

**UNIVERSIDAD BOLIVIANA DE INFORMÁTICA
SUB SEDE CIUDAD DE LA PAZ
CARRERA INGENIERÍA DE SISTEMAS**



PROYECTO DE GRADO

**SISTEMA DE GESTIÓN PARA LA LOCALIZACIÓN DE
CORRESPONDENCIA MEDIANTE EL ALGORITMO DE BÚSQUEDA
SEMÁNTICA**

CASO: CASO: FEDERACIÓN DEPARTAMENTAL DE LA PAZ

***Postulante:** Alejandra Isabella Parraga Cruz*

**PROYECTO DE GRADO PARA OBTENER EL
GRADO ACADÉMICO DE LICENCIATURA
EN INGENIERÍA DE SISTEMAS**

La Paz – Bolivia

2024

ÍNDICE

CAPITULO I.....	1
1. FUNDAMENTOS DE LA INVESTIGACIÓN	1
1.1. Antecedentes Institucionales	1
1.1.2. Visión	2
1.1.3. Organigrama.....	2
1.1.4. Antecedentes Temáticos	3
1.2. Problema.....	5
1.2.1. Planteamiento del Problema.....	5
1.2.2. Formulación del Problema	7
1.3. Objetivos	7
1.3.1. Objetivo General.....	7
1.3.2. Objetivos Específicos	7
1.4. Justificación.....	7
1.4.1. Justificación Económica	7
1.4.2. Justificación Social	8
1.4.3. Justificación Técnica.....	8
1.5. Límites Y Alcances	9
1.5.1. Límites.....	9
1.5.2. Alcances.....	10
1.6. Novedad Científica	10
CAPÍTULO II.....	12
2. MARCO TEÓRICO	12
2.1. Conceptos básicos de gestión documental.....	12
2.1.1. Definición de documento	12
2.1.2. Archivística	12
2.1.3. Archivo	12
2.1.4. Documento de Archivo	13
2.2. Gestión Documental	13
2.3. Correspondencia	14
2.3.1. Gestión de Correspondencia	14
2.3.2. División de la correspondencia	15

2.3.3.	Tipos de Correspondencia Gestionada.....	15
2.3.4.	Proceso de Gestión de Correspondencia	16
2.4.	Tecnologías de la Información	19
2.4.1.	Sistema de Información	19
2.4.2.	Tecnología de Información al servicio de la archivística	20
2.4.3.	Importancia de la tecnología en la gestión documental.....	21
2.5.	Inteligencia Artificial y Algoritmos de Búsqueda Semántica	21
2.5.1.	Inteligencia Artificial	21
2.5.2.	Beneficios de la Inteligencia Artificial	22
2.5.3.	Machine Learning (ML).....	23
2.5.4.	Algoritmo	23
2.5.5.	Semántica	24
2.5.6.	Procesamiento del Lenguaje Natural	24
2.5.7.	Algoritmo de Búsqueda	26
2.5.8.	Búsqueda Semántica.....	26
2.5.9.	Reconocimiento Óptico de Caracteres.....	27
2.6.	Metodología ágil para el desarrollo del sistema	28
2.6.1.	Metodología Ágil Scrum.....	28
2.7.	Diseño	34
2.7.1.	Diagrama UML – Lenguaje Unificado de Modelamiento	34
2.8.	Seguridad de Software y Estándar ISO 27002.....	38
2.8.1.	Estándar ISO 27002	38
2.8.2.	Estandar ISO/IEC 25010	38
2.9.	Integración de Tecnología	38
2.9.1.	Python	38
2.9.2.	Django.....	39
2.9.3.	API	39
2.9.4.	Gestor de Base de Datos PostGreSQL.....	39
2.9.5.	React.....	39
2.9.6.	Visual Studio Code	40
2.9.7.	Git Hub	40
CAPITULO III.....		42
3.	MARCO METODOLÓGICO	42

3.1.	Tipo de Investigación	42
3.1.1.	Investigación Descriptiva	42
3.2.	Diseño de la investigación	42
3.2.1.	Cronograma Detallado.....	43
3.3.	Selección de la Población.....	44
3.3.1.	Universo	44
3.3.2.	Muestra	45
3.3.3.	Método	45
3.4.	Instrumento de Recolección de Datos	46
3.4.1.	Entrevistas.....	46
3.4.2.	Observación	46
3.4.3.	Encuestas.....	46
	Bibliografía.....	47

CAPITULO I

FUNDAMENTOS DE LA INVESTIGACIÓN

1.1. Antecedentes Institucionales

La Federación Departamental de La Paz desempeña un papel crucial en la defensa de los derechos laborales de sus afiliados, buscando constantemente mejorar sus condiciones de trabajo. Debido a la importancia de esta labor, surge la necesidad de adoptar soluciones innovadoras que optimicen la gestión de correspondencia y mejoren la operatividad de la organización porque es su principal medio de comunicación.

Fundada el 2 de octubre de 1991 bajo la Personería Jurídica Nro. 214798 en la ciudad de La Paz, la Federación es una organización sindical que representa a más de 1,000 trabajadores pertenecientes a empresas como ENDE SYC, DELAPAZ, AXS, COTEL e Hidroeléctrica Boliviana, cada una de estas entidades tiene su propio sindicato de trabajadores, afiliado a la Federación Departamental La Paz.

La Federación Departamental La Paz está afiliada a la Central Obrera Boliviana, la Confederación Nacional y la Central Obrera Departamental de La Paz, lo que refuerza su capacidad de representación en decisiones y negociaciones que afectan a sus afiliados en diferentes niveles, incluyendo negociaciones y acuerdos con el Estado. Esta red de alianzas no solo fortalece su presencia, sino que también le permite abordar de manera más efectiva los desafíos laborales que enfrentan sus miembros.

El área en la que se intervendrá con el proyecto planteado es la gestión de correspondencia, cuyos procesos actualmente se realizan de manera manual. Esto genera dificultades tanto en la gestión como en la localización de la correspondencia, especialmente cuando se requiere con prontitud, además de ocasionar retrasos en los procesos.

1.1.1. Misión

Proteger y promover los derechos laborales de nuestros afiliados, coordinando acciones efectivas con las autoridades laborales y fomentando condiciones de trabajo justas y dignas.

1.1.2. Visión

Ser la organización sindical líder en la defensa de los derechos laborales en el departamento de La Paz, reconocida por su capacidad de gestión, innovación y compromiso con el bienestar de nuestros afiliados.

1.1.3. Organigrama

La estructura organizacional de la Federación Departamental de La Paz se organiza de la siguiente manera: en el primer nivel se encuentra la Dirección General, seguida por la Sub Dirección General en el segundo nivel. En un tercer nivel, se encuentran las siguientes áreas: Finanzas, Relaciones Institucionales, Gestión de Conflictos y Registro, cada una con un solo personal. Además, en este nivel se ubica el área de Supervisión Normativa, compuesta por cinco miembros. Por otro lado, el área Administrativa también está conformada por un solo miembro de personal.

Cada área tiene personal asignado para cumplir con sus respectivas funciones, garantizando una distribución clara de responsabilidades. Esta estructura asegura una operación eficiente y ordenada dentro de la Federación.



Figura 1. Estructura jerárquica Organizativa, Federación Departamental La Paz (2024)

1.1.4. Antecedentes Temáticos

Tabla 1

Tecnologías semánticas para realizar búsquedas de documentos, empresa Lyris IT S.A.S.

AUTORES	Caliusco, María Laura Martínez, Agustín Brusa, Graciela María
TÍTULO	Módulo basado en tecnologías semánticas para realizar búsquedas de documentos en la plataforma Signar
OBJETIVO	El objetivo del presente trabajo es mostrar la incorporación de tecnologías semánticas para realizar búsquedas de documentos. En este caso, esa búsqueda de documentos es realizada en la plataforma Signar de la empresa Lyris IT S.A.S. Este trabajo muestra cómo investigaciones que se llevan a cabo en una Universidad aportan valor agregado a empresas para potenciar sus productos y hacerlos más competitivos en el mercado internacional, incorporando tecnologías innovadoras.
ORIGEN	Sociedad Argentina de Informática e Investigación Operativa - octubre 2022

Tabla 2

Motor de Búsqueda Inteligente Thaqhaña.

AUTORES	Felipe Cujar-Rosero David Pinchao Ortiz
TÍTULO	Thaqhaña: Un Motor de Búsqueda Inteligente Basado en Recursos Semánticos
OBJETIVO	El objetivo del trabajo es comparar el rendimiento de un motor de búsqueda semántica, que utiliza una ontología y un modelo de aprendizaje automático (Word2vec y Doc2vec), con un motor de búsqueda manual en el contexto de la búsqueda de proyectos de investigación en la Universidad de Nariño. A través de esta comparación, se busca demostrar que el motor de búsqueda semántica (THAQHAÑA) puede ofrecer resultados más satisfactorios y eficientes en la localización de información relacionada con los proyectos de investigación
ORIGEN	Universidad de Nariño, Colombia - 2021

Tabla 3

DoLaw buscador semántico para la legislación peruana de tecnologías de información.

AUTORES	Diego Augusto OTOYA PAZ
TÍTULO	DoLaw: buscador semántico especializado para la legislación peruana de tecnologías de información
OBJETIVO	Se desarrolla un programa especializado en la búsqueda de documentos relacionados con la legislación peruana de tecnología de la información, utilizando la interpretación semántica de las palabras clave ingresadas por el usuario. El objetivo es permitir búsquedas en el contenido completo de estos documentos, ofreciendo funcionalidades personalizadas que faciliten la búsqueda
ORIGEN	Universidad Nacional Mayor de San Marcos, Lima - Peru 2019

Tabla 4

Sistema Web con inteligencia artificial para el modelo de gestión documental

AUTORES	Fernandez Carranza, Rony Linarez Villodas, Ernesto Juan
TÍTULO	Sistema Web con inteligencia artificial para el modelo de gestión documental para el área de Ingeniería y SIG, Data Services, 2022.
OBJETIVO	Esta investigación tuvo como objetivo determinar la influencia de un sistema web en la gestión documental de Data Services en 2022. El estudio fue de tipo aplicado con un diseño pre-experimental, utilizando una muestra de 203 documentos del área de ingeniería y SIG. El sistema web, que incorpora inteligencia artificial, fue desarrollado con la metodología ágil SCRUM, utilizando PHP, MySQL y TensorFlow.js. La recolección de datos se realizó mediante observación y el uso de una ficha de registro, validada por juicio de expertos. Los resultados muestran un incremento del 39.105% en documentos tramitados digitalmente, un aumento del 21.675% en documentos atendidos oportunamente y una reducción del 52.58% en el consumo de materiales de oficina.
ORIGEN	Universidad César Vallejo, Lima Perú -2022

Tabla 5

Sistema Web de Gestión de Correspondencia

AUTORES	Santos Wilmer Yana Tarqui
TÍTULO	Sistema de Gestión de Correspondencia – Vía Web Caso: Área Ciencias Económicas, Financieras y Administrativas – Universidad Pública de El Alto
OBJETIVO	En un entorno globalizado y tecnológico, las universidades están adoptando software libre. En este contexto, se desarrolló un sistema de Gestión de Correspondencia vía web para el Área de Ciencias Económicas, Financieras y Administrativas de la Universidad Pública de El Alto. Este sistema mejora la gestión administrativa de la correspondencia con módulos para administración de usuarios, consultas y generación de reportes, utilizando PHP con CodeIgniter y MySQL. La metodología de desarrollo es UML Web Engineering (UWE), que modela las etapas del sistema para impulsar el desarrollo integral y sostenible de la sociedad.
ORIGEN	Universidad Pública de El Alto, El Alto Bolivia - 2020

1.2. Problema

1.2.1. Planteamiento del Problema

La Federación Departamental de La Paz enfrenta desafíos significativos en la gestión de su correspondencia, debido a que sus actividades en cuanto al manejo de documentos no son procesos automatizados.

