

2022

PROYECTO DE GRADO PRESENTADO AL DEPARTAMENTO DE INGENIERÍA DE SISTEMAS PARA LA OBTENCIÓN DEL GRADO ACADÉMICO DE LICENCIATURA EN INGENIERÍA DE SISTEMAS

Universidad Católica Boliviana “San Pablo”

Facultad de Ingeniería

Departamento de Ingeniería de Sistemas

##### La Paz - Bolivia

Realizado por:

Alvaro Miguel Flores Huacote

Tutor: MSc. Jesús Bernardo Ruíz Flores

Relator: Ing. Ernesto Campohermoso

“SISTEMA WEB DE MARCADO ESTEGANOGRÁFICO APLICANDO TECNOLOGÍAS DE EXPRESIONES REGULARES, ESTEGANOGRAFÍA Y OCR”



Índice

[1. Marco Referencial 1](#_Toc121154574)

[1.1. Antecedentes 1](#_Toc121154575)

[1.2. Descripción del objeto de estudio 2](#_Toc121154576)

[1.2.1. Planteamiento del problema 2](#_Toc121154577)

[1.3. Formulación del problema 2](#_Toc121154578)

[1.4. Objetivo del trabajo de grado 3](#_Toc121154579)

[1.4.1. Objetivo general 3](#_Toc121154580)

[1.4.2. Objetivos específicos 3](#_Toc121154581)

[1.5. Justificación de la investigación 3](#_Toc121154582)

[1.5.1. Justificación metodológica 3](#_Toc121154583)

[1.5.2. Justificación práctica 4](#_Toc121154584)

[1.5.3. Justificación social 4](#_Toc121154585)

[1.6. Límites Y Alcances 4](#_Toc121154586)

[1.6.1. Límites 4](#_Toc121154587)

[1.6.2. Alcances 4](#_Toc121154588)

[1.7. Modelos y herramientas por utilizar 5](#_Toc121154589)

[1.7.1. Modelos 5](#_Toc121154590)

[1.7.2. Herramientas 5](#_Toc121154591)

[1.8. Análisis preliminar del Proyecto 9](#_Toc121154592)

[1.9. Propuesta de solución 10](#_Toc121154593)

[1.9.1. Diagrama BPMN Solución. 11](#_Toc121154594)

[1.9.2. Diagrama de Componente 12](#_Toc121154595)

[1.9.3. Diagrama Idef0 13](#_Toc121154596)

[1.9.4. Diagrama Tecnológico. 14](#_Toc121154597)

[2. Marco Teórico 14](#_Toc121154598)

[2.1. Esteganografía 14](#_Toc121154599)

[2.2. Reconocimiento Óptico de Caracteres (OCR) 16](#_Toc121154600)

[2.3. OpenCV 20](#_Toc121154601)

[2.4. Expresiones Regulares 21](#_Toc121154602)

[2.5. Pytesseract 22](#_Toc121154603)

[3. Marco Practico 24](#_Toc121154604)

[3.1. Requerimientos 24](#_Toc121154605)

[3.1.1. Requerimientos funcionales 24](#_Toc121154606)

[3.1.2. Requerimientos No funcionales 26](#_Toc121154607)

[3.1.3. Lista de Requerimientos 27](#_Toc121154608)

[3.2. Interfaces 29](#_Toc121154609)

[3.2.1. Inicio Sesión 29](#_Toc121154610)

[3.2.2. Pagina Home del Sistema Web 29](#_Toc121154611)

[3.2.3. Marcado Esteganográfico 30](#_Toc121154612)

[3.2.4. Selección Alias 30](#_Toc121154613)

[3.2.5. Descargar documento marcado 31](#_Toc121154614)

[3.2.6. Listado Alias 31](#_Toc121154615)

[3.2.7. Creación de Alias 32](#_Toc121154616)

[3.2.8. Pantalla Crear Imagen 32](#_Toc121154617)

[3.2.9. Pantalla de Analizar Imagen 33](#_Toc121154618)

[3.2.10. Pantalla de Resultado de Análisis de Imagen 33](#_Toc121154619)

[3.2.11. Pantalla de Cargar PDF 34](#_Toc121154620)

[3.2.12. Pantalla de Analizar PDF 34](#_Toc121154621)

[3.2.13. Pantalla de Resultado de Análisis de PDF 35](#_Toc121154622)

[3.3. Diagrama Navegacional 36](#_Toc121154623)

[3.4. Diagramas de Caso de Uso 36](#_Toc121154624)

[3.4.1. Caso de uso Nro. 1: Ingreso al sistema 37](#_Toc121154625)

[3.4.2. Caso de uso Nro. 2: Salida del sistema 42](#_Toc121154626)

[3.4.3. Caso de uso Nro. 3: Marcado Esteganográfico 47](#_Toc121154627)

[3.4.4. Caso de uso Nro. 4: Carga de documentación al sistema 52](#_Toc121154628)

[3.4.5. Caso de uso Nro. 5: Análisis de Imagen capturada 57](#_Toc121154629)

[3.4.6. Caso de uso Nro. 6: Análisis de PDF 62](#_Toc121154630)

[3.4.7. Caso de uso Nro. 7: Alta alias 66](#_Toc121154631)

[3.4.8. Caso de uso Nro. 8: Modificación de alias 71](#_Toc121154632)

[3.4.9. Caso de uso Nro. 9: Baja de alias 76](#_Toc121154633)

[3.5. Diagrama de Despliegue 80](#_Toc121154634)

[3.6. Diseño Base de Datos Entidad-Relación 81](#_Toc121154635)

[3.7. Diagrama de Clases 82](#_Toc121154636)

[4. Modelo C4 83](#_Toc121154637)

[4.1. Diagrama de Contexto 83](#_Toc121154638)

[4.2. Diagrama de Contenedores 84](#_Toc121154639)

[4.3. Diagrama de Componentes 85](#_Toc121154640)

[5. Definición de criterios para el desarrollo del proyecto 86](#_Toc121154641)

**INTRODUCCIÓN**

Desarrollar un sistema web de forma que ayude a la empresa a identificar al personal que fue responsable de divulgar documentación confidencial, ya sean contratos, propuestas u otros. Haciendo uso de marcado esteganográfico en documentación confidencial.

El sistema realizará el marcado esteganográfico a documentos en formato PDF, de forma que las páginas de los documentos filtrados puedan ser identificados por el mismo sistema, ya sea si estos son manipulados en formato digital o en formato físico. Solo será necesario una foto de una página para que el sistema pueda identificar la procedencia del documento divulgado previamente marcado por el sistema. El marcado esteganográfico se guardará en el documento incluso si este es impreso y manipulado de forma física.

Para que el sistema pueda identificar la procedencia de cualquier documentación divulgada solo será necesario una fotografía en formato JPG o él mismo documento en formato PDF, el sistema evaluará la información obtenida haciendo uso de reconocimiento de caracteres (OCR), expresiones regulares y esteganografía. De esta forma se podrá identificar al personal responsable de divulgar documentación confidencial.

# Marco Referencial

## Antecedentes

**Esteganografía en archivos digitales - Digital File Steganography[[1]](#footnote-1)**

Es una aplicación que permite elegir entre diferentes métodos esteganográficos basados en la técnica de la modificación del bit menos significativo (LSB, en sus siglas en inglés) para codificar o decodificar un mensaje de texto en una imagen.

**Análisis de la viabilidad de un estándar de seguridad que implemente técnicas de esteganografía y criptografía para aumentar la seguridad de la información en las empresas[[2]](#footnote-2)**

Analiza la viabilidad de un estándar de seguridad de la información para empresas que implemente técnicas de esteganografía y criptografía. Se analizan imágenes marcadas previamente con un estego-algoritmo de forma que agrega a los binarios de la imagen. Para este proyecto, se desarrolló una investigación mixta, en donde la investigación cuantitativa se usó para el estudio, análisis y comparación de los algoritmos de esteganografía y criptografía.

## Descripción del objeto de estudio

### Planteamiento del problema



Figura. Diagrama Ishikawa

## Formulación del problema

**¿De qué manera será posible identificar al personal responsable de la divulgación de documentos confidenciales en las empresas?**

## Objetivo del trabajo de grado

### Objetivo general

Desarrollar un sistema web de marcado esteganográfico de forma que ayude a la empresa a mitigar los daños e identificar al personal responsable de la filtración de la información, aplicando tecnologías de expresiones regulares, esteganografía y reconocimiento de caracteres (OCR) para identificar el documento divulgado previamente marcado por el sistema.

### Objetivos específicos

* Realizar marcado esteganográfico en documentos Word para que el sistema pueda identificar el documento en caso de filtración de la información.
* Interpretar el marcado esteganográfico de documentos en formato PDF para identificar al personal responsable de la documentación filtrada.
* Interpretar el marcado esteganográfico de imágenes en formato JPG para identificar la procedencia de la documentación filtrada.
* Desarrollar el estego-algoritmo para realizar el marcado a los documentos seleccionados.
* Mejorar la calidad de imagen para obtener mejores resultados en la interpretación de caracteres.

## Justificación de la investigación

### Justificación metodológica

En el presente proyecto se realizará una investigación descriptiva, haciendo uso del método analítico para implementar las tecnologías de OCR (reconocimiento óptico de caracteres) y esteganografía, de forma que se pueda conseguir un método el cual ayude a las empresas a identificar al personal que es responsable de la divulgación de la información.

### Justificación práctica

Gran parte de las empresas tienen o han tenido problemas de filtración de documentación sensible por parte de su propio equipo de trabajo, es por eso que se plantea desarrollar un sistema web de marcado esteganográfico de modo que se logre identificar al personal responsable de divulgar dicha documentación, de esta forma se ayudará a las empresas a mitigar daños causados.

### Justificación social

Los usuarios de la aplicación podrán evaluar y analizar los datos obtenidos en el sistema para obtener información que ayude a mitigar el daño ocasionado por la filtración de documentos.

## Límites Y Alcances

### Límites

* La aplicación web solo realizará el marcado esteganográfico a documentos Word.
* La aplicación web no realizará el marcado esteganográfico a páginas del documento que contengan imágenes.
* La aplicación web solo realizará el marcado esteganográfico a documentos en formato PDF.
* La aplicación web solo interpretará imágenes en formato JPG.
* La aplicación web solo interpretará documentos en formato PDF.
* La aplicación web solo trabajará con documentos de 10 páginas como máximo.
* La aplicación web solo funcionara en Google Chrome.

### Alcances

* Se realizará un prototipo del sistema para demostrar su funcionamiento.
* El sistema web tendrá un diseño *responsive*.
* El sistema realizará un marcado al documento de tal manera que el marcado pueda identificarse si el documento es manipulado en formato digital o formato físico.
* El sistema identificará al personal responsable de la filtración de la documentación con solo la fotografía de una página del documento, esto solo será posible si la imagen pertenece a una página del documento marcado previamente por el sistema.
* El sistema identificará al personal responsable de la filtración de documentación con la obtención del documento en formato PDF, esto solo será posible si la imagen pertenece a una página del documento marcado previamente por el sistema.

## Modelos y herramientas por utilizar

### Modelos

Para el desarrollo de este proyecto se utilizará:

* UML: Se usará para modelar y visualizar el sistema.
* Kanban: Se usará para llevar el proceso de desarrollo del sistema web, **Cada tarea pasa por varios estados** desde que se planifica hasta que se lleva a cabo y eso se ve reflejado en el número de columnas que tiene el tablero Kanban. **Las columnas con las que se trabajaran las** tareas son: **por hacer, en progreso y hecho.**

### Herramientas

#### Python

Python es un lenguaje de programación ampliamente utilizado en las aplicaciones web, el desarrollo de software, la ciencia de datos y el machine learning (ML). Los desarrolladores utilizan Python porque es eficiente y fácil de aprender, además de que se puede ejecutar en muchas plataformas diferentes. El software Python se puede descargar gratis, se integra bien a todos los tipos de sistemas y aumenta la velocidad del desarrollo.



#### JavaScript

JavaScript se introdujo en 1995 como una forma de agregar programas a páginas web en el navegador Netscape Navigator. El lenguaje ha sido desde entonces adoptado por todos los otros navegadores webs principales. Ha hecho que las aplicaciones web modernas sean posibles: aplicaciones con las que puedes interactuar directamente, sin hacer una recarga de página para cada acción. JavaScript también es utilizado en sitios web más tradicionales para proporcionar diversas formas de interactividad e ingenio.

Un letrero de color blanco

Descripción generada automáticamente con confianza baja

#### Docker

Docker es una plataforma abierta para desarrollar, enviar y ejecutar aplicaciones. Docker le permite separar sus aplicaciones de su infraestructura para que pueda entregar el software rápidamente. Con Docker, puede gestionar su infraestructura de la misma manera que gestiona sus aplicaciones. Al aprovechar las metodologías de Docker para enviar, probar y desplegar el código rápidamente, puede reducir significativamente el tiempo que transcurre entre la escritura del código y su ejecución en producción.

Logotipo, nombre de la empresa

Descripción generada automáticamente

#### Git

Git es una herramienta que realiza una función del control de versiones de código de forma distribuida.



#### GitHub

GitHub es una plataforma de desarrollo colaborativo de software para alojar proyectos utilizando el sistema de control de versiones Git.

