

UNIVERSIDAD PÚBLICA DE EL ALTO

CARRERA INGENIERÍA DE SISTEMAS



TESIS DE GRADO

**“HERRAMIENTA DE METADATOS INTELIGENTE PARA LA
RECUPERACIÓN DE LA INFORMACIÓN DIGITALIZADA”**

**Para Optar al Título de Licenciatura en Ingeniería de Sistemas
MENCIÓN: GESTION Y PRODUCCION**

| | |
|----------------------------|---|
| Postulante: | Walter Giovanni Aroma Quispe |
| Tutor Metodológico: | M. Sc. Lic. Ing. Fanny Helen Pérez Mamani |
| Tutor Especialista: | M. Sc. Lic. Adrian Eusebio Quisbert Vilela |
| Tutor Revisor: | Lic. Katya Maricela Pérez Martínez |

ALTO – BOLIVIA

2024

Dedicatoria

Queridos padres,

A mis queridos padres, Hirene Quispe Castro y Flavio Aroma Osco, cuyo amor incondicional, sacrificios y apoyo constante han sido la base de mis logros. Gracias, mamá, por tu inagotable paciencia y por enseñarme a nunca rendirme ante las dificultades. Gracias, papá, por ser un ejemplo de integridad y trabajo duro. Sin su guía y fortaleza, no habría alcanzado mis metas.

A mi hermano, Favio Aroma Quispe, por su motivación continua. Has sido una fuente de inspiración y un recordatorio constante de la importancia de la perseverancia y el optimismo. Tu apoyo incondicional me ha impulsado a seguir adelante en los momentos más desafiantes.

Y en especial, a la memoria de mi tutora, Lic. María Luisa Alanoca, cuya guía y sabiduría dejaron una huella imborrable en mi formación académica. Tu dedicación y pasión por la enseñanza transformaron mi perspectiva y enriquecieron mi conocimiento. Aunque ya no estés físicamente, tu legado perdurará siempre en mi corazón y en cada paso de mi camino profesional. Te agradezco profundamente por haber sido una luz en mi vida y por haber creído en mí cuando más lo necesitaba.

Walter Giovanni Aroma Quispe

Agradecimiento

Primeramente, agradezco a Dios por haberme dado la fortaleza, sabiduría y perseverancia necesarias para llevar a cabo esta tesis. Sin Su guía y bendiciones, este logro no habría sido posible.

A mi querida familia, por su amor incondicional, apoyo constante y por ser mi fuente de inspiración. Gracias por creer en mí y motivarme a alcanzar mis metas.

Expreso mi profunda gratitud a mis tutores, cuyas contribuciones fueron esenciales para la realización de este trabajo:

Tutor Metodológico: *M. Sc. Lic. Ing. Fanny Helen Pérez Mamani. Gracias por su orientación meticulosa y por enseñarme a abordar la investigación con rigor y precisión, Estoy verdaderamente agradecido por su ayuda y compromiso constante.*

Tutor Revisora: *Lic. Katya Maricela Pérez Martínez. Agradezco su valiosa retroalimentación y por ayudarme a mejorar la calidad y claridad de este trabajo, y estoy muy agradecido por el tiempo y el esfuerzo que ha invertido en mi desarrollo académico.*

Tutor Especialista: *M. Sc. Lic. Adrian Eusebio Quisbert Vilela. Gracias por compartir su vasta experiencia y conocimientos especializados, los cuales fueron cruciales para el desarrollo técnico de esta investigación y estoy agradecido por el esfuerzo que ha invertido en mi tesis de grado.*

A todos ustedes, mi más sincero agradecimiento por su apoyo, paciencia y dedicación. Su influencia y enseñanza han sido fundamentales en mi formación académica y profesional.

Walter Giovanni Aroma Quispe

INDICE

| | | |
|-------|--------------------------------------|----|
| 1 | MARCO TEORICO..... | 1 |
| 1.1 | INTRODUCCION..... | 1 |
| 1.2 | ANTECEDENTES..... | 2 |
| 1.2.1 | Antecedentes Internacionales..... | 2 |
| 1.2.2 | Antecedentes Nacionales..... | 3 |
| 1.2.3 | Antecedentes Locales..... | 3 |
| 1.3 | PLANTEAMIENTO DEL PROBLEMA..... | 4 |
| 1.3.1 | Problema Principal..... | 4 |
| 1.3.2 | Problemas Secundarios..... | 5 |
| 1.3.3 | Formulación Del Problema..... | 5 |
| 1.4 | OBJETIVOS..... | 5 |
| 1.4.1 | Objetivo General..... | 5 |
| 1.4.2 | Objetivos Específicos..... | 6 |
| 1.5 | HIPOTESIS..... | 6 |
| 1.5.1 | Identificación de variables..... | 6 |
| 1.5.2 | Operacionalización de Variables..... | 7 |
| 1.6 | JUSTIFICACION..... | 8 |
| 1.6.1 | Justificación Científica..... | 8 |
| 1.6.2 | Justificación Técnica..... | 8 |
| 1.6.3 | Justificación Social..... | 8 |
| 1.7 | Metodología..... | 9 |
| 1.7.1 | Método Científico..... | 9 |
| 1.7.2 | Metodología Buchanan..... | 10 |
| 1.8 | Herramientas..... | 12 |
| 1.8.1 | Servidor Web..... | 12 |
| 1.8.2 | Gestor de Base de Datos..... | 13 |
| 1.8.3 | Lenguajes de Programación..... | 13 |
| 1.9 | Software..... | 13 |
| 1.9.1 | PHP (Hipertext Preprocessor)..... | 13 |
| 1.9.2 | SQL (Structured Query Language)..... | 14 |
| 1.9.3 | FRAMEWORKS..... | 14 |
| 1.9.4 | LARAVEL..... | 14 |
| 1.9.5 | SYMFONY..... | 15 |

| | | |
|--------|---|----|
| 1.9.6 | JAVASCRIPT | 15 |
| 1.9.7 | OpenStreetMap | 15 |
| 1.10 | Límites y Alcances | 16 |
| 1.10.1 | Limites..... | 16 |
| 1.10.2 | Alcances | 16 |
| 2 | MARCO TEORICO..... | 18 |
| 2.1 | METADATOS | 18 |
| 2.2 | DATOS | 19 |
| 2.3 | CALIDAD..... | 20 |
| 2.4 | CALIDAD DE LOS METADATOS | 20 |
| 2.5 | Protocolo OAI-PMH | 21 |
| 2.6 | METADATOS INTELIGENTES..... | 22 |
| 2.7 | TIPOS DE METADATOS | 23 |
| 2.7.1 | Metadatos Descriptivos | 23 |
| 2.7.2 | Metadatos Administrativos | 23 |
| 2.7.3 | Metadatos Estructurales..... | 24 |
| 2.8 | Recuperación De Metadatos | 24 |
| 2.9 | Métodos de recuperación de información | 25 |
| 2.10 | ExifTool..... | 25 |
| 2.11 | Metadatos Dublin Core | 25 |
| 2.12 | La Importancia De Recuperación De Información | 28 |
| 2.12.1 | Informática forense..... | 28 |
| 2.13 | Imágenes digitales | 29 |
| 2.14 | Tipos de Formatos de Imágenes..... | 29 |
| 2.14.1 | Formato JPEG (Joint Photographic Experts Group) | 30 |
| 2.14.2 | Formato GIF (Graphics Interchange Format) | 32 |
| 2.14.3 | Formato PNG (Portable Network Graphics) | 32 |
| 2.14.4 | Formato TIFF (Tagged Image File Format)..... | 32 |
| 2.14.5 | Formato BMP (Windows bitmap) | 32 |
| 2.15 | Comprensión del problema..... | 33 |
| 2.16 | Ingeniería de sistemas..... | 33 |
| 2.17 | INVESTIGACION..... | 34 |
| 2.17.1 | Tipos de investigación | 34 |
| 2.18 | Metodologías..... | 38 |
| 2.19 | Método científico | 38 |

| | | |
|---------|---|----|
| 2.19.1 | Concebir La Idea A Investigar | 38 |
| 2.19.2 | Plantear El Problema De Investigación | 39 |
| 2.19.3 | Elaborar El Marco Teórico | 40 |
| 2.19.4 | Definir Tipo Y Nivel De Investigación | 42 |
| 2.19.5 | Establecer Las Hipótesis Y Definir Las Variables | 43 |
| 2.19.6 | Seleccionar El Diseño Apropiado De Investigación..... | 46 |
| 2.19.7 | Seleccionar Los Sujetos A Estudiar | 48 |
| 2.19.8 | Recolectar Los Datos | 50 |
| 2.19.9 | Analizar Los Resultados..... | 51 |
| 2.19.10 | Presentar El Reporte De Resultados | 52 |
| 2.20 | Metodología de desarrollo buchanan | 53 |
| 3 | DISEÑO METODOLOGICO..... | 61 |
| 3.1 | Metodología de la Investigación | 61 |
| 3.1.1 | Tipo de Investigación..... | 61 |
| 3.2 | Método de investigación | 62 |
| 3.2.1 | Enfoque De Investigación | 62 |
| 3.2.2 | Variables De Investigación | 63 |
| 3.3 | Ambiente de investigación. Universo Población y Muestra | 63 |
| 3.3.1 | Universo Poblacional..... | 63 |
| 3.3.2 | Muestra | 64 |
| 3.4 | Descripción de la metodología a usar | 67 |
| 3.4.1 | Paradigma De Investigación | 67 |
| 3.5 | Herramientas..... | 68 |
| 3.5.1 | Método de ingeniería – Modelo Buchanan | 68 |
| 3.5.2 | Técnicas de Investigación e Instrumentos..... | 69 |
| 3.6 | METRICAS DE CALIDAD..... | 79 |
| 3.6.1 | ISO/IEC 25000..... | 79 |
| 4 | PRUEBAS Y RESULTADOS..... | 83 |
| 4.1 | DEmostracion de la herramienta metadatos..... | 83 |
| 4.2 | Maquetación..... | 86 |
| 4.3 | CODIGO PANTALLA DE INICIO | 88 |
| 4.4 | Implementación Y Despliegue | 89 |
| 4.4.1 | Requerimientos De Hardware | 89 |
| 4.4.2 | Requerimientos De Software | 89 |
| 4.5 | Seguridad..... | 90 |

| | | |
|-------|--|-----|
| 4.6 | Pruebas de software | 90 |
| 4.7 | Prueba de hipótesis | 94 |
| 4.7.1 | Formulación De La Prueba De Hipótesis..... | 94 |
| 4.7.2 | Estado De La Hipótesis | 95 |
| 5. | CONCLUSIONES Y RECOMENDACIONES | 101 |
| 5.1 | Conclusiones..... | 101 |
| 5.2 | Recomendaciones | 102 |
| | Bibliografía | 103 |
| | ANEXOS | 111 |

INDICE DE TABLAS

| | | |
|-----------------|---|----|
| Tabla 1 | Operacionalización de Variable Independiente | 7 |
| Tabla 2 | Operacionalización de Variable Dependiente..... | 7 |
| Tabla 3 | Métricas e indicadores para evaluar la calidad de los metadatos. | 21 |
| Tabla 4 | Metadatos Dublin Core: Definición y descripción de elementos..... | 26 |
| Tabla 5 | Explorando el potencial de Dublin Core: | 27 |
| Tabla 6 | Comparación de Metodologías..... | 57 |
| Tabla 7 | Datos Demográficos (Nivel Educativo)..... | 70 |
| Tabla 8 | Datos Demográficos (Edad)..... | 71 |
| Tabla 9 | Datos Demográficos (Genero)..... | 71 |
| Tabla 10 | Conocimiento sobre Metadatos P1..... | 71 |
| Tabla 11 | Conocimiento sobre Metadatos P2..... | 72 |
| Tabla 12 | Conocimiento sobre Metadatos P3..... | 72 |
| Tabla 13 | Evaluación del Herramienta de Metadatos P1 | 73 |
| Tabla 14 | Evaluación del Herramienta de Metadatos P2 | 73 |
| Tabla 15 | Evaluación del Herramienta de Metadatos P3 | 74 |
| Tabla 16 | Prueba de ingreso a la herramienta | 91 |
| Tabla 17 | Prueba de Registro de Metadatos | 92 |
| Tabla 18 | Herramienta Metadatos | 93 |
| Tabla 19 | Evaluación del Diagnostico..... | 96 |
| Tabla 20 | Varianza Estimada | 97 |

INDICE DE FIGURAS

| | | |
|------------------|--|----|
| Figura 1 | Ciclo de vida de la Metodología Buchanan | 11 |
| Figura 2 | Metodología de recursos estructurados bajo el protocolo OAI-PMH..... | 22 |
| Figura 3 | Secuencia de pasos que sigue el método JPEG | 31 |
| Figura 4 | Fases de la Metodología Buchanan | 54 |
| Figura 5 | Fases de la Metodología | 56 |
| Figura 6 | Fórmula para la muestra | 66 |
| Figura 7 | Método Buchanan aplicado al sistema de información metadatos..... | 69 |
| Figura 8 | ISO/IEC 25000 | 80 |
| Figura 9 | Validación del tipo de archivo | 83 |
| Figura 10 | Lectura de datos EXIF | 83 |
| Figura 11 | Verificación de la presencia de datos EXIF | 83 |
| Figura 12 | Definición de claves excluidas y traducción de claves..... | 83 |
| Figura 13 | Función para imprimir metadatos en tablas | 84 |
| Figura 14 | Funciones para obtener coordenadas GPS | 84 |
| Figura 15 | Validación y procesamiento de coordenadas GPS: | 85 |
| Figura 16 | Redirección si el archivo no es una imagen JPEG | 85 |
| Figura 17 | MAQUETACION 1 | 86 |
| Figura 18 | MAQUETACION 2..... | 86 |
| Figura 19 | MAQUETACION 3..... | 87 |
| Figura 20 | MAQUETACION 4..... | 87 |
| Figura 21 | MAQUETACION 5..... | 88 |
| Figura 22 | Tabla t-Student | 99 |
| Figura 23 | Gráfico de aceptación de hipótesis..... | 99 |

Resumen

En la era digital, la gestión y el acceso eficiente a los recursos digitales son esenciales. Los metadatos juegan un papel crucial en este proceso, permitiendo la localización precisa de archivos digitales y facilitando su organización y recuperación. Esta tesis aborda el desarrollo de una herramienta de metadatos inteligente para la recuperación de información digitalizada, enfocada en imágenes en formatos JPG, GIF y PNG. El objetivo principal es crear una herramienta que extraiga y gestione metadatos de manera eficiente, evitando la pérdida de información relevante y permitiendo la geolocalización precisa de archivos digitales.

La investigación incluyó el diseño y desarrollo de la herramienta, la implementación de algoritmos avanzados para la extracción de metadatos y la integración de técnicas de geolocalización. Los resultados demostraron que la herramienta cumple con los objetivos planteados, proporcionando una extracción eficiente de metadatos y una geolocalización precisa de las imágenes. En conclusión, la herramienta de metadatos inteligente desarrollada mejora significativamente la eficiencia en la organización y recuperación de información digitalizada, representando un avance significativo en la gestión de datos digitales.

Palabras Claves: Metadatos, geolocalización, Herramienta,

Abstract

In the digital era, efficient management and access to digital resources are essential. Metadata plays a crucial role in this process, enabling the precise location of digital files and facilitating their organization and retrieval. This thesis addresses the development of an intelligent metadata tool for the recovery of digitized information, focused on images in JPG, GIF, and PNG formats. The main objective is to create a tool that efficiently extracts and manages metadata, preventing the loss of relevant information and allowing the precise geolocation of digital files.

The research included the design and development of the tool, the implementation of advanced algorithms for metadata extraction, and the integration of geolocation techniques. The results demonstrated that the tool meets the established objectives, providing efficient metadata extraction and precise geolocation of images. In conclusion, the developed intelligent metadata tool significantly improves the efficiency in the organization and retrieval of digitized information, representing a significant advancement in digital data management.

Keywords: Metadata, geolocation, tool

GLOSARIO DE ABREVIATURAS

PHP. - Preprocesador de Hipertexto

HTML. - Lenguaje de Marcado de Hipertexto

JPGE. - Grupo de Expertos en Articulación Fotográfica

RDBMS. - Sistema de gestión de bases de datos relacionales

Exif. - Formato de archivo de imagen intercambiable

XMP. - Plataforma de metadatos extensible

OAI-PMH. - Protocolo de la Iniciativa de Archivos Abiertos para la recolección de metadatos

ISO/RDI. - Organización Internacional de Normalización / Interfaz de datos de referencia

GPS. - Sistema de Posicionamiento Global



**INGENIERÍA
DE SISTEMAS**
UNIVERSIDAD PÚBLICA DE EL ALTO

CAPÍTULO I

MARCO PRELIMINAR

1 MARCO TEORICO

1.1 INTRODUCCION

Los metadatos juegan un papel fundamental en la identificación y acceso a recursos educativos, facilitando su reutilización y optimizando la implementación de tecnologías educativas a gran escala y según (Duval & Robson, 2001), los metadatos son esenciales para identificar y acceder a contenidos educativos, permitir su reutilización y mejorar el despliegue generalizado de tecnologías educativas.

Esto ha provocado la inclusión de los metadatos como herramienta, si beneficia al no perder información relevante, esto respaldado por un estudio realizado por (Duffy & Meadow, 2016) este encontró que la utilización de metadatos descriptivos aumentó la tasa de recuperación de información relevante en un 30%. También es importante la calidad y disponibilidad de los metadatos ya que estos son fundamentales para la eficacia de los metadatos. Los resultados de la búsqueda pueden ser inexactos si los metadatos son incorrectos o incompletos, según (Ghani & Smeaton, 2005).

En cuanto a la búsqueda, una investigación realizada por (Larson & Smith, 2014), mostró que el uso de metadatos de búsqueda avanzados redujo el tiempo promedio de búsqueda en un 45%, lo cual es un indicativo significativo del impacto positivo que pueden tener los metadatos en la eficiencia y efectividad de los procesos de recuperación de información.

En una variedad de contextos, los metadatos son herramientas cruciales para la gestión y el acceso efectivo a los recursos digitales. Su utilidad se evidencia principalmente en la localización precisa de archivos digitales, lo que permite una organización eficiente y una recuperación rápida de datos relevantes.

1.2 ANTECEDENTES

1.2.1 Antecedentes Internacionales

J. Rodríguez (2017), "Servicio de Recolección de Metadatos genérico para documentos" Proyecto realizado en la Universidad Nacional de La Plata (UNLP). Tiene como objeto detallar la implementación de una herramienta para permitir y optimizar el intercambio de información estructurada sobre recursos académicos y científicos proveniente de contextos que no necesariamente cumplen estándares de intercambio o esquemas de catalogación normalizados, para esto se desarrolló etapas de desarrollo del trabajo donde se exponen lenguaje y frameworks Ruby es un lenguaje multiplataforma, interpretado y orientado a objetos y su implementación oficial es distribuida bajo licencia de software libre y Nokogiri Es un analizador de HTML, XML, SAX y Reader. Entre las muchas características de Nokogiri se encuentra la capacidad de buscar documentos a través de selectores XPath o CSS3.

J. Barrios (2018), "El valor agregado del uso de metadatos en videojuegos educativos" Proyecto realizado en la Universidad De La Republica Uruguay. Una Propuesta para solucionar estos inconvenientes son las anotaciones de los videojuegos con Metadatos. en lo cual utilizaron, el estándar Dublin Core es un esquema de metadatos elaborado por Dublin Core Metadata Initiative. Basado en el estándar ISO 15836-1:201712. Esta organización se basa en el diseño y buenas prácticas de metadatos para intercambiar información.

M. Arellano, F. Felipe (2017), "El papel de los metadatos en la Web Semántica" Proyecto realizado en la Universidad Nacional Autónoma de México. Para evaluar las necesidades de los usuarios y al buscar información en la Web es encontrar con precisión aquella información que necesitan, es decir, que, al efectuar una Web Semántica, en donde los sistemas de información puedan entender el significado del contenido, particularmente el temático, de los recursos u objetos de información, y el esquema de metadatos Dublín Core

se convirtió en la norma ISO 15836/2003 donde comprende pasos para el rescate de metadatos.

1.2.2 Antecedentes Nacionales

A. Castro (2009), “Metadatos Geográficos Aplicación e Importancia” Proyecto desarrollado en la Universidad Mayor de San Simón. Debido a la gran cantidad de información geográfica se hace necesario tener una estructura bien establecida que permita la documentación de los datos, como fueron creados, de donde se obtuvieron, fecha de creación de los mismos, etc. Por lo que los metadatos geográficos son de una gran importancia ya que describen las características de los datos y proveen un mecanismo de información de los datos esto es de gran relevancia sobre todo a organizaciones que producen grandes cantidades de información geográfica.

J. Duran (2007), “Biblioteca y museo virtual Antonio Paredes Candia del G.M.E.A.” Proyecto de Grado realizado en la Universidad Mayor de San Andrés, Este trabajo presenta un problema real en el que se quiere ofrecer el desarrollo del Sistema Virtual como medio de difusión y acceso, al material bibliográfico, colección de las obras de arte y publicaciones Unidad de Bibliotecas y Museo Antonio Paredes Candía, haciendo el uso de las herramientas para la gestión de los documentos electrónicos (Digitalización de Textos e imágenes), Para ello, se propone uso de la metodología de Diseño UWE (Ingeniería Web basada en UML), que es extensión a UML en el desarrollo Orientado a Objetos que proporciona guías para la construcción de modelos de forma sistemática y con pasos acertados, enfocadas en personalización y en estudio de casos de uso.

1.2.3 Antecedentes Locales

D. Choquehuanca (2020) plantea Proyecto de Grado Denominada “Sistema de metadatos para la infraestructura de datos espaciales”, Tiene como objetivo evaluar la calidad

del software se utilizó la ISO 9126 que permite conocer el nivel de la calidad del software a través de un proceso de evaluación de acuerdo con las métricas o indicadores que presenta el modelo de calidad, en seguridad de la información se recurrió a la norma ISO 27001 y finalmente para la estimación del costo se usó COCOMO II basado en el peso o líneas de código. tesis realizada en la Universidad Pública de El Alto (UPEA).

R. Mendoza (2010) plantea Tesis de Grado denominado “Agentes móviles para la recuperación de información en bibliotecas digitales” En este trabajo de investigación se desarrolló un modelo de agentes móviles para buscar y recuperar información semántica, basada en metadatos Dublin Core para la descripción de recursos digitales, de acuerdo a la consulta de información realizada por el usuario en una biblioteca digital ubicada en un ambiente distribuido, el prototipo se construyó bajo la metodología propuesta para la plataforma JADE (Armazón para el Desarrollo de Agentes con Java), desarrollada en Italia y proporcionada de manera libre.

1.3 PLANTEAMIENTO DEL PROBLEMA

El creciente volumen de información digital que se enfoca en el análisis del contenido visual de imágenes en formatos como JPG, GIF, PNG. se convierte en un importante reto para los usuarios al intentar localizar la información de los datos de un archivo digital de imagen relevante. Las herramientas de análisis de metadatos suelen ser limitados en su capacidad para proporcionar resultados precisos y contextualmente relevantes.

1.3.1 Problema Principal

La ausencia de metadatos en imágenes digitales, especialmente en formatos como JPG, GIF y PNG, limita considerablemente la capacidad de realizar búsquedas de metadatos y recuperaciones precisas y rápidas. Estos metadatos son información adicional sobre distintos tipos de datos o sobre un archivo en particular, son esenciales para contextualizar y

organizar las imágenes. Sin ellos, la búsqueda de información específica se convierte en un proceso laborioso y poco eficiente, con la consecuente pérdida de metadatos que tienen las imágenes.

1.3.2 Problemas Secundarios

- Dificultad en la obtención de imágenes representativas y suficientes en los formatos JPG, GIF y PNG, especialmente si se requieren para un análisis exhaustivo de metadatos.
- Complejidad en la interpretación de los resultados se debe a las variables y a la posibilidad de que los datos sean incorrectos.
- Dificultad al asegurar de que los conjuntos de imágenes utilizados en las pruebas sean representativos y completos, lo que podría influir en la generalización de los resultados.
- Definir métricas de desempeño de la herramienta adecuadas y significativas para evaluar el éxito de la herramienta en diversos contextos de uso.

1.3.3 Formulación Del Problema

¿Cómo una herramienta de metadatos inteligentes permitirá la recuperación de información digitalizada?

1.4 OBJETIVOS

1.4.1 Objetivo General

Demostrar la importancia de los metadatos para coadyuvar de forma eficiente en la cuantificación del efecto de pérdida de metadatos y los posibles resultados y así permita extraer y gestionar los metadatos de manera eficiente, optimizando la precisión y rapidez en las búsquedas y evitando la pérdida de información relevante.

1.4.2 Objetivos Específicos

- Reunir los datos necesarios para el adecuado procesamiento de metadatos en imágenes digitales incluyendo la recopilación de imágenes para la identificación de los tipos de metadatos que contienen
- Describir las variables a través de un análisis estadístico para interpretarlas adecuadamente.
- Validar la eficacia de la herramienta de metadatos mediante pruebas con diferentes conjuntos de imágenes digitales, evaluando su rendimiento en términos de precisión, y eficiencia en la extracción y gestión de metadatos.
- Establecer métricas de desempeño de la herramienta meta para evaluar el éxito del funcionamiento de la herramienta en diferentes entornos de trabajo.

1.5 HIPOTESIS

El uso de la herramienta de metadatos para la recuperación de información digitalizada, mediante la extracción de metadatos de imágenes digitales, permitirá una localización de información relevante contenida en las mismas.

H1 La herramienta de metadatos inteligente para la recuperación de información digitalizada mejorará significativamente la facilidad de acceso a la información de imágenes digitales.

H0 La herramienta de metadatos inteligente para la recuperación de información digitalizada no mejorará significativamente la facilidad de acceso a la información de imágenes digitales.

1.5.1 Identificación de variables

Variable Independiente

Herramienta de Metadatos Inteligente

Variable Dependiente

Proceso de Recuperación de Información Digitalizada

1.5.2 Operacionalización de Variables

Variable Independiente

Tabla 1
Operacionalización de Variable Independiente

| VARIABLE | DEFINICION | DIMENSION | INSTRUMENTO | INDICADOR |
|--------------------------------------|--|---|---------------|----------------------|
| Herramienta de Metadatos Inteligente | El proceso de diseño, desarrollo de una herramienta de metadatos para la recuperación de información | Código fuente e implementación de los metadatos <ul style="list-style-type: none"> • CONCEPTOS • REGLAS • COMPRESION | Lenguajes php | Funcional Preciso |

Variable Dependiente

Tabla 2
Operacionalización de Variable Dependiente

| VARIABLE | DEFINICION | DIMENSION | INSTRUMENTO | INDICADOR |
|---|---|---|---|--|
| Recolección de metadatos de imágenes digitales y precisión de los resultados de información recuperados | Resultados relevantes obtenidos en obtención a la consulta de búsqueda realizada. | Nivel de información de metadatos y nivel de imagen | Formato de imagen vectorial y en formato JPEG | Exactitud de georreferenciación y calidad de imagen y resolución |

1.6 JUSTIFICACION

La recuperación y búsqueda efectiva mediante metadatos de información digitalizada es un desafío importante en la era de la tecnología de la información. Con el crecimiento exponencial de la cantidad de datos disponibles, es fundamental contar con herramientas eficientes que permitan acceder y encontrar rápidamente la información relevante.

