiosos para los usuarios.

En el Escenario 2, se puede notar que los usuarios tienen acceso a 3 o más tipos diferentes de archivos disponibles para exportar a otros sistemas.

En el Escenario 3, evaluado con una media total de 2.33, se evidencia que, después de completar los escenarios 1 y 2, los usuarios son capaces de buscar información de su interés de manera considerablemente más sencilla.

En el Escenario 4, se obtuvo una media total de 2.68, se puede concluir que la interacción con los mapas en GeoNetwork es relativamente sencilla. Esto se respalda con los comentarios de los participantes, quienes expresaron su interés en el uso de esta herramienta.

En el análisis general realizado, se calcularon tanto la media como la desviación estándar de cada pregunta de la

Para este apartado, también se lo puede resumir mediante la siguiente tabla:

Tabla 5.5. El resultado de estos cálculos nos permite afirmar que las medias obtenidas son representativas. Esto se debe a que la desviación estándar de cada pregunta es cercana a 1, lo que indica una baja dispersión en las respuestas.

En particular, la media total resultó ser 2.22, lo que refleja una valoración positiva para el sistema. Esta puntuación sugiere que el sistema goza de una gran aceptación a nivel general. Además, al desglosar las puntuaciones en las tres secciones del cuestionario PSSUQ, observamos que las calificaciones son igualmente positivas. La sección "Utilidad del sistema" obtuvo una media de 2.33, la sección "Calidad de la Información" alcanzó 2.18, y la sección "Calidad de la Interfaz" registró una media de 2.17.

Estos resultados indican que el sistema es considerado útil por los usuarios y que goza de una favorable acogida por parte de las personas que trabajan con metadatos geográficos o datos geoespaciales. Por último, los usuarios, además de indicar estar cómodos haciendo uso del GeoNetwork, pudieron presenciar que algunos de los servicios recolectados no funcionan correctamente. Este hallazgo subraya la importancia de que cada entidad que se una al cosechador debe mantener sus servicios funcionales y completos. Es esencial que los servicios de metadatos geoespaciales estén en óptimas condiciones para garantizar un funcionamiento eficiente y una experiencia positiva para los usuarios.

6. Conclusiones

En esta sección, procederemos a presentar las conclusiones a las que hemos llegado a lo largo de esta investigación. A lo largo de este estudio, hemos profundizado en la problemática de la búsqueda eficiente de información geoespacial dentro de la IDE de la Universidad de Cuenca. Nuestro enfoque principal ha sido evaluar distintas herramientas de gestión de metadatos como posibles alternativas para implementar un cosechador de metadatos. Asimismo, hemos puesto en funcionamiento un cosechador de metadatos diseñado para recopilar IG de diversas instituciones gubernamentales. Esta implementación tiene como objetivo optimizar las búsquedas en el portal geográfico de la IDE UCuenca.

Uno de los logros sobresalientes de este estudio reside en el funcionamiento exitoso del cosechador, el cual ha demostrado un aumento significativo en la eficiencia de las búsquedas, además de una mayor precisión en la recuperación de información relevante. Para lograr este resultado, en primer lugar, se llevó a cabo una exhaustiva revisión de la literatura para establecer el estado del arte en las aplicaciones de un cosechador de metadatos con el propósito de mejorar las IDE. Esta revisión reveló una disponibilidad limitada de información específica sobre la técnica de cosechamiento de metadatos, mientras que existía una mayor

documentación de resultados relacionados con la mejora de IDE mediante la implementación de herramientas destinadas a abordar una variedad más amplia de servicios. Entre estas mejoras se incluyen la integración de mapas interactivos, la implementación de servicios de catalogación de IG y uso de modelo 3D en mapas.

Después de un exhaustivo análisis y evaluación de diversas herramientas de gestión de metadatos, se procedió a seleccionar aquellas que mejor se ajustaban a las necesidades del proyecto. Este proceso permitió identificar el potencial de cada herramienta para cumplir con los objetivos establecidos en esta investigación. A través de la implementación de estas herramientas, los usuarios obtienen la capacidad de visualizar, consultar, analizar y compartir datos geográficos de manera efectiva.