#### MySQL

MySQL es el sistema de gestión de bases de datos relacional más extendido en la actualidad al estar basada en código abierto. Desarrollado originalmente por MySQL AB, fue adquirida por Sun MicroSystems en 2008 y esta su vez comprada por Oracle Corporation en 2010, la cual ya era dueña de un motor propio InnoDB para MySQL. (Robledano, ¿Qué es MySQL, 2019)



#### Visual Studio Code

Visual Studio Code es un editor de código fuente ligero pero potente que se ejecuta en el escritorio y está disponible para Windows, macOS y Linux. Viene con soporte incorporado para JavaScript, TypeScript y Node.js y tiene un rico ecosistema de extensiones para otros lenguajes (como C++, C#, Java, Python, PHP, Go) y tiempos de ejecución (como .NET y Unity).



#### UML

El Lenguaje Unificado de Modelado (UML) fue creado para forjar un lenguaje de modelado visual común y semántica y sintácticamente rico para la arquitectura, el diseño y la implementación de sistemas de software complejos, tanto en estructura como en comportamiento. UML tiene aplicaciones más allá del desarrollo de software, p. ej., en el flujo de procesos en la fabricación.



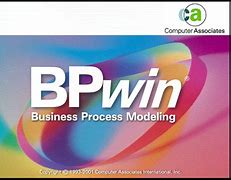
#### Bizagi

Bizagi es una plataforma que consiste en 3 componentes. Cada uno de ellos facilita un paso clave en la transformación y automatización de sus procesos de negocios.



#### BP Win

“BPwin es una potente herramienta de modelamiento utilizada para analizar, documentar y mejorar los procesos de negocio de nuestra organización” (Donozo, Tecnologías Emergentes 1, 2014).



#### Enterprise Architect

Enterprise Architect es una plataforma de modelado, diseño y administración, colaborativa, basada en UML 2.5 y estándares relacionados. Ágil, intuitiva y extensible, con poderosas características para dominios específicos totalmente integradas.



## Análisis preliminar del Proyecto

Figura. Modelado de Negocio – Análisis Preliminar

## Propuesta de solución

Desarrollar un sistema web de marcado esteganográfico basado en tecnologías web, expresiones regulares, inteligencia artificial y esteganografía que ayude a mitigar los daños e identificar al personal involucrado en la filtración de documentos confidenciales en la empresa. El sistema será capaz de realizar el marcado esteganográfico a un documento en formato PDF, el mismo sistema exportará el documento marcado por el estego-algoritmo, el cual ayudará a identificar la procedencia del documento filtrado. El sistema web será capaz de identificar el documento ya sea con el mismo documento en formato PDF o también con una foto legible de una página del documento filtrado.

### Diagrama BPMN Solución.

|  |
| --- |
| Figura. Diagrama BPMN – Modelo de Solución |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

Diagrama

Descripción generada automáticamente

### Diagrama Descripción generada automáticamenteDiagrama de Componente

Figura. Diagrama de Componentes

### Diagrama Idef0

Diagrama

Descripción generada automáticamente

Figura. Diagrama Idef-0

### Diagrama Tecnológico.

Figura. Diagrama Tecnológico

# Marco Teórico

## Esteganografía

La esteganografía estudia diferentes técnicas para la ocultación de datos en otros objetos, conocidos como objetos portadores. Actualmente, estos objetos portadores suelen ser medios digitales, como por ejemplo imágenes, videos o archivos de sonido. No obstante, sin lugar a duda, el medio más utilizado en la actualizad son las imágenes, por su amplia difusión en Internet. Uno de los métodos más usados para ocultar información en imágenes de mapas de bits es la sustitución del bit menos significativo (Least Significant Bit, LSB). Este método divide el mensaje original en bits y oculta cada uno de ellos en un píxel de la imagen (David Megías, 2014).

La esteganografía, aunque ha tardado más que la criptografía en darse a conocer (a pesar de tener también muchos siglos de antigüedad), ya es bastante popular en el mundo de la seguridad informática. Igual que dentro de la criptografía, la esteganografía también tiene variantes. De hecho, el término «esteganografía» se utiliza muchas veces para englobar lo que en realidad es un campo más general, que es la ciencia de la «ocultación de información» y que abarca otros campos relacionados, como las marcas de agua, la anonimia o los canales encubiertos (Jesús Diaz, 2015).

En general, existen 3 grandes de líneas de creación de algoritmos esteganográficos:

* La cubierta existe y la ocultación de la modificación no la modifica.
* La cubierta existe y la ocultación produce alteraciones.
* La generación automática de la cubierta incluye la información a ocultar.

Se puede insertar los mensajes en diferentes tipos de medios digitales, ya sea imágenes como lo hacían en la antigüedad, archivos de audio, como también se puede incluir mensajes en los ficheros de los discos duros o en el propio protocolo de comunicación TCP/IP.

Esteganografía lingüística

Según la clasificación que realizo *Friedrich I. Bawer* en *Decrypted Secrets. Methods and Maxims of Cryptology* en 1997, la esteganografía textual puede explicarse en dos grandes ramas códigos abiertos y semagramas (Serra y Lerch, p. 109).

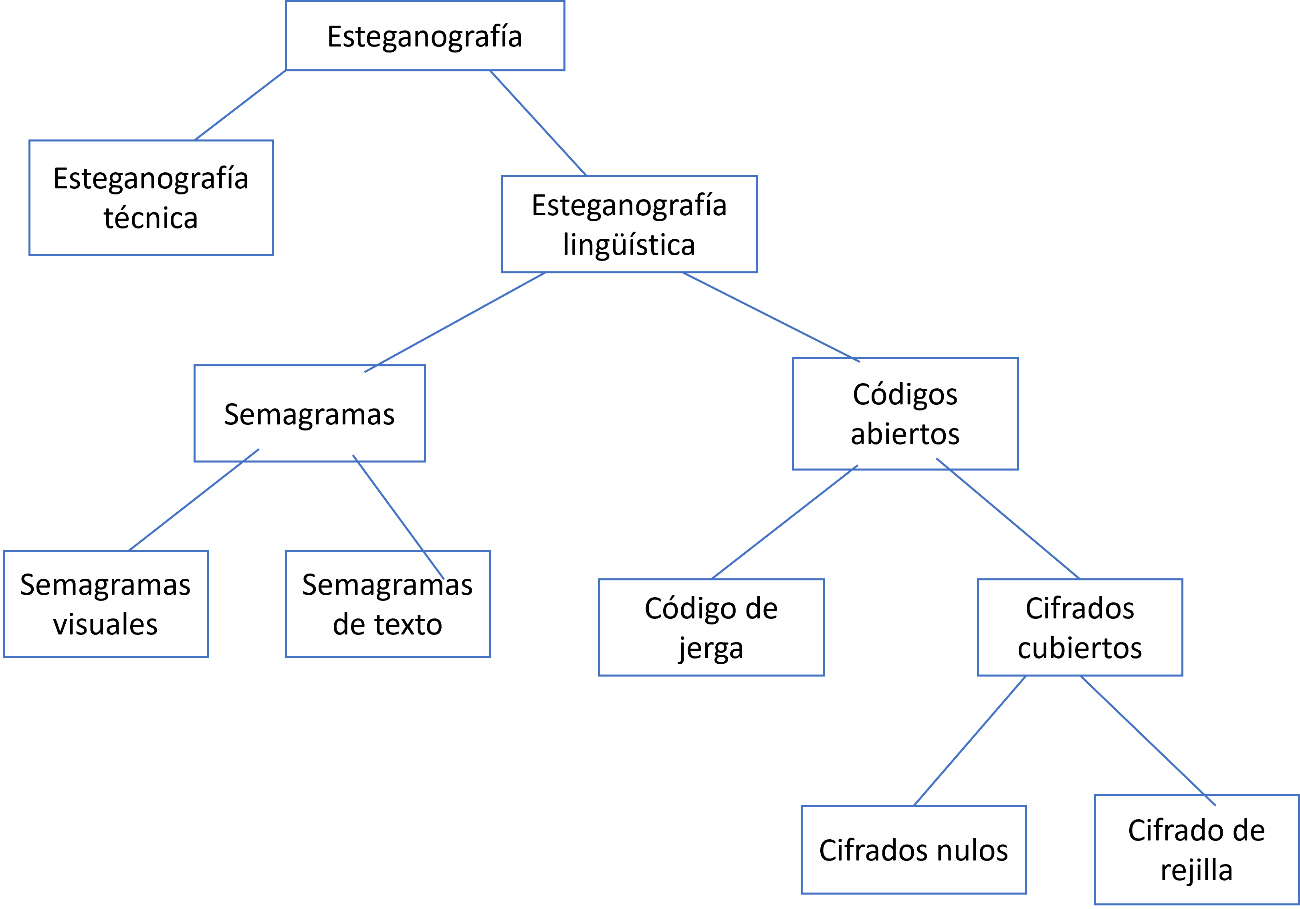


Figura. Posible clasificación de técnicas de esteganografía textual

Existen dos grandes líneas de investigación para proporcionar soluciones, la generación automática de textos estenografiados y la modificación de textos existentes.

**Generación automática de textos esteganografiados**

consiste en procedimientos que faciliten la generación de textos que ya lleven incluida la información que se desea ocultar (Serra y Lerch, p. 109).

**Modificación de textos existentes para crear textos esteganografiados.**

El mecanismo más tradicional de ocultación de información consiste en utilizar un texto existente para enmascarar información basándose en algunos elementos del texto o modificaciones del mismo. Los procedimientos actuales, no exentos de problemas, consisten en modificaciones léxicas, el mecanismo mas usado es la ocultación de información basada en la sustitución de palabras por sus sinónimos, modificaciones sintácticas y semánticas (Serra y Lerch, p. 109).

## Reconocimiento Óptico de Caracteres (OCR)

La tecnología de reconocimiento óptico de caracteres (OCR) es un proceso empresarial eficaz que ahorra tiempo, costes y otros recursos al utilizar la capacidad de extracción y almacenamiento de datos automatizados.

En 1974, Ray Kurzweil fundó Kurzweil Computer Products, Inc. cuyo producto de reconocimiento óptico de caracteres (OCR) omnicanal podía reconocer texto impreso en prácticamente cualquier fuente. Decidió que la mejor aplicación de esta tecnología sería un dispositivo de aprendizaje automático para ciegos, así que creó una máquina de lectura que podía leer el texto en voz alta en un formato de texto a voz. En 1980, Kurzweil vendió su empresa a Xerox, que estaba interesada en seguir comercializando la conversión de texto de papel a ordenador (IBM *Education*, 2022).

La tecnología OCR se popularizó a principios de la década de 1990 al digitalizar periódicos históricos. Desde entonces, la tecnología ha experimentado varias mejoras. Las soluciones actuales son capaces de ofrecer una precisión de OCR casi perfecta. Se utilizan métodos avanzados para automatizar los complejos flujos de trabajo de procesamiento de documentos. Antes de que existiera la tecnología OCR, la única opción para dar formato digital a los documentos era volver a escribir el texto manualmente. Esto no sólo requería mucho tiempo, sino que además conllevaba inevitables imprecisiones y errores de escritura (IBM *Education*, 2022).

El reconocimiento óptico de caracteres (OCR) utiliza un escáner para procesar la forma física de un documento. Una vez copiadas todas las páginas, el software de OCR convierte el documento en una versión en dos colores o en blanco y negro. La imagen escaneada o el mapa de bits se analizan en busca de áreas claras y oscuras, y las áreas oscuras se identifican como caracteres que necesitan ser reconocidos, mientras que las áreas claras se identifican como fondo. Las zonas oscuras se procesan para encontrar letras alfabéticas o dígitos numéricos. En esta etapa se suele seleccionar un carácter, una palabra o un bloque de texto a la vez. A continuación, los caracteres se identifican mediante uno de los dos algoritmos: reconocimiento de patrones o reconocimiento de rasgos (IBM *Education*, 2022).

El reconocimiento de patrones se utiliza cuando el programa de OCR recibe ejemplos de texto en varias fuentes y formatos para comparar y reconocer los caracteres en el documento o archivo de imagen escaneado

La detección de características se produce cuando el OCR aplica reglas relativas a las características de una letra o un número específico para reconocer los caracteres en el documento escaneado. Las características incluyen el número de líneas angulares, líneas cruzadas o curvas en un carácter. Por ejemplo, la letra "A" mayúscula se almacena como dos líneas diagonales que se juntan con una línea horizontal en el medio. Cuando se identifica un carácter, se convierte en un código ASCII (*American Standard Code for Information Interchange*) que los sistemas informáticos utilizan para realizar otras manipulaciones (IBM *Education*, 2022).

Un programa de OCR también analiza la estructura de la imagen de un documento. Divide la página en elementos como bloques de texto, tablas o imágenes. Las líneas se dividen en palabras y luego en caracteres. Una vez identificados los caracteres, el programa los compara con un conjunto de imágenes patrón. Tras procesar todas las coincidencias probables, el programa le presenta el texto reconocido (IBM *Education*, 2022).