1.6.1 Justificación Científica

El uso de un herramienta de metadatos inteligente para la recuperación y búsqueda de información digitalizada puede contribuir a ampliar y enriquecer el conocimiento científico en el campo de la gestión de la información al abordar la efectividad y precisión, esta herramienta busca la conservación de los datos, promueve la interoperabilidad y fomenta la reutilización de los datos en diferentes contextos y al proporcionar información sobre el formato, la estructura y las características de los datos, los metadatos permiten que diferentes herramienta y aplicaciones comprendan y utilicen los datos de manera coherente.

1.6.2 Justificación Técnica

La justificación técnica de esta herramienta se basa en el hecho de que tanto el sector público como el privado poseen recursos tecnológicos sofisticados para la implementación de herramienta de metadatos inteligentes. También para la recuperación de datos significativos como ser la geolocalización de archivos digitales. A medida que aumenta la cantidad de información digitalizada, es cada vez más difícil encontrar y acceder a los recursos adecuados de manera rápida y eficiente.

1.6.3 Justificación Social

Los metadatos inteligentes permiten a las personas acceder rápidamente a recursos relevantes al mejorar la recuperación de información digitalizada. Por lo tanto, permite a las

organizaciones, ya sean públicas o privadas, recuperar los datos importantes de estos archivos digitales de manera más efectiva, lo que reduce el tiempo y el esfuerzo.

1.7 METODOLOGÍA

1.7.1 Método Científico

Para (Koval, 2011), “Es posible llevar a cabo esta exploración a través de discusiones, reuniones o discusiones de ideas una vez seleccionado el tema de investigación, se aplicará el método científico, las propuestas iniciales sirven como base para las etapas que se proponen a continuación.”

a) Fase 1: Observación

En su inicio, el científico comienza su exploración con la observación del fenómeno o elemento a investigar. Esta observación ha de ser minuciosa y precisa, realizada de una manera lo más objetiva posible teniendo en cuenta las características y diferentes dimensiones de lo que observamos.

De esta forma, podremos captar el mayor número de características propias de lo observado para poder describirlo y recabar así la mayor información posible. Al observar, resulta casi inevitable que al observador le surjan interrogantes acerca del fenómeno, elemento u objeto fruto de la curiosidad. Estas preguntas serán la antesala de la fase de formulación de hipótesis.

b) Fase 2: Recolección de información

Es importante que a medida que vayamos realizando la fase anterior, se consulte a diferentes fuentes sobre el tema que estamos observando para recabar la mayor cantidad de datos posible. De esta manera se podrá extraer a partir de las observaciones, información suficiente para la formulación de hipótesis de la siguiente fase. A su vez, se registrará toda la

información recogida de las observaciones para poder estructurarla y ordenarla de manera que facilite la reflexión sobre los conocimientos obtenidos y ayude a consolidar contenidos.

c) Fase 3: Formulación de hipótesis

Esta tercera fase se caracteriza por la formulación de hipótesis, es decir, por la elaboración de posibles respuestas o explicaciones que pueden dar respuesta a las preguntas surgidas de lo observado. Estas hipótesis han de ser sometidas a experimentación (fase 4) para comprobar su verificación.

d) Fase 4: Comprobación y experimentación

Esta fase está dedicada a la comprobación de las hipótesis formuladas en la fase anterior dictaminando así si finalmente estas hipótesis son verdaderas o falsas. Para ello, se someterán a una serie de pruebas controladas las posibles predicciones surgidas de la observación. Para la ejecución de estas experimentaciones, se deben elegir las herramientas precisas que permitan una adecuada realización de las pruebas. En ocasiones, la resolución obtenida de estos experimentos puede derivar en una nueva vía para la investigación.

e) Fase 5: Conclusiones y resultados

Una vez finalizada la fase de experimentación, se pasará a analizar de manera detallada los datos obtenidos a lo largo del proceso experimental. Con este análisis, se podrá determinar si finalmente las hipótesis son verdaderas o falsas dando una explicación coherente a los fenómenos, elementos u objetos observados inicialmente.

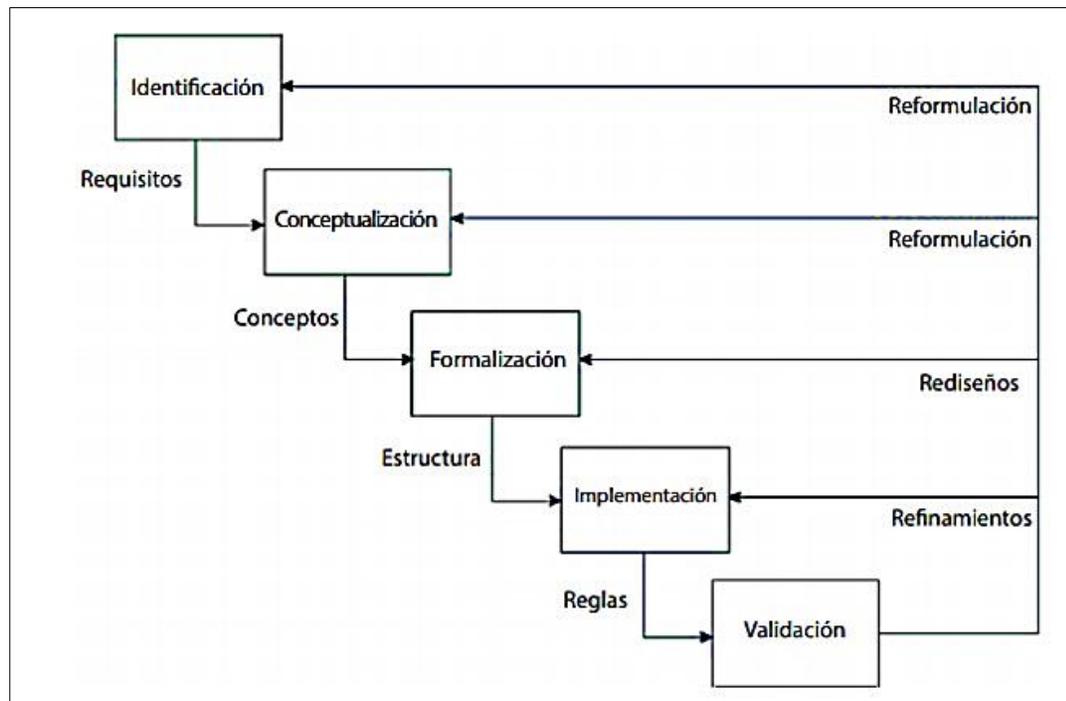
1.7.2 Metodología Buchanan

Para (Palma, José, & Marín, 2008), esta técnica se basa en el ciclo de vida en cascada común en el inicio de la ingeniería del software, de lo que se puede inferir que el proceso de

construcción de un sistema experto se plantea como un proceso de revisión casi constante que puede implicar la redefinición de conceptos, representaciones o refinamiento del sistema y se divide en cinco fases: identificación, conceptualización, formalización, implementación y prueba, tal como se puede observar en la figura a continuación:

Figura 1

Ciclo de vida de la Metodología Buchanan.



a) Fase 1: Identificación:

- En esta fase se identifican los conceptos y variables relevantes para el dominio del problema.
- Se desarrolla un vocabulario común para el desarrollador de la herramienta y el experto en el dominio.
- Se identifican las fuentes de conocimiento para la herramienta.

b) Fase 2: Conceptualización:

- En esta fase se desarrolla una representación conceptual del dominio del problema.
- Se identifican las relaciones entre los conceptos y variables.
- Se desarrollan modelos para representar el conocimiento del dominio.

c) Fase 3: Formalización:

- En esta fase se formaliza el conocimiento del dominio en un lenguaje de representación de conocimiento.
- Se selecciona un lenguaje de programación adecuado para el desarrollo de la herramienta.
- Se emplea la herramienta de metadatos para la verificación de datos.

d) Fase 4: Implementación:

- En esta fase se implementa la herramienta de ubicación de metadatos.

e) Fase 5: Prueba:

- En esta fase se prueba la herramienta para verificar que cumple con los objetivos del proyecto.
- Se evalúa el rendimiento de la herramienta de metadatos.

1.8 HERRAMIENTAS**1.8.1 Servidor Web**

Se están utilizando actualmente tecnologías web para proporcionar interfaces de usuario final en una variedad de entornos informáticos. Un servidor web, que se basa en el Protocolo de transferencia de hipertexto (HTTP), que funciona sobre TCP/IP, es el componente principal de estas soluciones web. Los servidores web deben responder a millones de

solicitudes de transacciones al día manteniendo un nivel aceptable de calidad de servicio (QoS) en términos de tiempo de respuesta del cliente y rendimiento del servidor (Reeser & Hariharan, 2002).

1.8.2 Gestor de Base de Datos

Es un conjunto de programas invisibles que el usuario final puede usar para administrar y administrar datos, que incluye una base de datos. Los gestores de datos, también conocidos como gestores de bases de datos, son diseñados para servir como una interfaz entre la base de datos, el usuario y las aplicaciones. Esto les permite gestionar todo acceso a la base de datos (Pérez, 2021).

1.8.3 Lenguajes de Programación

En la programación de computadoras, el lenguaje de programación se utiliza como lenguaje formal. Consiste en conjuntos de instrucciones que permiten a las computadoras usar algoritmos. Python, Java y C++ son lenguajes de alto nivel que no requieren hardware.

1.9 SOFTWARE

Los softwares a utilizar son:

1.9.1 PHP (Hipertext Preprocessor)

Para (Ullman, 2018), Es específicamente diseñado para el desarrollo web y puede integrarse en HTML para proporcionar características dinámicas como la integración de bases de datos, la autenticación de usuarios y otras. El código PHP se ejecuta en el servidor, produciendo HTML que luego se envía al navegador web del cliente. Es ideal para principiantes y desarrolladores experimentados que quieren crear aplicaciones web dinámicas e interactivas porque es simple, flexible y fácil de aprender.

1.9.2 SQL (Structured Query Language)

Structured Query Language (SQL), o Lenguaje de Consulta Estructurado, es un lenguaje de base de datos estándar que permite a los usuarios definir, manipular y controlar datos en un sistema de gestión de bases de datos relacionales (RDBMS). SQL proporciona un conjunto de comandos para realizar tareas como la consulta de datos, la actualización de datos, la inserción de datos y la eliminación de datos de una base de datos. También incluye comandos para definir y modificar el esquema de la base de datos, gestionar los derechos de acceso de los usuarios y controlar las transacciones (Connolly & Begg, 2014).

1.9.3 FRAMEWORKS

Una estructura de software conocida como marco de desarrollo web se utiliza para facilitar el desarrollo y mantenimiento de aplicaciones web. Proporciona una variedad de herramientas y librerías predefinidas que permiten a los desarrolladores crear aplicaciones de manera más eficiente, siguiendo un conjunto de convenciones y mejores prácticas. Estos marcos generalmente incluyen funcionalidades comunes como la gestión de rutas, el manejo de bases de datos, la autenticación de usuarios y la generación de interfaces de usuario, lo que permite a los desarrolladores concentrarse en la lógica específica de la aplicación (Sokambi, 2020).

1.9.4 LARAVEL

Laravel es un marco PHP útil que nos permite crear fácilmente aplicaciones web. Según (Subecz, 2021) Se basa en el sistema Symfony y utiliza el patrón de diseño conocido como MVC (modelo-vista-controlador). Como Laravel utiliza un sistema de paquetes modular, podemos agregar nuevos módulos a nuestra aplicación. Reutiliza componentes existentes de otros marcos para crear una aplicación operativa segura rápidamente. Algunas de sus características y ventajas incluyen: admite el acceso a múltiples bases de datos; ofrece herramientas que ayudan en el desarrollo de aplicaciones web; motor de enrutamiento simple

y rápido; la aplicación web se vuelve más escalable; y se ahorra un tiempo considerable en el diseño.

1.9.5 SYMFONY

Para (Yeates & Paris, 2023), Symfony es conocido por su enfoque modular, lo que lo hace altamente flexible y extensible. El framework se basa en el patrón de diseño Modelo-Vista-Controlador (MVC), que promueve la separación de preocupaciones y facilita el desarrollo de aplicaciones web mantenibles y escalables, es una buena opción para el desarrollo de API web, ya que es rápido, seguro y escalable.

Symfony tiene una curva de aprendizaje moderada, pero hay una gran cantidad de recursos disponibles para ayudar a los desarrolladores a aprender el framework.

1.9.6 JAVASCRIPT

JavaScript es un lenguaje de script, creado por Brendan Eich para el navegador Netscape en 1995 con el fin de posibilitar páginas web más dinámicas. Hoy en día es uno de los lenguajes de programación más importantes en el mundo de la informática, ya que se utiliza en la gran mayoría de los sitios web modernos.

Para (Flanagan, 2011), Una ventaja de JavaScript es que se incrusta fácilmente dentro de HTML, el lenguaje de marcado empleado para codificar páginas web. Además, como afirma (Subramaniam, 2018), es uno de los lenguajes de programación más poderosos, extendidos y flexibles que existen actualmente.

1.9.7 OpenStreetMap

El proyecto OpenStreetMap es un colectivo de conocimiento que proporciona mapas de calles generados por los usuarios. OSM sigue el modelo de producción entre pares que

creó Wikipedia; su objetivo es crear un conjunto de datos cartográficos de uso gratuito, editables y con licencia bajo nuevos esquemas de derechos de autor (Haklay & Weber, 2008).

1.10 LÍMITES Y ALCANCES

1.10.1 Limites

- Se deberá tener acceso a internet para resultado y pruebas que se realizará, en sitios web debido a la Geolocalización.
- La herramienta ubicara únicamente archivos digitales (imágenes).
- La herramienta no es compatible con la recuperación de información de imágenes que se hayan transferido a través de servicios y aplicaciones de terceros.

1.10.2 Alcances

- La herramienta generará metadatos automáticamente para cada imagen digitalizada ingresada directamente a la computadora.
- La herramienta permitirá la recuperación de documentos e imágenes basándose en diferentes criterios como nombres, fechas, e incluso ubicación y otras etiquetas relevantes.
- La interfaz permitirá visualizar los resultados de búsqueda de manera clara y organizada.
- Garantía de que la recuperación de información solo se realiza con las imágenes y documentos transferidos directamente a la computadora.



**INGENIERÍA
DE SISTEMAS**
UNIVERSIDAD PÚBLICA DE EL ALTO

CAPÍTULO II

MARCO TEÓRICO

2 MARCO TEORICO

2.1 METADATOS

Según (Calvo, 2019, pág. 9). "Los intentos por desarrollar algoritmos de IA que puedan tomar decisiones justas basadas en la objetividad de los datos y los metadatos en línea son otra clara manifestación del actual interés por la etificación y su aplicación práctica."

Esta es una de las cosas que tienen en común todas estas definiciones y categorizaciones, que los metadatos se crean para ser usados, con algunos propósitos, por personas o por aplicaciones informáticas. (Coyle, 2010, pág. 6) esquematiza cómo los metadatos son construidos, constructivos y accionables:

- **Los metadatos están contruidos:** son una creación artificial que no está presente en la naturaleza.
- **Los metadatos son constructivos:** se crean para un fin, para una actividad o para resolver un problema.
- **Los metadatos son accionables:** se pretende que sean útiles de alguna forma.

La creación de un conjunto de metadatos Este cambio hacia los "metadatos" en las comunidades académicas y profesionales de Biblioteconomía y Documentación, tanto en términos como en conceptos, es ilustrado por Dublin Core en 1995. El objetivo de Dublin Core, llamado así porque se creó en un taller en Dublin (Ohio) en marzo de 1995, era desarrollar un enfoque común para la descripción de recursos electrónicos que permitiera un mejor descubrimiento y recolección de recursos de la web (Weibel, 1995).

Según (Berners, 2019, pág. 6), se considera los metadatos como: "información inteligible para la computadora sobre recursos web u otras cosas y los metadatos son datos que pueden representar las características de un recurso de información para su búsqueda, recuperación, acceso y vinculación en diversas plataformas disponibles en el mundo digital."

Bajo esta premisa, la conexión de contenidos a través de perfiles semánticos de metadatos indica la creación de servicios de información digitales, que podrían apoyar los procesos de investigación. Además, mediante el establecimiento de grafos semánticos, es posible personalizar su búsqueda e identificar nuevas formas de conexiones mediante la extracción y lectura de metadatos (Russo, 2015).

Adicionalmente (Barreto, 2007), señala que: Cuando se trata de imágenes, el contenido visual se puede dividir en dos categorías: contenido primitivo de imágenes, que se refiere a los componentes básicos que componen las imágenes, y contenido complejo de imágenes, que se refiere a los estándares de imágenes que las personas percibieron como fuentes de significado. El contenido primitivo se refiere a características visuales que el computador puede reconocer y extraer automáticamente.

2.2 DATOS

(Furner, 2016, pág. 66) Los define como “instanciaciones concretas de representaciones simbólicas de proposiciones descriptivas, informadas por la observación empírica, sobre propiedades cuantitativas y cualitativas de fenómenos del mundo real”

Sin embargo, no es satisfactorio simplemente considerar que los metadatos son una subclase de datos, ya que aclara por qué llamar a una entidad específica "metadatos" en lugar de simplemente "datos". Los metadatos, también conocidos como clasificaciones o descriptores temáticos, son representaciones muy detalladas o descripciones de recursos de información, que no se basan en ciertos sistemas técnicos en los que los metadatos almacenados pueden ser mucho mayores que los datos (KLENSIN, 1995).

Presentan una técnica que utiliza la extracción de metadatos de fuentes pasivas de datos, que son fuentes sin motores de búsqueda para consultar. Su propuesta se centra en la creación de un esquema que capture las estructuras de los datos y las relaciones entre ellos

para crear herramientas que puedan extraer metadatos y almacenarlos en un repositorio. De esta manera, un usuario puede consultar directamente al repositorio de metadatos para encontrar datos relevantes en las fuentes pasivas (Martin, Powley, Weston, & Zyon, 2004).

Para (Vavouras, Gatzui, & Dittrich, 2002). presentan un método para modelar y ejecutar el proceso de refrescamiento de un almacén de datos en función de las especificaciones del proceso, que se almacenan en el repositorio de metadatos. Los autores describen un componente llamado "Gerente de refrescamiento de Data Warehouse" como parte de su propuesta. Este componente se utilizará para supervisar las tareas que deben realizarse durante el proceso de refrescamiento. El repositorio de metadatos y los subconjuntos de extracción, transformación y carga de datos componen este componente.

La creación de un conjunto de metadatos Este cambio hacia los "metadatos" en las comunidades académicas y profesionales de Biblioteconomía y Documentación, tanto en términos como en conceptos, es ilustrado por Dublin Core en 1995. El objetivo de Dublin Core, que se creó en un taller en Dublin (Ohio) en marzo de 1995, era desarrollar un enfoque común para la descripción de recursos electrónicos que permitiera un mejor descubrimiento y recolección de recursos de la web. (Weibel, 1995; Sugimoto, Baker y Weibel, 2002).

2.3 CALIDAD

En cualquier caso, las tres dimensiones funcionan como niveles de calidad interconectados. El nivel más alto se logra cuando los metadatos son consistentes, pero solo se puede lograr si antes cumplen con el nivel más básico, que es la suficiencia, y el nivel intermedio, que es la precisión. (Bellini & Nesi, 2013).

2.4 CALIDAD DE LOS METADATOS

Según el estudio planteado por (Parque, 2009), presenta el estado actual de la investigación y la práctica sobre la calidad de los metadatos centrándose en la perspectiva

funcional de la calidad, la medición y los criterios de evaluación de los metadatos, junto con mecanismos para mejorar la calidad de los metadatos.

Se necesita un modelo de datos común para mejorar la calidad de los metadatos en los repositorios digitales, siendo los criterios clave la precisión, la integridad y la coherencia.

Tabla 3

Métricas e indicadores para evaluar la calidad de los metadatos.

| Dimensión | Métrica | Indicador |
|--------------|---|----------------------------------|
| Suficiencia | Todos los registros tienen al menos título, autor y palabras clave. | Porcentaje. Valor esperado: 100% |
| Precisión | Los identificadores del documento (doi), la revista (issn) y autor (orcid) aparecen en el documento y en el registro. | Porcentaje. Valor esperado: 90% |
| Consistencia | Los autores se asientan siempre por apellido. | Porcentaje. Valor esperado: 100% |

Nota. Fuente: Elaboración Propia

2.5 PROTOCOLO OAI-PMH

(Gujardo, 2010, pág. 52), sostiene que “El protocolo OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting), también conocido como Iniciativa Abierta de Archivos, es una herramienta para la interoperabilidad que permite el intercambio de metadatos sobre cualquier material almacenado en forma electrónica.”

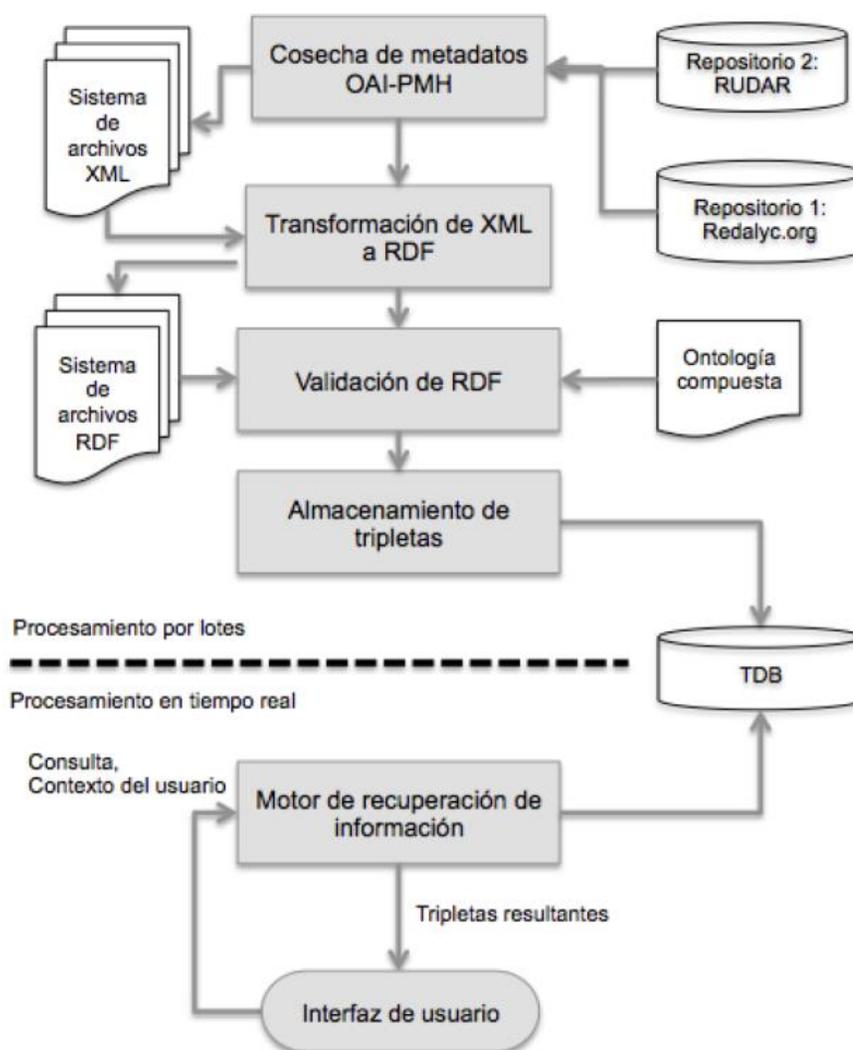
El Protocolo Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) es un protocolo de interoperabilidad para el intercambio y difusión de los metadatos de contenidos provenientes de distintas fuentes, plataformas y repositorios. Fue desarrollado por la Open

Archives Initiative (OAI) con el objetivo de facilitar la agregación y el descubrimiento de recursos digitales en la web (Arms & Turner, 2002).

Los archivos y bibliotecas nacionales pueden utilizar OAI-PMH para preservar sus colecciones digitales.

Figura 2

Metodología de recursos estructurados bajo el protocolo OAI-PMH



2.6 METADATOS INTELIGENTES

De acuerdo con (Jarke, Jeusfeld, Quix, & Vassiliadis, 2003), Explican cómo administrar la calidad de un almacén de datos a través de su repositorio de metadatos inteligentes. Los

autores sugieren incluir funciones de soporte a la calidad basadas en el enfoque genérico de gestión de la calidad del software.

Se basan en la utilización de tecnologías como la inteligencia artificial, el aprendizaje automático y el procesamiento del lenguaje natural para extraer y analizar información de diferentes fuentes, como el propio contenido del recurso, metadatos externos, ontologías y vocabularios controlados según (Bechtold & Buitelaar, 2004).

2.7 TIPOS DE METADATOS

Existen diferentes tipos de metadatos, que se pueden clasificar según el tipo de información que contienen:

2.7.1 Metadatos Descriptivos

Para (Senso & Piñero, 2003, pág. 5), Se utilizan para describir e identificar los datos en un recurso de información. “Incluye características físicas (medios, condición de dimensiones) y características bibliográficas (título, autor/creador, idioma, palabras claves).”

Sin embargo, para (Bates, 2005), los metadatos descriptivos son un conjunto de datos estructurados que proporcionan información esencial sobre un recurso digital, como su título, autor, fecha de creación, formato, idioma, tema, etc. Su objetivo principal es describir el contenido del recurso de manera precisa y concisa para facilitar su identificación, búsqueda, recuperación y gestión.

Al proporcionar detalles específicos y estructurados sobre el recurso, los metadatos permiten a los usuarios localizar rápidamente la información que necesitan.

2.7.2 Metadatos Administrativos

Según (Senso & Piñero, 2003), Se refieren a las características y características del recurso, lo que facilita la gestión, el procesamiento físico y tecnológico de las colecciones

digitales tanto a corto como a largo plazo. Incluye información sobre la calidad, la gestión de derechos, el control de acceso, el uso y las condiciones de preservación.

- **Técnicos.** Elementos o propiedades de datos que proporcionan información acerca del contexto técnico del recurso.
- **De derechos.** Elementos o propiedades de datos que describen los derechos y propiedades a los que se adhiere el recurso digital, como la propiedad, los derechos de autor o el copyright, u otros derechos de propiedad intelectual, así como las restricciones de usanza y seguridad.
- **De preservación.** Elementos o propiedades de datos que describen los requisitos para preservar el recurso a lo largo del tiempo y del cambio tecnológico.

2.7.3 Metadatos Estructurales

Proporcionan información sobre la estructura interna de los recursos electrónicos, que incluye páginas, secciones, capítulos, índices y tablas de contenido, y describen cómo se relacionan los materiales entre sí. Facilitan la navegación y presentación de los recursos y conectan las partes que los componen (TESTA & CEROTTO, 2012).

2.8 RECUPERACIÓN DE METADATOS

La principal necesidad de un usuario es encontrar la mayor cantidad de recursos de información sobre una temática específica, y el uso de este punto de acceso le permite recuperar en un sistema todos los recursos que abordan dicha temática. Esto demuestra la importancia de los metadatos utilizados para describir el contenido temático de los recursos de información (Arellano & Amaya, 2017).