La correspondencia recibida proviene principalmente de los sindicatos de las siguientes instituciones: COTEL, ENDE SYC, DELAPAZ, AXS e Hidroeléctrica Boliviana, todas afiliadas a la Federación Departamental de La Paz. La documentación entrante está mayormente relacionada con quejas y reclamos por el incumplimiento de derechos laborales por parte de las gerencias. Además, se gestiona correspondencia con otras instituciones sobre temas afines o generales. En respuesta a esta documentación, la Federación Departamental de La Paz emite cartas a las entidades involucradas, incluyendo instituciones gubernamentales como el Ministerio de Trabajo, Empleo y Previsión Social, el Ministerio de Energías, el Viceministerio de Hidrocarburos, la Central Obrera Departamental de La Paz, entre otras.

El proceso de gestión documental actualmente es manual. La documentación que ingresa y se remite se registra en un cuaderno dividido en dos secciones: una para los documentos recibidos y otra para los enviados. La sección de documentos recibidos incluye campos como fecha, código de registro, remitente, referencia y número de fojas; mientras que la sección de documentos enviados incluye campos como fecha, número de CITE, destino y referencia. Una vez registrados, los documentos se escanean y se almacenan en la carpeta de documentos del sistema de la computadora, además de archivarse físicamente. La codificación de los documentos que ingresan y se remiten es secuencial. Estos códigos se registran en una hoja numerada, en la cual se resalta el número correspondiente cada vez que se utiliza.

El método actual de gestión de correspondencia está expuesto a errores humanos, como la duplicación del código de registro y del CITE, además de un almacenamiento desorganizado debido a una clasificación poco rigurosa de la documentación, tanto en formato digital como físico. A esto se suma que los documentos ingresados suelen carecer de una referencia clara y específica sobre su contenido, limitándose a una referencia genérica como “para su conocimiento y atención,” lo que se transcribe en el cuaderno o en la descripción del documento escaneado. Esta situación genera demoras considerables cuando se requiere localizar un documento con urgencia, ya sea para reuniones no programadas, para preparar documentación para encuentros con instituciones externas, directivos, o para informar a un afiliado sobre el estado de un documento. Las demoras aumentan cuando los directivos, que no son los encargados directos del archivado, deben buscar documentos en ausencia de la Asistente Ejecutiva, responsable de esta tarea. Además, se observa la falta de seguimiento del estado de los documentos, ya sea si han sido atendidos, están pendientes o han sido archivados. Esto genera que no todos los directivos estén al tanto de la llegada de un documento o de su estado, debido a que, la comunicación interna o la actualización sobre estos procesos puede verse limitada por factores como la falta de un sistema centralizado de registro, dificultades en la comunicación, o la ausencia de notificaciones automáticas que aseguren que todos los directivos reciban la misma información en tiempo real.

Estas limitaciones pueden retrasar la toma de decisiones, la gestión de casos urgentes afectando negativamente la eficiencia de la Federación, como consecuencia, se generan molestias tanto para los afiliados como para los directivos.

Dado que la Federación tiene la responsabilidad de defender los derechos y bienestar laboral de sus afiliados, es crucial el desarrollo de un sistema de gestión documental, complementado

con un algoritmo de búsqueda semántica, que se presenta como una solución viable para optimizar el manejo y localización de la correspondencia.

1.2.2. Formulación del Problema

¿Cómo gestionar los procesos y localización de la correspondencia en la Federación Departamental La Paz?

1.3. Objetivos

1.3.1. Objetivo General

Desarrollar un sistema de gestión para la localización de la correspondencia mediante el algoritmo de búsqueda semántica en la Federación Departamental de La Paz, año 2024.

1.3.2. Objetivos Específicos

- Analizar los procesos actuales de gestión de correspondencia en la Federación Departamental de La Paz.
- Adoptar principios de una metodología ágil para el análisis, diseño y planificación del sistema de gestión de correspondencia.
- Desarrollar la funcionalidad para la digitalización y procesamiento de la información de los documentos.
- Implementar un algoritmo de búsqueda semántica para optimizar la localización de documentos.
- Desarrollar la funcionalidad para la generación de informes y reportes sobre la gestión de correspondencia.

1.4. Justificación

1.4.1. Justificación Económica

El desarrollo de un sistema de gestión de correspondencia para la Federación Departamental de La Paz facilitará el acceso rápido a la documentación. Al digitalizar los documentos, cada ejecutivo y demás miembros podrán acceder a ellos sin necesidad de imprimirlos o archivar copias físicas en carpetas o estanterías. Esto permitirá reducir significativamente los costos de papelería, tinta, mantenimiento de equipos de impresión y espacio de almacenamiento.

Dado que la Federación depende económicamente de los aportes de sus afiliados, la automatización del proceso de gestión de correspondencia, junto con la búsqueda a través de un algoritmo semántico, mejorará la atención al afiliado. Esto aumentará la satisfacción de los miembros y podría atraer a nuevos afiliados, lo que a su vez incrementará los ingresos.

1.4.2. Justificación Social

El desarrollo de un sistema de gestión de correspondencia, acompañado de un algoritmo semántico, permitirá una gestión más eficiente del tiempo del personal al disminuir la corrección de errores y el trabajo redundante. Esto facilitará que el personal se enfoque en tareas estratégicas y en la atención directa a los afiliados, mejorando así la experiencia general de estos últimos. Además, la automatización aliviará la carga laboral, reduciendo el estrés asociado con tareas administrativas, y contribuirá a un ambiente de trabajo más equilibrado y saludable.

La capacitación del personal en nuevas tecnologías no solo promoverá su inclusión digital, sino que también fortalecerá a los empleados al permitirles desarrollar nuevas habilidades. Al mismo tiempo, esta mejora en la atención y en la eficiencia del servicio fomentará la satisfacción de los afiliados, lo que puede fortalecer las relaciones entre la Federación y sus miembros.

1.4.3. Justificación Técnica

El desarrollo del sistema de gestión de correspondencia en la Federación Departamental de La Paz permitirá agilizar la gestión de la documentación recibida y emitida, optimizando su localización, minimizando errores humanos y brindando información de manera inmediata y precisa. Para su desarrollo, se empleará la siguiente tecnología:

Lenguaje de programación: Python, reconocido por su versatilidad y facilidad de mantenimiento.

Framework Backend: Django, que permite una rápida implementación de funcionalidades avanzadas.

Framework Frontend: React, ideal para crear interfaces de usuario dinámicas y modernas.

Base de datos: PostgreSQL: Seleccionada por su robustez y capacidades avanzadas de gestión de datos.

Entorno de desarrollo: Visual Studio Code, elegido por su flexibilidad y amplia gama de extensiones útiles para el desarrollo.

Control de versiones: GitHub, utilizado para el seguimiento, la colaboración y la gestión del código fuente del proyecto.

Entorno de desarrollo: Se orientará a computadoras de escritorio con sistemas Windows y se recomienda el uso de procesadores i5 o superiores, junto con un mínimo de 8 GB de RAM para asegurar un rendimiento óptimo.

1.5. Límites Y Alcances

1.5.1. Límites

Se establecen los siguientes límites para el sistema de gestión y localización de correspondencia mediante el algoritmo de búsqueda semántica, el sistema estará limitado a gestionar correspondencia a partir de la gestión 2021, los documentos anteriores a este año no serán procesados ni almacenados en el sistema.

La digitalización de documentos está limitada a aquellos que tengan una calidad de imagen suficiente para ser procesados correctamente por el sistema, por la precisión en la localización de documentos que puede verse afectada por la calidad del OCR, especialmente en documentos con texto borroso o fuentes poco legibles.

El sistema está diseñado para integrarse con los sistemas actuales dentro de la organización, pero no con la integración con sistemas externos no contemplados en el alcance del proyecto.

La búsqueda semántica tiene limitaciones en su capacidad para interpretar todos los contextos posibles, lo que podría generar resultados imprecisos en ciertos casos.

El sistema requiere de ciertos recursos de hardware (procesadores i5 o superiores, mínimo 8GB de RAM) y de infraestructura adecuada para funcionar de manera óptima, lo que podría ser una limitación en entornos con equipos obsoletos.

El sistema no incluye funcionalidades avanzadas como análisis predictivos o automatización de decisiones, ya que su alcance se limita a la gestión básica de la correspondencia.

1.5.2. Alcances

El alcance del proyecto se centrará en la gestión de correspondencia en la Federación Departamental de La Paz mediante el desarrollo de un sistema que facilite el registro, digitalización, organización, localización y almacenamiento de la correspondencia, tanto entrante como saliente. Para ello, el sistema incluirá los siguientes módulos.

El módulo de gestión de usuarios establecerá niveles de acceso basados en roles, garantizando la seguridad de la información y restringiendo el acceso no autorizado. Luego, el módulo de registro de correspondencia permitirá registrar toda la información relevante de los documentos, asegurando un seguimiento completo de los procesos tanto para correspondencia entrante como saliente.

También se desarrollará un módulo de almacenamiento y digitalización de documentos, que permitirá convertir los documentos físicos en archivos digitales y almacenarlos de manera organizada. El módulo de notificación facilitará la comunicación al informar a los usuarios sobre la correspondencia recién registrada o actualizaciones importantes.

Para apoyar la creación de documentos internos, el módulo de elaboración de documentos permitirá generar nueva documentación, la cual podrá compartirse para su revisión y aprobación. Además, se integrará un módulo de búsqueda semántica que utilizará un algoritmo avanzado para localizar documentos de forma rápida y precisa mediante palabras clave o frases relacionadas.

Por último, el módulo de generación de reportes ofrecerá herramientas para generar informes detallados sobre el flujo y estado de la correspondencia..

1.6. Novedad Científica

La novedad científica de este proyecto radica en la implementación de un sistema de búsqueda semántica aplicado a la gestión de correspondencia. A través del uso de tecnologías avanzadas como el reconocimiento óptico de caracteres (OCR) y el procesamiento de lenguaje natural (NLP), se digitalizan y extraen los textos de los documentos físicos, permitiendo su posterior indexación en una base de datos estructurada. Posteriormente, mediante un

algoritmo de búsqueda semántica, los documentos son localizados de manera rápida y precisa, no solo por coincidencias exactas de palabras, sino también por la relación de conceptos y contexto, lo que mejora significativamente la eficiencia en la recuperación de información. Este enfoque innovador optimiza la gestión de documentos, facilitando la búsqueda y organización de correspondencia dentro de la Federación Departamental de La Paz, y contribuye al avance de la automatización en la gestión documental.

CAPÍTULO II

MARCO TEÓRICO

2.1. Conceptos básicos de gestión documental

2.1.1. Definición de documento

Desde un punto de vista general, se define el documento como «escrito en el que constan datos fidedignos, que sirve de prueba o testimonio, o que proporciona una información, especialmente de carácter histórico, oficial o legal». En esencia, la definición se resume en la combinación de un soporte y la información registrada en él. Su contenido puede consistir tanto en textos escritos, fotografías, dibujos, películas, como en los nuevos soportes multimedia (páginas web, archivos informáticos, etc.) (Pública, 2015).

2.1.2. Archivística

Conjunto de actividades relacionadas con la totalidad del quehacer archivístico, que comprende desde la elaboración del documento hasta su eliminación o conservación permanente (Chinacota, 2019).

Archivística es la ciencia que se ocupa de los archivos en sus aspectos teóricos y prácticos, estableciendo principios inalterables y estudiando técnicas adecuadas de gestión de documentos, administración y tratamiento técnico de archivos, así como la función jurídica, administrativa y científica de los mismos, desde un punto de vista archivístico o de ciencias y técnicas diversas, y su relación con las entidades productoras de los conjuntos orgánicos de documentos, a fin de manejar y hacer accesible la información de los fondos documentales (Ruiz, 2001).

2.1.3. Archivo

Conjunto de documentos acumulados por una persona natural o jurídica, pública o privada que sirve como testimonio o información a la persona o institución que los produce y a ciudadanos u otras entidades como fuente de consulta (Chinacota, 2019).