No obstante, se reconoce que garantizar un rendimiento óptimo de estas herramientas representa un desafío fundamental. Es crucial asegurar que las herramientas sean capaces de manejar una variedad de tipos de datos geográficos, incluyendo tanto los formatos ráster como vectoriales que son comúnmente generados por instituciones de diversa índole. Además, se presta especial atención a la integración de estándares como CSW para la búsqueda y recuperación de metadatos, WMS para la visualización de datos geográficos, y WFS para la consulta y descarga de información geoespacial.

En este contexto, una parte fundamental del trabajo se enfoca en la selección de la herramienta más adecuada para asegurar el funcionamiento óptimo de un sistema de recolección de datos geoespaciales y la gestión de su IG. El objetivo es proporcionar a los usuarios una experiencia sólida y eficiente al interactuar con información geográfica y los recursos provistos por las instituciones involucradas. Además, es crucial evaluar el tiempo invertido en adquirir conocimientos acerca de las IDE y cómo estas ofrecen información geoespacial a un público general y la comunidad científica. Durante la investigación de este tema, la exploración de herramientas y la revisión de infraestructuras ya existentes han ocupado aproximadamente la mitad del tiempo dedicado al desarrollo de este proyecto.

Otro de los puntos clave del desarrollo es la configuración de las diferentes herramientas dentro de un servidor local para que su funcionamiento sea coordinado y evitar probables conflictos ya que las herramientas por lo general tienen configuraciones por defecto que deben ser revisadas y personalizadas.

El cosechamiento de metadatos a través de los catálogos de las instituciones seleccionadas resultó ser una técnica eficaz para obtener una amplia gama de información sobre los recursos disponibles. Estos resultados ofrecen una base sólida para futuras investigaciones y mejoras en la gestión y difusión de la información en el ámbito académico y cultural del país.

Además, sugieren la necesidad de promover el uso de estándares y buenas prácticas en la gestión de metadatos para mejorar la calidad y la accesibilidad de la información en los catálogos en línea. Considerando el significado de cosechamiento de metadatos utilizado en esta tesis, es necesario diferenciar de algunos trabajos donde se confunde el término con solamente la generación de metadatos. Es importante señalar que se sigue trabajando para articular a otros organismos y entidades con el fin de integrar sus datos geográficos a la plataforma. Estos hallazgos constituyen una base sólida para futuras investigaciones y mejoras en la infraestructura de metadatos de las instituciones mencionada, así como para el desarrollo de políticas y estrategias que promuevan la adecuada gestión y difusión de la información científica en el ámbito nacional.

Durante esta investigación, se ha constatado que la creación y gestión de metadatos en GeoNetwork, en el contexto de la administración de la IG de la Red Sísmica del Austro por parte de la Universidad de Cuenca, desempeñan un papel crítico en la mejora de la gestión y accesibilidad de los recursos geoespaciales. Este proceso, que involucra pasos sencillos desde el acceso a la plataforma hasta el almacenamiento de información esencial, proporciona una forma adecuada de búsqueda de geoinformación en los servidores como es el caso de la IDE Ucuena. Además, facilita la estandarización de los metadatos de información geoespacial, simplifica su descubrimiento y acceso, y garantiza la coherencia y calidad de los datos mediante la implementación de estándares como ISO 19139 e ISO 19115.

En conclusión, el geoportal IDE UCuenca cuenta con un gestor de metadatos, el cual ofrece una característica altamente destacable: un cosechador de metadatos a nivel nacional. Este logro representa un hito significativo en el ámbito de los geoportales, ya que, hasta la fecha, ningún otro geoportal ha conseguido realizar una cosecha de metadatos a esta escala en el país. Esta capacidad de cosechamiento de metadatos a nivel nacional resulta fundamental para la gestión eficaz de la IG de manera centralizada.

6.1. Conclusión del objetivo específico 1

El análisis efectuado sobre los sistemas de gestión de metadatos ha revelado que existen varias alternativas viables según las necesidades específicas del proyecto. Algunas herramientas se centran en la administración de los metadatos, mientras que otras se enfocan en la estructura y en la forma en que se gestionan los datos, ya sea a través de archivos descargables o bases de datos geoespaciales.