2.9 MÉTODOS DE RECUPERACIÓN DE INFORMACIÓN

(Bridges & Edmunson, 2011), señalan que los bibliotecólogos, archivistas y programadores informáticos de la biblioteca deben comprender cómo funcionan los motores de búsqueda como Google, especialmente en lo que respecta a la indización y recuperación de imágenes. De esta manera, estos profesionales pueden mejorar las técnicas de búsqueda de sus colecciones, compitiendo con las diversas herramientas en línea que utilizan los estudiantes para recuperar imágenes.

2.10 EXIFTOOL

El formato de metadatos abierto está implementado por ExifTool. Está diseñado para encapsular metainformación de varias fuentes en formato binario o textual y agruparla con cualquier tipo de archivo. Puede utilizarse como un archivo único que envuelve los datos actuales o como un archivo complementario que contiene metadatos Exif o XMP.

Para (Harvey, 2020), Exiftool puede mostrar información detallada sobre los metadatos de una imagen, como la fecha y hora de la toma, la cámara utilizada, la configuración de exposición, la ubicación GPS (si está disponible) y mucho más. También puede editar estos metadatos, como eliminar información de copyright o cambiar la fecha de la toma.

2.11 METADATOS DUBLIN CORE

El estándar internacional Dublin Core responde a la gran cantidad de recursos electrónicos publicados en internet. Los centros de información no podían tratar documentalmente todos los recursos que se publicaban, por lo que se creó un estándar sencillo con campos genéricos y de manejo fácil e intuitivo para que los propios autores pudieran catalogar sus propios recursos (ISO-15836, 2006).

(Barité & al., 2015, pág. 155) proponen la siguiente definición de Dublin Core: “Los metadatos son un conjunto estándar de componentes de información esenciales que se utilizan para describir recursos electrónicos”. Este estándar comprende quince elementos:

Tabla 4

Metadatos Dublin Core: Definición y descripción de elementos

| Elemento | Descripción |
|----------------------------|--|
| Título | El nombre del recurso. |
| Creador | La persona o entidad responsable de la creación del recurso. |
| Tópico o palabra clave | Términos que describen el tema o contenido del recurso. |
| Descripción o resumen | Breve explicación del contenido del recurso. |
| Editor o entidad | La entidad que publica o distribuye el recurso. |
| Otros contribuyentes | personas que colaboraron en la creación del recurso. |
| Fecha | La fecha de creación, publicación o modificación del recurso. |
| Tipo de recurso | La categoría del recurso, como libro, artículo, imagen, etc. |
| Forma | El formato del archivo o medio en que se presenta el recurso, como PDF, JPG, MP3, etc. |
| Identificador del recurso | Un código único que identifica, como URL, ISBN o DOI. |
| Fuente | La referencia u origen del recurso, como un sitio web o una publicación. |
| Idioma | El idioma en que está escrito o presentado el recurso. |
| Relación | La conexión del recurso con otros recursos relacionados. |
| Cobertura y/o localización | El alcance temporal o espacial del recurso. |
| Gestión de derechos | La información sobre los derechos de autor y uso del recurso. |

Nota. *Información estructurada que describe las características de un recurso digital.

Tabla 5*Explorando el potencial de Dublin Core:*

| Elemento | Ejemplo |
|------------------------------|--|
| Título | "El Señor de los Anillos" |
| Creador | J.R.R. Tolkien |
| Tópico o palabra clave | Fantasía, épica, literatura |
| Descripción o resumen | Una novela épica sobre un grupo de hobbits que emprenden una misión para destruir el Anillo Único. |
| Editor o entidad responsable | George Allen & Unwin |
| Otros contribuyentes | N/A |
| Fecha | 1954-1955 |
| Tipo de recurso | Libro |
| Forma | Impreso |
| Identificador del recurso | ISBN 978-0-618-05623-1 |
| Fuente | Biblioteca pública local |
| Idioma | Español |
| Relación | "El Hobbit", "El Silmarillion" |
| Cobertura y/o localización | La Tierra Media, una época ficticia |
| Gestión de derechos | Derechos de autor J.R.R. Tolkien, licencia Creative Commons |

Nota. * Ejemplo ilustrativo de la aplicación de Dublin Core en la descripción bibliográfica de recursos electrónicos.

2.12 LA IMPORTANCIA DE RECUPERACIÓN DE INFORMACIÓN

Según (Croft, 1987), define como un conjunto de tareas mediante las cuales el usuario localiza y accede a los recursos de información que son pertinentes para la resolución del problema planteado se conoce como recuperación de información. Estas tareas requieren lenguajes documentales, técnicas de resumen, descripciones de objetos documentales, etc.

Años antes, (Salton, 1971, pág. 14), propuso una definición amplia que plantea que el área de RI “Es un campo relacionado con la estructura, análisis, organización, almacenamiento, búsqueda y recuperación de información”

2.12.1 Informática forense

El área del análisis forense de imágenes puede dividirse en dos grandes ramas: autenticidad de las imágenes e identificación de la fuente de adquisición de la imagen según (Gloe, Kirchner, Winkler, & Böhme, 2007). A su vez las imágenes creadas con móviles presentan características especiales que deben tenerse en cuenta en la elaboración de técnicas y algoritmos para el análisis forense en cualquiera de sus ramas.

Para (Lanh, Chong, Emmanuel, & Kankanhalli, 2007) y (Thing, Ng, & Chang, 2010), se realiza un estudio de los potenciales elementos que pueden ser objeto de análisis forense en dispositivos móviles.

Para (Romero, y otros, 2008) existe una gran variedad de trabajos que hacen referencia a los distintos tipos de metadatos en las imágenes con fines de búsquedas y clasificación, En estos trabajos se busca obtener patrones de valores en metadatos para conjuntos de fotografías que puedan arrojar datos válidos para el análisis forense. En lo referente a esta parte es esencial tener en cuenta cuales son los estándares de metadatos utilizados en la industria.

(Richard, 2010), El estándar empleado de forma mayoritaria en las cámaras digitales de cualquier tipo es Exif. Es necesario realizar un análisis pormenorizado de otras alternativas como TIFF, JFIF, IPTC y XMP, independientemente del estándar utilizado, los metadatos tienen un gran problema con respecto al análisis forense: su facilidad de manipulación.

Existen aplicaciones que no sólo permiten ver todos los metadatos agregados por los distintos estándares, sino que posibilitan también su edición de una forma gráfica y sencilla. Se puede decir en cierta forma que los metadatos incluidos en las imágenes gozan de una alta "vulnerabilidad".

2.13 IMÁGENES DIGITALES

Definición de una imagen digital Una imagen se define como una función de dos dimensiones $f(x,y)$ donde x e y son las coordenadas de un plano que contiene todos los puntos de la misma, y $f(x,y)$ es la amplitud en el punto (x,y) a la cual se le llama intensidad o nivel de gris de la imagen en ese punto. En el caso de que tanto las coordenadas x e y como los valores de intensidad de la función f sean discretos y finitos, se habla de una imagen digital (González, Woods, & Eddins, 2003).

Una imagen digital está compuesta de un número de elementos y cada uno tiene una localidad y un valor particular. A estos elementos se les llama puntos elementales de la imagen o píxeles, siendo este último el término comúnmente utilizado para denotar la unidad mínima de medida de una imagen digital.

2.14 TIPOS DE FORMATOS DE IMÁGENES

Para (González & Wood, 2008), Un formato es una manera estándar de codificar datos de una imagen con la finalidad de almacenarla o transmitirla. Sin embargo, debe haber algún método para convertir a ceros y unos que es el lenguaje del computador. Se desarrollan las dos categorías de archivos de imágenes: mapa de bits y vectores, y se presentan los formatos

de imágenes más utilizados, en este último caso se presenta con más detalle la técnica JPEG (Joint Photographic Experts Group).

Existen cientos de formatos de imágenes, sin embargo, los más conocidos y que pertenecen a la categoría de mapas de bits son:

- JPEG (Joint Photographic Experts Group)
- GIF (Graphics Interchange Format)
- PNG (Portable Network Graphics)
- TIFF (Tagged Image File Format)
- BMP (Windows bitmap)

Los que pertenecen a la categoría de vectores se describen en los siguientes párrafos

2.14.1 Formato JPEG (Joint Photographic Experts Group)

El algoritmo JPEG tiene muchas características que no están cubiertas aquí. La intención de este trabajo es proporcionar una idea muy básica de cómo JPEG puede ser utilizado para comprimir una imagen digital. Hay cuatro pasos básicos en el algoritmo - preprocesamiento, transformación, cuantificación y codificación (Lezama, 2017).

Según (Gailly & Nelson, 1996), "JPEG utiliza una versión más sofisticada de la codificación de Hoffman para llevar a cabo la misma."

Etapas 1. Preparación de la imagen

- Se realiza una transformación del espacio de color. Las imágenes de color se transforman del espacio RGB al espacio YUV (luminancia y prominencia).
- Submuestreo, permite la reducción de color para obtener archivos de tamaños más pequeños que el original.

- Formación de bloques, se divide la imagen en bloques de 8x8 pixeles.

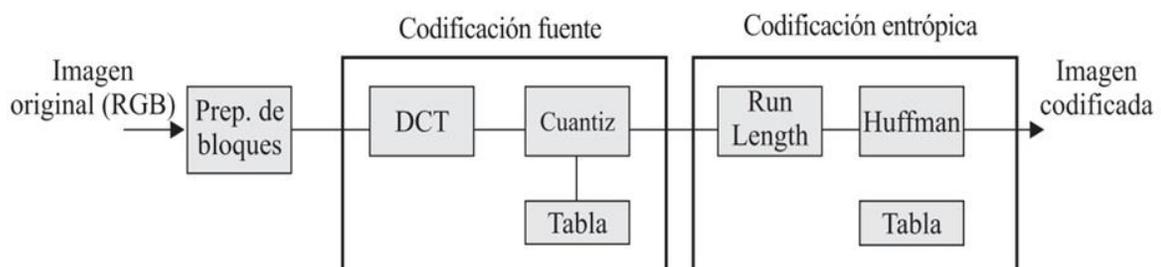
Etapa 2. Codificación fuente, codificación con pérdida

- Se aplica la transformación discreta de coseno o DCT, que se aplica a cada bloque de imagen, con lo que se obtiene un dominio de la frecuencia (matriz de coeficientes).
- Cuantificación, cada coeficiente de la matriz 8x8 es dividido por una constante de la matriz de cuantificación y se redondea a su número entero más cercano. Se eliminan los coeficientes menos representativos, produciéndose pérdida de información de la imagen y por consiguiente su calidad.

Etapa 3. Preparación de la imagen

- Se aplica la codificación Run-length, de todos los componentes de un bloque. Se hace un barrido “zig-zag” con el fin de agrupar todos los componentes nulos.
- La codificación Huffman se aplica a lo obtenido en la codificación anterior, con el fin de comprimir aún más la información. Este es el último paso del método para obtener la compresión de la imagen con el modelo JPEG.

Figura 3
Secuencia de pasos que sigue el método JPEG



2.14.2 Formato GIF (Graphics Interchange Format)

Según (Eppink, 2014), afirma que al igual que el kinetoscopio y el mutoscopio, el GIF ofrece imágenes en movimiento cortas y silenciosas como entretenimiento, pero hoy tiene una historia más larga, un contexto en evolución y un conjunto de estéticas que interactúan con las redes personales y sociales.

2.14.3 Formato PNG (Portable Network Graphics)

Para (Boutell, 1997), PNG es un formato de archivo extensible que se puede usar para almacenamiento portátil, comprimido y sin pérdidas de imágenes rasterizadas. Es un reemplazo sin patente para GIF y muchos de los tipos de archivos TIFF que se utilizan con frecuencia.

2.14.4 Formato TIFF (Tagged Image File Format)

TIFF es un formato de archivo de imagen estandarizado que admite varios formatos de imagen, incluidos JPEG, GIF, PNG y TGA (Poynton, 1992)

Es un formato de archivo de imagen etiquetado, pertenece a la categoría de mapa de bits. Creado originalmente por Aldus a mediados de los 80, actualmente es parte de Adobe Systems. Admite imágenes en la escala de grises y color real, incluso a 16 bits por píxel. Es un formato muy popular y usado en la mayoría de las aplicaciones de manipulación de imágenes, composición escaneada, fax, reconocimiento de caracteres ópticos (OCR), etc.

2.14.5 Formato BMP (Windows bitmap)

Para (Jian-fang, 2011), el algoritmo de visualización para archivos de mapa de bits en formato BMP es un método común y fácil de trasplantar para extraer información de imágenes, distinguir formatos de almacenamiento y obtener valores de color para cada píxel.

Los datos se codifican con compresión sin pérdidas, por lo que el tamaño de los archivos es considerable. Maneja colores de hasta 24 bits de profundidad. Son usados en los programas de Microsoft, en archivos de gráficos como íconos y papel tapiz.

2.15 COMPRENSIÓN DEL PROBLEMA

Se centra en la falta de metadatos en imágenes digitales, especialmente en formatos comunes como JPG, GIF y PNG. Esta deficiencia limita significativamente la capacidad para realizar búsquedas y recuperaciones de información de manera eficiente y precisa. Los metadatos son fundamentales ya que proporcionan información adicional y contextual sobre las imágenes, facilitando su organización y gestión.

Sin estos metadatos, la búsqueda de información específica se vuelve laboriosa y poco efectiva, resultando en la pérdida de información valiosa que las imágenes deberían proporcionar. En esencia, el problema radica en la ausencia de métodos o herramientas efectivas para integrar y gestionar metadatos en estos formatos, lo que impacta negativamente en la eficiencia y precisión de las búsquedas de información visual.

2.16 INGENIERIA DE SISTEMAS

La ingeniería de sistemas ha evolucionado desde el diseño de componentes individuales hasta el diseño de sistemas completos, y su desarrollo futuro se centrará en la mejora de las herramientas y la ampliación de su gama de aplicaciones, (Boaton & Ramo, 1984).

De acuerdo con (Elphick, 2000, pág. 9), “Para lograr el rendimiento y la confiabilidad deseados en grandes empresas dentro de los costos y plazos, la ingeniería de sistemas es el diseño del conjunto que se centra en optimizar los flujos de energía, masa y comunicaciones.”

Según (Bahill, 1995), la ingeniería de sistemas es un proceso interdisciplinario que garantiza que las necesidades del cliente se satisfagan durante todo el ciclo de vida de un sistema. Este proceso incluye comprender las necesidades del cliente, definir requisitos, realizar pruebas, validar requisitos, realizar revisiones de diseño, investigar ideas alternativas y analizar la descomposición funcional.

2.17 INVESTIGACION

Para (Deb, Dey, & Balas, 2018), “La investigación es un método sistemático para buscar conocimiento y formular teorías, contribuyendo a la expansión del conocimiento existente mediante la formulación de hipótesis, el análisis de datos y las deducciones.”

La investigación según (Monjes & Bailey, 2019, pág. 16), “El proceso de investigación incluye elegir una pregunta de investigación, elegir un enfoque de investigación, obtener la aprobación de una junta de revisión institucional, llevar a cabo la investigación y escribir sobre el proceso.”

2.17.1 Tipos de investigación

Existe una variedad de tipos de investigación para (Hernández, 2018), pueden clasificarse según el tipo de objeto de investigación, los objetivos o el nivel de conocimiento que se desea obtener. Se puede investigar desde una variedad de puntos de vista, con una variedad de objetivos o teniendo en cuenta una variedad de tipos de datos, procedimientos o métodos para obtenerlos.

2.17.1.1 Histórica, Descriptiva y Experimental

Tradicionalmente se presentan tres tipos de investigación de los cuales surgen los diversos tipos de investigaciones que se realizan y son: Histórica, Descriptiva, Experimental.

La investigación histórica se ocupa de las experiencias pasadas, describe lo que era y representa una búsqueda crítica de la verdad que sustenta los acontecimientos pasados. El investigador depende de fuentes primarias y secundarias, las cuales deben examinarse cuidadosamente para determinar su confiabilidad mediante críticas internas y externas. En las primeras dos situaciones, verifica la autenticidad de un documento o rastro, mientras que en las segundas, determina el significado y la validez de los datos que contiene el documento que se considera auténtico.

2.17.1.2 Básica y Aplicada

“La ciencia básica, también conocida como pura o fundamental, tiene como objetivo avanzar en el conocimiento científico y aumentar los conocimientos teóricos, sin preocuparse por sus posibles aplicaciones o repercusiones en la vida real. La ciencia más formal busca generalizaciones para el desarrollo de una teoría basada en principios y leyes.”

“La investigación aplicada está relacionada con la investigación básica porque se beneficia de los hallazgos y avances de la investigación básica y se enriquece con ellos. Sin embargo, la investigación aplicada se distingue por su interés en la aplicación, uso y efectos prácticos de los conocimientos. La investigación aplicada busca el conocimiento para hacer, actuar, construir y modificar.”

2.17.1.3 Documental, De campo o Mixta

A continuación, se refiere a otros tipos de investigación y en este caso se toma como criterio el lugar y los recursos donde se obtiene la información requerida.

La investigación documental es "aquella que se realiza mediante la consulta de documentos (libros, revistas, periódicos, memorias, anuarios, registros, códigos, constituciones, etc.). La investigación de campo, también conocida como investigación directa, es la que se lleva a cabo en el lugar y el tiempo en que ocurren los fenómenos objeto de

estudio. La investigación mixta es aquella que combina la investigación documental y de campo.

2.17.1.4 Exploratoria, Descriptiva y Explicativa

Se puede decir que esta clasificación usa como criterio lo que se pretende con la investigación, sea explorar un área no estudiada antes, describir una situación o pretender una explicación del mismo.

Los estudios exploratorios nos permiten aproximarnos a fenómenos desconocidos, con el fin de aumentar el grado de familiaridad y contribuyen con ideas respecto a la forma correcta de abordar una investigación en particular. El estudio exploratorio se centra en descubrir.

Los estudios descriptivos buscan desarrollar una imagen o fiel representación (descripción) del fenómeno estudiado a partir de sus características. Describir en este caso es sinónimo de medir. Miden variables o conceptos con el fin de especificar las propiedades importantes de comunidades, personas, grupos o fenómeno bajo análisis. En algunos casos los resultados pueden ser usados para predecir.

Los estudios explicativos pretenden conducir a un sentido de comprensión o entendimiento de un fenómeno. Apuntan a las causas de los eventos físicos o sociales. Pretenden responder a preguntas como: ¿por qué ocurre? ¿en qué condiciones ocurre? Son más estructurados y en la mayoría de los casos requieren del control y manipulación de las variables en un mayor o menor grado.

2.17.1.5 Experimental y No-experimental

La investigación puede ser clasificada como experimental o no experimental según el papel del investigador en los factores o características que se estudian.

Cuando se lleva a cabo un experimento, el investigador no solo identifica las características que se estudian, sino que también las controla, modifica o manipula para observar los resultados mientras intenta evitar que otros factores interfieran en la observación. Una investigación no experimental es aquella en la que el investigador se limita a observar los acontecimientos sin intervenir en ellos.

2.17.1.6 Transversales y Horizontales

La investigación puede ser de tipo horizontal o longitudinal cuando se extiende a través del tiempo dando seguimiento a un fenómeno o puede ser vertical o transversal cuando se enfoca en un momento y tiempo determinado. En el primer caso, la investigación podría incluir el seguimiento del desempeño intelectual de niños gemelos durante 35 años para determinar la influencia de factores genéticos y ambientales. Por otro lado, una investigación transversal sería un estudio de grado de desarrollo intelectual entre niños gemelos que tienen 9 años de edad, basándose en las teorías de Piaget.

2.17.1.7 De campo o Laboratorio

Esta clasificación distingue entre el lugar donde se desarrolla la investigación. Si las condiciones son naturales en el terreno de los acontecimientos, tenemos una investigación de campo, como observaciones en un barrio, encuestas a empleados de empresas, registro de datos sobre mareas, lluvias y temperaturas en condiciones naturales. En cambio, estamos ante una investigación de laboratorio si se crea un ambiente artificial para realizar la investigación, como un aula de laboratorio, un centro de simulación de eventos, etc.

2.17.1.8 Según la disciplina que se investiga

La investigación también puede clasificarse según el ámbito del quehacer humano en el que se realiza. Esto implica una separación entre varios campos científicos, el más conocido de los cuales es la separación entre las ciencias espirituales y las ciencias naturales. Sin

embargo, otros enfatizan el hecho desde el punto de vista lógico: cualquiera que sea el ámbito de aplicación de la investigación científica, no hay diferencias sustanciales, pero se admite que las diferencias se dan a nivel ontológico. Esto significa que las diferencias metodológicas no son radicales, sino que provienen del objeto.

2.18 METODOLOGÍAS

2.19 MÉTODO CIENTIFICO

El enfoque científico Cualquiera que se interese por la investigación y haya revisado algún libro escrito con ese propósito se habrá dado cuenta de que, en realidad, todos los libros dicen casi siempre lo mismo. Esto se debe al hecho de que los libros sobre metodología de investigación están basados en lo que se conoce como método científico. (Sampieri, Collado, & Lucio, 2000)

El método científico es simplemente una reconstrucción de cómo los científicos buscan el conocimiento. Es una de las cosas que distingue a la ciencia de otras disciplinas. Sin embargo, los filósofos de la ciencia están más interesados en el llamado método científico que en los científicos mismos. La disputa epistemológica entre inducción y deducción. Según Hernández, los pasos o etapas del proceso de investigación son:

PRIMER PASO

2.19.1 Concebir La Idea A Investigar

Generalmente con algunas instrucciones del profesor o por su cuenta, salvo que el tema haya sido dado previamente por quien encarga la investigación. Es común que al principio se tenga una idea vaga y poco precisa de lo que se desea o se puede lograr. Las ideas de investigación provienen de diferentes fuentes y niveles de conocimiento. Las ideas se pueden clasificar como totalmente nuevas, parcialmente investigadas o antiguas a las que se les da un nuevo enfoque según se precisen.

Las ideas de investigación tienen que ser atractivas para el investigador, deben alentar su interés por la investigación. Una vez concebida la idea de investigación y conocidos sus antecedentes, se debe avanzar al siguiente paso: el planteamiento del problema.

SEGUNDO PASO

2.19.2 Plantear El Problema De Investigación

El planteamiento del problema consiste en precisar y estructurar formalmente la idea de la investigación. El paso de la idea al planteamiento del problema unas veces es inmediato y automático y otras bastante más complejo; depende de lo que el investigador conozca el tema y de su experiencia y habilidad para la investigación. Dice Ackoff que "un problema correctamente planteado es un problema parcialmente solucionado". Un buen planteamiento nos acerca a la solución y evita pérdidas de tiempo, esfuerzos y recursos.

El planteamiento del problema debe expresar una relación entre dos o más variables; debe poderse formular claramente sin ambigüedades; y debe ser susceptible de prueba, porque la investigación científica aplicada estudia aspectos observables y medibles de la realidad. Para ello hay que realizar las siguientes tareas:

- a) Establecer los objetivos de la investigación. Estos objetivos deben elaborarse bajo los siguientes criterios:
 - i. Que sean explicitables y claramente expuestos.
 - ii. Que sirvan de guía para el estudio.
 - iii. Que sean alcanzables con el estudio.
 - iv. Que sean congruentes o compatibles entre sí.
- b) Desarrollar las preguntas de la investigación. Las preguntas deben resumir lo que va a ser la investigación. Las preguntas generales deben aclararse, esbozar el problema y

sugerir actividades pertinentes para el estudio. Las preguntas deben ser precisas, no generales. Con ellas se deben establecer los límites temporales y espaciales del estudio y esbozar un perfil de los sujetos a observar.

- c) Justificar la conveniencia y utilidad del estudio. Es necesario justificar los motivos del estudio y explicar por qué es conveniente realizar la investigación y qué beneficios se derivarán de ella. Los criterios para evaluar el valor potencial de una investigación son su conveniencia, su relevancia social, sus implicaciones prácticas, su valor teórico y su utilidad metodológica. En algunas investigaciones, aunque no sea necesario para fines estrictamente científicos, habrá que cuestionarse sobre sus consecuencias sobre los sujetos investigados y/o sobre toda o parte de la sociedad. Por ejemplo, las investigaciones que impliquen experimentos sobre personas.
- d) Demostrar la viabilidad o factibilidad de la investigación en función de la disponibilidad real de tiempo y recursos materiales, humanos y financieros para hacerla. esta disponibilidad es la que delimitará su alcance.

TERCER PASO

2.19.3 Elaborar El Marco Teórico

El marco teórico es el sustento teórico del estudio e incluye la exposición y análisis de las teorías, enfoques teóricos, investigaciones y antecedentes que se consideren válidos para el correcto encuadre del estudio. Las funciones del marco teórico son:

- Ayudar a prevenir errores cometidos en otros estudios
- Orientar sobre la forma de llevar a cabo el estudio
- Ampliar el horizonte y guiar al investigador en el planteamiento del problema
- Conducir el establecimiento de las hipótesis

- Inspirar nuevas líneas y áreas de investigación
- Proveer de un marco de referencia para interpretar los resultados del estudio.

Las etapas de elaboración del marco teórico son: revisión de la literatura y de las fuentes de información existentes y adopción de una teoría o desarrollo de una perspectiva teórica. La revisión de la literatura consiste en:

- **Identificar** las fuentes de información sobre el tema, tanto de los aspectos teóricos como de las aplicaciones ya existentes. Las fuentes pueden ser primarias (que proporcionan datos de primera mano o directos), secundarias (que proporcionan información procesada y/o reprocesada) y fuentes terciarias (que proporcionan fuentes primarias y secundarias).
- **Obtener** las informaciones, localizándolas y accediendo a ellas.
- **Seleccionar, leer y analizar** las fuentes.
- **Extraer y recopilar** la información de interés, para lo cual se fichan las fuentes y se seleccionan las ideas, datos, citas, etc. que se usarán, señalando correctamente las referencias.

En la construcción del marco teórico puede el investigador encontrar, entre otras opciones, que:

- Exista una teoría desarrollada y demostrada que se aplique al tema de la investigación. En este caso se toma la teoría como estructura del marco teórico de la investigación.
- Existan varias teorías aplicables al tema de la investigación. Se elige la que más se acerque al tema de investigación o se toman los elementos de otras teorías que sean útiles.

Criterios de evaluación de la utilidad de la teoría para la investigación: En general, las teorías son útiles porque describen, explican y predicen el fenómeno que se investiga, organiza el conocimiento y orienta la investigación. Los criterios de evaluación de la utilidad de las teorías son:

- Capacidad de descripción, explicación y predicción.
- Consistencia lógica.
- Perspectiva o nivel de generalidad.
- Fructificación (heurística).
- Parsimonia (sencillez).

