

# 

Integrantes: Nicolás Briceño  
 Sebastián Donoso  
 Andrés Lozano  
 Arnaldo Morales  
 Joseph Salas  
 Joaquín Santander

Empresa: Fundación para la Innovación Agraria.  
  
Entrega Final: Taller de Empresa I, Universidad San Sebastián.

**Sistema de información tipo ChatGPT a Biblioteca digital del FIA**

Indice

[Introducción 2](#_Toc170424396)

[Reseña Empresa 3](#_Toc170424397)

[Descripción del Desafío 4](#_Toc170424398)

[Objetivo General 4](#_Toc170424399)

[Objetivos Específicos 4](#_Toc170424400)

[Estado del arte del Desafío 5](#_Toc170424401)

[Investigación respecto del desafío 5](#_Toc170424402)

[Marco Teórico 7](#_Toc170424403)

[1. Tema tratado. 7](#_Toc170424404)

[2. Antecedentes. 7](#_Toc170424405)

[3. Conceptos Relevantes. 8](#_Toc170424406)

[4. Tecnologías relevantes para complementar el desarrollo. 9](#_Toc170424407)

[5. Resultados Esperados. 9](#_Toc170424408)

[Parámetros e información específica del Desafío 10](#_Toc170424409)

[Levantamiento de requerimientos de organización 10](#_Toc170424410)

[Requerimientos No funcionales 10](#_Toc170424411)

[Levantamiento de requerimientos del desafío 15](#_Toc170424412)

[Requerimiento Funcionales 15](#_Toc170424413)

[Definición y caracterización del habilitador tecnológico 17](#_Toc170424414)

[Inteligencia Artificial (IA) 17](#_Toc170424415)

[Propuesta de solución 19](#_Toc170424416)

[1. Idea principal 19](#_Toc170424417)

[2. Recursos 20](#_Toc170424418)

[3. Datos 20](#_Toc170424419)

[4. Herramientas y conceptos 20](#_Toc170424420)

[5. Metodología 22](#_Toc170424421)

[Diagrama de componentes 22](#_Toc170424422)

[Diagrama de Clases 23](#_Toc170424423)

[Arquitectura solución-diagrama de contexto 23](#_Toc170424424)

[Diseño de Solución 24](#_Toc170424425)

[Diagramas UML Solución 27](#_Toc170424426)

[Prototipo 30](#_Toc170424427)

[Alcance del Prototipo 30](#_Toc170424428)

[Principales Funcionalidades 30](#_Toc170424429)

[Aprendizaje 31](#_Toc170424430)

[Conclusión 31](#_Toc170424431)

# Introducción

La Fundación para la Innovación Agraria ha sido clave en el impulso de la investigación y la innovación en sectores agrícolas, pecuarios, forestales y acuícolas desde su creación en 1981. Ahora, se enfrenta al desafío de modernizar su biblioteca digital para mejorar la accesibilidad a su contenido. Con el objetivo de hacer que la búsqueda de información sobre los proyectos financiados sea más eficiente y efectiva, la fundación planea implementar un chatbot o asistente virtual impulsado por inteligencia artificial en el sitio web de su Biblioteca Digital.

Este chatbot no sólo facilitará la búsqueda de información sobre los proyectos financiados por la FIA, sino que también proporcionará resúmenes útiles y respuestas precisas a consultas realizadas por cualquier usuario a través del sitio web.

El informe detallado del proyecto aborda los objetivos generales y específicos, así como los desafíos técnicos y organizativos asociados con la implementación de esta solución tecnológica.

# Reseña Empresa

La Fundación para la Innovación Agraria (FIA), fue establecida en 1981 por el Ministerio de Justicia de Chile, ha desempeñado un papel importante en el impulso de la investigación y la innovación en el sector agrícola, pecuario, forestal y acuícola del país. Inicialmente nombrada Fondo de Investigación Agropecuaria, con el paso de los años llegó a FIA en 1996, y desde su creación ha marcado un compromiso continuo en apoyo a la transformación y modernización de la agricultura y la economía rural chilena.  
Su dedicación va enfocada a promover procesos de innovación bajo lineamientos estratégicos propios para el sector silvoagropecuario, ayudando a impulsar el desarrollo de capacidades y difusión tecnológica de iniciativas que contribuyen con el desarrollo sostenible y la competitividad de Chile.