La palabra archivo generalmente tiene tres acepciones: como contenido documental, como institución y como continente o lugar de conservación. El Diccionario de Terminología

Archivística nos presenta tres definiciones de archivo teniendo en cuenta cada uno de estos elementos conceptuales. "ARCHIVO: (1) Conjunto orgánico de documentos producidos y lo recibidos en el ejercicio de sus funciones por las personas físicas o jurídicas, públicas y privadas 5. (2) La institución cultural donde se reúnen, conservan, ordenan y difunden los conjuntos orgánicos de documentos para la gestión administrativa, la información, la investigación y la cultura. (3) El archivo también es el local donde se conservan y consultan los conjuntos orgánicos de documentos (Ruiz, 2001).

2.1.4. Documento de Archivo

Es el registro de la información producida o recibida por una persona o entidad pública o privada, en el desarrollo de sus funciones y actividades y que cuenta con valores administrativos, legales, fiscales, contables, económicos, científicos, históricos o culturales y que debido a su importancia son objeto de conservación (Chinacota, 2019).

El Diccionario de Terminología Archivística elaborado en 1993 por la Dirección 2 de Archivos Estatales Españoles lo define de esta manera: "Un documento de archivo es el testimonio material de un hecho o acto realizado en el ejercicio de sus funciones por personas físicas o jurídicas, públicas o privadas, de acuerdo con unas características de tipo material y formal (Ruiz, 2001).

2.2. Gestión Documental

La gestión de los documentos: entendida como el conjunto de normas, técnicas y conocimientos aplicados al tratamiento de los documentos desde su diseño hasta su conservación permanente. Podemos afirmar que bajo este rubro se agrupa el núcleo de nuestra ciencia, con aspectos tales como la clasificación, la ordenación la instalación, la descripción, la transferencia, la identificación, la valoración, la selección y la eliminación, sin ánimo exhaustivo (Documentos, 2011).

La gestión documental es el proceso consistente en garantizar la creación, el almacenamiento, el mantenimiento y el uso de las evidencias documentales creadas, mantenidas y utilizadas por una organización de forma tal que se preserve su autenticidad, su seguridad y su uso tanto en el presente como de cara al futuro (Mundial, 2020).

La gestión documental consiste en controlar de un modo eficiente y sistemático la creación, la recepción, el mantenimiento, la utilización y la disposición de los documentos (Catalunya, s.f.).

2.3. Correspondencia

Correspondencia es toda comunicación escrita dirigida de una persona a otra con el fin objeto de transmitir u obtener información (Irene, 2021).

La correspondencia es el conjunto de comunicaciones que mantiene la empresa con el mundo exterior. La persona que se encarga de la correspondencia realiza una serie de actividades coordinadas para que se cumpla la confidencialidad y la seguridad (Quirola, s.f.).

Se denomina correspondencia a un elemento de escritura importante, sobre todo, en el área de las organizaciones y empresas. Puede presentar información básica o importante acerca de un asunto determinado (Tafolla, s.f.).

Se denomina también correspondencia al trato recíproco que existe entre dos participantes mediante el intercambio de una carta, tarjetas, telegramas, folletos, etc. En las empresas se considera que la correspondencia es el alma del comercio y la industrial (Tafolla, s.f.).

2.3.1. Gestión de Correspondencia

La gestión de correspondencia es el conjunto de prácticas y herramientas destinadas a gestionar de manera eficiente los documentos e información que entran y salen de una organización desde diferentes canales. Este proceso no se limita únicamente a la gestión de documentos en papel, sino que también incluye la gestión de archivos digitales y electrónicos (Aguirre, 2024).

En un lenguaje no técnico, la correspondencia o mesa de entrada es como el buzón de correo de una organización. Donde llegan todos los mensajes, cartas, correos electrónicos y documentos de otras personas o empresas. Es como el punto de entrada de toda la información que necesita ser procesada por la empresa, ya sea para responder a preguntas de clientes, procesar facturas o coordinar proyectos. Esencialmente, es el lugar donde comienza el flujo de información hacia el interior de una organización (Aguirre, 2024).

2.3.2. División de la correspondencia

2.3.2.1. Correspondencia Comercial

Se refiere a la cruzada entre comerciantes, industrias, banqueros, etc. Y su finalidad es promover y agilizar las diversas transacciones comerciales (Luna, 2012).

Estos escritos se transmiten entre comerciantes, industrias, diferentes tipos de empresas, etc. Tienen como finalidad principal agilizar las transacciones y promover las acciones comerciales (Tipos, s.f.).

2.3.2.2. Correspondencia Oficial

Se refiere a la correspondencia cruzada entre los distintos organismos y oficinas de los gobiernos nacionales, municipales, estatales, etc. (Luna, 2012).

Las correspondencias oficiales son los escritos cruzados entre municipios estatales y distintos organismos gubernamentales, municipales, etc. (Tipos, s.f.)

2.3.3. Tipos de Correspondencia Gestionada

Dentro de una organización, se gestionan diferentes tipos de correspondencia, cada uno con su propia importancia y flujo de trabajo (Aguirre, 2024):

2.3.3.1. Correspondencia entrante

Se refiere a toda la información que llega a la organización desde fuentes externas, como clientes, proveedores, socios comerciales y otras entidades. Esto puede incluir cartas, correos electrónicos, facturas, formularios, entre otros. Es importante aclarar que la entrada de la información puede llegar desde diversos canales ya sea correo postal, correo electrónico, formularios públicos, WhatsApp, entre otros (Aguirre, 2024).

La correspondencia entrante es toda la información que la empresa recibe de sus clientes, socios comerciales y otras entidades. Puede ser tanto correspondencia tradicional como electrónica. Algunos ejemplos de correspondencia entrante incluyen cartas, faxes, correos

electrónicos, llamadas telefónicas, mensajes de texto e incluso mensajes de chat (Matuszczyk, 2022).

2.3.3.2. Correspondencia saliente

Por otro lado, la correspondencia saliente comprende la información que la organización envía hacia el exterior, dirigida a clientes, proveedores u otras partes interesadas. Esto puede incluir cartas, propuestas comerciales, informes, entre otros (Aguirre, 2024).

La correspondencia saliente es la transferencia de información de una empresa a sus clientes, socios comerciales y otras entidades. Puede ser tanto comunicación tradicional como electrónica. Algunos ejemplos de correspondencia saliente son las cartas, los faxes, los correos electrónicos, las llamadas telefónicas, los mensajes de texto y los mensajes de chat. (Matuszczyk, 2022).

2.3.3.3. Correspondencia Interna

Además de la correspondencia externa, las organizaciones también gestionan una gran cantidad de información de forma interna. Esto puede incluir comunicaciones entre departamentos, memorandos, informes internos, entre otros documentos utilizados para la coordinación y la toma de decisiones dentro de la empresa (Aguirre, 2024).

2.3.4. Proceso de Gestión de Correspondencia

2.3.4.1. Recepción

En este proceso, la organización recibe toda la correspondencia entrante, ya sea en forma de cartas físicas, correos electrónicos u otros medios digitales. Por ejemplo, un departamento de ventas puede recibir correos electrónicos de clientes interesados en sus productos o servicios, así como cartas físicas con consultas o quejas. Es fundamental que esta correspondencia se registre de manera precisa y se distribuya a los departamentos correspondientes para su procesamiento. Para garantizar un adecuado registro de la recepción de los documentos, es esencial llevar a cabo el proceso de radicación. Este procedimiento implica registrar el acuse de recibo de los documentos, ya sea de forma manual o mediante el uso de medios tecnológicos que permitan certificar que el documento llegó en una fecha y hora específicas (Aguirre, 2024).

2.3.4.2. Registro

El registro tiene como objetivo formalizar la incorporación de un documento (dejar constancia de que un documento ha sido creado o recibido) mediante un identificador único y una breve información descriptiva que facilite su posterior recuperación (metadatos). Los documentos se han de registrar en el momento de su incorporación (Baratz, 2021).

2.3.4.3. Clasificación

La clasificación es el conjunto de acciones técnicas y administrativas que nos permite agrupar los documentos relacionados de forma jerárquica, en función de algunas características preestablecidas, independientemente del origen, del destino, del soporte, etc (Baratz, 2021).

Una vez que la correspondencia se recepciona, se clasifica según su tipo y prioridad. Por ejemplo, los correos electrónicos pueden ser clasificados como consultas de clientes, solicitudes de información o quejas, mientras que las cartas físicas pueden ser clasificadas como facturas, formularios completados o correspondencia oficial. Esta clasificación permite a la organización priorizar y asignar recursos de manera eficiente para el procesamiento de la correspondencia (Aguirre, 2024).

2.3.4.4. Distribución

Después de clasificarla, la correspondencia se envía a las personas o departamentos adecuados para su procesamiento. Por ejemplo, los correos electrónicos de los clientes pueden ser enviados al equipo de atención al cliente para su respuesta, mientras que las facturas pueden ser enviadas al departamento de contabilidad para su procesamiento. Es importante que la correspondencia se distribuya de manera rápida y eficiente para garantizar una respuesta oportuna a los remitentes. El uso de tecnologías es punto importante para que dicha distribución se haga de manera automática minimizando tiempos y errores humanos (Aguirre, 2024).

2.3.4.5. Seguimiento

Una vez que la correspondencia ha sido distribuida, se realiza un seguimiento para asegurar que se esté tratando de manera adecuada y oportuna. Por ejemplo, se pueden establecer recordatorios para responder a ciertos correos electrónicos dentro de un plazo determinado o para hacer un seguimiento de las fechas límite para ciertos documentos. Este seguimiento

garantiza que ninguna solicitud quede sin respuesta y que los plazos se cumplan adecuadamente (Aguirre, 2024).

2.3.4.6. Archivo o Almacenamiento

Una vez que la correspondencia ha sido gestionada, se archiva de manera adecuada para futuras referencias. Por ejemplo, las copias físicas de las cartas pueden ser archivadas en carpetas designadas, mientras que lo que se recibe de manera digital puede ser archivado en carpetas específicas dentro del sistema de gestión de documentos. Este archivo permite a la organización acceder rápidamente a la correspondencia previa cuando sea necesario y mantener un registro histórico de todas las interacciones con los clientes y otras partes interesadas (Aguirre, 2024).

Este proceso tiene por objeto mantener y preservar los documentos asegurando su autenticidad, la fiabilidad, la integridad y la disponibilidad durante el periodo de tiempo necesario. Responde a uno de los principios fundamentales de un plan de gestión de documentos: garantizar que los documentos se conserven en un entorno seguro. Para ello, hay que controlar las condiciones de almacenamiento y las operaciones de manipulación, con el fin de proteger los documentos contra el acceso y la destrucción no autorizados; con el fin de prevenir su deterioro o pérdida; y de reducir los riesgos ante posibles robos o desastres (Baratz, 2021).

2.3.4.7. Acceso

En esta fase hay que regular a quién se permite llevar a cabo una operación relacionada con un documento (creación, consulta, modificación, eliminación, etc.) y en qué circunstancias, aplicando los controles previstos en la tabla de acceso y seguridad. Los derechos de acceso y de uso de los usuarios del sistema de gestión de documentos dependerán de los requisitos legales (por ejemplo, la privacidad de datos de carácter personal) y de las necesidades de la organización (por ejemplo, confidencialidad de documentos financieros o con información estratégica) (Baratz, 2021).

2.3.4.8. Trazabilidad

La trazabilidad de los documentos en un sistema de gestión documental es una medida de seguridad para las organizaciones. Garantiza que solo los usuarios con los permisos adecuados realicen tareas para las que han sido autorizados. Se debe controlar el uso y el

movimiento de los documentos de forma que se garantice, por un lado, que únicamente los usuarios con los permisos adecuados puedan llevar a cabo actividades que les han sido asignadas y, por otro lado, que los documentos puedan ser localizados siempre que se necesiten (Baratz, 2021).

El seguimiento del rastro de un documento permite mantener un control adecuado de los procesos documentales desde que es incorporado al sistema de gestión de documentos hasta que se aplica la disposición final (Baratz, 2021).

2.3.4.9. Disposición

Agotado el plazo de conservación establecido para un documento determinado, se aplica la disposición prevista en el calendario de conservación: eliminación, conservación permanente y transferencia a otro sistema archivístico (Baratz, 2021).