CUARTO PASO

2.19.4 Definir Tipo Y Nivel De Investigación

Los diferentes tipos de investigaciones hacen referencia al nivel de profundización de los estudios que implican y la elección de un tipo u otro de investigación depende esencialmente del estado de los conocimientos sobre el tema, el enfoque y los objetivos del estudio. Hay diferentes tipos de investigación, tales como:

- a) Estudios exploratorios, que se hacen cuando el objetivo es examinar un tema o problema de investigación poco o nada estudiado. Generalmente no tienen una finalidad propia, pero sirven de base para preparar los otros tipos de estudios
- b) Estudios descriptivos, que miden de manera independiente los conceptos o variables objeto de la investigación. Su objetivo es alcanzar la máxima precisión posible en la medición de los fenómenos. Pueden ofrecer posibilidades rudimentarias de predicción.
- c) Estudios correlacionales, que pretenden establecer y medir las relaciones entre dos o más conceptos o variables. Estudian las relaciones entre ellas y su comportamiento

conjunto. Sirven para predecir el comportamiento de una variable sabiendo cómo varían otras vinculadas con ella. Tienen un valor explicativo parcial. Se deben eliminar cuidadosamente las "correlaciones espúreas", aparentes o casuales.

- d) Estudios explicativos, que tratan de analizar las causas de los acontecimientos físicos y/o sociales. Tratan de explicar por qué ocurre un fenómeno, en qué circunstancias acaece y cómo y por qué se relacionan dos o más variables.

Una investigación puede contener elementos de todos o algunos de estos tipos. Todos ellos son igualmente válidos e importantes; su elección depende del nivel de conocimientos que se tenga del tema y del enfoque que se le vaya a dar al estudio.

QUINTO PASO

2.19.5 Establecer Las Hipótesis Y Definir Las Variables

Las hipótesis son proposiciones tentativas sobre el comportamiento de las variables y/o sus interrelaciones; se apoyan en conocimientos previos organizados y sistemáticos; con el estudio se trata de probar si la realidad analizada las confirma o no. Es habitual que en la formulación de las hipótesis, se incorporen los juicios de valor que tiene el investigador sobre el tema a investigar, por lo que se deberán identificar éstos. Las hipótesis pueden ser generales o específicas; surgen de los objetivos y preguntas de investigación y de los postulados y análisis de la teoría y su aplicación al objeto de estudio y sirven para guiar el estudio, describir y/o explicar lo que se estudia y probar y/o sugerir teorías. Los requisitos que deben reunir las hipótesis son:

- Referirse a una situación real.
- Expresarse en términos o variables comprensibles, precisas y concretas.
- Proponer relaciones claras y verosímiles (es decir, lógicas) entre las variables.
- Plantear los términos de la hipótesis y la relación entre ellos de manera que puedan

ser observados y medidos, es decir, tener referentes en la realidad.

- Estar relacionadas con las técnicas disponibles para probarlas.

Hay muchos tipos de hipótesis, entre las que destacan:

a) Hipótesis de investigación o hipótesis de trabajo (H1, H2,...) son las que cumplen los cinco requisitos señalados. Según el tipo de investigación a realizar pueden ser:

- **Hipótesis descriptivas del valor de las variables:** Describen los valores de éstas que se espera observar.
- **Hipótesis correlacionales:** Expresan la posible relación asociativa entre dos variables y la forma de esta asociación. El orden en que se colocan las variables no importa, ya que aquí aún no se habla de variables independientes y dependientes, ni de causas y efectos. Pueden ser bivariadas (entre dos variables) y multivariadas (entre más de dos variables).
- **Hipótesis de la diferencia entre grupos:** Comparan los comportamientos de las variables en dos o más grupos de sujetos. Sólo pretenden establecer diferencias entre los grupos y especificar a favor de qué grupo están las diferencias.
- **Hipótesis causales:** Expresan relaciones causa-efecto entre variables. Para que pueda darse causalidad debe haber correlación. Pueden ser hipótesis causales bivariadas (una variable independiente-una dependiente); hipótesis causales multivariadas (varias variables independientes - una dependiente; una independiente - varias dependientes; varias independientes - varias dependientes); un caso especial de éstas son las hipótesis causales con variables intervinientes (en las que se incluyen una o más variables modifican la relación causal).

- b) Hipótesis nulas (H_0) son proposiciones acerca de la relación entre variables que sólo sirven para refutar lo que afirman las hipótesis de trabajo. Su clasificación es similar a la de éstas.
- c) Hipótesis alternativas (H_a) son las que expresan posibles "alternativas" entre las hipótesis de trabajo y nula o explicaciones distintas a las que proporcionan éstas. Se suman a las hipótesis de trabajo originales.
- d) Hipótesis estadísticas son la transformación de las hipótesis de trabajo, nulas y alternativas en símbolos estadísticos. Estas hipótesis son de tres tipos:
- ❖ **Hipótesis estadística de estimación** que corresponde a las hipótesis descriptivas del valor de las variables y se basan en información previa.
 - ❖ **Hipótesis estadística de correlación** que traduce la correlación entre dos o más variables en términos estadísticos.
 - ❖ **Hipótesis estadística de la diferencia de medias u otros valores** en la que se comparan estadísticas de dos o más grupos.

En cada investigación se formulan hipótesis de todos o algunos de los tipos, según las características del estudio. Si los datos no aportan pruebas en favor de las hipótesis planteadas, esto puede significar que las hipótesis eran falsas (lo que implica un avance en el conocimiento) o la investigación tuvo errores (lo que ayuda a mejorar la investigación).

Una variable o atributo es una propiedad del objeto, numérica o no numérica, respectivamente, que puede variar (adquirir diversos valores o modalidades) y cuya variación es susceptible de medida. Al formular las hipótesis se deben definir las variables con objeto de tener un lenguaje común entre investigadores y usuarios, garantizar la evaluación real de

las variables, poder comparar con otras investigaciones y evaluar correctamente los resultados de la investigación.

Detectar las variables es el proceso de identificarlas y definir las conceptual y operacionalmente. Una definición conceptual es la explicación del término que expresa la variable en función de otros términos. Esta definición es necesaria, pero insuficiente para la investigación ya que no se relaciona directamente con la realidad. Una definición operacional es el conjunto de actividades u operaciones necesarias para medir la variable.

SEXTO PASO

2.19.6 Seleccionar El Diseño Apropriado De Investigación

El diseño de investigación se elige en función del problema a investigar, el contexto de la investigación, los objetivos del estudio, las hipótesis formuladas y, en algunos casos, de la disponibilidad de recursos. Los diseños más habituales en las Ciencias Sociales son:

- Los diseños experimentales que pueden ser experimentos verdaderos, cuasiexperimentos y preexperimental.
- Los diseños no experimentales que pueden ser transeccionales (descriptivos, correlacionales y causales) y longitudinales (análisis de tendencia, evolutivos y de panel).

Debido a las características de las Ciencias Sociales, las investigaciones en este campo no suelen ser experimentales puras. En caso concreto de las Ciencias Económicas sus investigaciones suelen ser no experimentales y sólo en algunas áreas concretas, semiexperimentales. La investigación no experimental posee un control menos riguroso que la experimental y es más difícil inferir relaciones causales. Pero la investigación no experimental es más natural y cercana a la vida cotidiana. Se utiliza cuando no se puede o no se debe hacer experimentos. Para muchos autores, el desarrollo de los modelos

econométricos de simulación con ordenadores de gran capacidad, ha introducido la posibilidad de simular experimentos en este ámbito. Sin embargo, para los objetivos de este documento, nos concentraremos en los diseños no experimentales que son los habituales en los estudios económicos.

La investigación no experimental es la que se realiza sin intervención del investigador sobre las variables independientes. Proporciona un enfoque retrospectivo (ex-post facto) y observa variables y relaciones en su contexto natural. Pueden ser:

- a) Diseños transversales: Son los que hacen las observaciones en un único momento del tiempo. Dentro de ellos, podemos distinguir:
 - i) Diseños descriptivos, que miden individualmente las variables y reportan esas mediciones
 - ii) Diseños correlacionales, que describen relaciones entre variables sin imputarles sentido de causalidad.
 - iii) Diseños correlacionales – causales, que describen relaciones causales entre las variables.

- b) Diseños longitudinales: Son los que realizan observaciones en dos o más momentos del tiempo. Dentro de ellos podemos distinguir:
 - i) Diseños tendenciales, que estudian toda la población.
 - ii) Diseños evolutivos de grupo, que estudian grupos o subpoblaciones.
 - iii) Diseños de panel o longitudinales puros, que estudian a los mismos sujetos a lo largo del tiempo.

El tipo de diseño a elegir está condicionado por el problema a investigar, el contexto que rodea a la investigación, el tipo de estudio a efectuar y las hipótesis formuladas.

SEPTIMO PASO

2.19.7 Seleccionar Los Sujetos A Estudiar

Hay dos fases que realizar en este paso: Determinar el universo o población a estudiar y seleccionar y extraer la(s) muestra(s), cuando sea necesario.

Los sujetos de estudio conforman las unidades de análisis y se identifican en función del problema y de los objetivos de la investigación. La población es el conjunto de sujetos de estudio que tienen o pueden tener las características contenido, lugar y tiempo de las unidades de análisis. La muestra es un subconjunto de la población determinado según ciertos criterios llamados parámetros muestrales. Los principales parámetros son el tamaño de la muestra y el error muestral admisible. Estos parámetros varían en sentido opuesto: a mayor tamaño de la muestra, menor error muestral y viceversa. La decisión sobre éstos depende de los objetivos, diseño y costo de la investigación. El tipo de muestra se selecciona según dichos criterios y el uso previsto de los resultados de la investigación.

Las muestras probabilistas son aquellas en cuya selección interviene el azar. Se usan en los estudios que pretenden generalizar sus resultados a toda la población. La hipótesis implícita de estos estudios es que los fenómenos a estudiar se distribuyen aleatoriamente entre los sujetos de la población. Para ello, todos los elementos de ésta deben tener al inicio la misma probabilidad de ser elegidos, de manera que las características de la muestra "representen" las de la población. Las muestras probabilistas se establecen por:

- ❖ Muestreo por racimos o segmentos: La población se divide en segmentos (grupos heterogéneos en sí y homogéneos entre sí) y se eligen los segmentos por muestreo simple; para segmentar la población generalmente se usan criterios de proximidad geográfica o física para ahorrar costes.

- ❖ Muestreo estratificado: Se divide la población en estratos, (grupos homogéneos internamente y heterogéneos entre sí) y se aplica un muestreo simple dentro de cada estrato.
- ❖ Muestreo simple: Se eligen los sujetos al azar entre todos los elementos de la población de los que se investigan.

La selección de los elementos de la muestra se efectúa a través de diferentes procedimientos estadísticos y se ve afectada por la disponibilidad de información y por el coste de la investigación. Algunos de ellos son el sorteo (tómbola o bombo), la tabla de números aleatorios (random variable) o selección sistemática (un elemento de cada x , empezando por uno elegido al azar). Para ello debe existir un listado exhaustivo de todos los elementos de la población (marco muestral) que pueden proceder de archivos, mapas, censos y directorios de diferente tipo.

Las muestras no probabilísticas o dirigidas son aquellas en que la selección de los sujetos se hace según el criterio subjetivo del investigador. Sus resultados no pueden generalizarse al conjunto de la población (aunque algunas veces lo hagan). Pueden ser muestras de sujetos voluntarios, que se usan en diseños experimentales; de expertos, que se usa en estudios exploratorios; de sujetos-tipo o estudio de casos, usada en investigaciones cualitativas y motivacionales; y por cuotas, que se usa en mercadotecnia y encuestas de opinión.

Cuando la muestra tiene más de 100 observaciones, se puede aplicar el teorema central del límite, según el cual los valores de las características observadas en la muestra siguen una curva de distribución normal (campana de Gauss). Esto permite hacer análisis de inferencia estadística en cuanto a la distribución de los valores de dicha variable en la población.

OCTAVO PASO

2.19.8 Recolectar Los Datos

Este paso comprende un conjunto de operaciones, tales como:

- ❖ Elaborar el instrumento de medición y calcular su validez y confiabilidad.
- ❖ Elaborar la codificación de los datos.
- ❖ Aplicar el instrumento y efectuar la colecta de datos.
- ❖ Crear los archivos que contengan los datos (base de datos).

Medir es el proceso de vincular conceptos abstractos con indicadores empíricos, mediante clasificación y/o cuantificación. Todo instrumento de medición debe reunir los requisitos de validez y confiabilidad. La validez se refiere al grado en que un instrumento de medición mide realmente la variable. Se pueden aportar tres tipos de evidencia para la validez: de contenido, de criterio y de constructo (un constructo es una variable medida en un contexto de esquema teórico o de una teoría). Entre los factores que afectan negativamente la validez están la improvisación, el uso de instrumentos no validados en el contexto de la investigación, la falta de empatía y los problemas en la aplicación del instrumento (que se suelen resolver en la prueba-piloto). Aunque no hay medición perfecta, el error de medición debe reducirse a límites tolerables.

La validez de contenido se obtiene comparando el universo de contenidos posibles con los incorporados al instrumento de medida; la de criterio se obtiene contrastando los resultados de aplicar el instrumento con los de un criterio externo; y la de constructo se determina a partir del análisis de factores. La validez total se da cuando existe validez de los tres tipos.

La confiabilidad se refiere al grado en que la aplicación repetida de un instrumento de medición al mismo sujeto u objeto produce iguales resultados. Se calcula mediante el coeficiente de confiabilidad, el que varía entre 0 (confiabilidad nula) y 1 (confiabilidad total). Para calcular la confiabilidad se usan diferentes métodos (medida de estabilidad, método de formas alternas, etc.). Puede ocurrir que un instrumento sea confiable pero no válido (o sea, consistente en las mediciones que proporciona, pero no mide las variables que realmente se quiere medir). En la investigación social se dispone de diferentes instrumentos de medición, escalas de actitudes, diversos tipos de cuestionarios, análisis de contenido, pruebas estandarizadas, entre otros.

Las respuestas recogidas en el instrumento de medición se codifican para facilitar el procesamiento de datos. Cada vez con mayor frecuencia, sobre todo en el caso de los cuestionarios, las respuestas están precodificadas. Hay que codificar los valores y/o atributos no precodificados, elaborar el libro de códigos, grabar y guardar los datos en un archivo permanente.

NOVENO PASO

2.19.9 Analizar Los Resultados

Este paso consiste en:

- Seleccionar las pruebas estadísticas y las tablas de resultados a analizar.
- Elaborar el problema de análisis.
- Realizar el análisis.

El análisis se efectúa a partir de la matriz de datos del archivo permanente, seleccionando las pruebas estadísticas y las tablas de resultados, elaborando y ejecutando el programa de análisis adecuado al problema investigado. El tipo de análisis a realizar depende

del nivel de medición de las variables, de las hipótesis a contrastar y del interés del investigador. Los principales tipos de análisis son las estadísticas descriptivas de una variable, la inferencia estadística, las pruebas paramétricas y no paramétricas y los análisis multivariados.

DECIMO PASO

2.19.10 Presentar El Reporte De Resultados

Para elaborar y presentar el reporte de investigación deben seguirse determinadas reglas, según quien sea el destinatario final del trabajo de investigación. Es decir, el informe de investigación debe adaptarse al usuario. El contexto determina el formato, naturaleza, extensión e, incluso, el lenguaje del informe. En un contexto académico, es usual que el reporte de investigación incluya portada, índice, sumario, introducción, marco teórico, método, resultados, conclusiones, bibliografía y anexos estadísticos u otros. En un contexto no académico, el marco teórico y los detalles técnicos del método suelen trasladarse a los anexos y el resumen hace énfasis en las conclusiones; el lenguaje a emplear debe huir de tecnicismos sin perder rigor ni precisión. La presentación del reporte, además de escrito, puede ser oral en cuyo caso conviene apoyarse en métodos audiovisuales de presentación.

En un contexto académico, lo más usual es presentar un reporte que comprende:

- ✓ Portada
- ✓ Índice
- ✓ Introducción
- ✓ Marco teórico
- ✓ Método
- ✓ Resultados

- ✓ Conclusiones
- ✓ Bibliografía
- ✓ Apéndices o anexos (estadísticos, gráficos, legislación, etc).

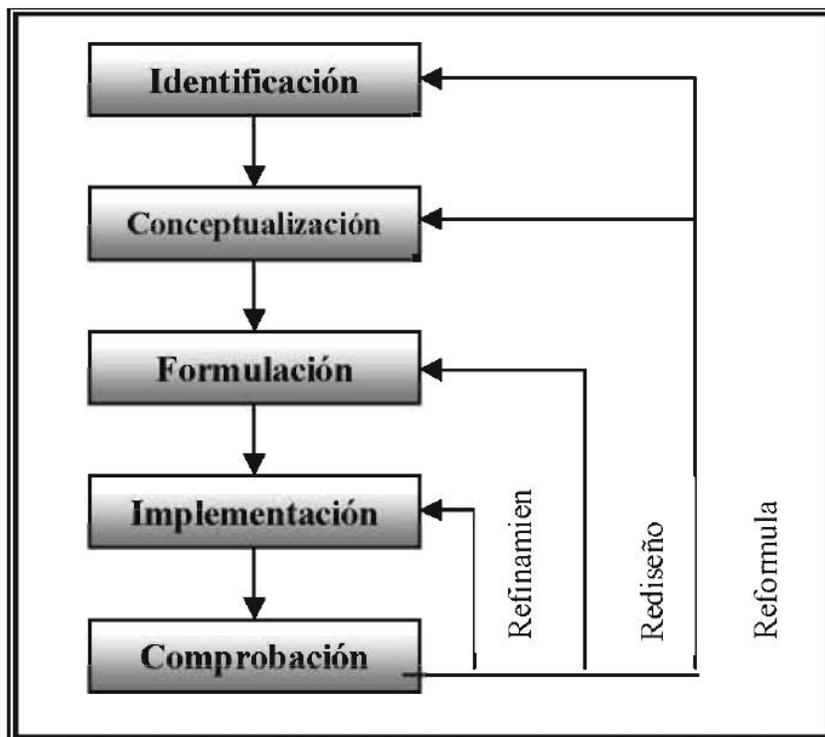
En un contexto no académico, el marco teórico, la bibliografía y los aspectos técnicos del método se traslada a los apéndices, presentando un resumen de la metodología utilizada, expresado en términos divulgativos, no técnicos. Los demás elementos se presentan en el orden señalado haciendo hincapié en los aspectos de aplicación práctica de los resultados cuando corresponda, con un resumen ejecutivo inmediatamente tras el índice.

Para la presentación del reporte de investigación se pueden utilizar diversos tipos de apoyos visuales y audiovisuales. El mejor apoyo audiovisual para las presentaciones de resultados de investigación es el que mejor domine el investigador que los presenta (Hernández, Fernández, & Baptista, 1991).

2.20 METODOLOGÍA DE DESARROLLO BUCHANAN

Uno de los primeros métodos de desarrollo estructurado de sistemas inteligentes fue el propuesto por Buchanan y otros autores en 1983. Según estos autores la adquisición del conocimiento de un sistema inteligente, y por extensión la construcción de todo el sistema, podía dividirse en las cinco fases de la figura: identificación, conceptualización, formalización, implementación y prueba.

Figura 4
Fases de la Metodología Buchanan



Sin embargo, el proceso real no está tan bien definido como puede sugerir la figura anterior, y más bien representa una aproximación a las distintas y complejas fases que se llevan a cabo a la hora de desarrollar un sistema inteligente, y que pueden variar de una situación a otra.

A. Fases

La descripción de cada una de estas fases es la siguiente:

Fase 1: Identificación. Fase mediante la que se reconocen aspectos importantes del problema, como son los participantes (expertos del dominio, ingenieros del conocimiento y futuros usuarios), las características del problema (tipo, sub tareas de que se compone, terminología a utilizar, aspectos fundamentales, etc.), los recursos disponibles (fuentes de

conocimiento, facilidades computacionales, tiempo de desarrollo, financiación, etc.), y las metas a alcanzar (formalizar conocimiento experto, distribuir experiencia, ayudar a la formación de nuevos expertos, etc.).

Fase 2: Conceptualización. Fase mediante la que se trata de organizar el conocimiento según un esquema conceptual. El experto y el ingeniero del conocimiento tratan de encontrar conceptos que representen el conocimiento del experto, al mismo tiempo que intentan determinar cómo es el flujo de información durante el proceso de resolución de problemas.

Fase3: Formalización. Esta fase consiste en traducir los conceptos clave, los sub problemas, y las características del flujo de información, identificados durante la fase anterior, en representaciones formales basadas en herramientas o esquemas de la ingeniería del conocimiento.

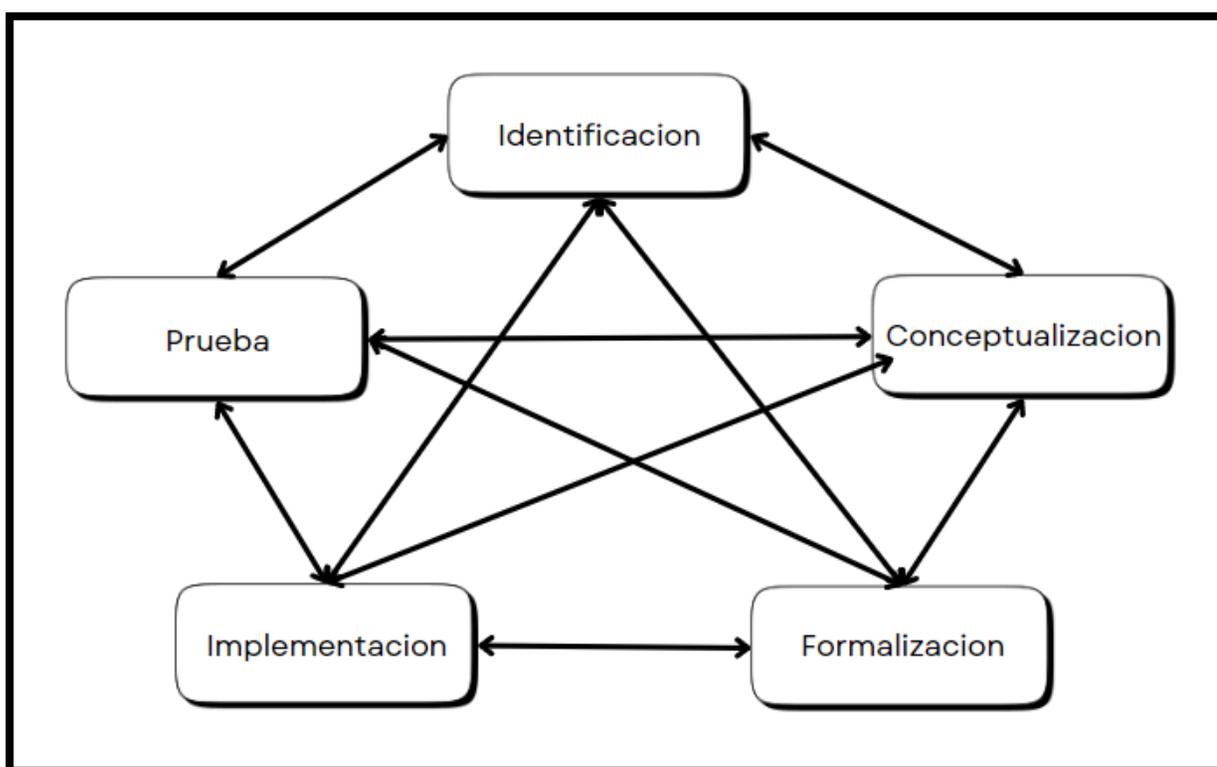
Fase 4: Elicitación. Aunque no aparece en el trabajo original de Buchanan, es común incluir una fase de Elicitación después de la fase de formalización. En esta fase se lleva a cabo la extracción del conocimiento mediante un soporte físico que es consistente con la información obtenida durante los procesos de identificación y conceptualización.

Fase 5: Implementación. En esta fase, el ingeniero de conocimiento formula reglas, y estructuras de control, que representan los conceptos y el conocimiento formalizado. El resultado es un programa prototipo que nos permite comprobar si hemos conceptualizado y formalizado bien el conocimiento que el experto tiene sobre el problema.

Fase 6: Prueba. Esta fase consiste en la evaluación del rendimiento de la herramienta construido para encontrar errores o anomalías en la base de conocimientos o en los mecanismos de inferencia.

Después de la fase de prueba, Buchanan sitúa los lazos de realimentación, pero también indica que el proceso no debe seguir estrictamente la secuencia que se muestra en la figura anterior. Según Mayrhauser y otros autores, las retroalimentaciones pueden surgir en cualquiera de los dos pasos de la metodología. Por ejemplo, si el ingeniero del conocimiento no puede encontrar las reglas correctas durante la implementación, puede ser necesario volver atrás y reformular el problema. La siguiente figura muestra una red completamente comunicada que sería la nueva representación del ciclo de vida de los sistemas inteligentes.

Figura 5
Fases de la Metodología



Las estructuras de este tipo son muy complejas de controlar y de manejar, ya que el número de iteraciones entre las fases es desconocido, y los objetivos pueden cambiar a medida que avanza el desarrollo. También es difícil llevar a cabo un control de los progresos realizados.

- **Síntesis Del Método De Buchanan**

Tabla 6
Comparación de Metodologías

| | | |
|-----------------------|---|--|
| F A C E 1 | Identificación | Seleccionar al experto, fuentes y medios de conocimiento y clara definición del problema |
| | Conceptualización | Encontrar los conceptos claves y las relaciones necesarias para 30 caracterizar el problema. |
| F A C E 2 | Formalización Implementación Control | reformulación permanente de conceptos, rediseño de la forma de representación, refinamiento del sistema implementado, críticas y sugerencias de los expertos permiten una mejora y un control del funcionamiento del sistema |

Nota. comparación de la metodología Buchanan en sus fases.

B. Etapas

Este método puede esquematizarse en seis etapas:

Etapa 1: Familiarizarse con el Problema y el Dominio.

Abarca Buscar el problema desde la lectura de libros o artículos, las entrevistas o charlas con las personas familiarizadas con el tema y la búsqueda de un experto que esté dispuesto a colaborar en la construcción del sistema.

Etapa 2: Delimitar el Sistema.

El objetivo es identificar y caracterizar el problema informalmente. El experto de campo y el ingeniero de conocimiento definen el alcance del sistema experto, es decir, que problemas va a resolver concretamente el sistema experto.

Etapas 3: Obtener la Estructura de Inferencia del Sistema Experto.

Con el problema adecuadamente definido el ingeniero de conocimiento empieza a determinar los principales conceptos del dominio que se requieren para realizar cada una de las tareas que va a resolver el sistema, es útil para la tarea de diseño, construcción y para posteriores modificaciones del sistema.

El ingeniero de conocimiento debe prestar atención al experto de campo para encontrar la estructura básica que el experto utiliza para resolver el problema.

El ingeniero de conocimiento además debe reconocer las estrategias básicas que usa el experto cuando desarrolla su tarea, que hechos establece primero, que tipos de preguntas realiza primero, si define supuestos inicialmente sin bases con información tentativa, como determina el experto que pregunta debe usar para refinar sus suposiciones y en qué orden el experto prosigue con cada sub tarea y si ese orden varía según el caso. La estructura del conocimiento indica que tareas y términos está usando y la estrategia indica cómo y cuándo el sistema experto debe establecerlas.

Etapas 4: Definir el Sistema Experto Prototipo.

El ingeniero de conocimiento debe formalizar el conocimiento obtenido del experto. Esta tarea implica definir que arquitectura permitirá una mejor organización del conocimiento. Es necesario elegir la organización, lenguaje y medio ambiente de programación adecuados para la aplicación particular.