Esta institución es reconocida por el Sistema Nacional de Innovación, y su misión es “Contribuir a la solución eficiente de desafíos estratégicos del sector silvoagropecuario nacional y/o de la cadena agroalimentaria asociada, por medio del fomento, articulación y difusión tecnológica de procesos de innovación orientados al desarrollo sustentable”.

# Descripción del Desafío

## Objetivo General

La Fundación para la Innovación Agraria posee una Biblioteca Digital que contiene una gran cantidad de documentos e información en libros, documentos, presentaciones, iniciativas y videos, entre otros. Referentes a iniciativas en el sector silvoagropecuario y la cadena agroalimentaria asociada. Ante esto, el objetivo del proyecto es implementar mediante el uso de inteligencia artificial, un chat o asistente virtual para su sitio web que sirva de ayuda para buscar o resumir información respecto a sus proyectos, de manera que estas consultas realizadas por usuarios externos o internos sean mucho más fluidas y eficaces.

## Objetivos Específicos

* Interiorizarse y comprender el funcionamiento actual de la Biblioteca.
* Desarrollar un Producto Mínimo Viable (MVP) de solución que luego sea mejorable para cumplir el desafío.
* Entrenar el modelo de inteligencia artificial con datos relevantes de la Biblioteca Digital para mejorar su capacidad de comprensión y respuesta a las consultas de los usuarios.
* Garantizar la seguridad y la privacidad de los datos de la Biblioteca Digital mediante la implementación de medidas de protección adecuadas en el sistema de chat o asistente virtual.
* Evaluar periódicamente en su fase de prueba el rendimiento del sistema de chat o asistente virtual, y realizar actualizaciones y optimizaciones para mantener su eficacia y relevancia a lo largo del tiempo.

# Estado del arte del Desafío

## Investigación respecto del desafío

En la actualidad, cada vez se hace más común encontrarnos con inteligencias artificiales incorporadas en páginas web, normalmente son chatbot, que permiten una interacción básica con el cliente, y que, si se complica la interacción, el chatbot deriva su función a una persona que esté capacitada.

Para este desafío, también se espera contar con un chatbot que pueda ayudar a recolectar la información disponible entre los documentos de proyectos con los que cuenta FIA en su Biblioteca Digital, para así, simplificar el esfuerzo que hace el usuario a la hora de buscar información.

Para esto, la IA se basaría en los siguientes pasos:

1. Recolección de datos.
2. Preprocesamiento del texto.
3. Análisis del texto.
4. Resumen del texto.
5. Extracción de la información.

Esto beneficiaría a los usuarios, mejorando el acceso a la información, realizando un mejor análisis de esta e identificando información clave.

En base a lo comentado anteriormente, existen algunas IAs que funcionan de manera similar a lo que se pretende crear en este proyecto, como por ejemplo las siguientes inteligencias:

* ChatGPT: Probablemente la IA más famosa en la actualidad, que, con su capacidad de procesamiento, puede concretar fácilmente todo lo que se propone en este proyecto.
* Gemini: La nueva IA de Google, que tiene un funcionamiento similar a la IA nombrada anteriormente.
* SciSpace: Una plataforma especializada en el ámbito académico que utiliza IA, y que permite analizar y comprender publicaciones científicas.
* Browse.ai: Otra herramienta de extracción de datos, pero que encuentra información de páginas web.
* Amazon Comprehend: Un servicio de procesamiento de lenguaje natural que ofrece funcionalidades como: clasificación de temas, extracción de identidades y análisis de sentimientos.

En base a lo investigado, podemos concluir que existen varias herramientas disponibles que pueden servir como guía, o derechamente implementarlas de alguna manera dentro de la Biblioteca Digital FIA, para poder cumplir con los objetivos propuestos.