6.2. Conclusión del objetivo específico 2

El diseño arquitectónico para la implementación del cosechador resultó invaluable para visualizar con claridad todas las actividades y entidades fundamentales en el funcionamiento de un cosechador de metadatos. Esta arquitectura no solo facilita la identificación de las fuentes de catálogos que pueden ser accedidas para la recolección de la Información IG, sino que también establece las bases para un proceso eficiente de recolección.

La arquitectura propuesta en este trabajo se fundamenta en el enfoque presentado por Chaglla y Morocho en su artículo "Cosechador de metadatos en GeoNetwork para mejorar el acceso a información en una IDE Subnacional". Este enfoque ha servido como un sólido punto de partida para identificar las conexiones interiores que pueden establecerse entre las herramientas de GeoNetwork. Dicha colaboración optimiza la compartición de información generada en diversas instituciones, con el objetivo de beneficiar tanto al público en general como a grupos de investigación.

Mediante esta arquitectura, se abre la posibilidad de integrar una variedad de servicios adicionales que enriquecerían el portal de la IDE de la Universidad de Cuenca. Entre estas mejoras se incluyen el potencial aprovechamiento de un Asistente Virtual y de visores actuales, lo que contribuiría a una experiencia más completa y útil para los usuarios.

6.3. Conclusión del objetivo específico 3

El resultado de la implementación del cosechador mediante la herramienta GeoNetwork ha sido exitoso en su despliegue, a pesar de enfrentar ciertos desafíos durante la fase de configuración. También se puede recalcar que el uso de contenedores, resultó ser crucial para permitir la colaboración fluida entre todos los programas, facilitando la comunicación a través de los puertos asignados a cada componente, como Elasticsearch, Kibana y PostgreSQL.

Como se menciona en la sección de 5.2. La versión final electa de GeoNetwork para asegurar su funcionamiento sin problemas con las demás herramientas y servicios de la IDE fue la 4.2.2. Por otro lado, las buenas prácticas implementadas en el proyecto quedan descritas en la Sección 3, específicamente en el análisis y síntesis de los resultados de los documentos, enfocadas en la calidad de creación y gestión de metadatos, mientras que las desarrolladas luego de culminar el mismo quedan descritas en el Anexo C, lo que permitirán que se dé continuidad al trabajo con metadatos en el Grupo de Investigación en IDE.

6.4. Conclusión del objetivo específico 4

La evaluación considerada se enfocó en medir los tiempos de respuesta a medida que la cantidad de metadatos recolectados aumentaba. Esto se debió a que, a diferencia de otras herramientas, que se centran principalmente en la visualización de información, GeoNetwork prioriza el rendimiento al llevar a cabo búsquedas con una mayor cantidad de metadatos almacenados.

Es relevante destacar que existen documentos que respaldan en mayor medida la utilidad de la herramienta GeoNetwork para la gestión de metadatos. Esto se debe a la facilidad en su administración y a su interfaz de usuario altamente interactiva, la cual resulta amigable tanto para usuarios generales como para aquellos con experiencia en datos geoespaciales.

6.5. Recomendaciones

La gestión eficiente de recursos en un servidor destinado a un recolector de información resulta fundamental para garantizar el adecuado funcionamiento de esta herramienta. Esto cobra especial relevancia debido a la necesidad de administrar volúmenes significativos de información. Para lograr un rendimiento óptimo de la herramienta, es imperativo contar con recursos de software robustos. Durante el desarrollo de este trabajo, se han enfrentado desafíos derivados de los recursos empleados, particularmente en la obtención de diversos tipos de documentos, ya sean imágenes u otros archivos portadores de información adicional junto con sus correspondientes metadatos.

6.6. Trabajos futuros

Es fundamental tener en cuenta los metadatos obtenidos de las instituciones seleccionadas para este estudio. La información recopilada se almacena en una base de datos en el servidor del sistema de recolección y se mantiene actualizada en caso de que las instituciones realicen cambios. Sin embargo, en situaciones donde una institución elimine o retire información de su catálogo, ya sea debido a actualizaciones en su portal o a la decisión de no mostrar ciertos recursos, es necesario que el sistema de recolección identifique estos cambios. Esto garantizará que la información se mantenga consistente y refleje con precisión las acciones llevadas a cabo por la institución en relación a su contenido.