El ingeniero de conocimiento deberá a medida que se desarrolla el prototipo, puede presentar las reglas definidas y en ocasiones los resultados obtenidos al usar las reglas, para que el experto manifieste su opinión sobre la representación y soluciones.

Etapa 5: Depurar el Sistema Prototipo.

Se refina el sistema prototipo, depurado la base de conocimientos, refinando reglas, rediseñando la estructura del conocimiento, o reformulando conceptos básicos, con el objetivo de capturar información adicional que haya proporcionado el experto. También se consultan en esta etapa otros expertos para corroborar, controlar, ampliar y refinar el prototipo.

Etapa 6: Optimizar el Sistema Experto Prototipo.

Cuando el sistema prototipo ha crecido tanto que resulta difícil de manejar el ingeniero de conocimiento rediseña un sistema más eficiente. Este nuevo sistema deberá refinarse y extenderse a fin de completar así el desarrollo del sistema experto. Esto es transformar efectivamente el sistema prototipo en un sistema experto aplicable.



**INGENIERÍA
DE SISTEMAS**
UNIVERSIDAD PÚBLICA DE EL ALTO

CAPÍTULO III

DISEÑO METODOLÓGICO

3 DISEÑO METODOLOGICO

3.1 METODOLOGÍA DE LA INVESTIGACIÓN

3.1.1 Tipo de Investigación

Según Hernández, la investigación es un proceso sistemático, que utiliza métodos científicos para generar nuevos conocimientos que pueden utilizarse para resolver una consulta o mejorar el sistema existente. Cualquier investigación en seres humanos está asociada con diversos grados de riesgo para el individuo participante y es importante salvaguardar el bienestar y los derechos de los participantes.

3.1.1.1 Investigación Descriptiva

La investigación descriptiva tiene como objetivo principal describir las características de un fenómeno o una población en particular. No busca establecer relaciones causales entre variables, sino simplemente observar y documentar cómo son las cosas en un momento dado. Se centra en proporcionar una imagen detallada de lo que se está estudiando.

Características:

- **Objetividad:** Se basa en la observación directa de los fenómenos.
- **No experimental:** No manipula variables, simplemente las observa y registra.
- **Uso de encuestas y cuestionarios:** Comúnmente utiliza herramientas como encuestas, entrevistas y cuestionarios para recolectar datos.
- **Estadísticas descriptivas:** Utiliza métodos estadísticos básicos para resumir los datos recolectados, como medias, medianas, modos y frecuencias.

3.1.1.2 Investigación Correlacional

La investigación correlacional tiene como objetivo identificar y medir la relación entre dos o más variables. No implica manipulación de variables, sino que busca determinar si existe una asociación o correlación entre ellas, y la fuerza y dirección de dicha relación.

Características:

- **Identificación de relaciones:** Examina cómo dos variables se relacionan entre sí.
- **Coefficiente de correlación:** Utiliza el coeficiente de correlación (r) para medir la fuerza y dirección de la relación. Este coeficiente varía entre -1 y 1.
- **Dirección de la relación:** Puede identificar si la relación es positiva (ambas variables aumentan o disminuyen juntas) o negativa (una variable aumenta mientras la otra disminuye).
- **No establece causalidad:** Aunque puede mostrar que dos variables están relacionadas, no puede determinar si una causa la otra.

3.2 MÉTODO DE INVESTIGACIÓN

Según la investigación, se eligió el método científico conforme a lo propuesto por Hernández.

El método es la ruta planificada o la estrategia que se debe seguir para obtener un resultado. Este método opera con conceptos, definiciones, hipótesis, variables e indicadores, que son los elementos fundamentales que proporcionan los recursos e instrumentos intelectuales necesarios para construir el sistema teórico de la ciencia, permitiendo así alcanzar el objetivo de la investigación.

3.2.1 Enfoque De Investigación

El enfoque de investigación es esencial para dirigir el estudio y elegir la metodología adecuada. Los enfoques cuantitativo, cualitativo y mixto son los tres principales.

Para Hernandez el enfoque cuantitativo es objetivo y utiliza datos estadísticos para probar hipótesis. Por el contrario, el enfoque cualitativo utiliza métodos más subjetivos y adaptables para comprender los fenómenos desde la perspectiva de los participantes. Para ofrecer una visión más completa y rica del fenómeno estudiado, el enfoque mixto integra

elementos de ambos, combinando técnicas cuantitativas y cualitativas. El problema de investigación y los objetivos del estudio determinan el enfoque a utilizar.

3.2.2 Variables De Investigación

Variable Independiente

Implementación de un sistema de metadatos inteligente para la recuperación de información digitalizada.

Variable Dependiente

Recolección de metadatos de imágenes digitales y precisión de los resultados de información recuperados en relación con la consulta realizada.

La investigación experimental se considera exitosa cuando el investigador ratifica que un cambio en la variable dependiente se debe a la manipulación de la variable independiente.

3.3 AMBIENTE DE INVESTIGACIÓN. UNIVERSO POBLACIÓN Y MUESTRA

3.3.1 Universo Poblacional

(Sampieri, Fernández, & Baptista, 2010) afirma que población de estudio es un conjunto de casos, definido, limitado y accesible, que formará el referente para la elección de la muestra, y que cumple con una serie de criterios predeterminados. Es necesario aclarar que cuando se habla de población de estudio, el término no se refiere exclusivamente a seres humanos, sino que también puede corresponder a animales, muestras biológicas, expedientes, hospitales, objetos, familias, organizaciones, etc.; para estos últimos, podría ser más adecuado utilizar un término análogo, como universo de estudio.

Es importante especificar la población de estudio porque al concluir la investigación a partir de una muestra de dicha población, será posible generalizar o extrapolar los resultados obtenidos del estudio hacia el resto de la población o universo.

3.3.2 Muestra

El autor (Wilks, 1962, pág. 16) define a una “muestra como estadística es un subconjunto de casos o personas en una población”.

Para varias aplicaciones es importante que una muestra sea representativa, por lo que es necesario elegir una técnica de muestra adecuada que produzca una muestra aleatoria adecuada (se obtiene una muestra sesgada cuyo interés y utilidad son más limitados dependiendo del grado de sesgo que se produzca).

presente). Como un subconjunto o subgrupo representativo de la población seleccionado mediante algún método de muestreo. Siempre se utiliza una muestra representativa de la población. Las muestras serán diferentes si hay varias poblaciones. Para que se pueda extraer toda la información deseada, la muestra debe estar bien seleccionada y trabajada con cuidado.

3.3.2.1 Muestras No Probabilísticas

- **Muestras por bola de nieve.** Este es uno de los métodos de muestreo más sencillos y económicos. Implica dirigir las encuestas a contactos directos o personales. Posteriormente, estos se referirán a otros contactos y así continuamente.

En ocasiones, aunque algún contacto no responda a la encuesta, contribuye al darle difusión. Generalmente, se hace uso de los contactos que se tienen en redes sociales como LinkedIn, Telegram, Facebook, WhatsApp, Twitter y otras.

- **Diseños por cuotas.** Este tipo de muestreo establece lotes de entrevistas según las variables a considerar. Un ejemplo puede ser la frecuencia de consumo de un producto o marca, sexo y edad.

Las cuotas no guardarían proporción con las características de la población. En cambio, estarían relacionadas con la necesidad de comparar los resultados entre las categorías de las diversas variables escogidas.

Los resultados obtenidos se probarían mediante herramientas de significación estadística, analizando la diferencia de porcentajes entre las distintas respuestas.

3.3.2.2 Muestras Probabilísticas

- **Muestras a través de bases de datos o listados.** Contar con bases de datos o listados facilita mucho el proceso de reclutar la muestra; y mucho más si se tienen datos específicos como dirección, números telefónicos o emails.

Este tipo de muestras son las más favorables al momento de hacer un estudio online, por la calidad de la muestra. Igualmente, permite prescindir del trabajo de los captadores, por lo que el trabajo se hace más económico y ágil.

- **Muestras a base de un panel.** Estas muestras están conformadas por un grupo de personas que han sido entrevistadas y reclutadas con antelación; puede ser a través de entrevistas personales o por teléfono.

Esto ayuda a evitar errores de cobertura o de no respuesta, ya que los paneles incluyen personas con interés en participar en el estudio. Una vez hecho el convenio, se establecen los canales de contacto más convenientes.

- **Muestras por cuotas con similitud distributiva de la población.** En este tipo de muestreo se emplean captadores que se responsabilizan de controlar las cuotas. De esta forma, es posible lograr representatividad distributiva.

Sin embargo, para que sea así, es necesario aplicar ciertos ajustes por ponderación. Es decir, considerar variables para controlar las cuotas. Las más convenientes suelen ser: ocupación, nivel de estudios, edad, sexo, ubicación, etc.

Son muy diversas las maneras de calcular el tamaño de una muestra, según los datos a los que se puede tener acceso. En este caso, hablaremos del modo de calcular una muestra para un universo o población finita.

Figura 6

Fórmula para la muestra

$$n = \frac{Z^2 \sigma^2 N}{e^2 (N-1) + Z^2 \sigma^2}$$

Nota. Podría tratarse de la cantidad de habitantes de una ciudad en particular. Para obtener el tamaño ideal de la muestra, aplicaremos la fórmula planteada por Murray y Larry (2005).

En esta fórmula:

n = tamaño de la muestra que se busca.

N = tamaño del universo a estudiar.

σ = desviación estándar de la población. Cuando este dato es desconocido, se utiliza una constante que equivale a 0.5

Z = valor que se obtiene a través de niveles de confianza. Es un número constante, que generalmente adquiere dos valores según el nivel de confianza que se quiera. El 99% es el valor más elevado (equivale a 2.58) y el 95% (equivale a 1.96) es el valor mínimo admitido para que la investigación se considere confiable.

e = es el límite de error muestral admisible. Comúnmente oscila entre el 1% (0.01) y el 9% (0.09). El 5% (0.05) es el valor estándar que se usa en investigaciones.

3.4 DESCRIPCIÓN DE LA METODOLOGÍA A USAR

Métodos de investigación hipotética/deductivo: Consiste en un procedimiento que intenta dar respuesta a los distintos problemas que se plantea la ciencia a través de la postulación de hipótesis que se toman como verdaderas, no habiendo ninguna certeza acerca de ellas. Son conjeturas que anticipan una solución a esos problemas, y van configurando estructuras jerárquicamente ordenadas de hipótesis donde algunas son fundamentales, otras son derivadas y otras cumplen una función auxiliar.

De esos sistemas hipotéticos se derivan deductivamente consecuencias que deberán ser empíricamente contrastadas. Por último, habrá una instancia de contrastación que permitirá evaluar las hipótesis sobre los resultados obtenidos. Si el resultado es desfavorable, la hipótesis será considerada refutada y en principio deberá abandonársela. Si en cambio el resultado fuera favorable, la hipótesis será considerada corroborada. Esta noción de corroboración difiere de la de verificación.

3.4.1 Paradigma De Investigación

Basarse en el positivismo es aceptar conocimientos que procedan de la experiencia del sujeto, el empirismo. Mediante el principio de verificación de las proposiciones, sólo tienen validez los conocimientos que existen ante la experiencia y observación; todo debe ser comprobado para ser válido para la ciencia. En este paradigma la experimentación ha constituido la principal forma para generar teoría formal (Hernández, Fernández, & Baptista, 2010).

Como se afirmó previamente, la ubicación de un investigador en un determinado paradigma es producto de la respuesta a tres preguntas. Por tanto, se responderá a cada una de ellas mediante la postura del positivismo. En primer lugar, la pregunta ontológica. Para el positivismo la realidad es absoluta y totalmente aprehensible por el ser humano, es regida por

las leyes y mecanismos naturales. Desde este paradigma se pueden determinar los diferentes factores que se encuentran alrededor de un fenómeno de estudio, sean éstos causales, mediadores o moderadores (Field, 2009).

3.5 HERRAMIENTAS

3.5.1 Método de ingeniería – Modelo Buchanan

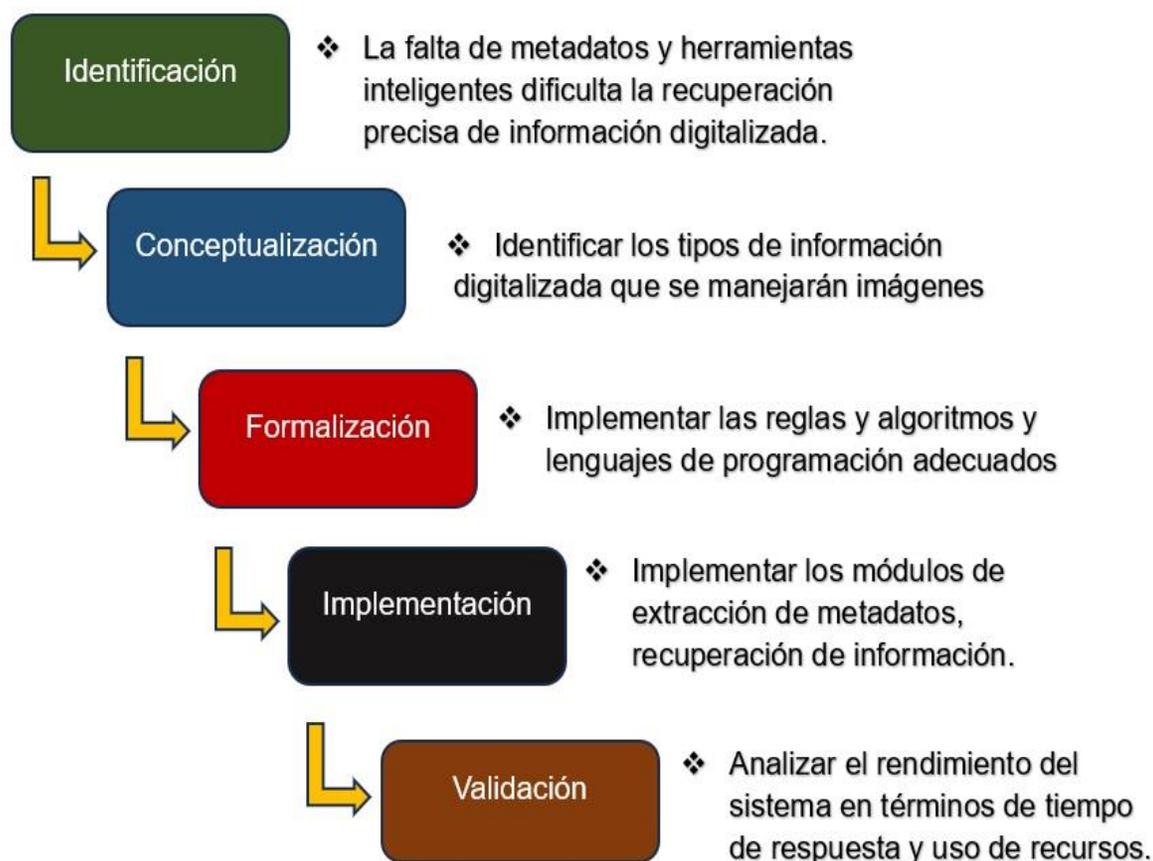
El Modelo Buchanan, también conocido como Método de Adquisición de Conocimiento por Reglas, es una metodología de ingeniería de software para el desarrollo de sistemas expertos. Fue desarrollado por Bruce G. Buchanan en la década de 1970 y se ha utilizado ampliamente en diversos campos, como la medicina, la ingeniería y la informática.

Etapas del Modelo Buchanan: El Modelo Buchanan se compone de seis etapas clave:

- a) Identificación del problema y análisis de necesidades**
- b) Adquisición de conocimiento**
- c) Formalización del conocimiento**
- d) Implementación del sistema experto**
- e) Evaluación y refinamiento**
- f) Mantenimiento y actualización**

Figura 7

Método Buchanan aplicado al sistema de información metadatos.



Nota. Modelo de ciclo de vida Buchanan, Elaboración Propia.

3.5.2 Técnicas de Investigación e Instrumentos

Para (Vindrola & Johnson, 2020), Los investigadores cualitativos se enfrentan a una mayor demanda de tiempo para recopilar, analizar y difundir rápidamente los hallazgos de sus investigaciones. Para permitir el uso de datos cualitativos oportunos y pertinentes en múltiples entornos disciplinarios, es posible que se requieran ajustes a los métodos cualitativos. El objetivo de esta revisión es examinar brevemente cómo las técnicas de recolección y análisis de datos se han adaptado para abordar períodos de estudio cortos en la investigación cualitativa.

3.5.2.1 Método de investigación

3.5.2.2 Encuesta

Para (Stratton, 2015), La investigación por encuesta es un enfoque que utiliza preguntas para recopilar información de los participantes del estudio para determinar sus creencias, actitudes, experiencias y expectativas.

Para la encuesta se tomó como muestras estudiantes de la carrera de ingeniería de sistemas donde se encontró profesionales de la materia tanto como estudiantes de primer ingreso con poco conocimiento y se demostró la importancia sobre los metadatos dando resultados a los datos demográficos de cada persona, así mismo se realizo preguntas acorde a los parámetros establecidos para encuestas a estudiantes y para poder determinar la importancia de los metadatos y la importancia de no perder los mismos.

Sección 1: Datos Demográficos

Tabla 7
Datos Demográficos (Nivel Educativo)

| Categoría | Subcategoría | Número de Encuestados | Porcentaje |
|------------------------|---------------------|------------------------------|-------------------|
| Nivel Educativo | Profesional | 140 | 40% |
| | Universitario | 105 | 30% |
| | Estudiante | 105 | 30% |
| Total | | 350 | 100% |

Nota: Esta tabla muestra los datos demográficos (nivel educativo) de las personas encuestadas para este estudio.

Tabla 8
Datos Demográficos (Edad)

| Categoría | Subcategoría | Número de Encuestados | Porcentaje |
|------------------|---------------------|------------------------------|-------------------|
| Edad | De 20 a 25 años | 119 | 34% |
| | De 25 a 30 años | 170 | 48.57% |
| | De 30 a 35 años | 61 | 17.43% |
| Total | | 350 | 100% |

Tabla 9
Datos Demográficos (Genero)

| Categoría | Subcategoría | Número de Encuestados | Porcentaje |
|------------------|---------------------|------------------------------|-------------------|
| Sexo | Masculino | 180 | 51.4% |
| | Femenino | 170 | 48.6% |
| Total | | 350 | 100% |

Nota: Esta tabla muestra los datos demográficos (sexo de los encuestados) de las personas encuestadas

Sección 2: Conocimiento sobre Metadatos

Tabla 10
Conocimiento sobre Metadatos P1

| Pregunta | Respuesta | Número de Encuestados | Porcentaje |
|--|------------------|------------------------------|-------------------|
| ¿Ha escuchado hablar de los metadatos? | Sí | 280 | 80% |
| | No | 70 | 20% |
| Total | | 350 | 100% |

Nota: Esta tabla muestra los conocimientos sobre los metadatos de la primera pregunta de la encuesta.

Tabla 11
Conocimiento sobre Metadatos P2

| Pregunta | Respuesta | Número de Encuestados | Porcentaje |
|--|------------------------|------------------------------|-------------------|
| En caso de haber respondido "Sí" a la pregunta anterior, ¿podría decir si son necesarios los metadatos? | Sí son muy necesarios | 220 | 78.6% |
| | No son nada necesarios | 60 | 21.4% |
| | Total | 350 | 100% |

Nota: Esta tabla muestra los conocimientos sobre los metadatos de la segunda pregunta de la encuesta.

Tabla 12
Conocimiento sobre Metadatos P3

| Pregunta | Respuesta | Número de Encuestados | Porcentaje |
|--|------------------|------------------------------|-------------------|
| ¿Tiene experiencia en la búsqueda y recuperación de información digitalizada? | Sí | 210 | 60% |
| | No | 140 | 40% |
| | Total | 350 | 100% |

Nota: Esta tabla muestra los conocimientos sobre los metadatos de la tercera pregunta de la encuesta.

Sección 3: Evaluación del Herramienta de Metadatos Inteligente

Tabla 13

Evaluación del Herramienta de Metadatos P1

| Pregunta | Respuesta | Número de Encuestados | Porcentaje |
|---|-----------|-----------------------|------------|
| ¿Ha utilizado alguna vez una herramienta de metadatos para la recuperación de información digitalizada? | Sí | 50 | 14.29% |
| | No | 300 | 85.71% |
| Total | | 350 | 100% |

Nota: Esta tabla muestra la utilización sobre los metadatos.

Tabla 14

Evaluación del Herramienta de Metadatos P2

| Pregunta | Respuesta | Número de Encuestados | Porcentaje |
|--|--------------------|-----------------------|------------|
| Si ha utilizado un sistema de metadatos, ¿son de utilidad según lo que conoce? | Sí son de utilidad | 35 | 70% |
| | No son de utilidad | 15 | 30% |
| Total | | 350 | 100% |

Nota: Esta tabla muestra la utilidad sobre los metadatos.

Tabla 15
Evaluación del Herramienta de Metadatos P3

| Pregunta | Respuesta | Número de Encuestados | Porcentaje |
|--|-----------|-----------------------|------------|
| ¿Le parece interesante la aplicación de la recuperación de información digitalizada? | Sí | 315 | 90% |
| | No | 35 | 10% |
| Total | | 350 | 100% |

Nota: Esta tabla muestra lo interesante de el uso de los metadatos para los encuestados.

Perfil de los Encuestados

Para esta investigación, se encuestaron a 350 personas al azar, todas ellas mujeres y hombres de 25 a 30 años de edad, que cursan la carrera de Ingeniería de Sistemas en la Universidad Pública de El Alto a continuación, se presenta un resumen de las características demográficas y de conocimiento sobre metadatos de los encuestados.

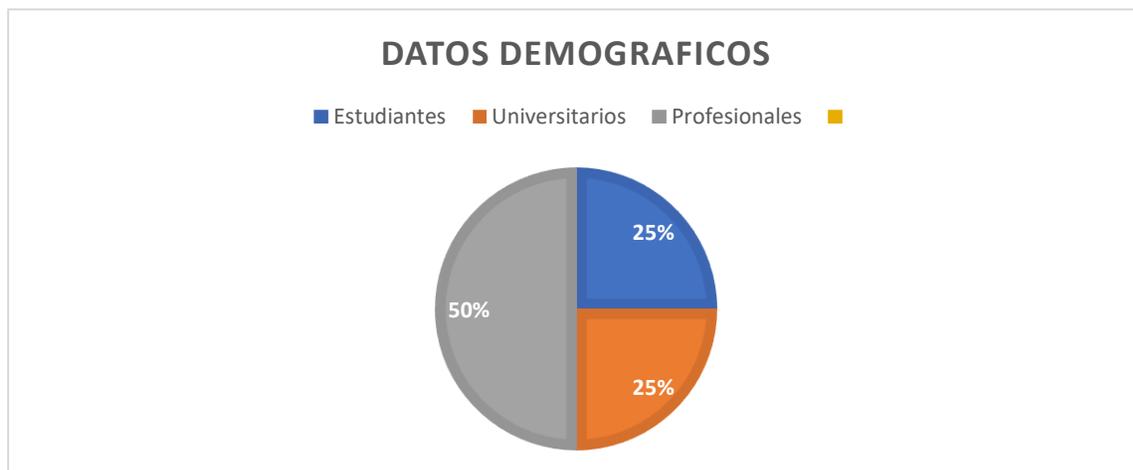
Conclusión

Sobre la Sección 1: Datos Demográficos

Datos Demográficos

- **Nivel Educativo:** La mayoría de los encuestados tienen un nivel educativo diverso, siendo profesionales el 40%, seguidos por universitarios y estudiantes, ambos con un 30%. Esto indica una muestra variada en términos de preparación académica, lo cual puede aportar diferentes perspectivas al estudio.

Gráfico 1
Datos Demográficos (Nivel Educativo)



- **Edad:** La mayoría de los encuestados (48.57%) se encuentra en el rango de edad de 25 a 30 años, seguido por un 34% de personas de 20 a 25 años y un 17.43% de 30 a 35 años. Esta distribución refleja una concentración significativa en adultos jóvenes y jóvenes adultos, quienes probablemente están familiarizados con tecnologías digitales y sistemas de información.

Gráfico 2
Datos Demográficos (Edad)



- **Sexo:** La muestra está bastante equilibrada en cuanto al género, con un ligero predominio de encuestados masculinos (51.4%) frente a femeninos (48.6%). Esto asegura una representación diversa en términos de género para las conclusiones del estudio.

Gráfico 3
Datos Demográficos (Sexo)



Sobre la Sección 2: Conocimiento sobre Metadatos

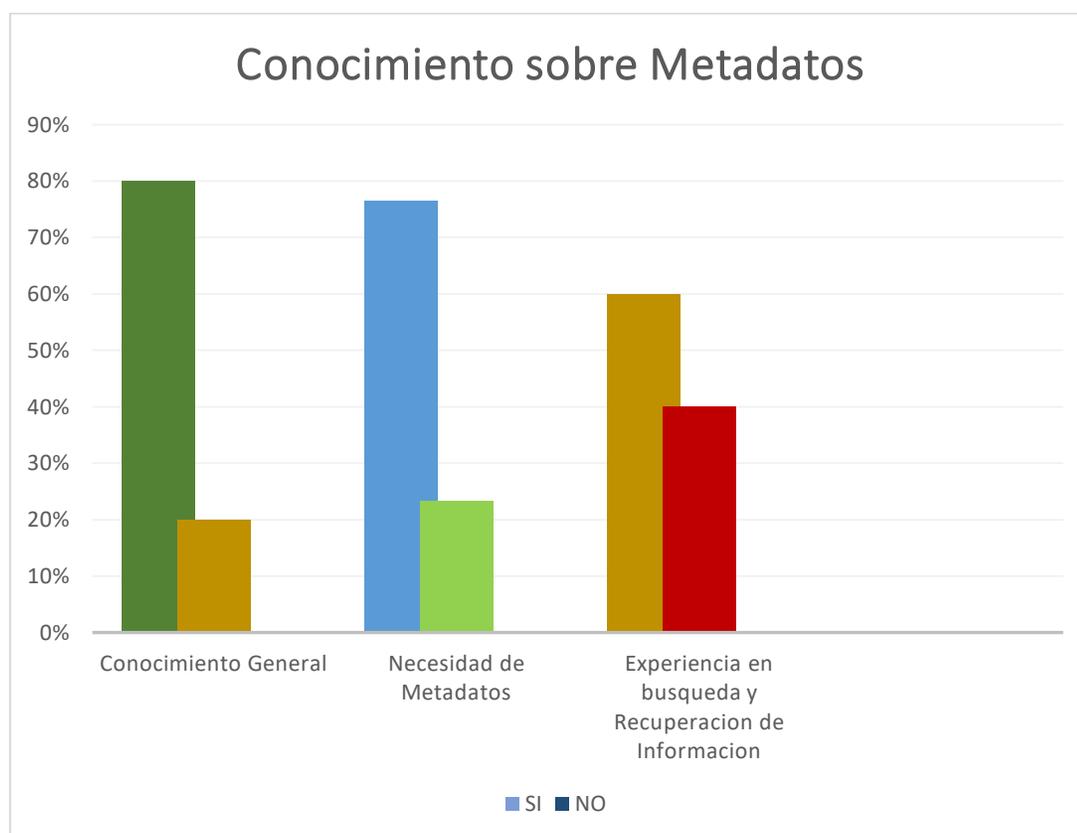
Conocimiento sobre Metadatos

- **Conocimiento General:** El 80% de los encuestados ha escuchado hablar de metadatos, lo cual indica un nivel generalizado de familiaridad con el término y su relevancia en la gestión de información digital.
- **Necesidad de Metadatos:** De aquellos que conocen los metadatos, el 78.6% considera que son muy necesarios. Esto subraya la percepción positiva y la importancia atribuida a los metadatos en la organización y recuperación de información.