El caso de uso más conocido del reconocimiento óptico de caracteres (OCR) es la conversión de documentos impresos en papel en documentos de texto legibles por máquina. Una vez que un documento en papel escaneado pasa por el proceso de OCR, el texto del documento puede editarse con un procesador de textos como Microsoft Word o Google Docs (IBM *Education*, 2022).

El OCR se utiliza a menudo como una tecnología oculta, impulsando muchos sistemas y servicios bien conocidos en nuestra vida diaria. Entre los casos de uso importantes, pero menos conocidos de la tecnología OCR se encuentran la automatización de la entrada de datos, la asistencia a personas ciegas y con problemas de visión y la indexación de documentos para motores de búsqueda, como pasaportes, matrículas, facturas, extractos bancarios, tarjetas de visita y el reconocimiento automático de matrículas (IBM *Education*, 2022).

El OCR permite optimizar el modelado de *big data* mediante la conversión de documentos en papel y de imágenes escaneadas en archivos PDF legibles por máquina y con capacidad de búsqueda. El procesamiento y la recuperación de información valiosa no pueden automatizarse sin aplicar primero el OCR en los documentos en los que no hay capas de texto (IBM *Education*, 2022).

Con el reconocimiento de texto OCR, los documentos escaneados pueden integrarse en un sistema de *big data* que ahora es capaz de leer los datos de los clientes a partir de extractos bancarios, contratos y otros documentos impresos importantes. En lugar de que los empleados examinen innumerables documentos de imagen y alimenten manualmente las entradas en un flujo de trabajo de procesamiento de big data automatizado, las organizaciones pueden utilizar el OCR para automatizar en la etapa de entrada de la minería de datos. El software de OCR puede identificar el texto en la imagen, extraer el texto en las imágenes, guardar el archivo de texto y admitir los formatos jpg, jpeg, png, bmp, tiff, pdf y otros (IBM *Education*, 2022).

**Funciones del OCR en el sistema web de marcado esteganográfico**

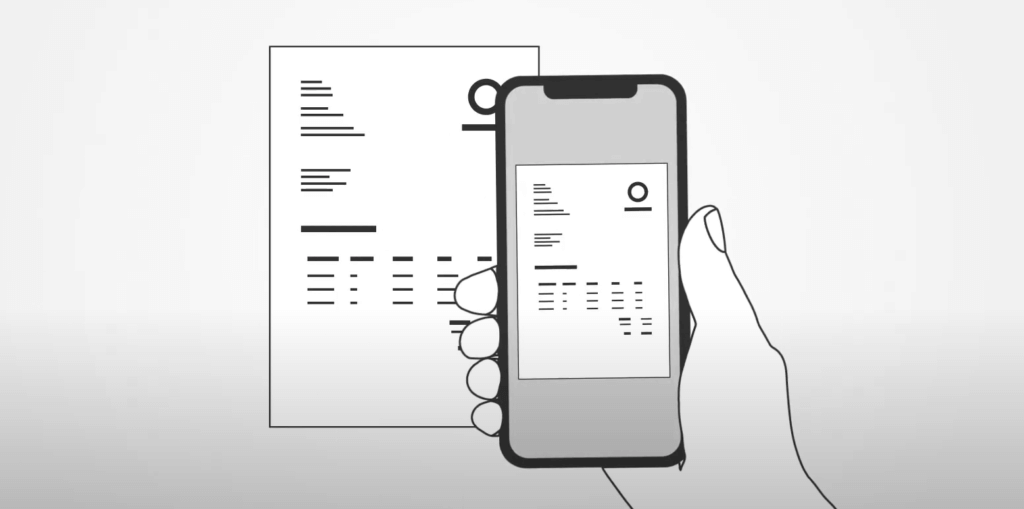


Figura. Tecnología OCR

Para comprender la tecnología OCR, como funciona de una forma más clara, veamos cómo funciona en práctica. Primero, necesitas una imagen a escanear y un software que pueda convertir la imagen en texto usando reconocimiento de texto. Para el proyecto utilizaremos la librería de Python Pytesseract.

El proceso OCR conllevará una serie de cuatro pasos para el sistema de marcado esteganográfico interprete el contenido de la imagen:

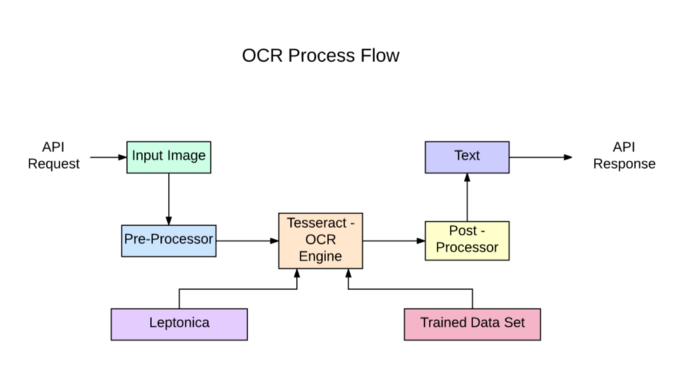


Figura. Flujo de Proceso OCR

Fuente Balaaji Parthasarathy

* Subir la imagen al sistema web
* Pre-procesamiento de imagen
* Reconocimiento de texto
* Post-procesamiento

**Subir la imagen al sistema web**

En el primer paso necesitamos tomar una foto del documento que se escaneara, posteriormente se carga la imagen al sistema web.

**Pre-procesamiento de imagen**

En este paso el sistema web escanea la imagen y optimiza la calidad de imagen. Ya que el sistema solo trabajara con documentos aplicamos un filtro con la ayuda de la librería OpenCV aumentando la escala de grises y reduciendo el ruido en la imagen para obtener una mejor salida, es de gran importación pre-procesar la imagen ya que si se cuenta con una imagen borrosa o marcas en la imagen pueden causar errores en la interpretación de caracteres.

**Reconocimiento de texto**

En este paso es donde se realiza el reconocimiento de caracteres. El sistema web analiza los patrones de los bits en la imagen que componen los caracteres para transformar la imagen escaneada en texto, haciendo uso de la librería de Pytesseract.

**Post-procesamiento**

En el último paso, el sistema web convierte los datos reconocidos en texto plano y también podemos obtener los resultados en formato XML, el cual utilizaremos para ejecutar posteriormente otras funcionalidades del sistema web.

## OpenCV

OpenCV (Open Source Computer Vision) comenzó como un proyecto de investigación en Intel. Actualmente es la biblioteca de visión por computadora más grande en términos de funciones poseídas. Actualmente, OpenCV contiene implementaciones de más de 2500 algoritmos. Además, está disponible de forma gratuita para fines comerciales y académicos. Esta gran biblioteca tiene interfaces para múltiples lenguajes, incluidos Python, Java y C++. La primera versión de OpenCV, 1.0, se lanzó en 2006 y la comunidad de OpenCV ha crecido a pasos agigantados desde entonces. OpenCVPython es la API de Python para OpenCV. Además, es compatible con las plataformas más utilizadas, Windows, Mac OS y Linux (Marin, 2020).

El sistema utiliza algunas funcionalidades proporcionadas por la librería OpenCV, para realizar el pre-procesamiento de las imágenes, con esto logramos obtener un mejor resultado del análisis realizado por el sistema web. En el pre-procesamiento de las imágenes utilizamos las funciones de escalado de grises, reducción de ruido en las imágenes y obtenemos los siguientes resultados con se muestran en la figura X.

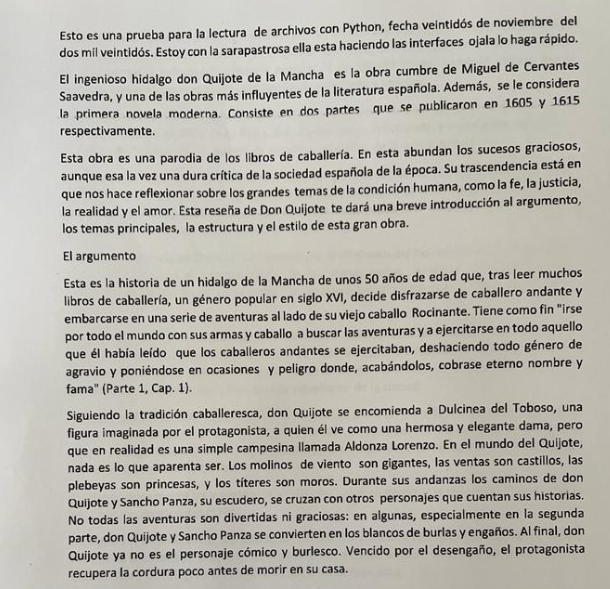


Figura. Imagen Original

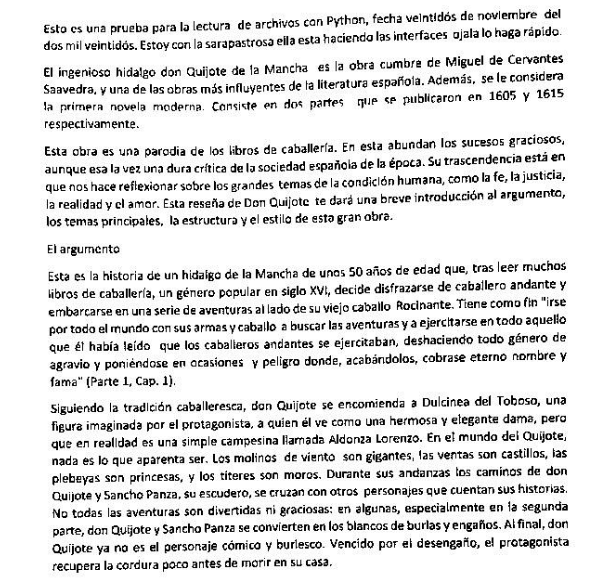


Figura. Imagen Preprocesada para ingresar al sistema

## Expresiones Regulares

Una expresión regular o RegEx es una secuencia de caracteres que forman un patrón de búsqueda. En Google Analytics, explicado de una forma más simple, podríamos decir que son comodines que nos ayudan a simplificar expresiones para lograr resultados.

Las Expresiones Regulares nos sirven para que podamos establecer filtros, para buscar ciertos grupos con ciertas combinaciones de caracteres específicas, por ejemplo, buscar números telefónicos, que tengan o no una clave de país, que tengan o no espacios entre los números, que tengan o no un símbolo de # al inicio, y fácilmente ignorar el resto de información que podamos tener, con esto puedes crear por ejemplo un Web Scrapper de teléfonos, puedes crear un filtro para validad números de teléfonos (tan precisos cómo tú expresión regular lo exija) entre otras cosas (Celis, 2018).

## Pytesseract

Python-tesseract es una herramienta de reconocimiento óptico de caracteres (OCR) para Python. Es decir, reconocerá y "leerá" el texto incrustado en las imágenes.

Python-tesseract es una envoltura para el motor Tesseract-OCR de Google. También es útil como un script de invocación independiente a tesseract, ya que puede leer todos los tipos de imágenes soportados por las bibliotecas de imágenes Pillow y Leptonica, incluyendo jpeg, png, gif, bmp, tiff, y otros. Además, si se utiliza como un script, Python-tesseract imprimirá el texto reconocido en lugar de escribirlo en un archivo (Smith, 2019).

**Funciones Pytesseract:**

* get\_languages Devuelve todos los idiomas actualmente soportados por Tesseract OCR.
* get\_tesseract\_version Devuelve la versión de Tesseract instalada en el sistema.
* image\_to\_string Devuelve el resultado sin modificar como cadena del procesamiento de Tesseract OCR.
* image\_to\_boxes Devuelve el resultado con los caracteres reconocidos y los límites de sus casillas.
* image\_to\_data Devuelve el resultado con los límites de las cajas, las confidencias y otra información. Requiere Tesseract 3.05+. Para más información, consulte la documentación de Tesseract TSV.
* image\_to\_osd Devuelve el resultado que contiene información sobre la orientación y la detección de guiones.
* image\_to\_alto\_xml Devuelve el resultado en formato ALTO XML de Tesseract.
* run\_and\_get\_output Devuelve el resultado sin procesar del OCR de Tesseract. Da un poco más de control sobre los parámetros que se envían a tesseract.

**Parámetros Pytesseract**

* **image Object or String** - PIL Image/NumPy array o file path de la imagen a procesar por Tesseract. Si se pasa el objeto en lugar de la ubicación de la imagen, pytesseract convertirá implícitamente la imagen a modo RGB.
* **lang String** - Cadena de código del idioma de Tesseract. Por defecto es Ingles si no se especifica. Ejemplo para múltiples idiomas: lang='eng+fra'
* **config String** - Cualquier configuración adicional que no esté disponible a través de la función pytesseract. Por ejemplo: config='--psm 6'
* **nice Integer** - Modifica la prioridad del procesador para la ejecución de Tesseract. No es compatible con Windows. nice ajusta la amabilidad de los procesos tipo unix.
* **output\_type Atributo de clase** - especifica el tipo de la salida, por defecto string. Para la lista completa de todos los tipos soportados, por favor revise la definición de la clase pytesseract.Output.
* **timeout Integer or Float** - duración en segundos para el procesamiento OCR, después del cual, pytesseract terminará y lanzará RuntimeError.
* **pandas\_config Dict** - sólo para la clase Output.DATAFRAME. Diccionario con argumentos personalizados para pandas.read\_csv. Permite personalizar la salida de image\_to\_data.