Referencias

- About EPA's Geospatial Program* | US EPA. (n.d.). Retrieved August 6, 2023, from <https://www.epa.gov/geospatial/about-epas-geospatial-program>
- Agrawal, S., Tripathi, A. K., & Gupta, R. D. (2022). Development and implementation of automatic metadata generation framework for SDI using OSS: a case study of Indian NSDI. *Arabian Journal of Geosciences* 2022 15:5, 15(5), 1–17. <https://doi.org/10.1007/S12517-022-09635-W>
- Aguilera, U., Augusta, M., Quinatoa, I. E. E., & Latacunga -Ecuador, A. (2020). *Estudio de metadatos y protocolos de interoperabilidad para la plataforma científica Ecucienca de la Universidad Técnica de Cotopaxi*. <http://repositorio.utc.edu.ec/handle/27000/8607>
- Aníbal, F., Castillo, J., Paola, M., Ordoñez, Z., Luis, E., Cueva, L., Magaly, N., & Mora, L. (2018). Metodología para la catalogación de metadatos espaciales basada en normas ISO. *Alternativas*, ISSN 1390-1915, Vol. 19, N°. 1, 2018, Págs. 63-70, 19(1), 63–70. <https://doi.org/10.23878/alternativas.v19i1.205>
- ArcGIS Resource Center. (n.d.). Retrieved August 6, 2023, from <https://resources.arcgis.com/es/help/getting-started/articles/026n00000014000000.htm>
- Astudillo Vargas, Y. Y. (2018). *Implementación de una plataforma de Infraestructura de Datos Espaciales*. <http://dspace.utpl.edu.ec/jspui/handle/20.500.11962/23092>
- At These Coordinates*. (2020). <https://atcoordinates.info/2020/07/07/geospatial-metadata-with-dublin-core-and-geoblacklight-standards/>
- Atapuma, R., Dasiel, R., Salazar Martínez, A., & Fernando, R. J. (2020). *Diagnóstico para definir una propuesta de perfil ecuatoriano de catastro mediante la aplicación de la norma ISO 19152 (LADM). Caso: Barrios Miranda y La Católica*.
- Carrascosa Bastera, J. (2018). *IDE y geoportal de Praga*. <https://riunet.upv.es/handle/10251/108381>
- CatMDEdit*. (n.d.). Retrieved August 6, 2023, from <https://catmdedit.sourceforge.io/>
- Criollo Carlosama, & Beatriz Alexandra. (2017). *Sistema web de gestión de actividades económicas georeferenciado para el Municipio del cantón Montúfar ciudad de San Gabriel*. <http://repositorio.utn.edu.ec/handle/123456789/7661>
- DCAT-AP 2.0.1*. (n.d.). Retrieved August 7, 2023, from <https://datos.gob.es/es/blog/dcat-ap-201-como-ha-evolucionado-la-especificacion-de-referencia-europea-para-la-descripcion-de>
- Del, P., De, S., & De, R. (2018). *Programación del servicio de recolección de desechos sólidos domiciliarios mediante herramientas S.I.G, en la zona urbana del cantón Riobamba, provincia de Chimborazo*. <http://dspace.espe.edu.ec/handle/123456789/10139>
- Eafit, U. (2017). *ECOCAMPUS EAFIT SISTEMA DE INFORMACIÓN GEOGRÁFICA NUBIA INES USUGA USUGA*.