- **Experiencia en Búsqueda y Recuperación de Información:** Un 60% de los encuestados tiene experiencia en la búsqueda y recuperación de información digitalizada, lo cual sugiere una base sólida de usuarios activos en este ámbito.

Gráfico 4

Conocimiento sobre Metadatos



Sobre la Sección 3: Evaluación del Herramienta de Metadatos Inteligente

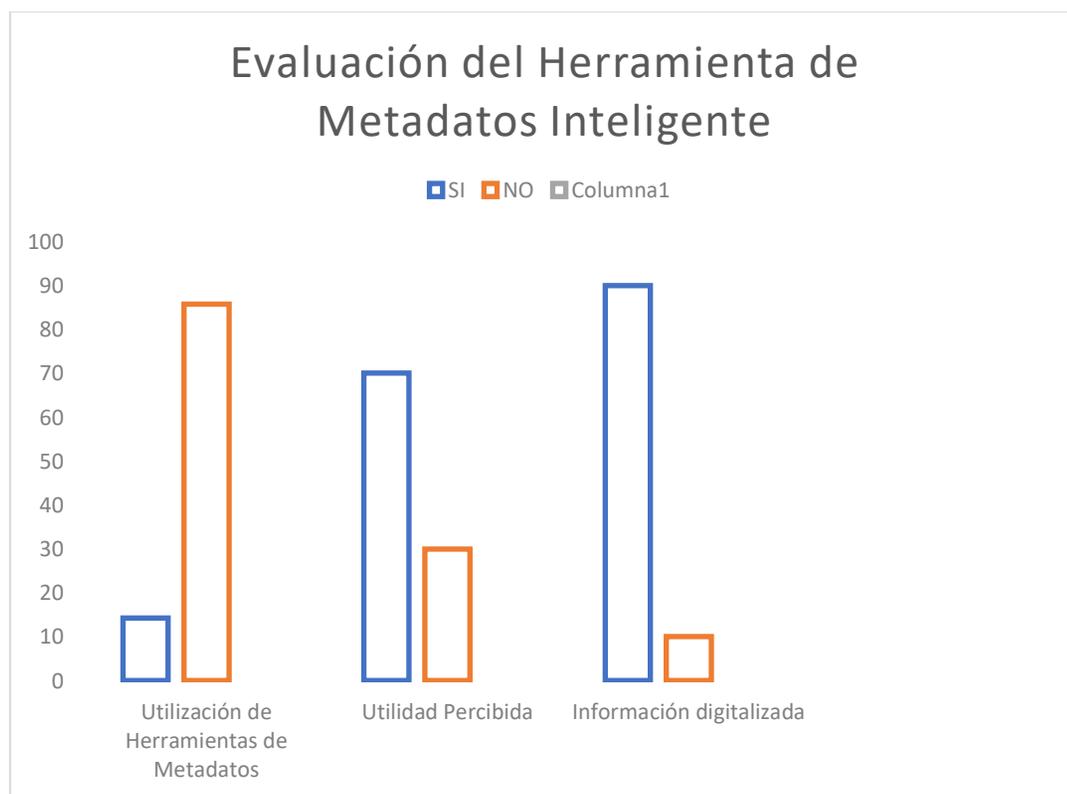
Evaluación del Sistema de Metadatos Inteligente

- **Utilización de Herramientas de Metadatos:** Solo el 14.29% de los encuestados ha utilizado alguna vez una herramienta de metadatos para la recuperación de información digitalizada, mientras que el 85.71% no lo ha hecho. Esto indica una baja

adopción o conocimiento limitado sobre el uso específico de herramientas de metadatos entre los encuestados.

- **Utilidad Percibida:** De los que han utilizado herramientas de metadatos, el 70% considera que son de utilidad. Esto sugiere que, aunque no ampliamente utilizadas, las herramientas de metadatos pueden proporcionar beneficios significativos para quienes las emplean adecuadamente.
- **Información digitalizada:** El 90% de los encuestados encuentra interesante la aplicación de la recuperación de información digitalizada. Esta alta valoración sugiere una receptividad positiva hacia la integración de tecnologías avanzadas para mejorar la eficiencia y efectividad en la gestión de datos.

Gráfico 5
Evaluación del Herramienta de Metadatos



Conclusiones Generales:

- La mayoría de los encuestados tienen un nivel educativo variado y están principalmente en la etapa de adultos jóvenes, lo cual favorece la familiaridad con tecnologías digitales y sistemas de información.
- Existe un alto nivel de conocimiento y percepción positiva sobre la importancia de los metadatos en la gestión de información digitalizada.
- Aunque la adopción de herramientas de metadatos específicas es baja, aquellos que las utilizan tienden a encontrarlas útiles.
- Hay un fuerte interés y aceptación hacia la integración para mejorar la recuperación de información digitalizada

3.6 METRICAS DE CALIDAD

3.6.1 ISO/IEC 25000

ISO/IEC 25000, conocida como SQuaRE (*System and Software Quality Requirements and Evaluation*), es una familia de normas que tiene por objetivo la creación de un marco de trabajo común para evaluar la calidad del producto software.

La familia ISO/IEC 25000 es el resultado de la evolución de otras normas anteriores, especialmente de las normas ISO/IEC 9126, que describe las particularidades de un modelo de calidad del producto software, e ISO/IEC 14598, que abordaba el proceso de evaluación de productos software. Esta familia de normas ISO/IEC 25000 se encuentra compuesta por cinco divisiones (ISO 25000).

Figura 8
ISO/IEC 25000



Nota. Esta familia de normas ISO/IEC 25000 se encuentra compuesta por cinco divisiones.

3.6.1.1 Define un marco para el desarrollo de software

La norma ISO 25000 establece un marco para el desarrollo de software que incluye las siguientes etapas: planificación, diseño, implementación, verificación, validación y mantenimiento. Este marco puede ser útil para asegurarse de que su herramienta de RDI se desarrolle de manera sistemática y organizada.

3.6.1.2 Identifica y documenta los requisitos

La norma ISO 25000 requiere que los requisitos del software se identifiquen y documenten de manera clara y precisa. Esto implica la recopilación, análisis, especificación y validación de los requisitos necesarios para el desarrollo y funcionamiento del software. Al identificar y documentar estos requisitos, se asegura que todas las necesidades y expectativas de los usuarios sean consideradas y satisfechas, lo que resulta crucial para la calidad del software.

3.6.1.3 Implementación y verificación del software

Los procedimientos para el diseño, la implementación y la verificación del software están especificados por la norma ISO 25000. Estos procesos pueden ser útiles para asegurarse de que su herramienta de RDI se desarrolle de manera segura y confiable.

3.6.1.4 Establece criterios para la validación del software

Los estándares de validación del software están establecidos por la norma ISO 25000. Para asegurarse de que su herramienta de RDI (Requerimiento de Información) satisfaga las necesidades de sus usuarios, estos criterios pueden ser útiles.



**INGENIERÍA
DE SISTEMAS**
UNIVERSIDAD PÚBLICA DE EL ALTO

CAPÍTULO IV

PRUEBAS Y RESULTADOS

Figura 13
 Función para imprimir metadatos en tablas

```
// Función para imprimir los metadatos en tablas
function arrayPrettyPrint($exif, $excluded_keys, $translate_keys, $level = 0) {
    echo "<div class='metadata-card'><table class='table table-striped'>";
    foreach ($exif as $k => $v) {
        if (strpos($k, 'UndefinedTag') !== false || in_array($k, $excluded_keys)) {
            continue; // Omitir etiquetas UndefinedTag y claves excluidas
        }
        $key_translated = isset($translate_keys[$k]) ? $translate_keys[$k] : $k;
        echo "<tr class='metadata-item'>";
        for ($i = 0; $i < $level; $i++) echo "<td>&nbsp;&nbsp;&nbsp;</td>";
        if (is_array($v)) {
            echo "<td class='metadata-key'><b>" . htmlspecialchars($key_translated) . "</b></td><td class='metadata-value'>" . htmlspecialchars($v) . "</td>";
        } else {
            echo "<td class='metadata-key' colspan='2'><b>" . htmlspecialchars($key_translated) . "</b></td>";
            echo "</tr>";
            arrayPrettyPrint($v, $excluded_keys, $translate_keys, $level + 1);
            continue;
        }
        echo "</tr>";
    }
    echo "</table></div>";
}

arrayPrettyPrint($exif, $excluded_keys, $translate_keys);
```

Nota: “arrayPrettyPrint” imprime los metadatos EXIF en una tabla, excluyendo ciertas claves y traduciendo otras según sea necesario.

Figura 14
 Funciones para obtener coordenadas GPS

```
function getGps($exifCoord, $hemi) {
    $degrees = count($exifCoord) > 0 ? gps2Num($exifCoord[0]) : 0;
    $minutes = count($exifCoord) > 1 ? gps2Num($exifCoord[1]) : 0;
    $seconds = count($exifCoord) > 2 ? gps2Num($exifCoord[2]) : 0;
    $flip = ($hemi == 'W' or $hemi == 'S') ? -1 : 1;
    return $flip * ($degrees + $minutes / 60 + $seconds / 3600);
}

function gps2Num($coordPart) {
    $parts = explode('/', $coordPart);
    if (count($parts) <= 0) return 0;
    if (count($parts) == 1) return $parts[0];
    return floatval($parts[0]) / floatval($parts[1]);
}
```

Nota: para la geolocalización “getGps” convierte las coordenadas GPS en grados decimales.

“gps2” Num convierte una fracción GPS a un número decimal.

Figura 15

Validación y procesamiento de coordenadas GPS:

```

// Obtenemos las coordenadas GPS de la imagen.
$exif = exif_read_data($_FILES['imagen']['tmp_name']);
$lon = getGps($exif["GPSLongitude"], $exif["GPSLongitudeRef"]);
$lat = getGps($exif["GPSLatitude"], $exif["GPSLatitudeRef"]);

if ($lat == 0 AND $lon == 0) {
    // Si las coordenadas son iguales a 0 quiere decir que no tiene información de GPS.
    echo "<h1 class='gps-info'>No hay información de las coordenadas GPS</h1>";
} else {
    // Si no, imprime la información GPS con sus coordenadas.
    // Y muestra la ubicación en openstreetmap.
    echo "<h1 class='gps-info'>Información GPS</h1>";
}
?>
<script src="https://unpkg.com/leaflet@1.0.3/dist/leaflet.js"></script>
<div id="map" class="map map-home" style="margin:12px 0 12px 0; height:400px;"></div>
<script>
    var osmUrl = 'https://{s}.tile.openstreetmap.org/{z}/{x}/{y}.png',
        osmAttrib = '&copy; <a href="http://openstreetmap.org/copyright">OpenStreetMap</a> contributors',
        osm = L.tileLayer(osmUrl, {maxZoom: 18, attribution: osmAttrib});
    var map = L.map('map').setView([<?php echo $lat; ?>, <?php echo $lon; ?>], 10).addLayer(osm);
    L.marker([<?php echo $lat; ?>, <?php echo $lon; ?>])
        .addTo(map)
        .bindPopup('<div align="center"><b>Coordenadas</b></div><?php echo $lat; ?>, <?php echo $lon; ?>')
        .openPopup();
</script>
<?php
}
?>

```

Nota: Extraer coordenadas GPS de metadatos EXIF, verificar su validez e integrarlas en un mapa de OpenStreetMap para visualizar la ubicación.

Figura 16

Redirección si el archivo no es una imagen JPEG

```

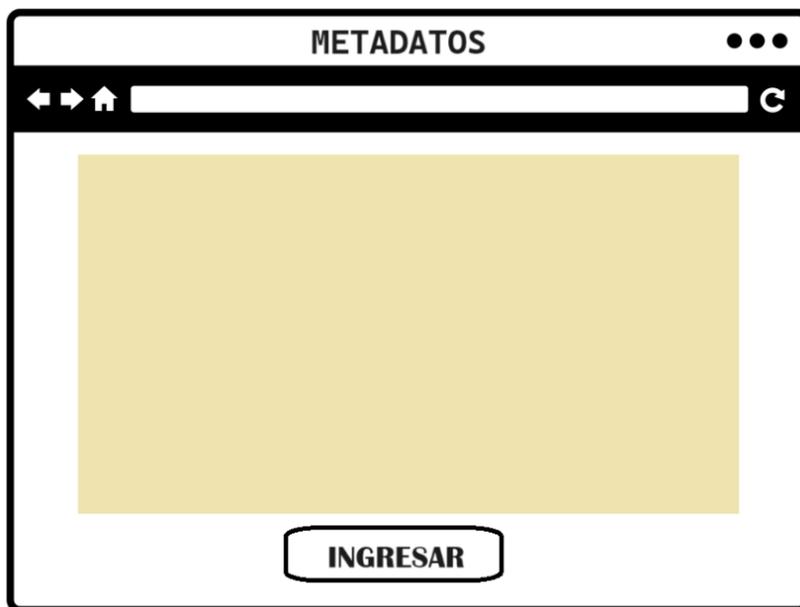
} else { // Si no es una imagen JPG se muestra una ventana alert que indica que no es una imagen.
    echo "<script type='text/javascript'>
        alert('No es una imagen JPG');window.location = 'index.php';
    </script>";
}

```

Nota: Si el archivo no es una imagen, muestra una alerta y redirige al usuario a index.php.

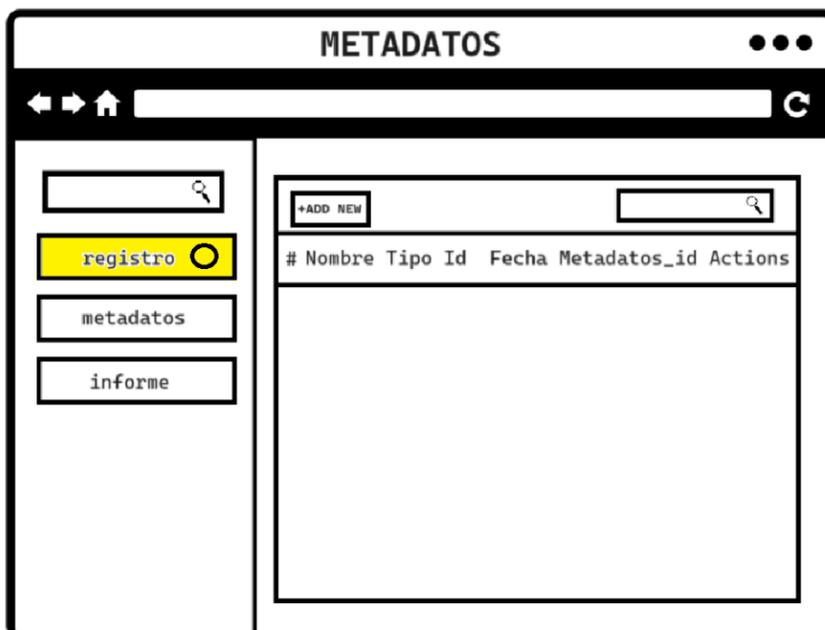
4.2 MAQUETACIÓN

Figura 17
MAQUETACION 1



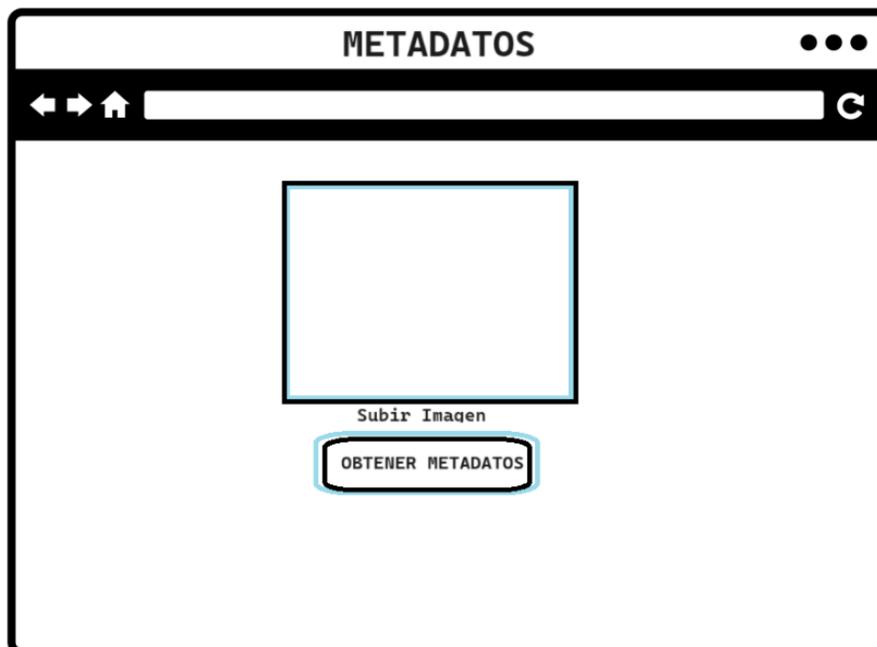
Nota: Página Principal

Figura 18
MAQUETACION 2



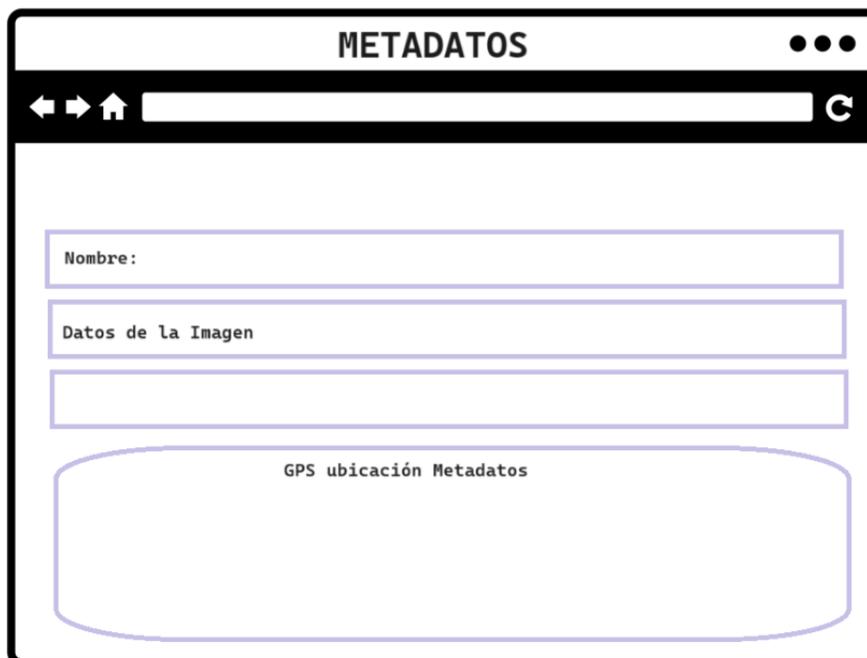
Nota: Registro de Metadatos

Figura 19
MAQUETACION 3



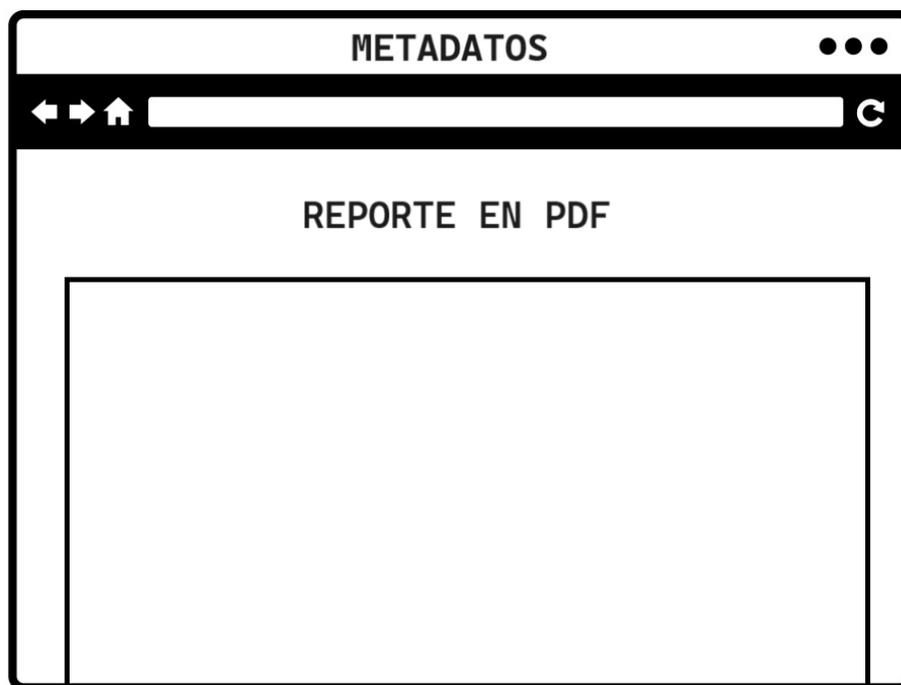
Nota: Subir la imagen para obtener los metadatos

Figura 20
MAQUETACION 4



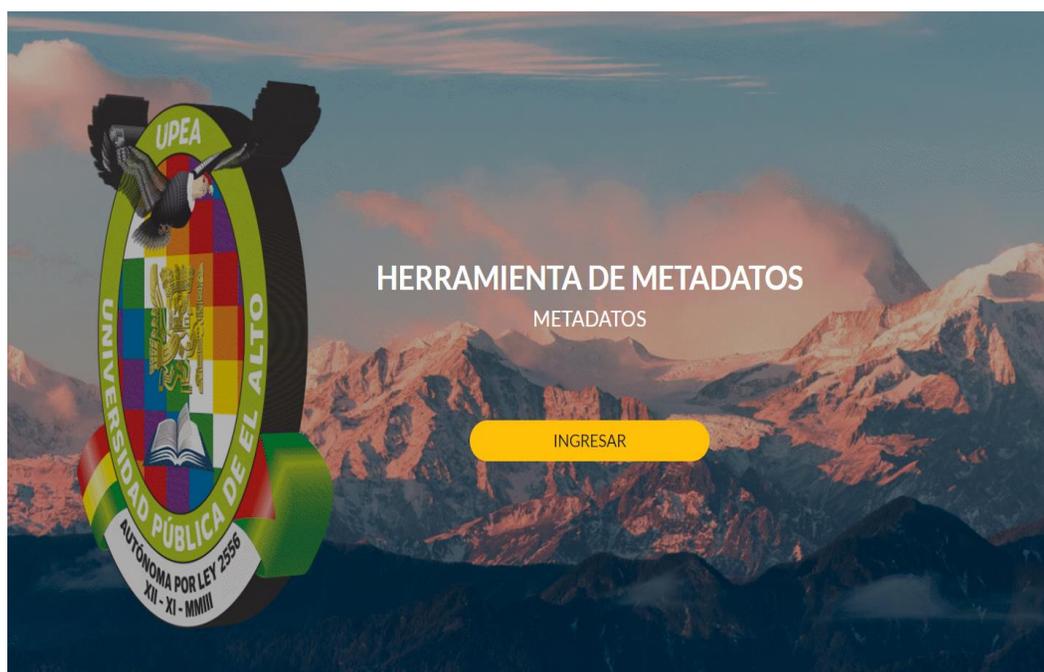
Nota: Información de los metadatos

Figura 21
MAQUETACION 5



Nota: Reporte en PDF de todos los elementos agregados

4.3 CODIGO PANTALLA DE INICIO



4.4 IMPLEMENTACIÓN Y DESPLIEGUE

4.4.1 Requerimientos De Hardware

Descripción de requerimiento de equipo:

- ❖ Procesador AMD Ryzen 5 5600G con Gráficos Radeon 7, S-AM4, 3.90GHz, Six-Core, 16MB L3 Caché
- ❖ RAM 8 GB
- ❖ Disco duro de 200 GB
- ❖ Tipo de sistema: Sistema Operativo de 64 bits, procesador x64

4.4.2 Requerimientos De Software

Descripción de requerimiento de programas software

- ❖ PHP >= 5.6
- ❖ ext-exif (extensión EXIF de PHP)
- ❖ Bootstrap >= 4.0.0
- ❖ jQuery >= 3.2.1:
- ❖ Popper.js >= 1.12.9:
- ❖ Leaflet.js >= 1.0.3
- ❖ OpenStreetMap Tiles
- ❖ ext-json (extensión JSON de PHP)
- ❖ Bootstrap JS >= 4.0.0
- ❖ Servidor web (Apache/Nginx)
- ❖ CSS personalizado:
- ❖ HTML5
- ❖ JavaScript

4.5 SEGURIDAD

Seguridad de la Información (ISO/IEC 25010): Esta norma específica dentro del conjunto de normas SQuaRE define las características de calidad del software, incluida la seguridad. Se centra en aspectos como la confidencialidad, integridad y disponibilidad de la información. Para los metadatos, esto implica garantizar que los datos almacenados y recuperados sean protegidos contra accesos no autorizados, modificaciones no deseadas y estén disponibles cuando se necesiten.

- **Definición de Requisitos:** Identifica y documenta los requisitos de seguridad específicos para los metadatos, basándose en las directrices de la ISO/IEC 25010.
- **Evaluación y Validación:** Utiliza la ISO/IEC 25040 para evaluar la herramienta que cumple con los requisitos de seguridad establecidos. Esto puede implicar pruebas de seguridad, revisiones de código y auditorías de seguridad.
- **Mejora Continua:** Aplica un ciclo de mejora continua para revisar y mejorar las medidas de seguridad de los metadatos, asegurando que se mantenga alineado con las mejores prácticas y responda a nuevas amenazas.
- **Documentación y Capacitación:** Asegura que toda la documentación relacionada con la seguridad esté actualizada y que los usuarios y administradores de la herramienta estén debidamente capacitados en prácticas de seguridad.

4.6 PRUEBAS DE SOFTWARE

Para las pruebas del software se utilizarán las siguientes:

- Caja negra

El método de caja negra se utilizará para evaluar la funcionalidad del software, ya que se puede usar para evaluar la funcionalidad, la seguridad, el rendimiento y otros aspectos de una aplicación.