# Marco Practico

## Requerimientos

### Requerimientos funcionales

|  |  |  |  |
| --- | --- | --- | --- |
| **Categoría** | | **Requerimiento** | **Descripción** |
| **ACCESO A LA APLICACIÓN** | Req-1 | Inicio de sesión | Los usuarios ingresaran con su usuario y contraseña |
| Req-2 | Aceptación de términos de uso | Los usuarios deberán de aceptar los términos de uso de la aplicación para poder usar sus funciones |
| Req-3 | Registro de Usuarios | El administrador podrá ver el historial de documentos marcados |
| Req-4 | Cierre de Sesión | El usuario podrá cerrar sesión |
| **PANTALLA DE INICIO** | Req-5 | Menú superior | El sistema contará con un menú superior que mostrará el botón de perfil del usuario. |
| Req-6 | Opciones de Perfil |  |
| Req-7 | Pantalla de inicio | Pantalla de inicial deberá mostrar botones de: Marcado Esteganográfico, Botón de Escaneo de Imágenes, botón de escaneado de documentos PDF. |
| Req-8 | Opción de Marcado Esteganográfico | El botón de marcado Esteganográfico estará en la pantalla de inicio del sistema |
| Req-9 | Opción de Escaneo de Imágenes | El botón de Escaneo de imágenes estará en la pantalla de inicio del sistema. |
| Req-10 | Opción de Escaneo de Documentos PDF. | El botón de Escaneo de Documentos PDF estará en la pantalla de inicio del sistema |
| **PANTALLA MARCADO ESTEGANOGRAFICO** | Req-11 | El sistema debe tener una Pantalla de listado de documentos | La pantalla mostrara el listado de todos los documentos cargados al sistema sin el marcado esteganográfico |
| Req-12 | El sistema debe tener un botón para cargar documentos al sistema | El sistema contara con un botón en la pantalla de listado de documentos para cargar documentos al sistema. |
| Req-13 | El sistema debe tener un botón de marcado | El sistema contará con un botón en el listado del documento, que dirigirá a la siguiente pantalla para realizar el marcado esteganográfico |
| Req-14 | Pantalla de formulario de Marcado | El sistema contará con una pantalla de formulario de marcado, en la que será asignado el alias para realizar el posterior marcado esteganográfico al documento. |
| Req-15 | El sistema debe tener un ABM de Alias para el marcado | El sistema contara con las funcionalidades para crear, modificar o eliminar los alias. |
| Req-16 | El sistema debe tener un botón de confirmación de marcado | El botón de confirmación de marcado estará ubicado en la pantalla del formulario de Marcado, el cual disparará el evento para que el sistema realice el marcado esteganográfico con todos los datos llenados en el formulario. |
| Req-17 | El sistema debe tener un botón de descarga | El sistema deberá tener un botón para descargar el documento marcado. |
| **PANTALLA ANALISIS IMAGEN** | Req-18 | El sistema debe tener un botón con la función de cargado de Imágenes | El sistema deberá tener un botón para cargar imágenes para ser analizadas por el sistema. |
| Req-19 | El sistema debe tener un botón con la función de Análisis de Imagen Cargada | El sistema deberá tener un botón para realizar el Análisis de la Imagen cargada |
| Req-20 | El sistema debe contar con una pantalla de Resultado | El sistema debe contar con una pantalla que muestre el resultado del Análisis de la Imagen Cargada |
| **PANTALLA ANALISIS DOCUMENTO** | Req-21 | El sistema debe tener un botón con la función de cargado de documentos | El sistema deberá tener un botón para cargar el documento para ser analizada por el sistema |
| Req-22 | El sistema debe tener un botón con la función de Análisis de Documento cargado | El sistema deberá tener un botón para realizar el Análisis del Documento cargado |
| Req-23 | El sistema debe contar con una pantalla de Resultado | El sistema debe contar con una pantalla que muestre el resultado del Análisis del documento cargado |

Tabla. Requerimientos Funcionales

### Requerimientos No funcionales

|  |  |  |  |
| --- | --- | --- | --- |
| **Categoría** | | **Requerimiento** | **Descripción** |
| **INTEGRIDAD** | Req-NF-1 | Los permisos al sistema estarán sujetos a un Super Usuario | El sistema deberá estar implementado de manera que cualquier modificación a sus datos se deberá realizar con la autorización de un Super Usuario. |
| **SEGURIDAD** | Req-NF-2 | El sistema se mantendrá activo por tiempo limitado | El sistema pasara a cerrar sesión automáticamente después de 2 min de inactividad. |
| **DISPONIBILIDAD** | Req-NF-3 | Navegador web | El sistema trabajara de manera eficaz y correcta en el navegador Chrome. |
| Req-NF-4 | Conexión a Internet | Se deberá tener conexión a internet para que el sistema funcione. |
| **ADAPTABILIDAD** | Req-NF-5 | Diseño Responsivo | El sistema tendrá un diseño responsive de forma que pueda hacer uso del mismo en cualquier tamaño de pantalla. |
| Req-NF-6 | Modificabilidad | El sistema debe estar preparado para poder realizar cambios a nuevos requerimientos futuros. |
| **USABILIDAD** | Req-NF-7 | Aprendizaje del sistema | El uso del sistema deberá ser implementado de tal manera, que el aprendizaje de su uso sea rápido, eficaz y efectivo. |
| Req-NF-8 | Lenguaje del Sistema | El sistema permitirá el uso de lenguajes que utilizan el alfabeto Ingles. En el caso de letras y símbolos especiales, solo se permitirá algunas cosas especiales como los usados en el idioma español. |
| Req-NF-9 | Manual del Sistema | El sistema tendrá un manual de usuario. |

Tabla. Requerimientos No Funcionales

### Lista de Requerimientos

Detalle de los requerimientos y su prioridad:

|  |  |  |  |
| --- | --- | --- | --- |
| **N°** | **Código** | **Descripción** | **Prioridad** |
| **Requerimientos Funcionales** | | | |
| 1 | Req-1 | Inicio de sesión | Alta |
| 2 | Req-2 | Aceptación de términos de uso | Media/Baja |
| 3 | Req-3 | Registro de Usuarios | Alta |
| 4 | Req-4 | Cierre de Sesión | Alta |
| 5 | Req-5 | Menú superior | Media/Baja |
| 6 | Req-6 | Opciones de Perfil | Media/Baja |
| 7 | Req-7 | Pantalla de inicio | Alta |
| 8 | Req-8 | Opción de Marcado Esteganográfico | Alta |
| 9 | Req-9 | Opción de Escaneo de Imágenes | Alta |
| 10 | Req-10 | Opción de Escaneo de Documentos PDF. | Alta |
| 11 | Req-11 | El sistema debe tener una Pantalla de listado de documentos | Alta |
| 12 | Req-12 | El sistema debe tener un botón para cargar documentos al sistema | Alta |
| 13 | Req-13 | El sistema debe tener un botón de marcado | Alta |
| 14 | Req-14 | Pantalla de formulario de Marcado | Alta |
| 15 | Req-15 | El sistema debe tener un ABM de Alias para el marcado | Alta |
| 16 | Req-16 | El sistema debe tener un botón de confirmación de marcado | Alta |
| 17 | Req-17 | El sistema debe tener un botón de descarga | Alta |
| 18 | Req-18 | El sistema debe tener un botón con la función de cargado de Imágenes | Alta |
| 19 | Req-19 | El sistema debe tener un botón con la función de Análisis de Imagen Cargada | Alta |
| 20 | Req-20 | El sistema debe contar con una pantalla de Resultado | Alta |
| 21 | Req-21 | El sistema debe tener un botón con la función de cargado de documentos | Alta |
| 22 | Req-22 | El sistema debe tener un botón con la función de Análisis de Documento cargado | Alta |
| 23 | Req-23 | El sistema debe contar con una pantalla de Resultado | Alta |
| **Requerimientos No Funcionales** | | | |
| 24 | Req-NF-1 | Los permisos al sistema estarán sujetos a un Super Usuario | Alta/Media |
| 25 | Req-NF-2 | El sistema se mantendrá activo por tiempo limitado | Alta/Media |
| 26 | Req-NF-3 | Navegador web | Alta/Media |
| 27 | Req-NF-4 | Conexión a Internet | Alta/Media |
| 28 | Req-NF-5 | Diseño Responsivo | Alta/Media |
| 29 | Req-NF-6 | Modificabilidad | Alta/Media |
| 30 | Req-NF-7 | Aprendizaje del sistema | Alta/Media |
| 31 | Req-NF-8 | Lenguaje del Sistema | Alta/Media |
| 32 | Req-NF-9 | Manual del Sistema | Alta/Media |

Tabla. Prioridad de Requerimientos

## Interfaces

### Inicio Sesión



Figura. Pantalla de Inicio de Sesión

La figura 8 nos presenta la interfaz de inicio de sesión al sistema, solicitando el correo electrónico del usuario y la contraseña para ingresar al sistema.

### Pagina Home del Sistema Web



Figura. Pantalla Inicio

La figura 9 presenta la interfaz del menú de inicio del sistema, la cual nos presenta tres opciones: marcado esteganográfico, analizar imagen y analizar documento.

### Marcado Esteganográfico

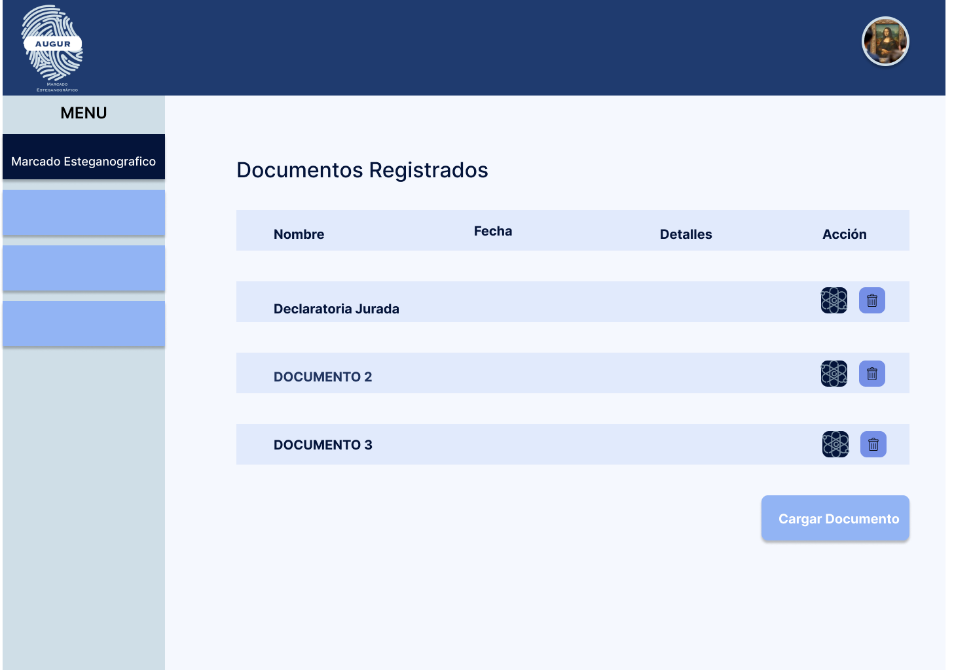


Figura. Pantalla Lista de Documentos – Marcado Esteganográfico

La figura 10 presenta la interfaz para seleccionar el documento que se desea marcar, la interfaz presenta un listado de los documentos cargados al sistema, también nos presenta los botones de selección de documento, eliminar documento y cargar documento.

### Selección Alias



Figura. Pantalla Selección Alias

La figura 11 presenta la interfaz para seleccionar el documento que se desea marcar, la interfaz presenta un listado de los documentos cargados al sistema, también nos presenta los botones de selección de documento, eliminar documento y cargar documento.

### Descargar documento marcado



Figura. Pantalla Documento Marcado

La figura 12 presenta la interfaz de usuario para descargar el documento marcado por el estego algoritmo.

### Listado Alias



Figura. Pantalla Listado Alias

La figura 13 presenta la interfaz del listado de alias creados en el sistema.

### Creación de Alias



Figura. Pantalla Crear Alias

La Figura 14 presenta la interfaz para crear un nuevo alias en el sistema. La interfaz presenta un formulario el cual solicita los datos para crear el nuevo usuario.

### Pantalla Crear Imagen



Figura. Pantalla Cargar Imagen

La figura 15 presenta la interfaz para cargar la imagen para analizar y verificar si la imagen del documento es un documento que fue filtrado.

### Pantalla de Analizar Imagen



Figura. Pantalla Analizar Imagen

La figura 16 presenta la interfaz para ejecutar la función de análisis de la imagen y verificar si es un documento filtrado.