2.4. Tecnologías de la Información

Actualmente, las Tecnologías de la Información y la Comunicación (TIC) han sobresalido en todas las áreas del conocimiento por la infinidad de aplicaciones y beneficios que brindan en la ejecución de actividades diarias que solían tomar mucho tiempo y esfuerzo (Reducindo, 2022).

En este punto cabe mencionar que es necesario hacer una diferencia entre las Tecnologías de la Información (TI) y las Tecnologías de la Comunicación (TC). Aunque ambos conceptos en la actualidad se usan de manera indistinta tanto juntos como separados, dentro de este texto debemos tener claro que las TI son aquellas que apoyan en el almacenamiento, protección, procesamiento y transmisión de la información administrada como conjunto de datos, señales o conocimientos, registrados en diversos tipos de soportes documentales. Por otro lado, las TC son aquellas que apoyan en la ejecución de prácticas comunicativas, es decir, en la transmisión de información de un emisor a un receptor (Bloom et al., 2014, como se cita en (Reducindo, 2022).

2.4.1. Sistema de Información

Es necesario pensar en herramientas de TI que puedan apoyar en la administración del currículum, con la finalidad de almacenar y gestionar aquella información útil para las actividades cotidianas de las instituciones, y que agilicen la GD y sus procesos, y es aquí

donde podemos encontrar los Sistemas de Información (SI), los cuales permiten a las organizaciones e instituciones sistematizar y en consecuencia agilizar procesos administrativos que repercuten directamente en la mejora y el rendimiento de sus funciones, en la precisión de sus actividades y en el tiempo que toma realizar cada una de ellas. Entonces, podemos establecer que los SI son la herramienta adecuada para ser utilizada en la mejora de la GC (Reducindo, 2022).

Los SI son un conjunto de procesos y de herramientas como hardware, software y bases de datos; usuarios e información que se conjuntan para gestionar, resguardar, almacenar y recuperar cualquier tipo de información (Prasanna et al., 2014). Actualmente existen diferentes tipos de SI, todos creados con la finalidad de gestionar distintos dominios de información, permitiendo a las organizaciones llevar un control de sus actividades en la toma de decisiones. Entre estos sistemas podemos encontrar a aquellos que apoyan en el control de calidad, en la operación de actividades administrativas de una organización, en la gestión de los recursos humanos de cierta institución, en la administración de las finanzas de alguna empresa o bien, en la producción de materiales o servicios de cualquier industria (Reducindo, 2022).

2.4.2. Tecnología de Información al servicio de la archivística

La tecnología de la información al servicio de la Archivística debe ser observada a la luz de las posibilidades ofrecidas por una serie de herramientas tecnológicas para agilizar muchos procesos que se dan en el contexto de los archivos. Pero también es necesario considerar a estas tecnologías dentro del marco de la sociedad de la información, donde las irrupciones de nuevos productos informáticos deben conducirnos a meditar acerca de su potencial para el quehacer de los archivos, aunque sin perder de vista su carácter utilitario (Fernández, 2001).

El mayor impacto que las tecnologías de la información ejercen sobre el campo de los archivos se relaciona con las posibilidades que brindan las redes de comunicación para acceder e intercambiar información, con lo cual tanto el usuario interno como el externo resultan beneficiados; en igual proporción se incrementan las posibilidades de la difusión de información a un público más amplio, lo que incide en un mayor dinamismo de los archivos (Fernández, 2001).

La industria de las computadoras, la industria de las telecomunicaciones y la de la comunicación se han integrado; de esta alianza estratégica, los archivos en particular y la Archivística en general, deben extraer el mayor provecho de estas y ponerlas al servicio de

múltiples procesos, aunque sin perder de vista su carácter instrumental. Asimismo, no debe olvidarse que la tecnología ha dado origen a nuevos soportes, muchos de los cuales han venido a sustituir a los soportes tradicionales, con lo que se abren nuevos retos para los archivistas (Fernández, 2001)

Por lo tanto, no solo debemos apreciar la tecnología informática al servicio de la Archivística, en términos del mejoramiento de muchos procesos, especialmente de la descripción, sino que también debemos prestar atención a la producción de un número cada vez mayor de documentos de soporte informático, en los que la tecnología se torna una herramienta imprescindible para tener acceso a la información que estos contienen (Fernández, 2001).

2.4.3. Importancia de la tecnología en la gestión documental

En el mundo empresarial actual, donde la información fluye constantemente, la gestión documental se ha vuelto crucial. La tecnología juega un papel fundamental en este proceso, ya que ofrece herramientas y soluciones que optimizan los flujos de trabajo y permiten un acceso rápido y sencillo a la información. Además, la tecnología también ahorra tiempo y recursos al automatizar tareas que antes requerían una gran cantidad de trabajo manual (Archivos Empresariales del Caribe, s.f.).

La implementación de herramientas y soluciones tecnológicas ha simplificado el proceso de gestión documental y ha mejorado la productividad en las empresas. Hoy en día, es posible asociar etiquetas y palabras clave a los documentos, lo que facilita la búsqueda y recuperación de información. Además, también se puede acceder a los documentos desde cualquier lugar y en cualquier momento, gracias a las soluciones basadas en la nube (Archivos Empresariales del Caribe, s.f.).

2.5. Inteligencia Artificial y Algoritmos de Búsqueda Semántica

2.5.1. Inteligencia Artificial

En mis seminarios, intento simplificar el tema definiendo la IA como «la habilidad de los ordenadores para hacer actividades que normalmente requieren inteligencia humana». Pero, para brindar una definición más detallada, podríamos decir que la IA es la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano. Sin embargo, a diferencia de las personas, los

dispositivos basados en IA no necesitan descansar y pueden analizar grandes volúmenes de información a la vez. Asimismo, la proporción de errores es significativamente menor en las máquinas que realizan las mismas tareas que sus contrapartes humanas (Rouhiainen, 2018).

La inteligencia artificial es la habilidad de una máquina de presentar las mismas capacidades que los seres humanos, como el razonamiento, el aprendizaje, la creatividad y la capacidad de planear (Europeo, 2021).

La IA permite que los sistemas tecnológicos perciban su entorno, se relacionen con él, resuelvan problemas y actúen con un fin específico. La máquina recibe datos (ya preparados o recopilados a través de sus propios sensores, por ejemplo, una cámara), los procesa y responde a ellos (Europeo, 2021).

2.5.2. Beneficios de la Inteligencia Artificial

Precisión mejorada: Los sistemas de IA pueden analizar grandes volúmenes de datos de manera rápida y precisa, lo que ayuda a tomar decisiones basadas en información precisa y actualizada (Uribe, 2023).

Análisis profundo: La IA es capaz de identificar patrones y tendencias ocultas en conjuntos de datos complejos. Esto que proporciona una comprensión más profunda y detallada de la información relevante para la toma de decisiones (Uribe, 2023).

Velocidad y eficiencia: Los algoritmos de IA pueden procesar datos en tiempo real, lo que permite una toma de decisiones más rápida y eficiente en comparación con los enfoques tradicionales (Uribe, 2023).

Personalización: Estos sistemas pueden adaptarse a las preferencias y necesidades individuales, de manera que la toma de decisiones sea más personalizada y adaptada a cada usuario o situación específica (Uribe, 2023).

Automatización de tareas rutinarias: La IA puede encargarse de tareas repetitivas y rutinarias, liberando tiempo y recursos para que los humanos se enfoquen en tareas más estratégicas y de mayor valor (Uribe, 2023).

Mayor eficiencia operativa: Permite optimizar procesos y recursos, lo que conduce a una mayor eficiencia operativa en diversos sectores y organizaciones (Uribe, 2023).

Mejor detección de anomalías y riesgos: A través de los sistemas de IA, pueden analizarse datos buscando patrones anormales y riesgos potenciales. Esto nos ayuda a detectar problemas y tomar medidas preventivas de manera oportuna (Uribe, 2023).

Mejora en la precisión del pronóstico: Al usarse modelos predictivos sofisticados, pueden realizarse pronósticos precisos sobre el futuro, lo que permite tomar decisiones más informadas y estratégicas (Uribe, 2023).

Mayor capacidad de aprendizaje: Como hemos mencionado anteriormente, la Inteligencia Artificial es capaz de aprender de experiencias anteriores y mejorar continuamente a medida que se le proporciona más información y datos (Uribe, 2023).

Reducción de errores humanos: Al automatizar ciertos aspectos de la toma de decisiones, la IA nos puede ayudar a reducir los errores humanos y mejorar la consistencia en la toma de decisiones (Uribe, 2023).

2.5.3. Machine Learning (ML)

Machine Learning (ML) es una disciplina de la Inteligencia Artificial, que, a través de ciertos algoritmos, dota a los sistemas de la capacidad para analizar una gran cantidad de datos pasados y predecir futuros estados. Es decir, se ocupa del proceso de aprendizaje (Gallardo, 2020).

El aprendizaje máquina o automático (machine learning, ML, en inglés) es la disciplina dentro de la ciencia de datos que permite que las máquinas aprendan sin ser programadas con reglas específicas. Aplica la estadística para inferir propiedades y otros métodos matemáticos para detectar patrones en los datos y, a partir de ahí, hacer predicciones e incluso tomar decisiones (Olivas, 2023).

2.5.4. Algoritmo

Un algoritmo es un conjunto de instrucciones que especifica la secuencia de operaciones a realizar, en orden, para resolver un problema específico; en otras palabras, un algoritmo es

una fórmula para la resolución de un problema. La definición de un algoritmo debe describir tres partes: Entrada, Proceso y Salida, así (Sonora, 2016):

Entrada: Información dada al algoritmo, o conjunto de instrucciones que generen los valores con que ha de trabajar (Sonora, 2016).

Proceso: Cálculos necesarios para que a partir de un dato de entrada se llegue a los resultados (Sonora, 2016).

Salida: Resultados finales o transformación que ha sufrido la información de entrada a través del proceso (Sonora, 2016).

De forma general, se puede definir un algoritmo como una secuencia de instrucciones o pasos que al ejecutarlos recibe datos de entrada, de manera directa o indirecta y los transforma en una salida. También se lo puede definir con un conjunto de instrucciones ordenadas de manera lógica tal que al ser ejecutadas por una computadora o persona dan solución a un determinado problema para el cual fue diseñado (Bel, 2020).

2.5.5. Semántica

La Semántica es la disciplina que estudia el significado de las expresiones lingüísticas. Como estrategia metodológica, es común separar el estudio del significado de las palabras y el de las expresiones más complejas, como sintagmas y oraciones. Distinguimos, en consecuencia, entre Semántica léxica y Semántica composicional. La Semántica léxica se ocupa de todo lo relativo al significado de las palabras; la Semántica composicional parte del supuesto de que las palabras tienen significado y se centra en descubrir cuál es la contribución que la estructura y las relaciones sintácticas aportan a la construcción de las expresiones complejas (Vidal, 2008).

2.5.6. Procesamiento del Lenguaje Natural

El procesamiento de lenguaje natural da la posibilidad de establecer una comunicación entre una persona y una máquina a través de lenguajes naturales como el español. El procesamiento del lenguaje es definido como una rama de la inteligencia artificial la cual pretende, mediante entrenamiento y distintos ámbitos de la programación, replicar la facultad del lenguaje humano. Y su uso supone una evolución enorme, debido a que nos permite hacer una amplia cantidad de tareas con una finalidad concreta, como puede ser la traducción, establecer una

conversación, y obtener contenidos de las redes sociales entre otras tantas posibilidades. Dicho de una manera más general, este procesamiento se puede aplicar con cualquier faceta en la que intervenga el lenguaje humano (Hernandez, 2021).

2.5.6.1. Modelos para procesamiento el Lenguaje Natural

Lógicos: estos modelos están basados en determinadas formas gramaticales, de manera que los lingüistas especializados definen diferentes patrones estructurales de la lingüística para que las máquinas puedan identificarlo. Si a este concepto le añadimos la información existente en los distintos diccionarios computacionales damos la posibilidad de definir modelos que son capaces de resolver determinadas tareas como pueden ser buscar información, traducir, resumir etc (Hernandez, 2021).