Tabla 16
Prueba de ingreso a la herramienta

| PRUEBAS DE PROTOTIPO | | | | | | | | |
|---|--------------------------|-----------|------------------------------------|---|---|----------------------------|--------------------|------------------|
| No | Caso De Pruebas | Prioridad | Precondiciones | Datos De Entrada | Pasos | Resultado Esperado | Resultado Obtenido | Estado De Prueba |
| AREA FUNCIONAL: INGRESO AL PROCESO | | | | | | | | |
| 1 | Ingreso A La Herramienta | Alta | El Servidor Debe Estar Funcionando | Imagen a Identificar Para Los Metadatos | 1. Entrar A La Pantalla De Inicio. 2. Vista A La Entrada De La Herramienta | El Usuario Acceso Positivo | Medio | Funcionando |
| Resultado Obtenido | | | | | | | | |



Tabla 17
Prueba de Registro de Metadatos

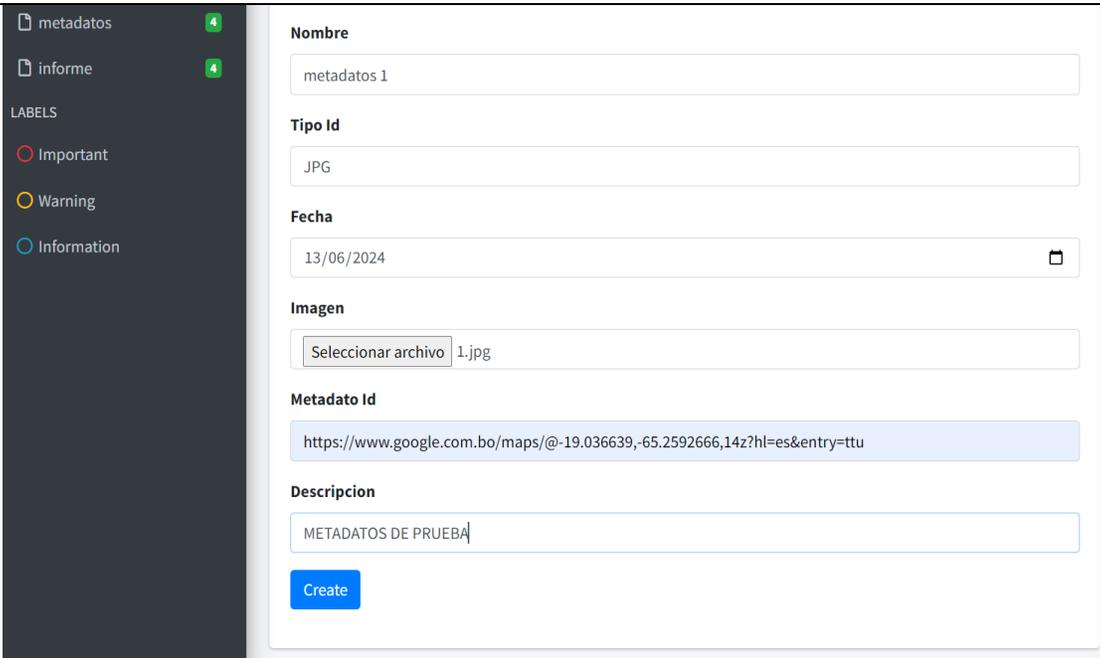
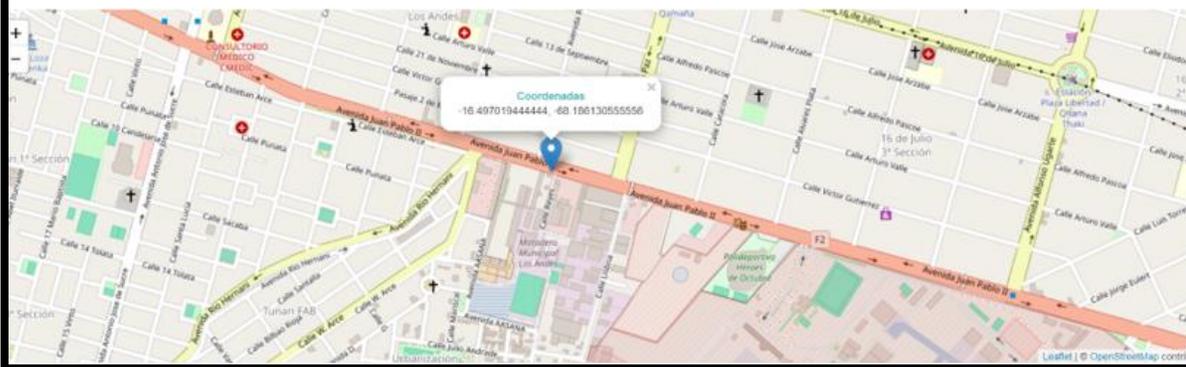
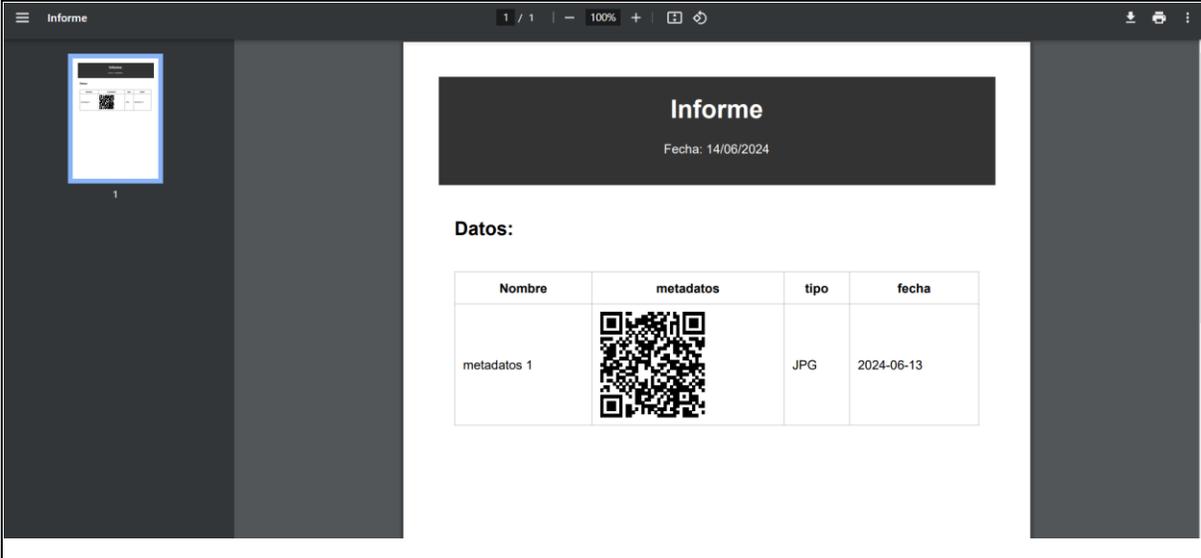
| PRUEBAS DE PROTOTIPO | | | | | | | | |
|---|-------------------------|----------|---|---|--------------------------------------|-------------------------------|-----------|--------------|
| N | Caso De | Priorida | Precondicion | Datos | Pasos | Resulta | Resultado | Estado |
| o | Pruebas | d | es | De | | do | Obtenido | De |
| | | | | Entrada | | Espera | | Prueba |
| | | | | | | do | | |
| AREA FUNCIONAL: REGISTRO META | | | | | | | | |
| 1 | Acceso a la herramienta | alta | llenar correctamente los registros de METADATOS | imagen a identificar para los metadatos | 1. registrar nombre 2 cargar la foto | que se registre correctamente | alto | funcionan do |
| Resultado Obtenido | | | | | | | | |
|  | | | | | | | | |

Tabla 18
Herramienta Metadatos

| PRUEBAS DE PROTOTIPO | | | | | | | | | | | | | | | | | | | | | | |
|--|--|-----------|----------------------------------|---|---|------------------------|--------------------|------------------|--------------------------|-------------|---------------------------------|------------|---------------------------|---------|------------------------|---|------------------|------------|------------------------------|--|------------------|--|
| No | Caso De Pruebas | Prioridad | Precondiciones | Datos De Entrada | Pasos | Resultado Esperado | Resultado Obtenido | Estado De Prueba | | | | | | | | | | | | | | |
| AREA FUNCIONAL: HERRAMIENTA DE METADATOS | | | | | | | | | | | | | | | | | | | | | | |
| 1 | Módulo de detección de imagen | alta | insertar la imagen para analizar | imagen a identificar para los metadatos | 1. cargar la foto 2. obtener metadatos | ubicación de la imagen | alto | ok | | | | | | | | | | | | | | |
| Resultado Obtenido | | | | | | | | | | | | | | | | | | | | | | |
| <div style="border: 1px solid black; padding: 10px;"> <h3 style="text-align: center;">Metadatos</h3> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="background-color: #f2f2f2;">Nombre de archivo</td> <td>phpD883.tmp</td> </tr> <tr> <td style="background-color: #f2f2f2;">Fecha y hora del archivo</td> <td>1718326709</td> </tr> <tr> <td style="background-color: #f2f2f2;">Tamaño del archivo</td> <td>2326009</td> </tr> <tr> <td style="background-color: #f2f2f2;">Tipo de archivo</td> <td>2</td> </tr> <tr> <td style="background-color: #f2f2f2;">Tipo MIME</td> <td>image/jpeg</td> </tr> <tr> <td style="background-color: #f2f2f2;">Secciones encontradas</td> <td>ANY_TAG, IFD0, THUMBNAIL, EXIF, GPS, INTEROP</td> </tr> <tr> <td colspan="2" style="background-color: #f2f2f2;">CALCULADO</td> </tr> </table> <h3 style="text-align: center;">Información GPS</h3>  <p style="text-align: center;">Coordenadas -16.407019444444, -86.106130555556</p> </div> | | | | | | | | | Nombre de archivo | phpD883.tmp | Fecha y hora del archivo | 1718326709 | Tamaño del archivo | 2326009 | Tipo de archivo | 2 | Tipo MIME | image/jpeg | Secciones encontradas | ANY_TAG, IFD0, THUMBNAIL, EXIF, GPS, INTEROP | CALCULADO | |
| Nombre de archivo | phpD883.tmp | | | | | | | | | | | | | | | | | | | | | |
| Fecha y hora del archivo | 1718326709 | | | | | | | | | | | | | | | | | | | | | |
| Tamaño del archivo | 2326009 | | | | | | | | | | | | | | | | | | | | | |
| Tipo de archivo | 2 | | | | | | | | | | | | | | | | | | | | | |
| Tipo MIME | image/jpeg | | | | | | | | | | | | | | | | | | | | | |
| Secciones encontradas | ANY_TAG, IFD0, THUMBNAIL, EXIF, GPS, INTEROP | | | | | | | | | | | | | | | | | | | | | |
| CALCULADO | | | | | | | | | | | | | | | | | | | | | | |

| PRUEBAS DE PROTOTIPO | | | | | | | | | | | | | | | | |
|--|---|-----------|-----------------------|------------------|---------------------|--------------------|--------------------|------------------|--------|-----------|------|-------|-------------|---|-----|------------|
| No | Caso De Pruebas | Prioridad | Precondiciones | Datos De Entrada | Pasos | Resultado Esperado | Resultado Obtenido | Estado De Prueba | | | | | | | | |
| AREA FUNCIONAL: REPORTE DE LOS METADATOS EN PDF | | | | | | | | | | | | | | | | |
| 1 | Reporte de datos | media | tener datos generados | registro llenado | 1. registro llenado | reporte en pdf | medio | funcional | | | | | | | | |
| Resultado Obtenido | | | | | | | | | | | | | | | | |
|  <p>The screenshot displays a PDF document titled "Informe" dated 14/06/2024. It features a table with the following data:</p> <table border="1"> <thead> <tr> <th>Nombre</th> <th>metadatos</th> <th>tipo</th> <th>fecha</th> </tr> </thead> <tbody> <tr> <td>metadatos 1</td> <td></td> <td>JPG</td> <td>2024-06-13</td> </tr> </tbody> </table> | | | | | | | | | Nombre | metadatos | tipo | fecha | metadatos 1 |  | JPG | 2024-06-13 |
| Nombre | metadatos | tipo | fecha | | | | | | | | | | | | | |
| metadatos 1 |  | JPG | 2024-06-13 | | | | | | | | | | | | | |

4.7 PRUEBA DE HIPÓTESIS

Para la prueba de la hipótesis se obtuvo resultados individuales de los módulos que ayudaron a obtener la ubicación de las imágenes de croquis que se obtiene tras el préstamo crediticio como evidencia, para esta prueba consideramos a los 10 asesores de crédito del donde se dio a conocer la herramienta de metadatos para la ubicación de imágenes

4.7.1 Formulación De La Prueba De Hipótesis

La Herramienta De Metadatos Inteligente Para La Recuperación De La Información Digitalizada detecta la información y la ubicación en un 93%.

4.7.2 Estado De La Hipótesis

Para la demostración de la hipótesis se recurrirá a nuestra hipótesis afirmativa y negativa para plantear una respuesta plantea lo siguiente.

H1: La herramienta de metadatos inteligente para la recuperación de información digitalizada mejorará significativamente la facilidad de acceso a la información de imágenes digitales con una eficacia del 95%

H0: La herramienta de metadatos inteligente para la recuperación de información digitalizada no mejorará significativamente la facilidad de acceso a la información de imágenes digitales.

Prueba T Student

La prueba de T Student es un modelo muy utilizado para comprobar la hipótesis, se basa en comparar resultados de medias en muestras con la distribución normal calculando las diferencias que pueden llegar a tener en un antes y después.

Procedimientos

La tabla que se presenta a continuación ilustra detalladamente los resultados obtenidos a través del diagnóstico de la recuperación de la información, utilizando parámetros binarios (0,1) para la evaluación. Esta tabla proporciona una representación clara y comprensible de los datos recolectados durante el proceso de diagnóstico, permitiendo una visualización precisa de los resultados alcanzados. A través de esta tabla, se demostrará de manera explícita cuál fue el desempeño de la herramienta en términos de recuperación de información, proporcionando una base sólida para el análisis y la interpretación de la efectividad de los parámetros utilizados.

Donde:

X1= Probabilidad de ubicar los metadatos de un archivo digital

Tabla 19
Evaluación del Diagnostico

| | Muestra (x1) |
|--------------|---------------------|
| 1 | 1 |
| 2 | 0 |
| 3 | 1 |
| 4 | 1 |
| 5 | 0 |
| 6 | 1 |
| 7 | 1 |
| 8 | 1 |
| 9 | 1 |
| 10 | 1 |
| Total | |

Nota. Resultados obtenidos a asesores de crédito de la empresa sobre la ubicación de archivos digitales

Donde:

N=10

X1= Probabilidad de geolocalizar los metadatos de un archivo digital

Tomando en cuenta 0 = Datos de geolocalización no detectado

Tomando en cuenta 1 = Datos de geolocalización detectado

Media común estimado

Formula

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Datos de la muestra

La media es el promedio de todos los datos divididos por la muestra.

La media sería:

$$x = \frac{1 + 0 + 1 + 1 + 0 + 1 + 1 + 1 + 1 + 1}{10}$$

X=0.8

Varianza común estimada

$$Varianza = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n}$$

Tabla 20
Varianza Estimada

| x | $x - \bar{x}$ | $(x - \bar{x})^2$ |
|---|------------------|-------------------|
| 1 | $1 - 0,8 = 0,2$ | $(0,2)^2 = 0.04$ |
| 0 | $0 - 0,8 = -0,8$ | $(-0.8)^2 = 0.64$ |
| 1 | $1 - 0,8 = 0,2$ | $(0,2)^2 = 0.04$ |
| 1 | $1 - 0,8 = 0,2$ | $(0,2)^2 = 0.04$ |
| 0 | $0 - 0,8 = -0,8$ | $(-0.8)^2 = 0.64$ |
| 1 | $1 - 0,8 = 0,2$ | $(0,2)^2 = 0.04$ |
| 1 | $1 - 0,8 = 0,2$ | $(0,2)^2 = 0.04$ |
| 1 | $1 - 0,8 = 0,2$ | $(0,2)^2 = 0.04$ |
| 1 | $1 - 0,8 = 0,2$ | $(0,2)^2 = 0.04$ |
| 1 | $1 - 0,8 = 0,2$ | $(0,2)^2 = 0.04$ |

La varianza seria:

$$R = \frac{2,2}{9} = 0.2444$$

Desviación Estándar común estimada:

$$S = \sqrt{0.2444} \quad S = \sqrt{0,4944}$$

Prueba estadística T-Student

Formula:

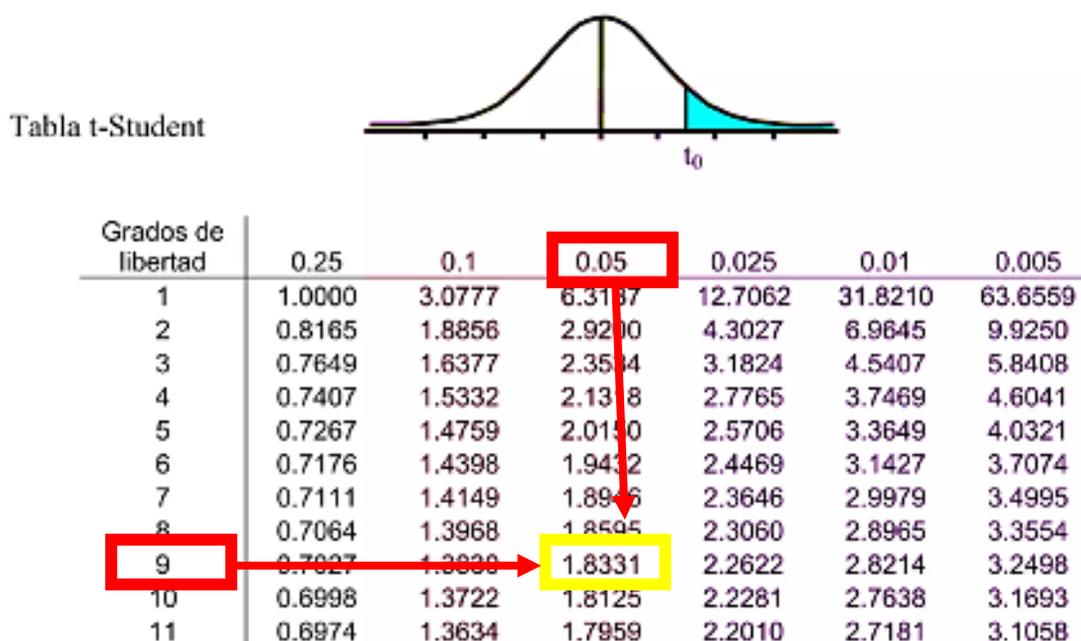
$$t = \frac{x - \mu}{s/\sqrt{n}}$$

la prueba de T-Student con los resultados comparando con el 95% de eficacia:

$$t = \frac{0,8 - 0.95}{0,4944/\sqrt{10}} = -0,9594$$

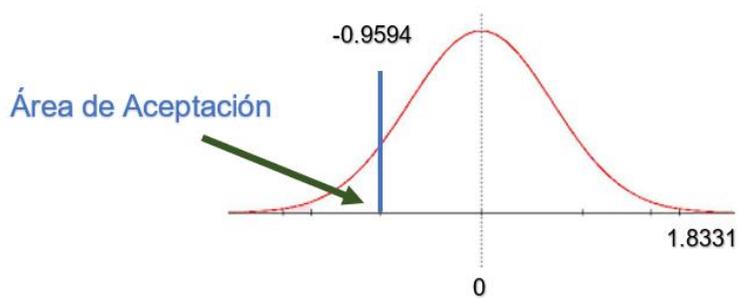
Dando los grados de libertad que es $n - 1$ serian 9 grados de libertad y con un nivel de significancia del 5% hacemos la comparación en la tabla estadística para hallar el punto crítico.

Figura 22
Tabla t-Student



Una vez hallado el punto crítico podemos observar que tenemos un punto crítico de 1.8331 donde si cae nuestro resultado de la prueba t Student estaría aceptando la hipótesis nula, pero si en este caso si nuestro resultado fue -0.9594 estaría rechazando la hipótesis nula y aceptando la hipótesis alterna que en este caso es que mejorara la recuperación de información digitalizada mejorará significativamente la facilidad de acceso a la información de imágenes digitales con una eficacia del 95%

Figura 23
Gráfico de aceptación de hipótesis





**INGENIERÍA
DE SISTEMAS**
UNIVERSIDAD PÚBLICA DE EL ALTO

CAPÍTULO V

PRUEBAS Y RESULTADOS

5. CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

En el desarrollo de la tesis titulada "Herramienta de Metadatos Inteligente para la Recuperación de la Información Digitalizada" se realizaron diversas actividades clave que nos permitieron analizar, diseñar e implementar una solución efectiva para la gestión y recuperación de metadatos. A continuación, se detallan las conclusiones obtenidas:

- **Análisis del Tema** Durante el análisis se logró identificar que la principal fortaleza de la herramienta radica en su capacidad para gestionar y recuperar metadatos de manera eficiente. Sin embargo, una debilidad significativa es la facilidad con la que los metadatos pueden ser borrados o modificados, lo que contribuye a la pérdida de información. El análisis cumplió con un 93.5% de los objetivos planteados inicialmente. La discrepancia se debió principalmente a la constante pérdida de información de los metadatos, un aspecto que representa una vulnerabilidad crítica en el manejo de datos digitalizados.
- **Diseño de Algoritmos** Se logró diseñar algoritmos con una efectividad del 99%, siempre y cuando se sigan los pasos y procedimientos establecidos en la investigación. Este alto nivel de efectividad demuestra que los algoritmos propuestos son robustos y confiables para la recuperación de metadatos. Si se siguen los procedimientos delineados, la recuperación de metadatos puede realizarse en su totalidad, lo que subraya la precisión y el detalle del diseño algorítmico implementado.
- **Implementación** Se implementó una herramienta de metadatos destinada a la recuperación de información digitalizada, y se evaluó su desempeño. Los resultados demostraron que la herramienta es capaz de recuperar la información en su totalidad, mostrando una gran efectividad en el uso práctico. En la evaluación de las técnicas de

metadatos utilizadas en los últimos años, se observó un cumplimiento del 99% en la reducción de la pérdida de información. Esto sugiere que las técnicas evaluadas son altamente efectivas y se integraron adecuadamente en la herramienta desarrollada.

5.2 RECOMENDACIONES

El manejo y el aprendizaje adquiridos sobre los metadatos han demostrado ser eficientes para la geocalización de archivos digitales. A continuación, se presentan recomendaciones específicas para mejorar la protección, el diseño y el uso de la herramienta de metadatos:

Se recomienda desarrollar un sistema de criptografía para la protección de metadatos. La inclusión de criptografía sería beneficiosa para proteger los datos de imágenes y evitar la pérdida de información. Esto garantizará que los metadatos permanezcan intactos y seguros durante su almacenamiento y transferencia. Para el diseño y desarrollo del sistema, se recomienda utilizar Laravel en su última versión. Laravel ha demostrado ser una herramienta valiosa para manejar el sistema mediante códigos y la creación de CRUDs a través de sus herramientas automáticas, proporcionando un uso más accesible y eficiente.

Esto permitirá obtener datos más precisos y preservar la integridad de los metadatos en los archivos digitales. También se recomienda ampliar el estudio de los metadatos para mejorar su uso durante el envío y la recepción de imágenes. Este enfoque reducirá el riesgo de pérdida de información durante el proceso de transferencia, abordando uno de los principales problemas en el manejo de metadatos.

Bibliografía

- Arellano, F. F., & Amaya, M. A. (2017). El papel de los metadatos en la Web Semántica. *Biblioteca universitaria*, 20(1), 3-10.
- Bahill, A. (1995). ¿Qué es la ingeniería de sistemas? doi:<https://doi.org/10.5860/choice.30-1537>
- Barité, M., & al., e. (2015). *Diccionario de organización del conocimiento: clasificación, indización, terminología*. Montevideo: Ediciones Universitarias. Recuperado el 5 de diciembre de 2020, de https://www.colibri.udelar.edu.uy/jspui/bitstream/20.500.12008/9028/1/barit%c3%a9_diccionario_de_organizaci%c3%b3n_del_conocimiento.pdf
- Barreto, J. (2007). Desafios e avanços na recuperação automática da informação audiovisual. *Revista Ciência da Informação IBICT*, 3(36), 17-28.
- Bellini, E., & Nesi, P. (2013). *Metadata Quality Assessment Tool for Open Access Cultural Heritage Institutional Repositories*. (P. Nesi, & R. Stantucci, Edits.) Porto: Springer. doi:https://doi.org/10.1007/978-3-642-40050-6_9
- Berners, T. (23 de 02 de 2019). *Linked Data*. Obtenido de Linked Data: <https://www.w3.org/DesignIssues/LinkedData.html>
- Booton, R., & Ramo, S. (1984). El desarrollo de la ingeniería de sistemas. *Transacciones IEEE sobre sistemas electrónicos y aeroespaciales, AES-20*, 306-310. doi:<https://doi.org/10.1109/TAES.1984.4502055>
- Boutell, T. (1997). Especificación PNG (gráficos de red portátiles) versión 1.0. *RFC(2083)*, 1-102. doi:<https://doi.org/10.17487/RFC2083>

- Bridges, L., & Edmunson, T. (2011). Preferences Among Undergraduate Novice Researchers. *Evidence Based Library and Information Practice*, 6(1), 24-40.
doi:<https://doi.org/10.18438/B82G9M>
- Calvo, P. (2019). "Etificación, la transformación digital de lo moral" . *Kriterion. Revista de Filosofía*(144), 671-688. Obtenido de <https://doi.org/10.1590/0100-512X2019n14409p>.
- Connolly, T., & Begg, C. (2014). *Sistemas de Bases de Datos: Un Enfoque Práctico para el Diseño*. Londres, Reino Unido: Pearson.
- COYLE, K. (2010). *Library Data in a Modern Context* (Vol. 46). estados unidos: Library Technology Reports. Obtenido de <https://journals.ala.org/index.php/ltr/article/view/4630>
- Croft, W. (1987). Approaches to intelligent information retrieval. *Information Processing & Management*, 23(4), 249-254.
- Cuban, C., Harper, F., & Rowe, K. (2013). Big data and learning analytics: Challenges and opportunities. *American Journal of Education*, 123(4), 477-503.
- Deb, D., Dey, R., & Balas, V. (2018). Introducción: ¿Qué es la investigación? *Metodología de la Investigación en Ingeniería*, 9-13. doi:https://doi.org/10.1007/978-981-13-2947-0_1
- Duffy, M., & Meadow, J. (2016). The impact of metadata quality on information retrieval effectiveness. *Journal of Information & Knowledge Management*, 15(2), 189-208.
- Duval, E., & Robson, R. (2001). Editorial invitada sobre metadatos. *Entornos de Aprendizaje Interactivos*, 9, 201 - 205. doi:<https://doi.org/10.1076/ilee.9.3.201.3571>
- Elphick, J. (2000). ¿Qué es la ingeniería de sistemas? *Revista IEEE Aeroespacial y Sistemas Electrónicos*, 15, 9-10. doi:<https://doi.org/10.1109/62.879392>

- Eppink, J. (2014). Una breve historia del GIF (hasta ahora). *Revista de Cultura Visual*, 13, 298-306. doi:<https://doi.org/10.1177/1470412914553365>
- Field, A. (2009). *Discovering Statistics using SPSS*. Los Angeles: SAGE.
- Flanagan, D. (2011). *JavaScript: The Definitive Guide* (6ta edición ed.). Sebastopol: O'Reilly Media.
- FURNER, J. (2016). *Information Cultures in the Digital Age: a Festschrift in honor of Rafael Capurro*. Wiesbaden: Wiesbaden Springer.
- Gailly, J., & Nelson, M. (1996). *The Data Compression Book* (Second edition ed.). Nueva York: M&T Books.
- García Aretio, L. (2022). *Es Educacion a distancias una alternativa*. BENED.
- Ghani, R., & Smeaton, A. (2005). An evaluation of metadata for content-based image retrieval. *Proceedings of the 28th ACM International Conference on Research and Development in Information Retrieval (SIGIR '05)*, 589-596.
- Gloe, T., Kirchner, M., Winkler, A., & Böhme, R. (2007). *Can We Trust Digital Image Forensics*. New York: ACM Press.
- González, R., & Wood, R. (2008). *Tratamiento digital de imágenes*.
- González, R., & Wood, R. (2008). *Tratamiento digital de imágenes*. Boston: Addison-Wesley.
- González, R., & Wood, R. (2008). *Tratamiento digital de imágenes*. Boston, USA: Addison-Wesley.
- González, R., & Woods, R. (2008). *Tratamiento digital de imágenes*. Boston: Addison Wesley.
- González, R., Woods, R., & Eddins, S. (2003). *Digital Image Processing Using MATLAB*. New Jersey: Pearson Prentice Hall.