- Eclipse Jetty*. (n.d.). Retrieved August 7, 2023, from <https://projects.eclipse.org/projects/rt.jetty>
- Espinoza Martínez, M. P. (2017). *Análisis, diseño e implementación de una infraestructura de datos espaciales eólicos*. <https://hdl.handle.net/20.500.14352/19829>
- ESRI. (n.d.). *ArcGIS Resource Center*. Retrieved July 19, 2023, from <https://resources.arcgis.com/es/help/getting-started/articles/026n00000014000000.htm>
- Federal Geographic Data Committee*. (2023). <https://www.fgdc.gov/metadata/iso-standards>
- Fernández Eguiarte, A., Castelán Hernández, C., & Bautista Durán, Á. V. (2019). *REPOSITORIO DE RIESGOS ASOCIADOS A FENÓMENOS NATURALES Y ANTRÓPICOS EN MÉXICO*. 50–57.
- Freddo, B. V., Cristina, D., & Massera, B. (2021). *Geotecnologías y aprendizajes. La mapoteca virtual como recurso educativo*. <https://rea.unpa.edu.ar/handle/123456789/2316>
- Geodetic, G. (2018). *The 14 Global Fundamental Geospatial Data Themes*. http://ggim.un.org/meetings/2018-Addis_Ababa/
- GeoNetwork Documentation OGC Services*. (n.d.). Retrieved August 3, 2023, from <https://geonetwork-opensource.org/manuals/3.12.x/en/user-guide/harvesting/harvesting-ogcwx.html>
- GeoNode documentation*. (n.d.). Retrieved August 6, 2023, from <https://docs.geonode.org/en/master/about/index.html>
- GeoNode documentation GC Web Services*. (n.d.). Retrieved August 3, 2023, from <https://geonode-docs.readthedocs.io/en/latest/tutorials/devel/api/ogc.html>
- GitHub - Official Docker image GeoNetwork*. (n.d.). Retrieved August 7, 2023, from <https://github.com/geonetwork/docker-geonetwork>
- Graciana Pombo, D. (2019). Conocimiento, acceso, derechos y democratización de los datos: Infraestructura de Datos Espaciales (IDE). *Huellas*, 23(1), 71–85. <https://doi.org/10.1080/13658816.2012.743664>
- IBM. (2023). *IBM*. <https://www.ibm.com/mx-es/topics/geospatial-data>
- Ideca. (2023). *Metadato Geográfico | Ideca*. <https://www.ideca.gov.co/recursos/glosario/metadato-geografico>
- IEDG. (2010). *Perfil Ecuatoriano de Metadatos*. https://iedg.sni.gob.ec/wp-content/uploads/2022/06/perfil_ecuatoriano_metadatos_pem.pdf
- Ingeniería Sistemas Proyecto De Grado, C. DE. (2020). *SISTEMA DE METADATOS PAR LA INFRAESTRUCTURA DE DATOS ESPACIALES*. <http://repositorio.upea.bo/handle/123456789/142>
- Kibana: Explora, visualiza y descubre datos*. (n.d.). Retrieved August 7, 2023, from <https://www.elastic.co/es/kibana/>

- Lopez-Vazquez, C., Gonzalez-Campos, M. E., Bernabe-Poveda, M. A., Moctezuma, D., Hochsztain, E., Barrera, M. A., Granell-Canut, C., Leon-Pazmino, M. F., Lopez-Ramirez, P., Morocho-Zurita, V., Moya-Honduvilla, J., Manrique-Sancho, M. T., Montiveros, M. E., Narvaez-Benalcazar, R., De Perez-Alcazar, J., Resnichenko, Y., & Seco, D. (2022). Building a Gold Standard Dataset to Identify Articles About Geographic Information Science. *IEEE Access*, 10, 19926–19936. <https://doi.org/10.1109/ACCESS.2022.3150869>
- Meneses, L. U., Iguasnia, F. O., Iguasnia, J. O., & Calderón, F. C. (2018). Sistema de información geográfica para la integración de información Geo referenciada de entidades públicas y privadas en la ciudad de Santo Domingo. *Revista Científica y Tecnológica UPSE*, 4(3), 13–27. <https://doi.org/10.26423/RCTU.V4I3.294>
- Metavist - editor for the Federal Geographic*. (n.d.). Retrieved August 6, 2023, from <https://metavist.software.informer.com/>
- Morocho, V. (2014). Mejorando la Usabilidad en las Plataformas de Infraestructura de Datos Espaciales. *Congreso de Ciencia y Tecnología ESPE*, 9(1), 106–110. <https://doi.org/10.24133/CCTESPE.V9I1.90>
- Morocho, V., Achig, R., Bustamante, J., & Mendieta, F. (2022). Virtual Assistants to bring geospatial information closer to a smart citizen. *6th IEEE Ecuador Technical Chapters Meeting, ETCM 2022*. <https://doi.org/10.1109/ETCM56276.2022.9935761>
- Narváez, D., & Andrés, C. (2020). *Generación y evaluación de la cartografía ambiental para la plantación florícola FIORI DI MIWI, ubicada en el sector de Convalecencia, parroquia Juan Montalvo, Cayambe, Ecuador, mediante la aplicación de la información publicada por la SENPLADES - CONAGE*. <http://www.dspace.uce.edu.ec/handle/25000/21663>
- Ortiz González, T. (2021). *Modelo de IDE con complementos 3D para el mejoramiento del IDESC de Santiago de Cali - Colombia*.
- Otero, J. Á., & De Lázaro Y Torres, M. L. (2019). Las infraestructuras de datos espaciales: un reto y una oportunidad en la docencia de la Geografía. *Boletín de La Asociación de Geógrafos Españoles*, 0(82). <https://doi.org/10.21138/bage.2787>
- PgMetadata*. (n.d.). Retrieved August 6, 2023, from <https://docs.3liz.org/qgis-pgmetadata-plugin/>
- PyCSW Documentation*. (n.d.). Retrieved August 6, 2023, from https://live.osgeo.org/archive/10.5/es/overview/pycsw_overview.html
- Ramos Chango, K. P. (2017). *DESARROLLO DE UN GEOPORTAL PARA EL INSTITUTO NACIONAL DE PATRIMONIO CULTURAL EN LA CIUDAD DE RIOBAMBA*.
- Revista Cartográfica*. (n.d.). Retrieved June 30, 2023, from <https://www.revistasipgh.org/index.php/rcar>
- Secretaría Nacional de Planificación. (n.d.). *Dirección Zonal de Planificación 6 – Secretaría Nacional de Planificación*. Retrieved June 28, 2023, from <https://www.planificacion.gob.ec/coordinacion-zonal-planificacion-6/>
- Secretaría Nacional de Planificación y Desarrollo. (2016). *Guía Técnica para la Generación de Metadatos de Geoinformación*.