Probabilístico: A diferencia del modelo lógico en este no se valoran las reglas gramaticales en ningún momento. Se basa en la recolección por parte de los lingüistas de una gran cantidad de datos que se emplearán como foco principal en el análisis. A partir de estos datos se analizan y localizan la frecuencia de distintos elementos lingüísticos como pueden ser letras o palabras. A partir de este análisis se puede predecir la siguiente unidad que aparecerá en el texto basándonos en la probabilidad de esta. En el ámbito de la inteligencia artificial esto se conoce como aprendizaje automático (Hernandez, 2021).

2.5.6.2. Análisis para Procesamiento del Lenguaje Natural

Análisis morfológico: es el análisis más interno que se le realiza a las palabras, y su finalidad es obtener su significado y categoría (Hernandez, 2021).

Análisis sintáctico: analiza el texto para obtener la estructura de cada una de las frases que lo componen (Hernandez, 2021).

Análisis semántico: su finalidad es conseguir una correcta deducción de las frases (Hernandez, 2021).

Análisis pragmático: pretende obtener una correcta conclusión del contexto (Hernandez, 2021).

2.5.7. Algoritmo de Búsqueda

Un algoritmo de búsqueda es un conjunto de instrucciones que están diseñadas para localizar un elemento con ciertas propiedades dentro de una estructura de datos; por ejemplo, ubicar el registro correspondiente a cierta persona en una base de datos, o el mejor movimiento en una partida de ajedrez (Ponce, 2021).

Los algoritmos de búsqueda son procedimientos sistemáticos utilizados para encontrar un elemento deseado dentro de un conjunto de datos. Estos algoritmos siguen una serie de pasos predefinidos para buscar y localizar la información necesaria de manera eficiente. En esencia, son como una guía que nos ayuda a encontrar una aguja en un pajar (TecnoDigital, 2023).

2.5.8. Búsqueda Semántica

La "semántica" es una disciplina lingüística que se enfoca en el estudio del significado y la comprensión del lenguaje. Analiza cómo las palabras, frases y oraciones representan diferentes conceptos y significados en distintos contextos (Aplyca, 2023).

Las búsquedas semánticas tratan de capturar el significado y la intención de una búsqueda en lugar de simplemente coincidir con las palabras clave (Aplyca, 2023).

Las búsquedas semánticas permiten encontrar similitudes relevantes en un conjunto de datos. En lugar de buscar coincidencias exactas de palabras clave – como en una búsqueda de bolsa de palabras – las búsquedas semánticas permiten a los motores de búsqueda identificar elementos que son semánticamente similares a la consulta de búsqueda del usuario (Aplyca, 2023).

La semántica trata sobre el significado de palabras y frases y cómo encajan entre sí. Va más allá de las búsquedas tradicionales basadas en palabras clave al descifrar no solo el significado de lo que usted dice, sino también su intención, gracias a tecnologías avanzadas como el procesamiento del lenguaje natural (NLP), machine learning (ML), gráficos de conocimiento e inteligencia artificial (MongoDB, s.f.).

En el corazón de la búsqueda semántica se encuentra la poderosa combinación del procesamiento del lenguaje natural y el machine learning. Estas tecnologías permiten al motor de búsqueda comprender las complejidades del lenguaje humano. El NLP, junto con el aprendizaje automático, permite al sistema no sólo reconocer palabras, sino también captar

los matices del lenguaje, incluidos sinónimos, significados contextuales de frases y palabras clave, e incluso peculiaridades lingüísticas específicas del usuario. Esto significa que puede hablar con su motor de búsqueda como si fuera un humano (MongoDB, s.f.).

2.5.9. Reconocimiento Óptico de Caracteres

El OCR (Reconocimiento Óptico de Caracteres) es un conjunto de algoritmos desarrollados para convertir documentos impresos y escritos a mano en texto digital. Esta tecnología se utiliza para descifrar texto que está escrito a mano o en soportes impresos para su posterior procesamiento digital. Esta tecnología de digitalización utiliza algoritmos de reconocimiento de patrones para recopilar y procesar información basándose en los caracteres presentes, como mayúsculas, minúsculas, números y símbolos (NORMADAT, 2023).

La tecnología OCR es un sistema que permite transformar documentos que contienen texto, como imágenes, fotografías o archivos en PDF, en un formato de texto editable. Esta tecnología es una herramienta muy poderosa para las empresas ya que es capaz de convertir información que antes solo podía comprender un humano, en algo que los ordenadores pueden procesar y archivar (Ayerdi, 2024).

El OCR (Optical Character Recognition) es una tecnología que trata de emular la capacidad del ojo humano para reconocer objetos. Concretamente es un software que permite el reconocimiento óptico de los caracteres contenidos en una imagen (documento escaneado o fotografía), de forma que estos se vuelven comprensibles o reconocibles para un ordenador, obteniendo como resultado final un archivo en un formato de texto editable. El formato del archivo de salida (txt, pdf, etc.) dependerá de las posibilidades que ofrezca el software (Kultura Salla, 2011).

2.5.9.1. Reconocimiento Óptico de Caracteres en la gestión de datos

La tecnología OCR ha tenido un impacto significativo en la gestión de documentos en diversas industrias. Aquí hay algunas formas en las que está transformando este proceso (Grupo ASD, 2023):

Automatización de la Captura de Datos: Antes de la era del OCR, la entrada manual de datos a partir de documentos físicos era un proceso lento y propenso a errores. El OCR ha permitido la automatización de esta tarea, acelerando la captura de datos y reduciendo los errores humanos (Grupo ASD, 2023).

Búsqueda y Acceso Rápido: El texto digital generado por el OCR se puede buscar de manera eficiente, lo que facilita la recuperación de documentos y la extracción de información específica. Ya no es necesario revisar manualmente montañas de documentos para encontrar lo que se necesita (Grupo ASD, 2023).

Colaboración en Tiempo Real: La digitalización de documentos a través del OCR permite compartir información de manera instantánea y colaborar en tiempo real. Esto es especialmente útil en entornos de trabajo remoto y distribuido (Grupo ASD, 2023).

Reducción de Espacio de Almacenamiento: Al convertir documentos físicos en formato digital, las empresas pueden reducir la necesidad de espacio de almacenamiento físico. Esto no solo ahorra costos, sino que también contribuye a una gestión más sostenible de los recursos (Grupo ASD, 2023).

Mayor Seguridad y Cumplimiento: Los sistemas de gestión de documentos basados en OCR suelen contar con medidas de seguridad y funciones de seguimiento de auditoría, lo que ayuda a garantizar la privacidad y el cumplimiento normativo (Grupo ASD, 2023).

2.6. Metodología ágil para el desarrollo del sistema

2.6.1. Metodología Ágil Scrum

Desarrollada por Ken Schwaber, Jeff Sutherland y Mike Beedle. Define un marco para la gestión de proyectos, que se ha utilizado con éxito durante los últimos 10 años. Está especialmente indicada para proyectos con un rápido cambio de requisitos. Sus principales características se pueden resumir en dos. El desarrollo de software se realiza mediante iteraciones, denominadas sprints, con una duración de 30 días. El resultado de cada sprint es un incremento ejecutable que se muestra al cliente. La segunda característica importante son las reuniones a lo largo del proyecto. Éstas son las verdaderas protagonistas, especialmente la reunión diaria de 15 minutos del equipo de desarrollo para coordinación e integración (Patricio Letelier, 2006).

La metodología Scrum permite abordar proyectos complejos desarrollados en entornos dinámicos y cambiantes de un modo flexible. Está basada en entregas parciales y regulares

del producto final en base al valor que ofrecen a los clientes. Dicho en otras palabras: Scrum sirve para mejorar el trabajo colaborativo entre equipos (Hurtado, 2021).

2.6.1.1. Equipo Scrum

Son los responsables de entregar el producto de forma iterativa e incremental, un incremento valioso y útil en cada Sprint. El Equipo Scrum debe ser auto-gestionado y multifuncional, debe estar diseñado para optimizar la flexibilidad, creatividad y la productividad, aprovechando los escenarios de retroalimentación para mejorar los procesos (Zalimben, 2022).

2.6.1.1.1. Product Owner

Es responsable de maximizar el valor del producto resultante de los developers, debe tener un entendimiento claro de la organización y del negocio. Es una persona, no un comité y debe ser respetado por toda la organización. Debe ayudar a los interesados a plasmar las ideas en requerimientos y construirlos como elementos del product backlog; lista de las características, funcionalidades y requerimientos que implican cambios a realizarse en el producto, para que sean desarrollados en el Sprint. Único responsable de gestionar el product backlog y capaz de cancelar un Sprint (Zalimben, 2022).

2.6.1.1.2. Scrum Master

Facilitador que obra para que todos los eventos de Scrum se realicen correctamente y de forma eficiente, trabaja para promover y apoyar la metodología Scrum, ayudando a entender la teoría, prácticas, reglas y valores, y busca maximizar las interacciones entre el equipo. Debe proteger al Sprint y los developers, es decir, el compromiso que el equipo ha tomado al comienzo del Sprint. No debería trabajar con los interesados de forma directa (Zalimben, 2022).

2.6.1.1.3. Developers

Es el equipo encargado de desarrollar los elementos que se encuentran dentro del Sprint Backlog, que es la lista de elementos que fueron seleccionados para ser desarrollados en un Sprint. El equipo debe ser autogestionado y multifuncional, a su vez que debe gozar de cierta autonomía. Todos los miembros deberían ser capaces de realizar todas las tareas de desarrollo (Zalimben, 2022).

Son responsables de entregar un incremento terminado del producto, listo para ser presentado a los interesados al final de cada Sprint. Son los únicos responsables del incremento y nadie puede forzarlos a trabajar en un conjunto diferente de requisitos (Zalimben, 2022).

Al hablar de un equipo es importante mencionar que Scrum no reconoce títulos dentro del equipo, ni tampoco reconoce sub-equipos dentro del mismo (Zalimben, 2022).

2.6.1.2. Eventos Scrum

2.6.1.2.1. Sprint Planning

Reunión donde se planifica el trabajo que será realizado durante el Sprint por los developers, este plan se crea mediante el trabajo del Equipo Scrum (Zalimben, 2022).

El Product Owner se asegura de que cada participante esté preparado para discutir los elementos más importantes del Product Backlog y cómo se relacionan con el Objetivo del Producto. También puede invitar a otras personas a asistir al Sprint Planning para brindar asesoramiento. El Scrum Master debe velar por la realización de este evento y deben participar el Product Owner y los developers (Zalimben, 2022).

En general se recomienda un bloque de tiempo de 8 horas para un Sprint de 4 semanas, en caso que el Sprint tenga una duración menor, el Sprint Planning también debe tener una duración menor (Zalimben, 2022).

2.6.1.2.2. Daily Scrum

Son reuniones diarias que se realizan cada 24 horas, o bien una vez al día con el fin de conocerla salud del Sprint, inspeccionando el trabajo realizado, los impedimentos y la proyección del mismo, evaluando el progreso hacia el cumplimiento del objetivo del Sprint, y adaptar el sprint backlog según sea necesario. Es una reunión con una duración máxima de 15 minutos y se recomienda realizarlo al comienzo del día, pero también es posible consensuar un horario para realizarlo. Los developers son responsables de llevar a cabo esta reunión y de seleccionar la estructura y las técnicas que deseen, ya sea utilizando preguntas y/o discusiones, tales como ¿Qué hice ayer? ¿Qué haré hoy? ¿Existen impedimentos? (Zalimben, 2022).

Estas reuniones mejoran la comunicación, identifican impedimentos, promueven la toma rápida de decisiones y, en consecuencia, eliminan la necesidad de otras reuniones. Si hubiese

necesidad después del Daily Scrum se deben desarrollar discusiones más detalladas sobre cómo adaptar o volver a planificar el resto del trabajo del Sprint (Zalimben, 2022).

2.6.1.2.3. Sprint Review

El propósito de esta reunión es inspeccionar el resultado del Sprint, el incremento del producto y adaptar el product backlog si fuese necesario para el siguiente Sprint. En esta reunión participan los interesados y el Equipo Scrum con el fin de facilitar la retroalimentación (Zalimben, 2022).