### Pantalla de Resultado de Análisis de Imagen



Figura. Pantalla Resultado de Análisis de Imagen

La figura 17 presenta la interfaz en donde podemos apreciar el resultado del análisis realizado por el sistema, ya sea si se analizó una imagen o un documento PDF

### Pantalla de Cargar PDF

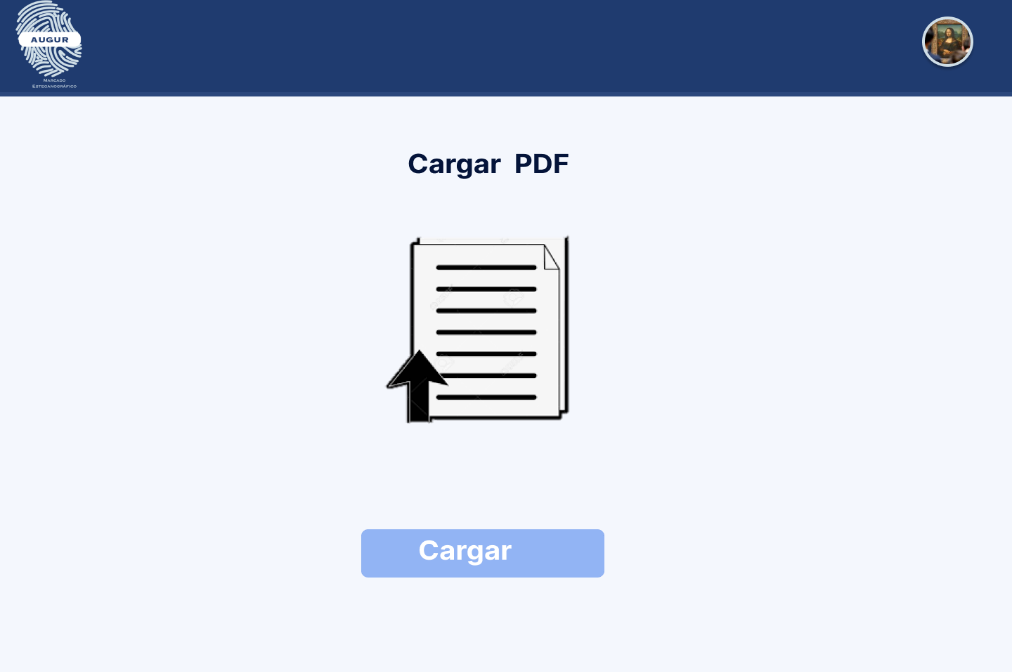


Figura. Pantalla Cargar PDF

La figura 18 presenta la interfaz de carga de documentos PDF para su posterior análisis.

### Pantalla de Analizar PDF



Figura. Pantalla Analizar PDF

La figura 19 presenta la interfaz de usuario para ejecutar el análisis del sistema y verificar si dicho documento pertenece al sistema.

### Pantalla de Resultado de Análisis de PDF



Figura. Pantalla Resultado de Análisis de PDF

La figura 20 presenta la interfaz en donde podemos apreciar el resultado del análisis realizado por el sistema, ya sea si se analizó una imagen o un documento PDF.

## Diagrama Navegacional

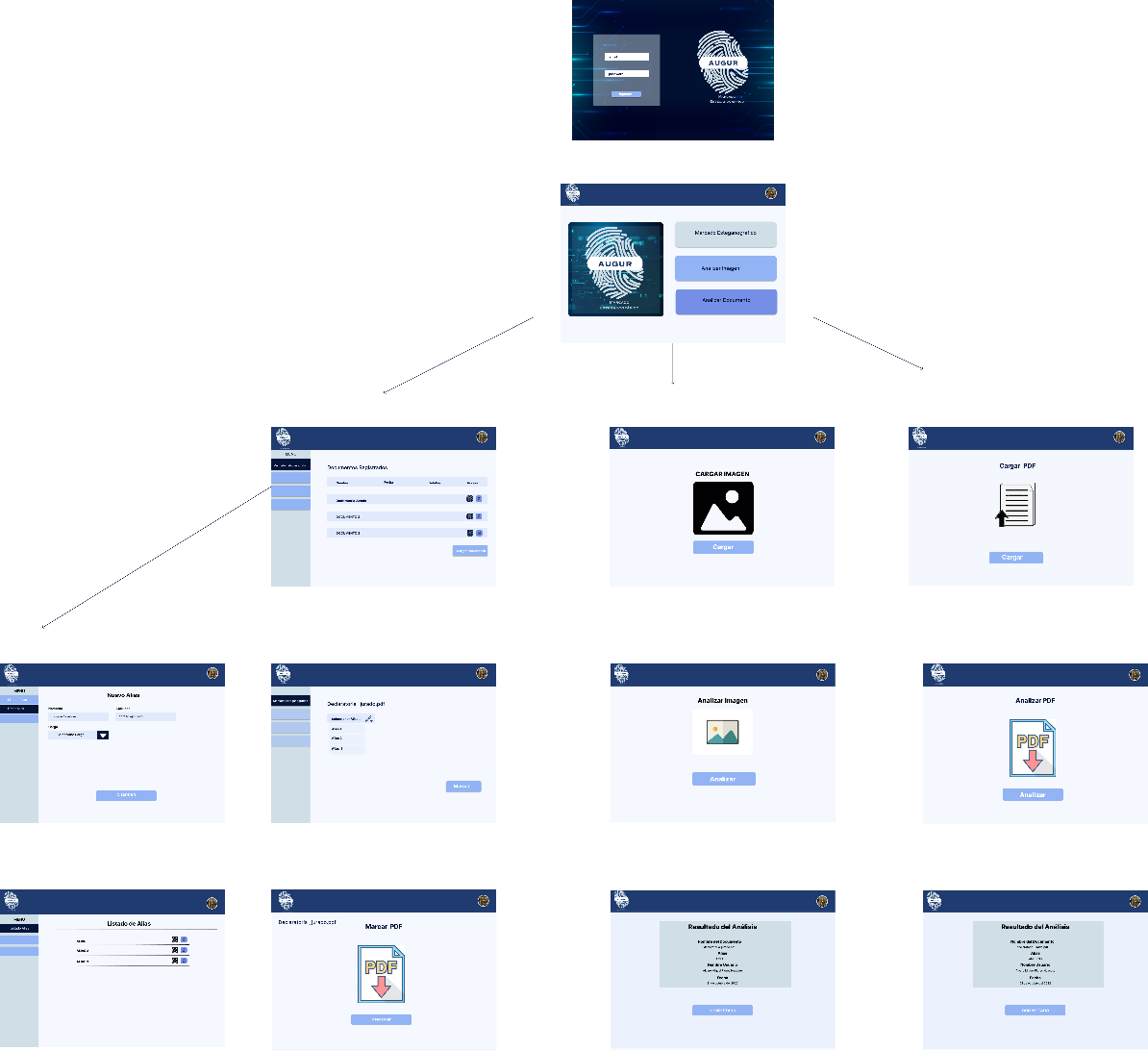
****

Figura. Diagrama Navegacional del Sistema Web

## Diagramas de Caso de Uso

Un caso de uso es un artefacto que define una secuencia de acciones que da lugar a un resultado de valor observable. Los casos de uso proporcionan una estructura para expresar requisitos funcionales en el contexto de procesos empresariales y de sistema, pueden representarse como un elemento gráfico en un diagrama y como una especificación de caso de uso en un documento textual. Un caso de uso de sistema es una secuencia de acciones que un sistema lleva a cabo que da lugar a un resultado de valor observable para un actor particular (IBM, 2021).

A continuación, se muestran los casos de uso existentes en el sistema web.

### Caso de uso Nro. 1: Ingreso al sistema

**Descripción:**

El caso de uso describe los pasos a seguir para que el usuario pueda ingresar al sistema.

**Identificación de los actores y casos de uso:**

**Actores**

* Usuario
* Sistema web

**Casos de Uso**

* Ingresar nombre de usuario
* Validación del sistema
* Ingresar contraseña de usuario
* Pantalla de home del sistema

**Diagrama de Caso de Uso**

****

Figura. Caso de Uso Ingreso ala Sistema

**Flujo de caso de uso**

|  |  |
| --- | --- |
| **CASO DE USO** | Inicio de Sesión |
| **ACTORES** | Usuario Sistema Web |
| **PROPOSITO** | Iniciar Sesión en el Sistema |
| **TIPO** | Alta |
| **CURSO NORMAL DE EVENTOS** | |
| ACCIONES DE ACTORES | RESPUESTA DEL SISTEMA |
| 1. El caso de uso inicia cuando el actor ingresa al sistema.  3. El actor ingresa los datos requeridos para el inicio de sesión. | 2. El sistema pedirá que se ingrese usuario y contraseña.  4. El sistema recibe la información y valida los datos con la base de datos.  5. El sistema muestra la página home de la aplicación web. |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

**Diagrama de Secuencia: Inicio de Sesión**

Diagrama de secuencia que muestra la interacción del usuario y los componentes para el inicio de sesión.



Figura. Diagrama de Secuencia – Inicio de Sesión

**Detalle de Diagrama de Secuencias: Inicio de Sesión**

|  |  |  |  |
| --- | --- | --- | --- |
| Origen del Mensaje | Destino del Mensaje | Mensaje | Descripción |
| Usuario | Pantalla de Login | Ingresar Datos (Usuario, contraseña) | Usuario ingresa credenciales para el inicio de Sesión |
| Pantalla Login | Controlador Login | Sigin | Se envían las credenciales a través de solicitud mediante una solicitud ORM |
| Controlador Login | Base de Datos del Sistema | Obtener estado del usuario ingresado | Envía los datos ingresados a la BD para generar una consulta acerca de los datos |
| Base de Datos del sistema | Controlador Login | Envía autorización | Responde con el estado del usuario |
| Controlador Login | Pantalla Home | Recibe el estado del usuario | Recibe el estado del usuario y envía la pantalla home |
| Pantalla Home | Usuario | Ingreso Sesión Correcto| Incorrecto | Muestra la pantalla Home si los datos fueron correctos |

Tabla. Detalle de Diagrama de Secuencias: Inicio de Sesión

**Diagrama de Actividades: Inicio de Sesión**

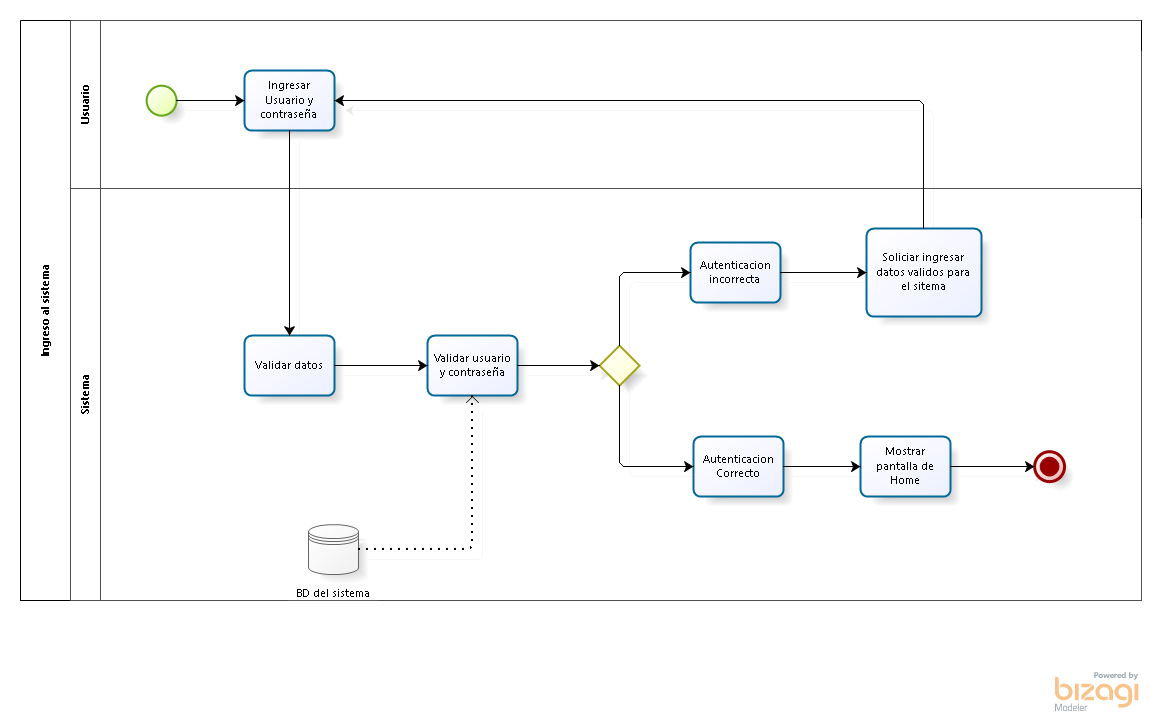


Figura. Diagrama de Actividades – Inicio de Sesión

El diagrama de actividades para el caso de uso Inicio de sesión cuenta con dos carriles:

* Sistema: Carril que representa todo es sistema web y las actividades que serán realizados por el mismo.
* Usuario: Carril que muestra las actividades realizadas por el usuario en el sistema web.

El usuario podrá ingresar al sistema mediante un usuario y contraseña, datos que serán únicos para el usuario.

### Caso de uso Nro. 2: Salida del sistema

**Descripción:**

El caso de uso describe los pasos a seguir para que el usuario pueda salir del sistema.