- Specka, X., Gärtner, P., Hoffmann, C., Svoboda, N., Stecker, M., Einspanier, U., Senkler, K., Zoader, M. A. M., & Heinrich, U. (2019). The BonaRes metadata schema for geospatial soil-agricultural research data – Merging INSPIRE and DataCite metadata schemes. *Computers & Geosciences*, 132, 33–41. <https://doi.org/10.1016/J.CAGEO.2019.07.005>
- Tk Metadata Editor. (n.d.). Retrieved August 6, 2023, from <https://geology.usgs.gov/tools/metadata/tools/doc/tkme.html>
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., ... Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467–473. <https://doi.org/10.7326/M18-0850>
- Vedor De Paula, E., Lopes De Souza Da Paz, O., & Pereira Da Silva, J. (2019). *Elaboración de Bases de Datos Geográficos. Para la Planificación y Gestión de Áreas Protegidas Elaboration of Geographic Database For Planning and Management of Protected Areas.*
- Wang, S., Wang, J., Zhan, Q., Zhang, L., Yao, X., & Li, G. (2023). A unified representation method for interdisciplinary spatial earth data. *Big Earth Data*, 7(1), 136–155. https://doi.org/10.1080/20964471.2022.2091310/SUPPL_FILE/TBED_A_2091310_SM2881.DOCX

Anexos

Anexo A: Cuadro con las búsquedas en cada base de datos, los resultados totales luego de aplicar los criterios.

BASE DE DATOS	Cod	Palabras clave	Filtro Fecha	Filtro Area	Filtro Idioma	Filtro Type Docs	Todos Filtros	Condicion NumDoc
Scopus OK	1	geospatial AND metadata	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	157	< 120
	2	geospatial AND harvesting	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	39	ok
	3	metadata AND spatial AND harvesting	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	2	ok
	4	geospatial AND metadata AND harvesting	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	1	ok
	5	geospatial AND ide	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	1	ok
	6	geospatial AND metadata AND ide	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	1	ok
	8	tools AND managment AND metadata	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	1	ok
	TOTAL							45
IEEE Xplore OK	1	geospatial AND metadata	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	38	ok
	2	geospatial AND harvesting	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	12	ok
	3	metadata AND spatial AND harvesting	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	7	ok
	4	geospatial AND metadata AND harvesting	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	1	ok
	5	geospatial AND ide	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	3	ok
	6	geospatial AND metadata AND ide	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	1	ok
	8	tools AND managment AND metadata	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	0	ok
								62
ACM Digital Library	1	geospatial AND metadata	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	447	< 120
	2	geospatial AND harvesting	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	125	< 120
	3	metadata AND spatial AND harvesting	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	467	< 120
	4	geospatial AND metadata AND harvesting	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	71	ok
	5	geospatial AND ide	2018-2023	Computer Science, Engineering	Ingles	Conference paper, articulo	56	ok

Anexo B: Cuadro con los documentos seleccionados de cada base de datos.