El Equipo Scrum presenta los resultados de su trabajo a los interesados y se discute el progreso hacia el Objetivo del Producto. El Product Owner realiza una breve introducción, un resumen del Sprint que finalizó y los elementos que fueron desarrollados, mientras que los developers son los responsables de presentar los elementos que fueron terminados durante la ejecución del Sprint para ser validados por los interesados. Deben presentar el incremento, realizar una demostración del mismo y explicar los detalles de implementación si son requeridos, así como los problemas presentes durante el Sprint (Zalimben, 2022).

En general se recomienda un bloque de tiempo de 4 horas para un Sprint de 4 semanas, en caso que el Sprint tenga una duración menor, el Sprint Review también debe tener una duración menor (Zalimben, 2022).

El Product Owner realiza un cierre mostrando el estado actual del product backlog, ilustrando las proyecciones de los siguientes Sprints y el estado actual del producto. El resultado de esta reunión es un product backlog revisado y priorizado (Zalimben, 2022).

2.6.1.2.4. Sprint Retrospective

Reunión para inspeccionar al equipo y planificar formas de aumentar la calidad y la efectividad, tiene lugar después del Sprint Review y antes del Sprint Planning del Sprint entrante (Zalimben, 2022).

En esta reunión se presentan todos los inconvenientes presentes durante la ejecución de un Sprint, con el fin de aprender de ellos e ir mejorando en cada iteración, se debe diseñar un plan de mejora y aplicarlo durante la ejecución del siguiente Sprint (Zalimben, 2022).

Se inspecciona cómo fue el último Sprint con respecto a las personas, las interacciones, los procesos, las herramientas y su Definición de Terminado. Se identifican los supuestos que los

llevaron por mal camino y se exploran sus orígenes. El Scrum Team analiza qué salió bien durante el Sprint, qué problemas encontró y cómo se resolvieron (o no) esos problemas (Zalimben, 2022).

Con esta reunión concluye el Sprint. Se recomienda un bloque de tiempo de 3 horas para un Sprint de 4 semanas, en caso que el Sprint tenga una duración menor, el Sprint Review debe durar menos (Zalimben, 2022).

2.6.1.3. Artefactos Sprint

Son elementos útiles para proporcionar transparencia y oportunidades para la inspección y adaptación del trabajo y el Equipo Scrum, brindan información sobre diferentes aspectos del producto durante su desarrollo (Zalimben, 2022).

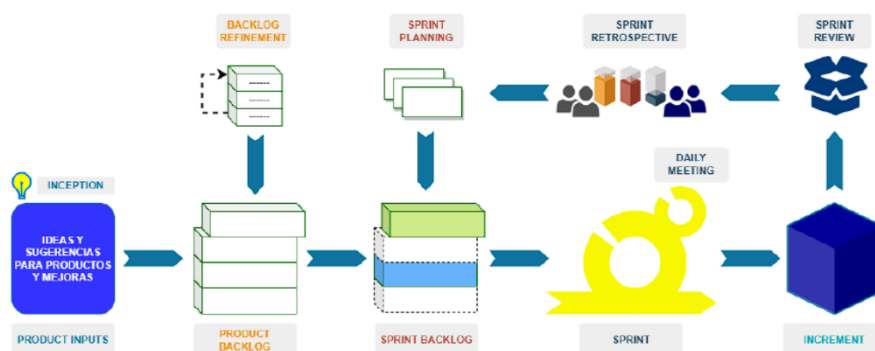


Figura 2: Eventos y Artefactos de Scrum. Zalimben (2022).

2.6.1.3.1. Product Backlog

Es la lista de elementos que enumera todas las características, funcionalidades, requerimientos, mejoras y correcciones que constituyen cambios a realizarse en el producto en entregas futuras. Es la única fuente para cualquier cambio a realizarse, el Product Owner es el responsable del contenido, disponibilidad y orden (Zalimben, 2022).

Evoluciona a medida que el producto y el entorno que lo usará lo hacen, es volátil acorde a las necesidades, ideas o requisitos del producto o del entorno. Cambios en los requisitos de negocio, condiciones de mercado o de tecnología podrían ocasionar cambios y/o ajustes en el product backlog (Zalimben, 2022).

Los elementos deben estar listos para formar parte del sprint backlog, deben tener cierto grado de detalle, transparencia e independencia. Cada elemento debería de tener, mínima-mente, descripción, orden, estimación y valor agregado, y si es posible demostraciones de la completitud del mismo (Zalimben, 2022).

Es posible que varios equipos trabajen en el mismo producto, pero de igual manera se debería de mantener un solo product backlog pero agrupando los elementos por algún atributo común entre ellos (Zalimben, 2022).

Los elementos deben ser descompuestos con el fin de que puedan ser desarrollados dentro de un Sprint, es decir, puedan ser terminados en el tiempo límite del Sprint. Generalmente los elementos con mayor prioridad son los que están mejor detallados, que facilita la estimación de los mismos (Zalimben, 2022).

El product backlog es desarrollado por los developers en Sprints, agrupando los elementos en sprint backlogs. Los developers son responsables de las estimaciones de dichos elementos (Zalimben, 2022).

2.6.1.3.2. Sprint Backlog

Se compone del Objetivo del Sprint (el por qué), el conjunto de elementos del product backlog seleccionados para ser desarrollados en el Sprint (el qué), así como un plan de acción para entregar el incremento del producto (el cómo). El sprint backlog es un plan realizado por y para los developers. Es una imagen visible y en tiempo real del trabajo que los developers planean realizar durante el Sprint para lograr el objetivo del Sprint. Solo los developers pueden cambiar los elementos del Sprint durante la ejecución del mismo (Zalimben, 2022).

2.6.1.3.3. Incremento

Un incremento es el conjunto de todos los elementos del Sprint terminados y el valor de todos los incrementos de los Sprints anteriores. Debe estar listo para ser utilizado independiente de si el Product Owner desea ponerlo o no a disposición de los interesados (Zalimben, 2022).

Un elemento no puede considerarse parte de un incremento a menos que cumpla con la definición de Terminado (Zalimben, 2022).

2.7. Diseño

2.7.1. Diagrama UML – Lenguaje Unificado de Modelamiento

Un diagrama UML es una representación gráfica parcial (vista) de un modelo de un sistema. Es una herramienta que permite a los creadores de sistemas generar diseños que capturen sus ideas en una forma convencional y fácil de comprender y así poder comunicárselas a otras personas. (Cotaña, 2012)

Es un lenguaje de modelado visual que se usa para especificar, visualizar, construir y documentar un sistema de software. Se usa para entender, diseñar, configurar, mantener y controlar la información sobre los sistemas a construir (Universidad Nacional Autonoma de Mexico, s.f.).

UML capta la información sobre la estructura estática y el comportamiento dinámico de un sistema. Un sistema se modela como una colección de objetos que interactúan para realizar un trabajo que finalmente beneficia a un usuario externo. El lenguaje de modelado pretende unificar la experiencia pasada sobre técnicas de modelado e incorporar las mejores prácticas actuales en un acercamiento estándar (Universidad Nacional Autonoma de Mexico, s.f.).

2.7.1.1. Diagrama de Casos de Uso

Un Diagrama de Casos de Uso muestra la relación entre los actores y los casos de uso del sistema. Representa la funcionalidad que ofrece el sistema en lo que se refiere a su interacción externa (Grau, 2011).

2.7.1.1.1. Elementos

Los elementos que pueden aparecer en un Diagrama de Casos de Uso son: actores, casos de uso y relaciones entre casos de uso (Grau, 2011).

2.7.1.1.2. Actores

Un actor es una entidad externa al sistema que realiza algún tipo de interacción con el mismo. Se representa mediante una figura humana dibujada con palotes. Esta representación sirve tanto para actores que son personas como para otro tipo de actores (otros sistemas, sensores, etc.) (Grau, 2011).

2.7.1.1.3. Casos de Uso

Un caso de uso es una descripción de la secuencia de interacciones que se producen entre un actor y el sistema, cuando el actor usa el sistema para llevar a cabo una tarea específica. Expresa una unidad coherente de funcionalidad, y se representa en el Diagrama de Casos de Uso mediante una elipse con el nombre del caso de uso en su interior. El nombre del caso de uso debe reflejar la tarea específica que el actor desea llevar a cabo usando el sistema (Grau, 2011).

2.7.1.1.4. Relaciones

Entre dos casos de uso puede haber las siguientes relaciones (Grau, 2011):

Extiende: Cuando un caso de uso especializa a otro extendiendo su funcionalidad.

Usa: Cuando un caso de uso utiliza a otro.

Se representan como una línea que une a los dos casos de uso relacionados, con una flecha en forma de triángulo y con una etiqueta <<extiende>> o <<usa>> según sea el tipo de relación (Grau, 2011).

En el diagrama de casos de uso se representa también el sistema como una caja rectangular con el nombre en su interior. Los casos de uso están en el interior de la caja del sistema, y los actores fuera, y cada actor está unido a los casos de uso en los que participa mediante una línea (Grau, 2011).

2.7.1.2. Diagrama de Secuencias

Un diagrama de Secuencia muestra una interacción ordenada según la secuencia temporal de eventos. En particular, muestra los objetos participantes en la interacción y los mensajes que intercambian ordenados según su secuencia en el tiempo (Grau, 2011).

El eje vertical representa el tiempo, y en el eje horizontal se colocan los objetos y actores participantes en la interacción, sin un orden prefijado. Cada objeto o actor tiene una línea vertical, y los mensajes se representan mediante flechas entre los distintos objetos. El tiempo fluye de arriba abajo (Grau, 2011).

Se pueden colocar etiquetas (como restricciones de tiempo, descripciones de acciones, etc.) bien en el margen izquierdo o bien junto a las transiciones o activaciones a las que se refieren (Grau, 2011).

2.7.1.3. Diagrama de Estado

Un Diagrama de Estados muestra la secuencia de estados por los que pasa un caso de uso o un objeto a lo largo de su vida, indicando qué eventos hacen que se pase de un estado a otro y cuáles son las respuestas y acciones que genera (Grau, 2011).

En cuanto a la representación, un diagrama de estados es un grafo cuyos nodos son estados y cuyos arcos dirigidos son transiciones etiquetadas con los nombres de los eventos (Grau, 2011).

Un estado se representa como una caja redondeada con el nombre del estado en su interior. Una transición se representa como una flecha desde el estado origen al estado destino (Grau, 2011).

La caja de un estado puede tener 1 o 2 compartimentos. En el primer compartimento aparece el nombre del estado. El segundo compartimento es opcional, y en él pueden aparecer acciones de entrada, de salida y acciones internas (Grau, 2011).

Una acción de entrada aparece en la forma entrada/acción_asociada donde acción_asociada es el nombre de la acción que se realiza al entrar en ese estado. Cada vez que se entra al estado por medio de una transición la acción de entrada se ejecuta (Grau, 2011).

Una acción de salida aparece en la forma salida/acción_asociada. Cada vez que se sale del estado por una transición de salida la acción de salida se ejecuta (Grau, 2011).

Una acción interna es una acción que se ejecuta cuando se recibe un determinado evento en ese estado, pero que no causa una transición a otro estado. Se indica en la forma nombre_de_evento/acción_asociada (Grau, 2011).

2.7.1.4. Diagrama de Paquetes

En UML el paquete es un mecanismo de propósito general para organizar elementos de modelado en grupos. Los paquetes se utilizan para organizar los elementos de modelado en partes mayores que se pueden manipular como un grupo (Cotaña, 2012).

La visibilidad de estos elementos puede controlarse para que algunos sean visibles fuera del paquete mientras que otros permanecen ocultos (Cotaña, 2012).

Los paquetes bien estructurados son cohesivos y poco acoplados, estando muy controlado el acceso a su contenido (Cotaña, 2012).