**Identificación de los actores y casos de uso:**

**Actores**

* Usuario
* Sistema web

**Casos de Uso**

* Seleccionar opción de cierre de sesión
* Confirmar cierre de sesión
* Pantalla de inicio de sesión

**Diagrama de caso de Uso**



Figura. Caso de Uso – Salida del Sistema

**Flujo de caso de uso 2**

|  |  |
| --- | --- |
| **CASO DE USO** | Salida del Sistema |
| **ACTORES** | Usuario Sistema Web |
| **PROPOSITO** | Cerrar Sesión del Sistema |
| **TIPO** | Alta |
| **CURSO NORMAL DE EVENTOS** | |
| ACCIONES DE ACTORES | RESPUESTA DEL SISTEMA |
| 1. El caso de uso inicia cuando el actor desea salir del sistema. | 2. El sistema pedirá confirmación del actor para el cierre de sesión.  3. El sistema mostrara la página de inicio de sesión. |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

**Diagrama de Secuencia: Salida del Sistema**

Diagrama de secuencia que muestra la interacción del usuario y los componentes para la salida del sistema.



Figura. Diagrama de Secuencia – Salida de Sistema

**Detalle de Diagrama de Secuencias: Salida del Sistema**

|  |  |  |  |
| --- | --- | --- | --- |
| Origen del Mensaje | Destino del Mensaje | Mensaje | Descripción |
| Usuario | Pantalla de Cierre | mostrar opciones de Usuario | Usuario Selecciona la acción deseada |
| Pantalla de cierre | Usuario | Seleccionar Logout | El usuario selecciona la opción de Logout |
| Usuario | Controlador de Login | Confirmación de cierre de sesión | Enviar solicitud de cierre de sesión |
| Controlador de Login | Base de Datos del sistema | Validar Cierre | El controlador hace una solicitud ORM para el cierre de sesión |
| Base de Datos del sistema | Pantalla de Cierre | Enviar Autorización | La base de datos del sistema envía la autorización de Logout |
| Pantalla de cierre | Controlador de Login | Logout  correcto| incorrecto | Muestra la pantalla de Login si el cierre de sesión fue correcto |

Tabla. Diagrama de Secuencias: Salida del Sistema

**Diagrama de Actividades: Salida del Sistema**

Diagrama

Descripción generada automáticamente

Figura. Diagrama de Actividades – Salida del Sistema

El diagrama de actividades para el caso de uso salida del sistema cuenta con dos carriles:

* Sistema: Carril que representa todo es sistema web y las actividades que serán realizados por el mismo.
* Usuario: Carril que muestra las actividades realizadas por el usuario en el sistema web.

El usuario podrá salir del sistema.

### Caso de uso Nro. 3: Marcado Esteganográfico

**Descripción:**

El caso de uso describe los pasos a seguir para que el usuario pueda realizar el marcado Esteganográfico.

**Identificación de los actores y casos de uso:**

**Actores**

* Usuario
* Sistema web

**Casos de Uso**

* Seleccionar opción de marcado Esteganográfico
* Selección del documento a marcar
* Seleccionar uso de alias
* Marcar documento
* Pantalla de descarga
* Descarga de documento

**Diagrama de caso de Uso 3**



Figura. Caso de Uso – Marcado Esteganográfico

**Flujo del Caso de Uso 3**

|  |  |
| --- | --- |
| **CASO DE USO** | Marcado Esteganográfico |
| **ACTORES** | Usuario Sistema Web |
| **PROPOSITO** | Marcado de Documentos |
| **TIPO** | Alta |
| **CURSO NORMAL DE EVENTOS** | |
| ACCIONES DE ACTORES | RESPUESTA DEL SISTEMA |
| 1. El caso de uso inicia cuando el actor use el marcado Esteganográfico.  3. Seleccionar el documento a marcar.  5. Seleccionar el alias para el marcado y marcar.  7. Descargar el Documento. | 2. El sistema mostrara la pantalla del listado de documentos para que el actor seleccione el documento que quiera marcar.  4. El sistema mostrara la pantalla del listado de alias.    6. El sistema marcara el documento y mostrara la pantalla con el documento marcado listo para descargar. |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

**Diagrama de Secuencia: Marcado Esteganográfico**

El siguiente diagrama muestra la interacción entre los componentes para realizar el marcado Esteganográfico.



Figura. Diagrama de Secuencias – Marcado Esteganográfico

**Detalle de Diagrama de Secuencias: Marcado Esteganográfico**

|  |  |  |  |
| --- | --- | --- | --- |
| Origen del Mensaje | Destino del Mensaje | Mensaje | Descripción |
| Usuario | Pantalla Home | Boton Marcado Esteganográfico | Usuario selecciona el boton de Marcado Esteganográfico |
| Pantalla Home | Pantalla de Listado de Documentos | Seleccionar Documento | Selección de documento para ser marcado con el estego-algoritmo |
| Pantalla de Listado de Documentos | Pantalla de Listado Alias/ Marcado Esteganográfico | Seleccionar Alias para el Marcado Esteganográfico | Selección de Alias para usar en el marcado del documento con el estego-algoritmo |
| Pantalla de Listado Alias/Marcado Esteganográfico | Controlador Marcado | Validar Marcado | Solicitud ORM |
| Controlador Marcado | Base de Datos del Sistema | Registrar Marcado | Con las características enviadas por la solicitud ORM realizar el marcado con el estego-algoritmo |
| Base de Datos del Sistema | Controlador Marcado | Notificar al sistema | Notificar al sistema que el marcado se realzo con éxito |
| Controlador Marcado | Pantalla de Descarga de Documento marcado | Mostrar Pantalla de descarga | Sistema envía el Documento marcado |
| Pantalla de Descarga de Documento Marcado | Usuario | Mostrar Botón de descarga del documento | Selección del Botón de descarga |
| Usuario | Pantalla de Descarga | Seleccionar Botón de Descarga del PDF | Descargar PDF marcado con el estego-algoritmo |

Tabla. Detalle de Diagrama de Secuencias: Marcado Esteganográfico

**Diagrama de Actividades: Marcado Esteganográfico**

Diagrama

Descripción generada automáticamente

Figura. Diagrama de Actividades – Marcado Esteganográfico

El diagrama de actividades para el caso de uso Marcado Esteganográfico cuenta con dos carriles:

* Sistema: Carril que representa todo es sistema web y las actividades que serán realizados por el mismo.
* Usuario: Carril que muestra las actividades realizadas por el usuario en el sistema web.

El usuario podrá realizar el marcado Esteganográfico a los documentos de su preferencia.

### Caso de uso Nro. 4: Carga de documentación al sistema

**Descripción:**

En este caso de uso el usuario podrá realizar la carga de documentación para su posterior marcado.

**Identificación de los actores y casos de uso:**

**Actores**

* **Usuario**
* **Sistema web**

**Casos de Uso**

* Seleccionar opción de marcado Esteganográfico
* Seleccionar botón de carga de documentos
* Seleccionar documento
* Mostrar listado de documentos

**Diagrama de caso de Uso 6**



Figura. Caso de Uso – Carga de Documentación al Sistema

**Flujo del Caso de Uso**

|  |  |
| --- | --- |
| **CASO DE USO** | Cargado de documentos |
| **ACTORES** | Usuario Sistema Web |
| **PROPOSITO** | Realizar la carga de documentación al sistema |
| **TIPO** | Alta |
| **CURSO NORMAL DE EVENTOS** | |
| ACCIONES DE ACTORES | RESPUESTA DEL SISTEMA |
| 1. El actor selecciona la opción de Marcado Esteganográfico.  3. El actor cargara el documento al sistema. | 2. El sistema mostrara la pantalla con un listado de documentación cargada al sistema    4. El sistema almacenara el documento cargado en la base de datos.  5. El sistema mostrara loa lista de documentos cargados. |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

**Diagrama de Secuencia: Carga de Documentos**

El siguiente diagrama muestra la interacción entre los componentes para el cargado de documentos al sistema.



Figura. Diagrama de Secuencias – Carga de Documentación al Sistema

**Detalle de** **Diagrama de Secuencias: Carga de Documentos**

|  |  |  |  |
| --- | --- | --- | --- |
| Origen del Mensaje | Destino del Mensaje | Mensaje | Descripción |
| Usuario | Pantalla Home | Botón Marcado Esteganográfico | Usuario selecciona el botón de Marcado Esteganográfico |
| Pantalla Home | Pantalla de Listado de Documentos | Botón de Carga de Documentos | Usuario selecciona el botón de Carga de Documentos |
| Pantalla de Listado de Documentos | Control de Documentos | Seleccionar Documento | Usuario selecciona el documento que desea cargar al sistema |
| Controlador de Documentos | Base de Datos del Sistema | Enviar Validación | El controlador envía una solicitud de validación ORM a la base de datos |
| Base de Datos del Sistema | Base de Datos del Sistema | Registrar Documento | La Base de Datos del Sistema almacena el documento en la BD |
| Base de Datos del Sistema | Pantalla de Listado de Documentos | Notificar Registro | Notificar al usuario que el documento fue guardado correctamente |
| Pantalla de Listado de Documentos | Usuario | Mostrar Pantalla de listado de Documento | Sistema muestra pantalla de Listado de documentos con el documento del usuario |

Tabla. Detalle de Diagrama de Secuencias: Carga de Documentos

**Diagrama de Actividades: Carga de Documentos**

Diagrama

Descripción generada automáticamente

Figura. Diagrama de Actividades – Carga Documentación al Sistema

El diagrama de actividades para el caso de uso carga de documentos cuenta con dos carriles:

* Sistema: Carril que representa todo es sistema web y las actividades que serán realizados por el mismo.
* Usuario: Carril que muestra las actividades realizadas por el usuario en el sistema web.

El usuario podrá realizar la carga de documentos al sistema para su posterior marcado.

### Caso de uso Nro. 5: Análisis de Imagen capturada

**Descripción:**

En este caso el usuario podrá realizar el análisis de la imagen en el sistema.

**Identificación de los actores y casos de uso:**

**Actores**

* Usuario
* Sistema web

**Casos de Uso**

* Seleccionar opción de análisis de imagen
* Cargar imagen para analizar
* Analizar Imagen filtrado
* Resultado del análisis

**Diagrama de caso de Uso 4**



Figura. Caso de Uso – Análisis de Imagen

**Flujo de caso de uso**

|  |  |
| --- | --- |
| **CASO DE USO** | Análisis de imagen |
| **ACTORES** | Usuario Sistema Web |
| **PROPOSITO** | Realizar el análisis de una imagen |
| **TIPO** | Alta |
| **CURSO NORMAL DE EVENTOS** | |
| ACCIONES DE ACTORES | RESPUESTA DEL SISTEMA |
| 1.El actor selecciona la opción de Analizar imagen en el sistema.  3.El actor cargara la imagen para ser analizada. | 2.El sistema mostrara la pantalla para cargar la imagen.  4.El sistema analizara y verificara si la imagen pertenece a un documento que fue divulgado sin autorización.  5.El sistema mostrara una pantalla con la información obtenida del análisis. |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

**Diagrama de Secuencias: Análisis de Imagen**

El siguiente diagrama muestra la interacción entre el usuario y los componentes para realizar el análisis de una imagen, para verificar si el documento fue divulgado sin autorización.



Figura. Diagrama de Secuencias – Análisis de Imagen

**Detalle de Diagrama de Secuencias: Análisis de Imagen**

|  |  |  |  |
| --- | --- | --- | --- |
| Origen del Mensaje | Destino del Mensaje | Mensaje | Descripción |
| Usuario | Pantalla Home | Botón Análisis de Imagen | Usuario selecciona el botón de Análisis de Imagen |
| Pantalla Home | Pantalla de Carga de Imagen | Botón de Carga de Imagen | Usuario selecciona el botón de Carga de Imágenes |
| Pantalla de Carga de Imagen | Controlador de Marcado | Seleccionar Imagen | Usuario selecciona la imagen que desea cargar al sistema |
| Controlador de Marcado | Base de Datos del Sistema | Enviar Verificación | El controlador envía una solicitud de verificación ORM a la base de datos |
| Base de Datos del Sistema | Base de Datos del Sistema | Validación de imagen con el estego-algoritmo | La Base de Datos del Sistema almacena la imagen en la BD |
| Base de Datos del Sistema | Pantalla de Resultados | Notificar Resultados | Notificar al sistema de los resultados del análisis de la imagen con el estego-algoritmo |
| Pantalla de Resultados | Usuario | Mostrar Pantalla de Resultado del Análisis de la Imagen Cargada | Sistema muestra pantalla de Resultados con la información correspondiente del análisis de la imagen |

Tabla. Detalle de Diagrama de Secuencias: Análisis de Imagen

**Diagrama de Actividades: Análisis de Imagen**

Diagrama

Descripción generada automáticamente

Figura. Diagrama de Actividades – Análisis de Imagen

El diagrama de actividades para el caso de uso análisis de imagen cuenta con dos carriles:

* Sistema: Carril que representa todo es sistema web y las actividades que serán realizados por el mismo.
* Usuario: Carril que muestra las actividades realizadas por el usuario en el sistema web.

El usuario podrá realizar el análisis de una imagen para verificar si el documento fue divulgado sin autorización.

### Caso de uso Nro. 6: Análisis de PDF

**Descripción:**

En este caso el usuario podrá realizar el análisis del PDF en el sistema.