Nro	Titulo	Autores	Idioma	Año	Tipo	Objetivo	Resultado	Herramientas
1	Sistema de información geográfica para la integración de información Geo referenciada de entidades públicas y privadas en la ciudad de Santo Domingo	Luis Ulloa Meneses, Fausto Orozco Iguasnia, Jaime Orozco Iguasnia, Frankiz Carrera Calderón	Español	2017	Artículo	Desarrollar un sistema de información geográfica para la integración de información geo referenciada de entidades públicas y privadas en la ciudad de Santo Domingo	Desarrollo de un visor y portal web usando archivos SHAPE para mostrar puntos georeferenciados, además un sitio web con información de las instituciones participantes	GrassGIS, gvSIG y QuantumGIS.
2	Diagnóstico para definir una propuesta de perfil ecuatoriano de catastro mediante la aplicación de la norma ISO 19152 (LADM). Caso: Barrios Miranda y La Católica	Romo Atapuma, Rocío Dasiel, Ambiente Saizar Martínez, Rodolfo Jaime Fernando	Español	2020	Tesis	Analizar la aplicación de la ISO 19152:2012 LADM en los barrios "Miranda" y "La Católica", de la ciudad de Quito, para la creación de una propuesta de perfil nacional que defina las entidades y los atributos que se deben incorporar a una base de datos de administración territorial, como la primera versión de un estándar de catastro nacional, utilizando herramientas geomáticas.	Tabla comparativa entre atributos DMQ y atributos Norma ISO 19152. Perfil Ecuatoriano de Catastro Sistema de administración catastral en STD M Geospatial. Análisis de Resultados	Uso de normas y estándares
3	Estudio de metadatos y protocolos de interoperabilidad para la plataforma científica Ecuencia de la Universidad Técnica de Cotopaxi.	Ulloa Aguilera, María Augusta	Español	2020	Tesis	Análisis de los tipos de metadatos que existen y realizar una comparación de cual es el más eficiente para gestionarlo, compartirlo en una biblioteca digital. Caso de uso de artículos científicos y usando los metadatos de los artículos.	La selección del metadato DC es óptimo para trabajar en la parte descriptiva de documentos utilizados para difundir conocimiento.	Uso de normas y estándares
4	Elaboración de Bases de Datos Geográficas Para la Planificación y Gestión de Áreas Protegidas	Eduardo Vedor De Paula, Otaclio Lopes De Souza Da Paz, Josemar Pereira Da Silva	Español	2019	Artículo	Uso de datos geográficos almacenados en los años 2014-2017, para el control de áreas de reserva y otros	La captura de datos se realiza con ciertas excepciones. En algunos casos, faltan informaciones relevantes como el linaje, fecha e incluso autoría. Una dificultad fue la inexistencia de metadatos en datos secundarios bajados de servidores de mapas y sitios web de instituciones públicas.	ArcGIS
5	Gestión de Riesgos ante Fenómenos Hidrometeorológicos Extremos	Agustín Fernández Equiarte, Carolina Castelañ Hernández, Anel V. Raulista Durán	Español	2019	Artículo	El Repositorio de riesgos asociados a fenómenos naturales y antrópicos en México constituye una nueva opción dentro de las facilidades del Atlas Climático Digital de México el cual muestra información de metadatos que se accede a una descripción de los	Creación de un repositorio usando una herramienta de código abierto Geonetwork, para gestionar la datos geográficos y conocimiento científico generado por especialistas en geociencias y ciencias sociales enfocados a	Geonetwork

Anexo C: Enlace al documento de buenas prácticas desarrollado para esta tesis.

<https://n9.cl/nkui1>