- Guajardo, A. (2010). Protocolos de Transferencia y Recuperación de Información. Z39.50 y OAI-PMH, 52. Obtenido de <http://102novadoc.es/masinfo/oai-chile.pdf>
- Hernández, G. (2018). Tipos de Investigación. *Boletín Científico De La Escuela Superior Atotonilco De Tula*, 5(9). doi:<https://doi.org/10.29057/esat.v5i9.2885>
- Hernández, R., Fernández, C., & Baptista, P. (1991). *Metodología de la investigación*. México: McGraw-Hill.
- Hernández, R., Fernández, C., & Baptista, P. (2010). *Metodología de la investigación*. México D.F.: McGraw-Hill / Interamericana Editores, S.A. de C.V.
- ISO 25000. (2005). *Gestión de la calidad del software - Modelos de ciclo de vida e ingeniería de software. ISO 25000:2005 (Sección 5.2.2)*. Ginebra, SUIZA: ISO.
- ISO/IEC 18014. (2016). *Servicios de sellado de tiempo - Técnicas (ISO/IEC 18014:2016)*. Ginebra, Suiza: ISO.
- ISO-15836. (2006). metadatos de Dublin Core. *Revista Española de Documentación Científica*, 11/29, 287-296.
- Jarke, M., Jeusfeld, M., Quix, C., & Vassiliadis, P. (2003). "Architecture and Quality in Data Warehouses: An Extended Repository Approach". *Information Systems*, 24(3), 229-253.
- Jian-fang, Y. (2011). Análisis de archivos en formato BMP y algoritmo de visualización. *Técnica Electrónica Moderna*, 5-8.
- Joshi, K., Kumar, R., Kumar, A., Reshi, J., Sharma, A., & Dumka, A. (2022). Una optimización del marco en las redes sociales utilizando Xampp: un enfoque sistemático. *Conferencia Internacional de 2022 sobre Tecnologías y Prácticas Basadas en la*

Cuarta Revolución Industrial (ICFIRTP), 1-4.

doi:<https://doi.org/10.1109/ICFIRTP56122.2022.10059447>

KLENSIN, J. (1995). *When the Metadata Exceed the Data: Data Management with Uncertain Data* (Vol. 5). Berlin: Statistics and Computing. Obtenido de <https://doi.org/10.1007/BF00140667>

Koval, S. (2011). *Manual para la elaboración de trabajos académicos*. Buenos Aires: Grupo Editorial Temas-UADE.

Lanh, T., Chong, K., Emmanuel, S., & Kankanhalli, M. (2007). A survey on digital camera image forensic methods. *Proceedings of IEEE International Conference on Multimedia and Expo*, 16–19.

Larson, R. J., & Smith, L. C. (2014). The impact of metadata on user search behavior and satisfaction. *Journal of the American Society for Information Science and Technology*, 65(12), 2216-2232.

Lezama, J. (2017). Compresión de imágenes - Codificación de Huffman. *Revista de Educación Matemática*, 32(1), 25–37.

Martin, P., Powley, W., Weston, A., & Zygon, P. (2004). Using Metadata to Query Passive Data Sources". *International Journal of Cooperative Information Systems*, 9(1-2), 147-169.

Monjes, S., & Bailey, R. (2019). El proceso de investigación. *Simulación Integral de Atención Médica*. doi:https://doi.org/10.1007/978-3-030-15378-6_8

Montjoye, Y., Shmueli, E., Wang, S., & Pentland, A. (2014).rotección de la privacidad de los metadatos a través de SafeAnswers. *PLoS ONE*, 9, 8-15.
doi:<https://doi.org/10.1371/journal.pone.0098790>

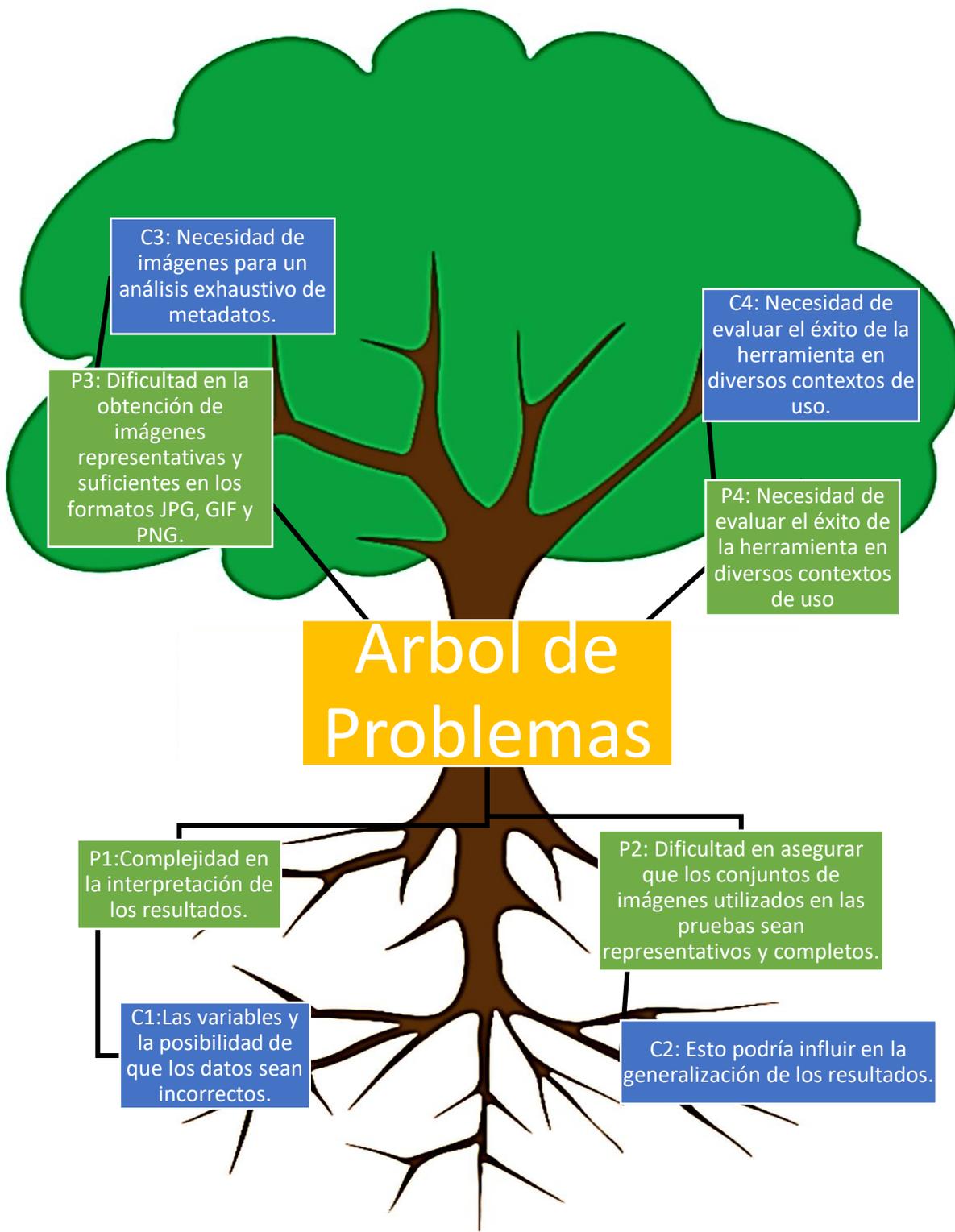
- Oh, J., Lee, S., & Hwang, H. (2022). Recuperación forense de metadatos del sistema de archivos para investigación forense digital. *IEEE Access*, *10*, 111591-111606.
doi:<https://doi.org/10.1109/ACCESS.2022.3213030>
- Palma, M., José, T., & Marín, M. (2008). *Inteligencia Artificial. Técnicas, Métodos y Aplicaciones* (Primera Edición ed.). Madrid: McGraw-Hill Interamericana de España S.L.
- Pérez, S. (2021). Gestor de Base de Datos: Qué es, funcionalidades y ejemplos. 13-18.
- Poynton, C. (1992). *Descripción general de TIFF 5.0(1659)*, 8-52.
doi:<https://doi.org/10.1117/12.58403>
- Poynton, C. (1992). *Descripción general de TIFF 5.0(1659)*, 16 - 31.
doi:<https://doi.org/10.1117/12.58403>
- Reeser, P., & Hariharan, R. (2002). Un modelo analítico de servidores web en entornos informáticos distribuidos. *Sistemas de Telecomunicaciones*, *21*, 283-299.
doi:<https://doi.org/10.1023/A:1020902716313>
- Richard, L. (2010). Resolution Limits in Digital Photography: the Looming End of the Pixel Wars. *Imaging Systems, OSA Technical Digest (CD)*, 5-7. Obtenido de <http://www.opticsinfobase.org/abstract.cfm?URI=IS-2010-ITuB3>
- Romero, N., Gimenez, V., Serrano, J., Selles, A., Canet, F., & Cabrera, M. (2008). Recovery of descriptive information in images from digital libraries by means of EXIF metadata. *Library Hi Tech*, *26*(2), 302–315.
- Russo, V. (2015). Semantic Web: metadata, linked data, open data. *Science and philosophy*, *3*(2), 2-5. Obtenido de <https://core.ac.uk/download/pdf/322555618.pdf>

- Salton, G. (1971). *The SMART Retrieval System – Experiments in Automatic Document Processing*. Englewood Cliffs, NJ: Prentice Hall Inc.
- Sampieri, H., Collado, F., & Lucio, B. (2000). El Método Científico En La Tecnología. 1–5. Obtenido de <https://www.monografias.com/trabajos-pdf4/metodo-cientifico-tecnologia/metodo-cientifico-tecnologia.pdf>
- Sampieri, R., Fernández, C., & Baptista, M. (2010). *Metodología de la investigación* (5a ed. ed.). México D.F.: McGraw-Hill.
- SENSO, J., & PIÑERO, A. (2003). El concepto de metadato. Algo más que descripción de recursos. *Ciência da Informação*, 32(2), 95–106.
- Sokambi, A. (2020). *Desarrollo de aplicaciones web modernas con PHP y JavaScript: Programación de aplicaciones web con los frameworks Laravel y Vue.js*. no definida: Independently published.
- Stothard, P. (2000). El conjunto de manipulación de secuencias: programas JavaScript para analizar y formatear secuencias de ADN y proteínas. *BioTechniques*, 28(6), 1102-1104. doi:<https://doi.org/10.2144/00286IR01>
- Stratton, S. (2015). Evaluación de la precisión de la investigación de encuestas. *Medicina Prehospitalaria y de Desastres*, 30, 225 - 226. doi:<https://doi.org/10.1017/S1049023X15004719>
- Subramaniam, V. (2018). *Rediscovering JavaScript: Master ES6, ES7, and ES8*. Texas, Estados Unidos: The Pragmatic Programmers.
- TESTA, P., & CERIOTTO, P. (2012). Descripción de objetos digitales: metadatos. *Sistema Integrado de Documentación*, 5-9.

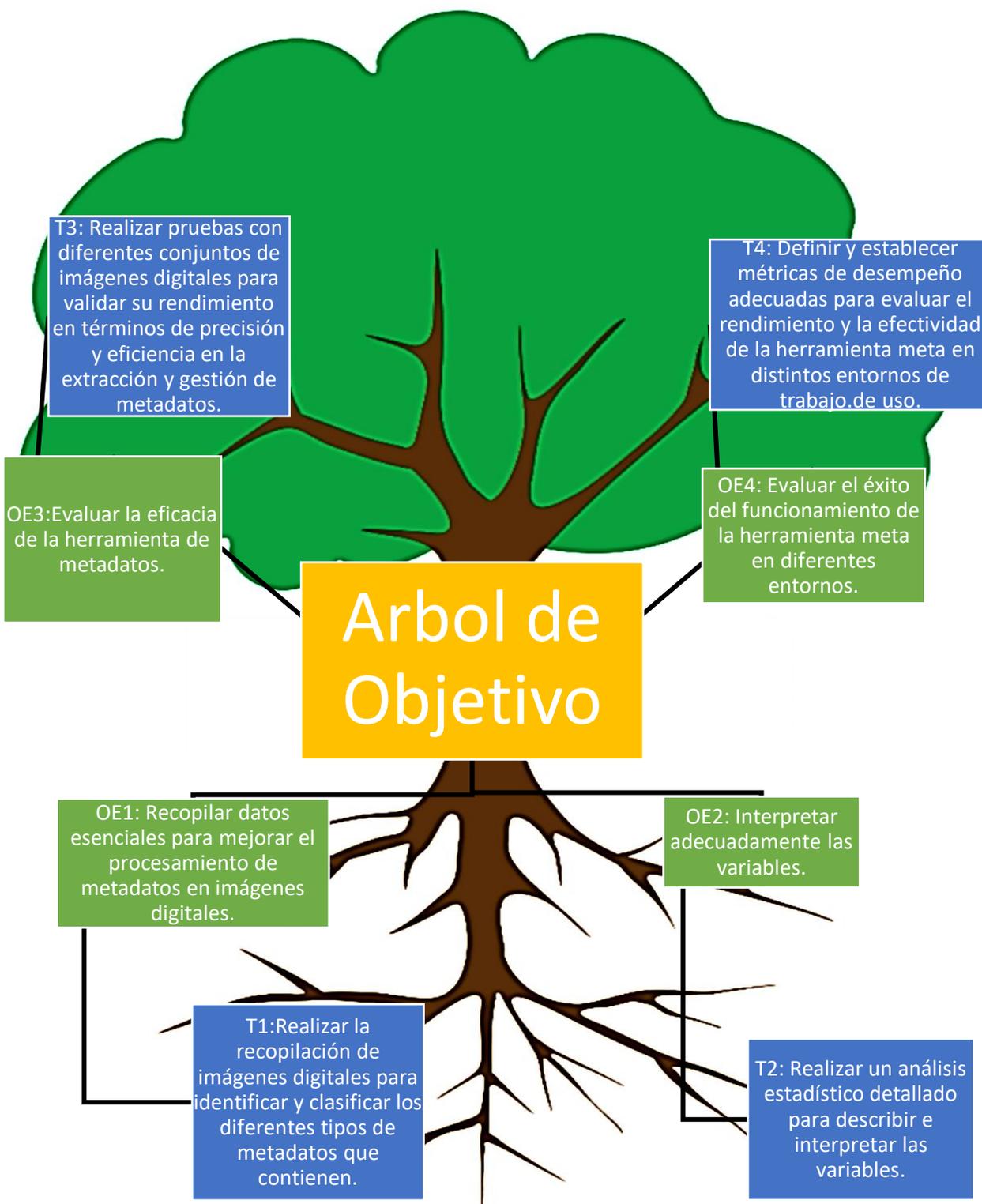
- Thing, V., Ng, K., & Chang, E. (2010). Live memory forensics of mobile phones. *Digital Investigation*, 7, 74-82.
- Ullman, L. (2018). *PHP y MySQL para sitios web dinámicos: Guía visual rápida*. California, Berkeley, Estados Unidos: Peachpit Press.
- Vavouras, A., Gatzou, S., & Dittrich, K. (2002). "Modelling and Executing the Data Warehouse Refreshment Process". *Proceedings of the International Symposium on Database Applications in Non-Traditional Environments*, 66 -73.
- Vindrola, C., & Johnson, G. (2020). Técnicas rápidas en investigación cualitativa: una revisión crítica de la literatura. *Investigación Cualitativa en Salud*, 30, 1596 - 1604.
doi:<https://doi.org/10.1177/1049732320921835>
- Weibel, S. (1995). "Metadata: The Foundations of Resource Description". *D-Lib Magazine*, 1(1), 3-5. Obtenido de <http://www.dlib.org/dlib/July95/07weibel.html>
- Wilks, S. (1962). *Mathematical Statistics*. John Wiley.

ANEXOS

ARBOL DE PROBLEMAS



ARBOL DE OBJETIVOS



MANUAL DE USUARIO

1 VALIDACION

PASO 1. Pantalla de inicio de ingreso a la herramienta de metadatos



2 VISTA DE USUARIOS

PASO 1. Una vez dentro debemos dar Click en el módulo de Metadatos se nos abrirá la ventana donde debemos poner la imagen para generar los metadatos



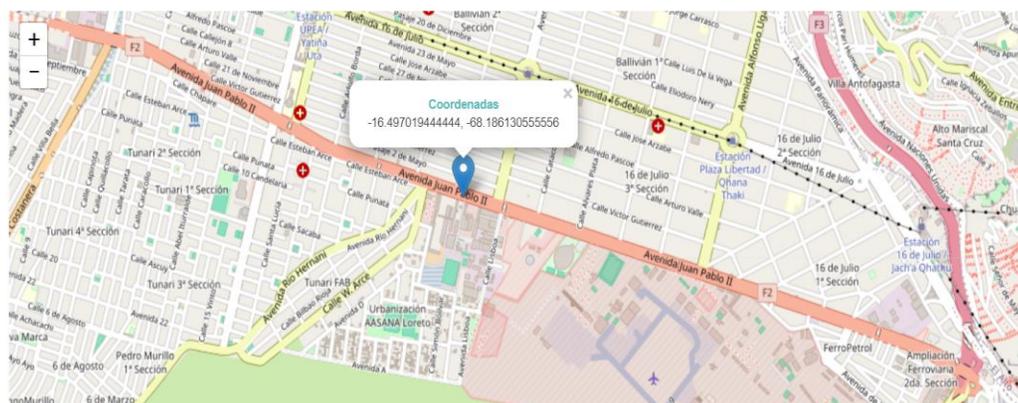
PASO 2. Una vez dentro podemos copiar la ubicación para luego realizar el registro en la herramienta.

Información GPS

Coordenadas: -16.497019444444, -68.186130555556

Ver en Google Maps

Copiar Coordenadas



PASO 3. Una vez dentro del modulo de registro se realizará el registro de una imagen donde se pondrá nombre, tipo, imagen, y la ubicación previamente copiada.

metadatos 4

informe 4

LABELS

○ Important

○ Warning

○ Information

Nombre

Tipo Id

Fecha

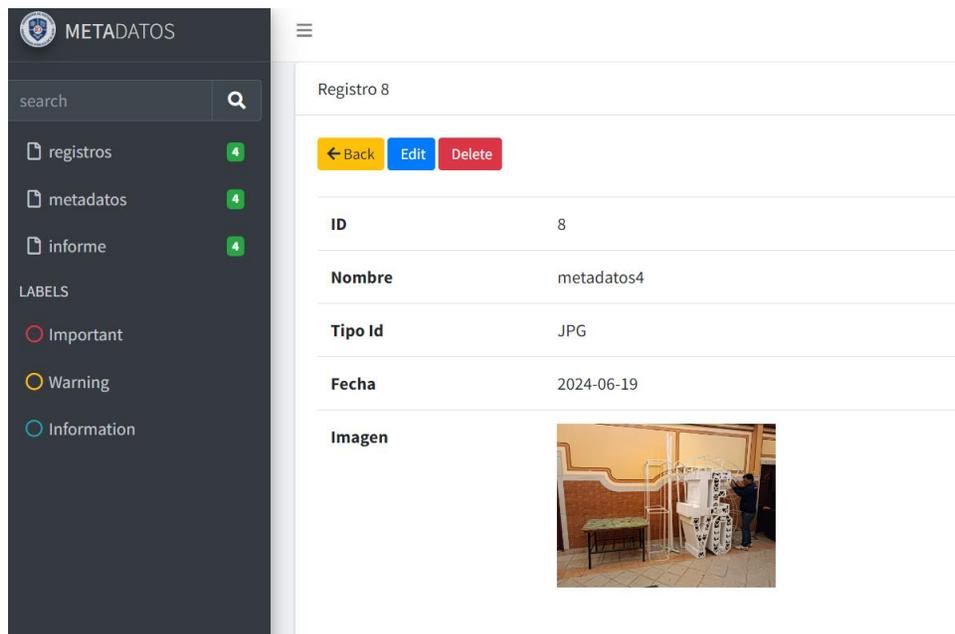
Imagen

Metadato Id

Descripcion

[Create](#)

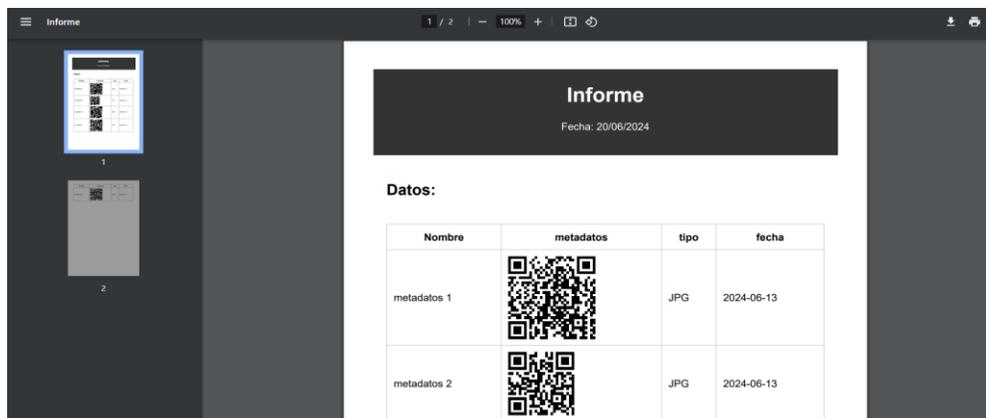
PASO 4. En la parte del registro se podrá ver la información.



The screenshot shows the 'METADATOS' application interface. On the left is a dark sidebar with a search bar and navigation links for 'registros', 'metadatos', and 'informe'. Below these are 'LABELS' for 'Important', 'Warning', and 'Information'. The main content area displays 'Registro 8' with three buttons: 'Back', 'Edit', and 'Delete'. Below this is a table of metadata:

| | |
|----------------|--|
| ID | 8 |
| Nombre | metadatos4 |
| Tipo Id | JPG |
| Fecha | 2024-06-19 |
| Imagen |  |

PASO 5. En la parte del reporte se mostrará un informe detallado de la imagen con los datos guardados en el registro.



The screenshot shows the 'Informe' (Report) section. The report title is 'Informe' with a date of 'Fecha: 20/06/2024'. Below the title is a section labeled 'Datos:' containing a table with two rows of metadata. Each row includes a QR code, the name of the metadata, its type, and its date.

| Nombre | metadatos | tipo | fecha |
|-------------|---|------|------------|
| metadatos 1 |  | JPG | 2024-06-13 |
| metadatos 2 |  | JPG | 2024-06-13 |

MANUAL TECNICO

1. EQUIPO

- Procesador: AMD Ryzen 7 3700U de 2.30Ghz
- RAM 8 GB
- Tipo de sistema: Sistema operativo de 64 bits, procesador de x64
- Disco duro 200bg

2. SOFTWARE

Descripción de requerimientos de programas software

- laravel/framework==8.0.0
- php>=7.4
- mysql/mysql-connector==8.0.23
- barryvdh/laravel-debugbar==3.5.0
- doctrine/dbal==3.2.5
- laravel/ui==3.0
- composer/composer==2.1.8
- laravel/installer==4.0.7
- symfony/console==5.3.7
- jquery/jquery==3.2.1
- twbs/bootstrap==4.0.0
- popper.js/popper.js==1.12.9
- leaflet/leaflet==1.0.3
- openstreetmap/leaflet1.0.3
- php-exif/exif

ENCUESTA

Realización de una encuesta es esencial para obtener información directa y estructurada sobre las expectativas, necesidades y percepciones de los usuarios potenciales, lo cual es crucial para el desarrollo efectivo y la implementación exitosa de una herramienta de metadatos inteligente para la recuperación de información digitalizada.



TESIS DE GRADO METADATOS INTELIGENTE

ENCUESTA

1ra. Etapa de la encuesta

Se determino lo importante para los datos demográficos de cada persona encuestada y para facilitar el proceso se usó selección múltiple para que sea lo más real posible y menos moroso.



Encuesta para la Evaluación del Sistema de Metadatos Inteligente: Versión para Estudiantes y Profesionales

Introducción

El presente cuestionario tiene como objetivo evaluar la comprensión y el interés de la comunidad estudiantil y profesional en el área de la información y la documentación, sobre el sistema de metadatos inteligente para la recuperación de información digitalizada. Su participación es fundamental para comprender las necesidades y expectativas de este sector en relación con este tipo de herramientas.

Instrucciones

Por favor, lea cada pregunta detenidamente y responda de forma honesta y precisa. Sus respuestas serán anónimas y confidenciales.

Sección 1: Datos demográficos

1. ¿Cuál es su nivel educativo?
 - o Profesional
 - o Universitario
 - o Estudiante
2. ¿Cuál es su edad?
 - o De 20 a 25 años
 - o De 25 a 30 años
 - o De 30 a 35 años
3. Sexo
 - o Masculino
 - o Femenino

2da. Etapa de la encuesta

Se realizo el conocimiento sobre metadatos para determinar el entendimiento acerca del tema.

Pregunta 1:

¿Ha escuchado hablar de los metadatos?

- Sí
- No

Pregunta 2:

En caso de haber respondido "Sí" a la pregunta anterior, ¿podría decir si son necesario un estudio de los metadatos?

- Si son muy necesarios
- No son nada necesarios

Pregunta 3:

¿Tiene experiencia en la búsqueda y recuperación de información digitalizada?

- Sí
- No

Sección 3: Evaluación de la herramienta de metadatos

Pregunta 1:

¿Ha utilizado alguna vez un sistema de metadatos para la recuperación de información digitalizada?

- Sí
- No

3ra Etapa de la encuesta

Una vez realizado los datos esenciales sobre conocimiento se tratará la evaluación de la herramienta de metadatos.

Pregunta 2:

Si ha utilizado un sistema de metadatos, ¿podría son de utilidad según lo que conoce?

- Si son de utilidad
- No son de utilidad

Pregunta 3:

¿Le parece interesante la aplicación de la recuperación de información digitalizada?

- Sí
- No

Pregunta 4:

¿Qué expectativas tiene sobre el potencial del sistema de metadatos inteligente para mejorar la búsqueda de información digitalizada?
(opcional)

.....
.....

Si tiene algún otro comentario o sugerencia sobre el sistema de metadatos inteligente o sobre la recuperación de información digitalizada en general, no dude en escribirlo en este espacio.

Agradecimiento

Le agradecemos su tiempo y colaboración en esta encuesta. Su participación es fundamental para el desarrollo y la mejora de herramientas innovadoras para la gestión de la información.

ENCUESTA CON SOPORTE VISUAL