**Identificación de los actores y casos de uso:**

**Actores**

* Usuario
* Sistema web

**Casos de Uso**

* Seleccionar opción de análisis de PDF
* Cargar PDF para analizar
* Analizar PDF filtrado
* Resultado del análisis

**Diagrama de caso de Uso 5**



Figura. Caso de Uso – Análisis de PDF

**Flujo de Caso de Uso**

|  |  |
| --- | --- |
| **CASO DE USO** | Análisis de PDF |
| **ACTORES** | Usuario Sistema Web |
| **PROPOSITO** | Realizar el análisis de un PDF |
| **TIPO** | Alta |
| **CURSO NORMAL DE EVENTOS** | |
| ACCIONES DE ACTORES | RESPUESTA DEL SISTEMA |
| 1. El actor selecciona la opción de Analizar PDF en el sistema.  3. El actor cargara el PDF para ser analizado. | 2. El sistema mostrara la pantalla para cargar el PDF.  4. El sistema analizara y verificara si el PDF es un documento que fue divulgado sin autorización.  5. El sistema mostrara una pantalla con la información obtenida del análisis. |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

**Diagrama de Secuencias: Análisis de PDF**

El siguiente diagrama muestra la interacción entre los componentes para realizar el análisis de un PDF, para verificar si el documento fue divulgado sin autorización.



Figura. Diagrama de Secuencias – Análisis de PDF

**Detalle de Diagrama de Secuencias: Análisis de PDF**

|  |  |  |  |
| --- | --- | --- | --- |
| Origen del Mensaje | Destino del Mensaje | Mensaje | Descripción |
| Usuario | Pantalla Home | Botón Análisis de PDF | Usuario selecciona el botón de Análisis de PDF |
| Pantalla Home | Pantalla de Carga de PDF | Botón de Carga de PDF | Usuario selecciona el botón de Carga de PDF |
| Pantalla de Carga de PDF | Controlador de Marcado | Seleccionar PDF | Usuario selecciona la PDF que desea cargar al sistema |
| Controlador de Marcado | Base de Datos del Sistema | Enviar Verificación | El controlador envía una solicitud de verificación ORM a la base de datos |
| Base de Datos del Sistema | Base de Datos del Sistema | Validación de PDF con el estego-algoritmo | La Base de Datos del Sistema almacena el PDF en la BD |
| Base de Datos del Sistema | Pantalla de Resultados | Notificar Resultados | Notificar al sistema de los resultados del análisis del PDF con el estego-algoritmo |
| Pantalla de Resultados | Usuario | Mostrar Pantalla de Resultado del Análisis del PDF Cargada | Sistema muestra pantalla de Resultados con la información correspondiente del análisis del PDF |

Tabla. Detalle de Diagrama de Secuencias: Análisis de PDF

**Diagrama de Actividades: Análisis de PDF**

Diagrama

Descripción generada automáticamente

Figura. Diagrama de Actividades – Marcado Esteganográfico

El diagrama de actividades para el caso de uso análisis de PDF cuenta con dos carriles:

* Sistema: Carril que representa todo es sistema web y las actividades que serán realizados por el mismo.
* Usuario: Carril que muestra las actividades realizadas por el usuario en el sistema web.

El usuario podrá realizar el análisis de un PDF para verificar si el documento fue divulgado sin autorización.

### Caso de uso Nro. 7: Alta alias

**Descripción:**

En este caso de uso el usuario podrá crear un alias de marcado en el sistema.

**Identificación de los actores y casos de uso:**

**Actores**

* Usuario
* Sistema web

**Casos de Uso**

* Seleccionar opción de crear alias
* Ingresar los datos requeridos
* Confirmar nuevo alias
* Mostrar lista de alias

**Diagrama de caso de Uso 7**



Figura. Caso de Uso – Alta de Alias

**Flujo de Caso de Uso**

|  |  |
| --- | --- |
| **CASO DE USO** | Alta de Alias |
| **ACTORES** | Usuario Sistema Web |
| **PROPOSITO** | Realizar alta de Alias |
| **TIPO** | Alta |
| **CURSO NORMAL DE EVENTOS** | |
| ACCIONES DE ACTORES | RESPUESTA DEL SISTEMA |
| 1. El actor selecciona la opción de crear alias.  3. Ingresar los datos requeridos.  4. Confirma nuevo alias | 2. El sistema mostrará la pantalla con un  formulario que requerirá datos para la creación del nuevo alias.  5. El sistema mostrara la lista de Alias |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

**Diagrama de Secuencia: Alta de Alias**

El siguiente diagrama muestra la relación de los componentes para realizar el alta de alias en el sistema necesario para el marcado Esteganográfico.



Figura. Diagrama de Secuencia – Alta Alias

**Detalle de Diagrama de Secuencias: Alta de Alias**

|  |  |  |  |
| --- | --- | --- | --- |
| Origen del Mensaje | Destino del Mensaje | Mensaje | Descripción |
| Usuario | Pantalla Home | Botón Alta Alias | Usuario desea dar Alta Alias |
| Pantalla Home | Pantalla de Listado de Alias | Botón de Alta de Alias | Usuario selecciona el botón de Listado de Alias |
| Pantalla de listado de Alias | Pantalla de Creación de Alias | Ingresar Datos al formulario para la creación del nuevo Alias | Usuario ingresa los datos necesarios para la creación del nuevo alias |
| Pantalla de creación de Alias | Controlador del Alias | Enviar Solicitud | \*\*\* |
| Controlador del Alias | Base de Datos del Sistema | El controlador envía una solicitud de validación ORM a la base de datos | La Base de Datos del Sistema almacena el nuevo Alias |
| Base de Datos del Sistema | Pantalla de Listado de Alias | Notificar registro de nuevo alias | Notificar al sistema que el nuevo alias ha sido registrado |
| Pantalla de Listado de Alias | Usuario | Mostrar Pantalla Listado de Alias | Sistema muestra pantalla de Listado de Alias con el nuevo alias como opción |

Tabla. Detalle de Diagrama de Secuencias: Alta de Alias

**Diagrama de Actividades: Alta de Alias**

Diagrama

Descripción generada automáticamente

Figura. Diagrama de Actividades – Alta de Alias

El diagrama de actividades para el caso de uso alta de alias cuenta con dos carriles:

* Sistema: Carril que representa todo es sistema web y las actividades que serán realizados por el mismo.
* Usuario: Carril que muestra las actividades realizadas por el usuario en el sistema web.

El usuario podrá realizar alta de un nuevo alias de acuerdo con las necesidades del usuario.

### Caso de uso Nro. 8: Modificación de alias

**Descripción:**

En este caso de uso el usuario podrá editar un alias de marcado en el sistema.

**Identificación de los actores y casos de uso:**

**Actores**

* Usuario
* Sistema web

**Casos de Uso**

* Seleccionar alias a editar
* Ingresar los datos a editar
* Confirmar cambios

**Diagrama de caso de Uso 8**



Figura. Caso de Uso – Modificación Alias

**Flujo de Caso de Uso**

|  |  |
| --- | --- |
| **CASO DE USO** | Modificación de Alias |
| **ACTORES** | Usuario Sistema Web |
| **PROPOSITO** | Realizar modificación de Alias |
| **TIPO** | Alta |
| **CURSO NORMAL DE EVENTOS** | |
| ACCIONES DE ACTORES | RESPUESTA DEL SISTEMA |
| 1. El actor selecciona la opción de editar alias.  3. Ingresar los datos requeridos.  4. Confirma la modificación del alias | 2. El sistema mostrará la pantalla con un  formulario que requerirá datos para la modificación del alias.  5. El sistema mostrara la lista de Alias |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

**Diagrama de Secuencia: Modificación de Alias**

El siguiente diagrama muestra la relación de los componentes para realizar la modificación de alias en el sistema necesario para el marcado Esteganográfico.



Figura. Diagrama de Secuencia – Modificación Alias

**Detalle de Diagrama de Secuencias: Modificación de Alias**

|  |  |  |  |
| --- | --- | --- | --- |
| Origen del Mensaje | Destino del Mensaje | Mensaje | Descripción |
| Usuario | Pantalla Home | Listado Alias | Usuario desea modificar Alias |
| Pantalla Home | Pantalla de Listado de Alias | Seleccionar Alias a modificar | Usuario elige que Alias a modificar. |
| Pantalla de listado de Alias | Pantalla de modificación de Alias | Enviar solicitud de cambio/modificación de Alias | Usuario ingresa los datos necesarios para la modificación del nuevo alias |
| Pantalla de modificación de Alias | Controlador del Alias | Enviar solicitud de cambio/modificación de Alias | Se envía la solicitud ORM |
| Controlador del Alias | Base de Datos del Sistema | El controlador envía una solicitud de validación ORM a la base de datos | La Base de Datos del Sistema registra el cambio/modificación del Alias |
| Base de Datos del Sistema | Pantalla de Listado de Alias | Notificar registro de la modificación del alias | Notificar al sistema el alias ha sido modificado correctamente |
| Pantalla de Listado de Alias | Usuario | Mostrar Pantalla Listado de Alias | Sistema muestra pantalla de Listado de Alias con el cambio del alias ya modificado |

Tabla. Detalle de Diagrama de Secuencias: Modificación de Alias

**Diagrama de Actividades: Modificación de Alias**

Diagrama

Descripción generada automáticamente

Figura. Diagrama de Actividades – Modificación Alias

El diagrama de actividades para el caso de uso modificación de alias cuenta con dos carriles:

* Sistema: Carril que representa todo es sistema web y las actividades que serán realizados por el mismo.
* Usuario: Carril que muestra las actividades realizadas por el usuario en el sistema web.

El usuario podrá realizar la modificación de cualquier alias de acuerdo con las necesidades del usuario.

### Caso de uso Nro. 9: Baja de alias

**Descripción:**

En este caso de uso el usuario podrá eliminar un alias en el sistema.

**Identificación de los actores y casos de uso:**

**Actores**

* Usuario
* Sistema web

**Casos de Uso**

* Seleccionar alias a eliminar
* Confirmar eliminación
* Mostrar lista de alias

**Diagrama de caso de Uso 9**

****

Figura. Caso de Uso – Baja Alias

**Flujo del Caso de Uso**

|  |  |
| --- | --- |
| **CASO DE USO** | Baja de Alias |
| **ACTORES** | Usuario Sistema Web |
| **PROPOSITO** | Realizar baja de Alias |
| **TIPO** | Alta |
| **CURSO NORMAL DE EVENTOS** | |
| **ACCIONES DE ACTORES** | **RESPUESTA DEL SISTEMA** |
| 1. El actor selecciona la opción de eliminar alias. 2. Confirma la modificación del alias | 1. El sistema mostrara la lista de Alias |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

**Diagrama de Secuencia: Baja de Alias**

El siguiente diagrama muestra la relación de los componentes para realizar la baja de alias en el sistema.



Figura. Diagrama de Secuencias – Baja Alias

**Detalle de Diagrama de Secuencias: Baja de Alias**

|  |  |  |  |
| --- | --- | --- | --- |
| Origen del Mensaje | Destino del Mensaje | Mensaje | Descripción |
| Usuario | Pantalla de Listado de Alias | Listado Alias | Usuario desea dar de baja un Alias |
| Pantalla de Listado de Alias | Controlador del Alias | Seleccionar Alias a borrar | Usuario elige que Alias desea borrar |
| Controlador del Alias | Base de Datos del Sistema | El controlador envía una solicitud de validación ORM a la base de datos | Se envía la solicitud ORM |
| Base de Datos del Sistema | Controlador del Alias | Registrar Cambio de estado en el alias correcto | La Base de Datos del Sistema registra la baja del Alias |
| Base de Datos del Sistema | Pantalla de Listado de Alias | Notificar de baja del alias | Notificar al sistema el alias ha sido borrado correctamente |
| Pantalla de Listado de Alias | Usuario | Mostrar Pantalla Listado de Alias | Sistema muestra pantalla de Listado de Alias con el cambio del alias ya borrado |

Tabla. Detalle de Diagrama de Secuencias: Baja de Alias

**Diagrama de Actividades: Baja de Alias**

Diagrama

Descripción generada automáticamente

Figura. Diagrama de Actividades – Baja Alias

El diagrama de actividades para el caso de uso modificación de alias cuenta con dos carriles:

* Sistema: Carril que representa todo es sistema web y las actividades que serán realizados por el mismo.
* Usuario: Carril que muestra las actividades realizadas por el usuario en el sistema web.

El usuario podrá realizar la baja de cualquier alias de acuerdo con las necesidades del usuario.

## Diagrama de Despliegue

Un Diagrama de Despliegue modela la arquitectura en tiempo de ejecución de un sistema. Esto muestra la configuración de los elementos de hardware (nodos) y muestra cómo los elementos y artefactos del software se trazan en esos nodos (Sparx System, 2021).

La siguiente imagen representa el diagrama de despliegue del sistema que cuenta con los siguiente componentes base de datos PostgreSQL, un ordenador en el que se ejecuta la aplicación del marcado Esteganográfico, servidor web Apache que contiene el framework Flask, función de marcado Esteganográfico y Pytesseract OCR.



Figura. Diagrama de Despliegue del Sistema Web

## Diseño Base de Datos Entidad-Relación

A continuación, se muestra el diseño de la Base de Datos de entidad-relación del sistema web que se aplicara para el manejo de los datos que ingresan al sistema, así como los resultados del marcado Esteganográfico.