2.7.1.5. Diagrama de Clases

Los diagramas de clases se utilizan para modelar la visión estática de un sistema. Esta visión soporta los requisitos funcionales del sistema, en concreto, los servicios que el sistema debería proporcionar a sus usuarios finales. Normalmente contienen: clases, interfaces y relaciones entre ellas: de asociación, de dependencia y/o de generalización (Vidal, s.f.).

Las clases describen un conjunto de objetos con características y comportamiento idénticos, es decir, objetos que comparten los mismos atributos, operaciones y relaciones (Vidal, s.f.).

Las clases se representan gráficamente por medio de un rectángulo con tres divisiones internas. Los tres compartimentos alojan el nombre de la clase, sus atributos y sus operaciones, respectivamente. En muchos diagramas se omiten los dos compartimentos inferiores. Incluso cuando están presentes, no muestran todos los atributos y todas las operaciones (Vidal, s.f.).

2.7.1.5.1. Relaciones en un diagrama de clases

Los diagramas de clases están compuestos por clases y por relaciones entre ellas. Las relaciones que se pueden usar son (Vidal, s.f.):

Relación de asociación: Una asociación es una conexión entre clases, una conexión semántica (enlace) entre los objetos de dichas clases. Un tipo especial de asociación es la relación de agregación (Vidal, s.f.).

Relación de dependencia: Una dependencia es una relación entre elementos, uno independiente y otro dependiente. Un cambio en el elemento independiente afectará al elemento dependiente (Vidal, s.f.).

Relación de generalización: Una generalización es una relación entre un elemento más general y otro más específico. El elemento más específico puede contener sólo información

adicional. Una instancia (un objeto es una instancia de una clase) del elemento más específico se puede usar si el elemento más general lo permite (Vidal, s.f.).

2.8. Seguridad de Software y Estándar ISO 27002

2.8.1. Estándar ISO 27002

La ISO 27001 es la normativa internacional que tiene como finalidad la protección de la confidencialidad de datos e información dentro de una organización. Esta norma proporciona el marco de trabajo para establecer un SGSI, que después puede ser certificado oficialmente por el organismo ISO (Estruga, 2020).

2.8.2. Estandar ISO/IEC 25010

El modelo de calidad representa la piedra angular en torno a la cual se establece el sistema para la evaluación de la calidad del producto. En este modelo se determinan las características de calidad que se van a tener en cuenta a la hora de evaluar las propiedades de un producto software determinado (25000, s.f.).

La calidad del producto software se puede interpretar como el grado en que dicho producto satisface los requisitos de sus usuarios aportando de esta manera un valor. Son precisamente estos requisitos (funcionalidad, rendimiento, seguridad, mantenibilidad, etc.) los que se encuentran representados en el modelo de calidad, el cual categoriza la calidad del producto en características y subcaracterísticas (25000, s.f.).

2.9. Integración de Tecnología

2.9.1. Python

Es un lenguaje de programación de alto nivel, con una sintaxis muy limpia que favorece un código legible y entendible por el ser humano. Posee una gran cantidad de librerías, con la cual se pueden crear programas rápidamente. Al ser multiplataforma se puede ejecutar en diferentes sistemas operativos y es de código abierto (Condor Tinoco & De La Cruz Roca, 2020).

2.9.2. Django

Es un marco web Python de alto nivel que promueve un desarrollo rápido y un diseño limpio y pragmático. Elimina la mayor parte de las molestias de desarrollo web para enfocarse en escribir aplicaciones sin tener que empezar de cero. Es gratuito y de código abierto (Django Project, 2024).

2.9.3. API

Es un conjunto de definiciones y protocolos para desarrollar e integrar sistemas de software de aplicaciones. Generalmente se piensa como un contrato entre el proveedor de información y el usuario que define lo que necesita el consumidor (la petición) y lo que necesita el productor (la respuesta) (Red Hat, 2023).

REST Estilo arquitectónico y una metodología de uso frecuente en el desarrollo de servicios de internet (Naeem, 2024).

2.9.4. Gestor de Base de Datos PostGreSQL

Es un sistema de base de datos objeto-relacional distribuido bajo licencia BSD y con su código fuente distribuido gratuitamente. Funciona según el modelo cliente/servidor, en lugar de utilizar un enfoque de multihilos, opera bajo multiprocesos para garantizar la estabilidad del sistema. Esto significa que una falla en uno de los procesos no afectará el funcionamiento de los demás procesos, permitiendo que el sistema continúe funcionando sin interrupciones (Zea, Ordoñez, Molina, Ríos, & Redrován, Castillo, 2017).

Es una base de datos que integra los principios ACID, mantiene una integridad referencia, posee una capacidad de realizar copias de seguridad en tiempo real, posee múltiples métodos de autenticación, acceso encriptado vía SSL asegurando que las operaciones sean confiables y consistentes (Zea, Ordoñez, Molina, Ríos, & Redrován, Castillo, 2017).

2.9.5. React

React una biblioteca de JavaScript que está enfocada para trabajar e interactuar con componentes. De este modo el pensamiento está muy en línea con el desarrollo de la interfaz de usuario. Se piensa con detenimiento en el concepto de estado y donde ubicarlo. Se caracteriza por el pragmatismo y ofrece a los desarrolladores varias opciones. Si el modelo de

React no satisface las necesidades del proyecto, se puede solventar con opciones de bajo nivel. Este enfoque de React hace que este desacoplado de la web. La biblioteca de React de Facebook es fácil de que se acople a diferentes dispositivos gracias al modelo de objetos de documento (DOM) (Vepsäläinen & Expósito, 2017).

2.9.6. Visual Studio Code

Visual Studio Code (VS Code) es un editor de código fuente desarrollado por Microsoft, diseñado para ser ligero, rápido y altamente personalizable. A diferencia de otros editores de código, como los IDE completos (Entornos de Desarrollo Integrados), VS Code se enfoca en ofrecer una experiencia ágil para la escritura y edición de código, sin sacrificar características avanzadas (Cuadrado, 2022).

Lo que lo diferencia es su capacidad de ser extendido con una gran cantidad de extensiones y herramientas, adaptándose a las necesidades específicas de cada desarrollador (Cuadrado, 2022).

Entre sus características clave, se incluyen la autocompletado inteligente mediante IntelliSense, la depuración integrada, y el control de versiones con Git, lo que convierte a VS Code en una opción versátil tanto para desarrolladores individuales como para equipos (Cuadrado, 2022).

Uno de los mayores puntos a favor de Visual Studio Code es que es completamente gratuito y multiplataforma. Esto significa que puedes descargar y utilizar VS Code en Windows, macOS y Linux sin costo alguno. Esta accesibilidad ha facilitado su adopción tanto por parte de desarrolladores principiantes como por equipos profesionales que buscan una herramienta eficiente sin tener que invertir en costosas licencias de software. Además, al ser un proyecto de código abierto, la comunidad puede contribuir al desarrollo del editor y crear nuevas extensiones para mejorar su funcionalidad (Cuadrado, 2022).

2.9.7. Git Hub

Github es un portal creado para alojar el código de las aplicaciones de cualquier desarrollador, y que fue comprada por Microsoft en junio del 2018. La plataforma está creada para que los desarrolladores suban el código de sus aplicaciones y herramientas, y que como usuario no

sólo puedas descargarte la aplicación, sino también entrar a su perfil para leer sobre ella o colaborar con su desarrollo (Fernandéz, 2019).

Como su nombre indica, la web utiliza el sistema de control de versiones Git diseñado por Linus Torvalds. Un sistema de gestión de versiones es ese con el que los desarrolladores pueden administrar su proyecto, ordenando el código de cada una de las nuevas versiones que sacan de sus aplicaciones para evitar confusiones. Así, al tener copias de cada una de las versiones de su aplicación, no se perderán los estados anteriores cuando se va a actualizar (Fernandéz, 2019).

Así pues, Git es uno de estos sistemas de control, que permite comparar el código de un archivo para ver las diferencias entre las versiones, restaurar versiones antiguas si algo sale mal, y fusionar los cambios de distintas versiones. También permite trabajar con distintas ramas de un proyecto, como la de desarrollo para meter nuevas funciones al programa o la de producción para depurar los bugs (Fernandéz, 2019).

CAPITULO III

MARCO METODOLÓGICO

3.1. Tipo de Investigación

3.1.1. Investigación Descriptiva

El presente proyecto utiliza un enfoque de investigación **Descriptiva**, orientado a detallar el funcionamiento actual de la gestión de correspondencia en la Federación Departamental de La Paz. Este tipo de investigación permite observar y documentar las características, procesos y desafíos específicos que presenta la gestión documental, con el objetivo de identificar las áreas que requieren optimización. Al aplicar un algoritmo de búsqueda semántica, se busca mejorar la eficiencia en la localización de documentos, contribuyendo a un sistema de gestión que sea más preciso y rápido en la recuperación de información. La investigación descriptiva permite, además, comprender cómo un sistema automatizado podría beneficiar la administración de la correspondencia, optimizando tanto la organización como el acceso a los documentos.

3.2. Diseño de la investigación

El diseño de investigación de este proyecto es **No Experimental** porque no se modificaron las variables de forma intencionada. En lugar de eso, se utilizaron datos reales sobre la gestión de correspondencia tal como ocurren en la práctica. Este enfoque permite observar los procesos tal cual son, sin alterar nada, lo que ayuda a obtener una visión más clara y real de la situación. Al trabajar con datos reales y no intervenir en el proceso, el estudio puede ofrecer una perspectiva más auténtica de cómo funcionan los sistemas de correspondencia en la Federación Departamental de La Paz.

El enfoque **Cualitativo** permite entender en profundidad los procesos y necesidades específicos del sistema de gestión de correspondencia actual. Este enfoque facilita la exploración de las percepciones y experiencias de los usuarios, centrándose en cómo la automatización de la gestión documental y la optimización en la localización de documentos pueden impactar y mejorar las prácticas existentes en la Federación Departamental de La Paz.

3.2.1. Cronograma Detallado

3.2.1.1.1. Planificación del Estudio y Diseño del Sistema

La planificación del estudio es un proceso crucial que establece las bases para el desarrollo del sistema de gestión de correspondencia que integra un algoritmo de búsqueda semántica. Esta fase incluye la identificación de los objetivos, el alcance del proyecto y la metodología a seguir para garantizar que el sistema se ajuste a las necesidades de los usuarios y optimice la gestión documental.

En primer lugar, se llevará a cabo un análisis exhaustivo de los procesos actuales de gestión de correspondencia en la Federación Departamental de La Paz. Esto permitirá identificar las áreas que requieren mejoras y definir los requisitos funcionales y no funcionales del nuevo sistema. Se realizarán entrevistas y encuestas a los usuarios para recopilar información sobre sus percepciones y experiencias con el sistema actual, lo que facilitará la comprensión de sus necesidades y expectativas.

El diseño del sistema se basará en la metodología de desarrollo ágil, específicamente utilizando principios de la metodología Scrum. Este enfoque permite una adaptación continua a los cambios y una colaboración efectiva entre los diferentes actores involucrados en el proyecto. Se crearán prototipos del sistema que se someterán a revisión y retroalimentación por parte de los usuarios, asegurando que el producto final sea intuitivo y eficaz en la automatización de la gestión documental.

3.2.1.1.2. Recolección de Datos

En esta fase, se llevará a cabo un proceso sistemático de recopilación de información relevante sobre los procesos actuales de gestión de correspondencia. Se utilizarán técnicas como entrevistas, encuestas y revisión de documentos para obtener una comprensión detallada de las necesidades y expectativas de los usuarios.

3.2.1.1.3. Análisis de los Datos

Una vez que se hayan recolectado los datos, se procederá a su análisis para identificar patrones, tendencias y áreas de mejora en la gestión documental. Este análisis permitirá transformar los datos recopilados en información útil que guiará el diseño y desarrollo del sistema.

3.2.1.1.4. Diseño del Sistema

En esta etapa, se desarrollará el diseño del sistema de gestión de correspondencia, teniendo en cuenta los requerimientos funcionales y no funcionales. Se crearán prototipos que reflejen la interfaz y la arquitectura del sistema, los cuales serán revisados y validados por los usuarios para asegurar su adecuación.

3.2.1.1.5. Desarrollo del Sistema

Esta fase implica la programación y construcción del sistema de acuerdo con el diseño previamente establecido. Se utilizarán herramientas y tecnologías apropiadas para implementar las funcionalidades del sistema, incluyendo el algoritmo de búsqueda semántica.

3.2.1.1.6. Fases de Prueba y Evaluación

Una vez desarrollado el sistema, se llevarán a cabo pruebas rigurosas para asegurar que cumple con los requisitos establecidos y funciona correctamente. Se realizarán evaluaciones tanto técnicas como de usabilidad, y se recogerá feedback de los usuarios para realizar las mejoras necesarias.

3.3. Selección de la Población

3.3.1. Universo

El proyecto abarca a todos los involucrados que están vinculados directamente con los procesos de gestión de correspondencia. A continuación, se detallará la población que comprende este universo:

3.3.1.1. Directorio

Los directivos son los encargados de la administración estratégica y operativa de la Federación Departamental de La Paz. Su responsabilidad principal es garantizar el cumplimiento de los objetivos organizacionales, la toma de decisiones clave y la supervisión de las actividades diarias de la institución.

- | | |
|--------------------------|----------------------------|
| ✓ Secretario Ejecutivo | ✓ Secretario de Relaciones |
| ✓ Secretario General | ✓ Secretario de Conflictos |
| ✓ Secretario de Finanzas | ✓ Secretario de Actas |

3.3.1.2. Comité Disciplinario

El Comité Disciplinario es un órgano clave dentro de la estructura organizativa, compuesto por cuatro miembros. Su principal responsabilidad es velar por el cumplimiento de las normativas internas y la ética sindical. Este comité tiene la función de investigar y tomar decisiones sobre posibles faltas o infracciones cometidas por miembros afiliados a la Federación Departamental La Paz, asegurando que se mantenga un ambiente profesional y respetuoso dentro de la organización.

3.3.1.3. Asistente Ejecutiva

La asistente ejecutiva tiene un papel fundamental en la operatividad diaria de la federación, brindando apoyo a los directivos y al equipo en general. Su responsabilidad incluye la gestión de agendas, la coordinación de reuniones, la preparación de informes, así como la atención a los requerimientos administrativos y logísticos de la organización. Además, la asistente ejecutiva colabora en tareas de comunicación interna y externa, facilitando el flujo de información y asegurando que las actividades administrativas se realicen de manera eficiente.

Tabla 6

Población

Nro.	Actores	Cantidad
1	Directivos	6
2	Comité Disciplinario	5
3	Asistente Ejecutiva	1
TOTAL		11

Fuente: Elaboración Propia

3.3.2. Muestra

Criterios para seleccionar la muestra

Relevancia: Los participantes seleccionados tienen un papel directo en la gestión de correspondencia, lo que asegura la validez de los datos recolectados.

Accesibilidad: Se seleccionó un grupo accesible en términos de disponibilidad de tiempo y disposición para participar en el estudio.

Diversidad de roles: Se incluyeron personas de diferentes cargos para obtener una perspectiva integral.

3.3.3. Método

Se utilizará el método de recolección de datos **No Probabilístico Por Conveniencia**, ya que permite seleccionar de manera intencional y práctica la muestra en función de la accesibilidad y relevancia de los participantes. La población total está compuesta por 11 personas: seis directivos, cinco miembros del comité de disciplinario y 1 asistente. Se optará por seleccionar a los directivos y la asistente, quienes son los responsables directos de la gestión de correspondencia.

Los directivos supervisan y toman decisiones clave sobre los procesos de recepción, distribución y almacenamiento de documentos. Por su parte, la asistente está involucrada diariamente en la gestión operativa de la correspondencia, lo que la convierte en una fuente valiosa de información. Ambos grupos, por su rol y cercanía con los procesos de correspondencia, son los más indicados para proporcionar datos relevantes para este estudio.

Tabla 7

Muestra

Nro.	Actores	Cantidad
1	Directorio	6
2	Asistente Ejecutiva	1
TOTAL		7

Fuente: Elaboración Propia

3.4. Instrumento de Recolección de Datos

Para la obtención de la información, se utilizarán diversos instrumentos de recolección de datos que permitirán obtener información precisa y relevante para cumplir con los objetivos planteados. Los datos recolectados serán utilizados para diseñar y verificar el funcionamiento del sistema de gestión de correspondencia, asegurando que se adapte a las necesidades y procesos de la Federación Departamental de La Paz.

3.4.1. Entrevistas

Se llevarán a cabo entrevistas con los directivos y la asistente ejecutiva, quienes tienen un conocimiento profundo y directo de los procesos de gestión de correspondencia. Las entrevistas permitirán obtener información cualitativa detallada sobre los procedimientos actuales, los problemas que enfrentan y las expectativas respecto al nuevo sistema, brindando una perspectiva más completa y contextualizada.

3.4.2. Observación

La técnica de observación permitirá registrar y analizar la información obtenida mediante la observación directa de las acciones y comportamientos de los actores clave involucrados en los procesos de gestión de correspondencia. Este enfoque será esencial para comprender de manera detallada cómo se desarrollan las actividades relacionadas con el manejo de correspondencia, identificando patrones, prácticas habituales y posibles áreas de mejora.

3.4.3. Encuestas

Se distribuirán encuestas a los mismos participantes para obtener datos cuantitativos sobre la eficiencia de los procesos actuales y la percepción que tienen sobre la gestión de correspondencia. Las encuestas permitirán recopilar respuestas estructuradas que facilitarán el análisis estadístico y la identificación de áreas específicas que requieren optimización.

Bibliografía

- Aguirre, M. (14 de 05 de 2024). *Athento*. Obtenido de Athento: <https://www.athento.com/es/como-digitalizar-y-optimizar-la-gestion-de-correspondencia-con-athento/>
- Aplyca. (04 de 07 de 2023). Obtenido de Aplyca: <https://www.aplyca.com/blog/Introduccion-a-la-busqueda-semantic>
- Archivos Empresariales del Caribe. (s.f.). Obtenido de Archivos Empresariales del Caribe: <https://www.archivocaribe.com/relacion-entre-tecnologia-y-gestion-documental-descubre-su-impacto-y-beneficios/>
- Archivos, C. d. (1995). Diccionario de Terminología Archivística. Madrid, España. Obtenido de CATÁLOGO COLECTIVO DE LA RED DE BIBLIOTECAS DE LOS ARCHIVOS ESTATALES.
- Ayerdi, A. (25 de 10 de 2024). *DocuWare*. Obtenido de DocuWare: <https://start.docuware.com/es/blog/tecnologia-ocr-docuware>
- Baratz. (18 de 03 de 2021). *Comunidad Baratz*. Obtenido de Comunidad Baratz: <https://www.comunidadbaratz.com/blog/los-7-procesos-de-la-gestion-documental-en-empresas-y-organizaciones/>
- Bel, W. (2020). *Algoritmos y estructuras de datos en Python: un enfoque ágil y estructurado*. Entre Rios: UADER.
- Catalunya, U. O. (s.f.). *Universitat Oberta de Catalunya*. Obtenido de Universitat Oberta de Catalunya: <https://www.uoc.edu/portal/es/arxiu/gestio-documental/que-es/index.html>
- Chinacota, A. d. (mayo de 2019). Manual de Archivo y Correspondencia. Chinacota, Colombia.
- Documentos, C. d. (2011). Administración de documentos y archivos. Madrid, España.
- España, G. d. (s.f.). *administracion.gob.es*. Obtenido de administracion.gob.es: https://administracion.gob.es/pag_Home/espanaAdmon/Publicaciones-oficiales-y-documentos-administrativos/Documentos-Administrativos-basica.html
- Estatales, S. G. (1995). Diccionario de terminología archivística. Madrid, Madrid, España.
- Europeo, P. (26 de 03 de 2021). ¿Qué es la inteligencia artificial y cómo se usa?
- Fernández, J. B. (s.f.). La Tecnología de la Información al Servicio de la Archivística.
- Fernández, J. B. (s.f.). La Tecnología de la Información al Servicio de la Archivística.
- Gallardo, A. J. (2020). Control Predictivo basado en Machine Learning de una planta de laboratorio. Sevilla, Españan.
- Grupo ASD. (20 de 12 de 2023). *Grupo ASD*. Obtenido de Grupo ASD: <https://www.grupoasd.com/como-la-tecnologia-ocr-esta-transformando-la-gestion-de-documentos/>
- Hernandez, J. H. (10 de 06 de 2021). Procesamiento del Lenguaje Natural. La Laguna, España.
- Hurtado, J. S. (3 de 12 de 2021). *IEBS*. Obtenido de IEBS: <https://www.iebschool.com/blog/metodologia-scrum-agile-scrum/>
- Irene, J. (2021). *Mecanografía*. Obtenido de <https://cpem25.wordpress.com/wp-content/uploads/2021/03/mecanografia-redacciones-comerciales-2020.pdf>
- Kultura Saila. (2011). OCR: tecnología para el reconocimiento óptico de caracteres en una imagen.
- Luna, Y. (15 de 11 de 2012). *Slideshare*. Obtenido de Slideshar: <https://es.slideshare.net/slideshow/correspondencia-15197721/15197721>
- Matuszczyk, T. (9 de 9 de 2022). *Correspondence Management System*. Obtenido de Correspondence Management System: <https://correspondence-software.com/incoming-and-outgoing-correspondence/>
- MongoDB. (s.f.). *MongoDB*. Obtenido de MongoDB: <https://www.mongodb.com/es/resources/basics/semantic-search>
- Mundial, B. (11 de 01 de 2020). Gestión Documental Principios y Prácticas.

NORMADAT. (04 de 09 de 2023). *NORMADAT*. Obtenido de NORMADAT: <https://www.normadat.es/ocr-como-functiona-el-sotware-de-reconocimiento-optico-de-caracteres/>

Olivas, E. S. (2023). *Sistemas de Aprendizaje Automático*. Bogotá: Ra-ma.

Patricio Letelier, M. C. (Enero de 2006). Metodologías ágiles para el desarrollo de software: eXtreme Programming (XP). Valencia, España.

Pilataxi, A. M. (11 de 2009). Sistema Web para la organización local miembro (OLM) de la Junior Chamber International (JCI) Quito Metropolitano. Quito, Ecuador.

Ponce, J. (21 de 02 de 2021). *Jahaziel Ponce*. Obtenido de Jahaziel Ponce: <https://jahazielponce.com/algoritmos-de-busqueda/>

Pública, E. d. (2015). Los documentos administrativos. Murcia, España.

Quirola, K. N. (s.f.). *SCRIBD*. Obtenido de SCRIBD: <https://es.scribd.com/document/706868216/Tema-7-Gestion-de-correspondencia-e-informacion-Resumen>

Reducindo, I. (01 de 2022). La gestión documental y las tecnologías de la información como apoyo a la gestión curricular y el currículum vivido. Mexico.

Rouhiainen, L. (2018). *Inteligencia Artificial 101 cosas que debes saber hoy sobre nuestro futuro*. España: Alienta.

Ruiz, F. F. (2001). Archivística, Archivo, Documento de Archivo. Lima, Perú.

Sonora, I. T. (03 de 2016). Pseudocódigo y PSEINT. Mexico.

Tafolla, C. (s.f.). *SCRIBD*. Obtenido de SCRIBD: <https://es.scribd.com/document/435177762/Que-Es-La-Correspondencia>

TecnoDigital. (06 de 07 de 2023). *Informatica y Tecnologia Digital*. Obtenido de Informatica y Tecnologia Digital: <https://informatecdigital.com/algoritmos/algoritmos-de-busqueda-que-son-y-como-funcionan/>

Tipos. (s.f.). Obtenido de Tipos: <https://www.tipos.co/tipos-de-correspondencia/>

Uribe, I. (20 de 06 de 2023). *Secmotic Smart Inside*. Obtenido de Secmotic Smart Inside: <https://secmotic.com/inteligencia-artificial-toma-decisiones/#gref>

Vidal, M. V. (2008). *Apuntes de Semántica Lexica*. Madrid: UNED.