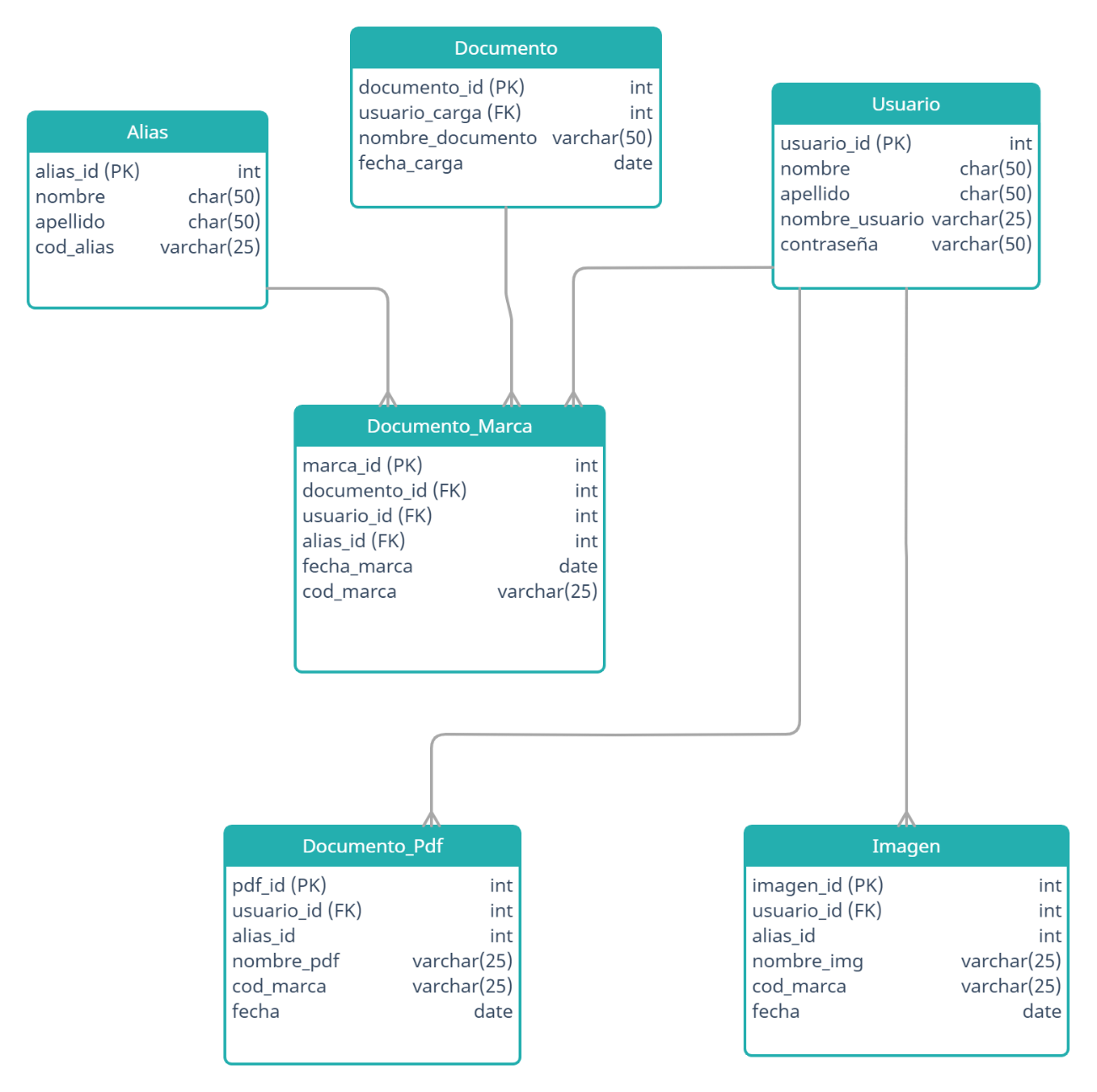


Figura. Diseño Base de Datos Entidad – Relación

.

## Diagrama de Clases

A continuación, se presenta el diagrama de clases correspondiente al sistema web.



Figura. Diagrama del Sistema Web.

# Modelo C4

El modelo C4 es un enfoque para diagramar la arquitectura del software, basado en abstracciones que reflejan cómo los arquitectos y desarrolladores de software piensan y construyen el software.

## Diagrama de Contexto

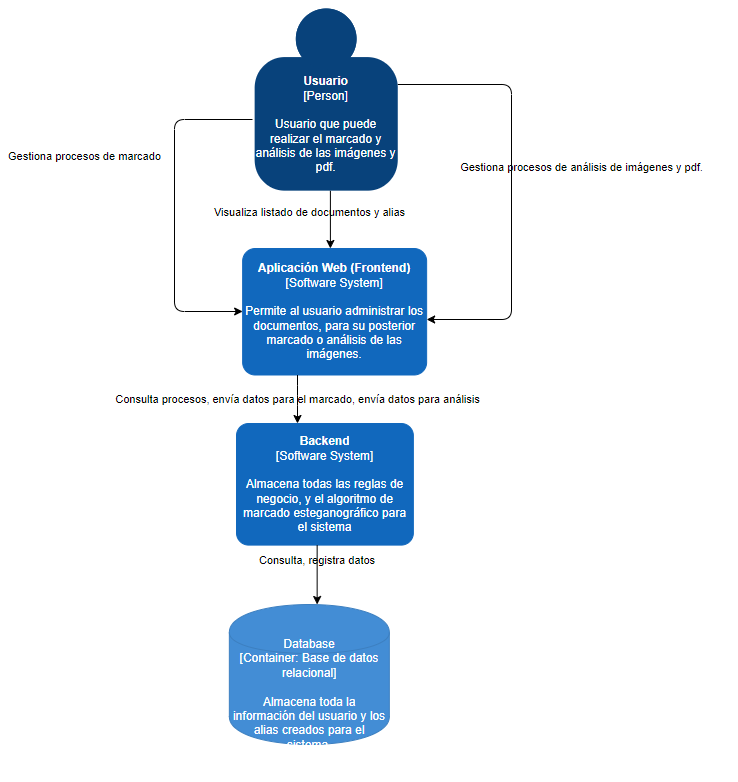


Figura. Diagrama de Contexto

El diagrama de contexto muestra cuales son las funcionalidades que puede utilizar el usuario en el sistema web, las cuales son la gestión del proceso de marcado, visualización del estado de documentos y alias, gestionar los procesos de imágenes y PDF. En el backend se almacenará todas las reglas de negocio del sistema es decir que se encontraran lo9s controladores.

## Diagrama de Contenedores

Diagrama

Descripción generada automáticamente

Figura. Diagrama de Contenedores

La aplicación web es una aplicación desarrollada con JavaScript, HTML5 y CSS3, la cual se ejecuta en el navegador web proporcionando todas las interfaces del sistema web. El backend es una aplicación desarrollada con el framework Flask construida con base en Python la cual se comunica con la aplicación web, además de obtener información del usuario de la base de datos relacional

## Diagrama de Componentes

Diagrama

Descripción generada automáticamente

Figura. Diagrama de Componentes

La aplicación del backend los cuales proporcionan puntos de acceso a la base de datos a través de los controladores del sistema.

# Definición de criterios para el desarrollo del proyecto

**BIBLIOGRAFÍA**

Bezral Fernando (2018), *“REDES NEURONALES & DEEP LEARNING”* editorial Fernando Bezral

Benítez, Escudero, Kanaan (2013), “*Inteligencia artificial avanzada”* Universitat Oberta de Catalunya.

Russell, Norvig (2004), *“INTELIGENCIA ARTIFICIAL UN ENFOQUE MODERNO”* Segunda Edición, PEARSON EDUCACION, S.A. Madrid.

LARMAN, C (2003). “*UML Y PATRONES. Una introducción al análisis y diseño orientado a objetos y al proceso unificado”*. Segunda edición PEARSON EDUCACIÓN, S.A., Madrid,

Alberto, Airas (2020) “Esteganografía en archivos digitales - Digital File Steganography” Recuperado 25 de agosto del 2022. https://eprints.ucm.es/id/eprint/68341/1/ARIAS%20DEL%20CERRO%2083236\_ALBERTO\_GERMAN\_ARIAS\_DEL\_CERRO\_Esteganografia\_en\_archivos\_digitales\_1006096\_860402980.pdf.

Ibarra, Herrera, Bojacá (2019) “Análisis de la viabilidad de un estándar de seguridad que implemente técnicas de esteganografía y criptografía para aumentar la seguridad de la información en las empresas” Recuperado 30 de agosto del 2022. http://repository.unipiloto.edu.co/handle/20.500.12277/5919?show=full

Deitel, P. y Deitel, H. Java Como programar. 9na ed. Pearson Educación, México 2012.

Java, Consultado 3 de septiembre del 2022 https://www.java.com/es/download/help/whatis\_java.html

React Documentación, Consultado 1 de septiembre del 2022: https://es.reactjs.org/docs/getting-started.html

Castillo, Luciano (2012) “Conociendo GitHub” Recuperado el 2 de septiembre del 2022https://conociendogithub.readthedocs.io/en/latest/data/introduccion/

Docker, (2018) “Docker Overview” Recuperado el 2 de septiembre 2022 https://docs.docker.com/get-docker/

Robledano, Angel (2019) “Qué es MySQL: Características y ventajas” Recuperado el 4 de septiembre del 2022: https://openwebinars.net/blog/que-es-mysql/

Microsoft, (2018) “Visual Studio Code” Recuperado el 5 de septiembre 2022: https://code.visualstudio.com/docs

Donozo, Alvaro, (2014) “Tecnologías Emergentes 1” Recuperado el 5 de septiembre del 2021: http://alvarodonozo.blogspot.com/2014/08/bpwin.html

Python, “Python” Recuperado el 5 de septiembre del 2022: <https://www.python.org/>

Diaz Jesús, (2015) “Esteganografía y estegoanálisis básicos” recuperado el 26 de septiembre 2022: <https://www.incibe-cert.es/blog/esteganografia-y-estegoanalisis-basicos>

Numerentur, (2020) “FA Softmax” Recuperado el 20 de septiembre de 2022: <https://numerentur.org/funcion-de-activacion-softmax/>

Numerentur, (2019) “Backpropagation-BP” Recuperado el 19 de septiembre de 2022: <https://numerentur.org/backpropagation/>

Gonzalo Alvaro, (2020), “¿Qué es el sobreajuste u overfitting y por qué debemos evitarlo?” Recuperado el 23 de septiembre de 2022: <https://machinelearningparatodos.com/que-es-el-sobreajuste-u-overfitting-y-por-que-debemos-evitarlo/>

Sparx Systems, (2021), “Tutoriales de UML” Recuperado el 5 de septiembre de 2022:

[Enterprise Architect - Recursos (sparxsystems.com.ar)](http://www.sparxsystems.com.ar/resources/index.php)

IBM Cloud Education, (2022), “What Is Optical Character Recognition (OCR)?” Recuperado el 3 de noviembre de 2022: <https://www.ibm.com/cloud/blog/optical-character-recognition>

Marin, Rafael (2020), *“¿Qué es OpenCV? Instalación en Python y ejemplos básicos*” Recuperado el 18 de noviembre de 2022: <https://www.inesem.es/revistadigital/informatica-y-tics/opencv/>

Parthasarathy, Balaaji (2018), “*Build your own OCR(Optical Character Recognition) for free*” Recuperado el 15 de noviembre de 2022: <https://medium.com/@balaajip/optical-character-recognition-99aba2dad314>

[Hoffstaetter](https://github.com/h), Samuel (2020), “Python Tesseract” Recuperado el 19 de noviembre de 2022: <https://github.com/madmaze/pytesseract>

Tesseract-ocr, (2020), “*Command Line Usage Tesseract documentation*” Recuperado el 19 de noviembre de 2022: <https://tesseract-ocr.github.io/tessdoc/Command-Line-Usage.html>

Klippa, (2022), *“¿Qué es OCR y para qué sirve?*” Recuperado el 22 de noviembre de 2022: <https://www.klippa.com/es/blog/informativo/que-es-ocr/>

Boiangiu, Costin-Anton & Ioanitescu, Radu & Dragomir, Razvan-Costin. (2016). “VOTING-BASED OCR SYSTEM.” Recuperado el 10 de octubre 2022: https://www.researchgate.net/publication/326016983\_VOTING-BASED\_OCR\_SYSTEM

OpenCv Org (2019), “*OpenCV Tutorials*” Recuperado el 12 de enero de 2022 <https://docs.opencv.org/>

Dar-Shyang Lee & Ray Smith (2012) “*Improving Book OCR by Adaptive Language and Image Models*” recuperado el 4 de septiembre de 2022: https://tesseract-ocr.github.io/docs/Improving\_Book\_OCR\_by\_Adaptive\_Language\_and\_Image\_Models.pdf

1. Alberto Arias, (2020) Disponible en: https://eprints.ucm.es/id/eprint/ARIAS [↑](#footnote-ref-1)
2. Ibarra, Herrera (2019) Disponible en: http://repository.unipiloto.edu.co/bitstream/handle/ [↑](#footnote-ref-2)