## Marco Teórico

### 1. Tema tratado.

La Fundación para la Innovación Agraria (FIA) quiere implementar un chatbot o asistente virtual en su Biblioteca Digital para facilitar la búsqueda y/o resumen de información sobre sus proyectos. Por lo que este marco teórico presenta los conceptos clave y las tecnologías relevantes para el desarrollo de este proyecto.

### 2. Antecedentes.

La Biblioteca Digital de la FIA ofrece un sistema de búsqueda que incluye un buscador tradicional y herramientas más avanzadas para cumplir con lo que necesiten los usuarios. El buscador general permite realizar búsquedas por palabras clave y categorías, mostrando una lista de resultados con documentos relevantes. Esta funcionalidad es accesible desde la página principal de la biblioteca.

Para consultas más detalladas, la Biblioteca Digital dispone de una búsqueda avanzada. Esta herramienta permite a los usuarios refinar sus búsquedas por título, autor, fecha, región, sector e iniciativa. Este filtrado permite obtener resultados más precisos y específicos.

Además, la biblioteca ofrece opciones de búsqueda directa por colecciones específicas, como iniciativas, libros, presentaciones y videos. Esto facilita la localización de información dentro de áreas temáticas particulares.

La Biblioteca Digital de la FIA organiza su contenido en cuatro colecciones principales: Iniciativas, Libros, Presentaciones y Videos.

### 3. Conceptos Relevantes.

#### 3.1 Chatbots y asistentes virtuales.

Los chatbots y asistentes virtuales son programas informáticos que simulan la conversación con usuarios humanos a través de interfaces de texto o voz. Estos sistemas utilizan técnicas de inteligencia artificial, como el procesamiento del lenguaje natural (PLN) y el aprendizaje automático, para comprender las consultas de los usuarios y proporcionar respuestas relevantes.

 Los chatbots y asistentes virtuales tienen un amplio rango de aplicaciones, incluyendo:

* Servicio al cliente: Brindar información y soporte a los clientes sobre productos, servicios o problemas técnicos.
* Búsqueda de información: Ayudar a los usuarios a encontrar información específica en sitios web o bases de datos.
* Resumen de información: Proporcionar resúmenes concisos de documentos o artículos extensos.
* Atención personalizada: Ofrecer recomendaciones o sugerencias personalizadas a los usuarios.

#### 3.2 Procesamiento del lenguaje natural (PLN)

El PLN es una rama de la inteligencia artificial que se enfoca en la interacción entre las computadoras y el lenguaje humano. El PLN se utiliza en una variedad de tareas, incluyendo:

* Análisis sintáctico: Comprender la estructura gramatical de las oraciones.
* Análisis semántico: Extraer el significado de las palabras y frases.
* Análisis pragmático: Comprender el contexto en el que se utiliza el lenguaje.
* Generación de lenguaje natural: Producir texto que sea gramaticalmente correcto y que tenga sentido para los humanos.

#### 3.3 Aprendizaje automático

El aprendizaje automático es una rama de la inteligencia artificial que permite a las computadoras aprender sin ser explícitamente programadas. El aprendizaje automático se utiliza en una variedad de tareas.

### 4. Tecnologías relevantes para complementar el desarrollo.

APIs e Integraciones: Pueden ser integradas diversas APIs y sistemas externos para ampliar la funcionalidad y obtener respuestas más útiles para los usuarios.

Interfaces de Usuario Conversacionales (CUI): Estas interfaces están diseñadas para imitar conversaciones humanas naturales y ofrecer una mejor experiencia. Ayudan con el diseño de diálogos, la gestión de contextos y la capacidad de manejar interacciones.

Seguridad y Privacidad: Como en todo desarrollo deben ser implementadas medidas de seguridad y privacidad para proteger la información confidencial de los usuarios y del sistema.

### 5. Resultados Esperados.

Se busca cumplir con los requerimientos de FIA, los cuales contemplan el desarrollo de un asistente virtual que permita resumir información almacenada dentro de la Biblioteca Virtual de la FIA de acuerdo con las necesidades de información que requieren los usuarios, de manera que este chatbot sea capaz interpretar, buscar y optimizar esta información, generando una respuesta clara y precisa a la consulta realizada.

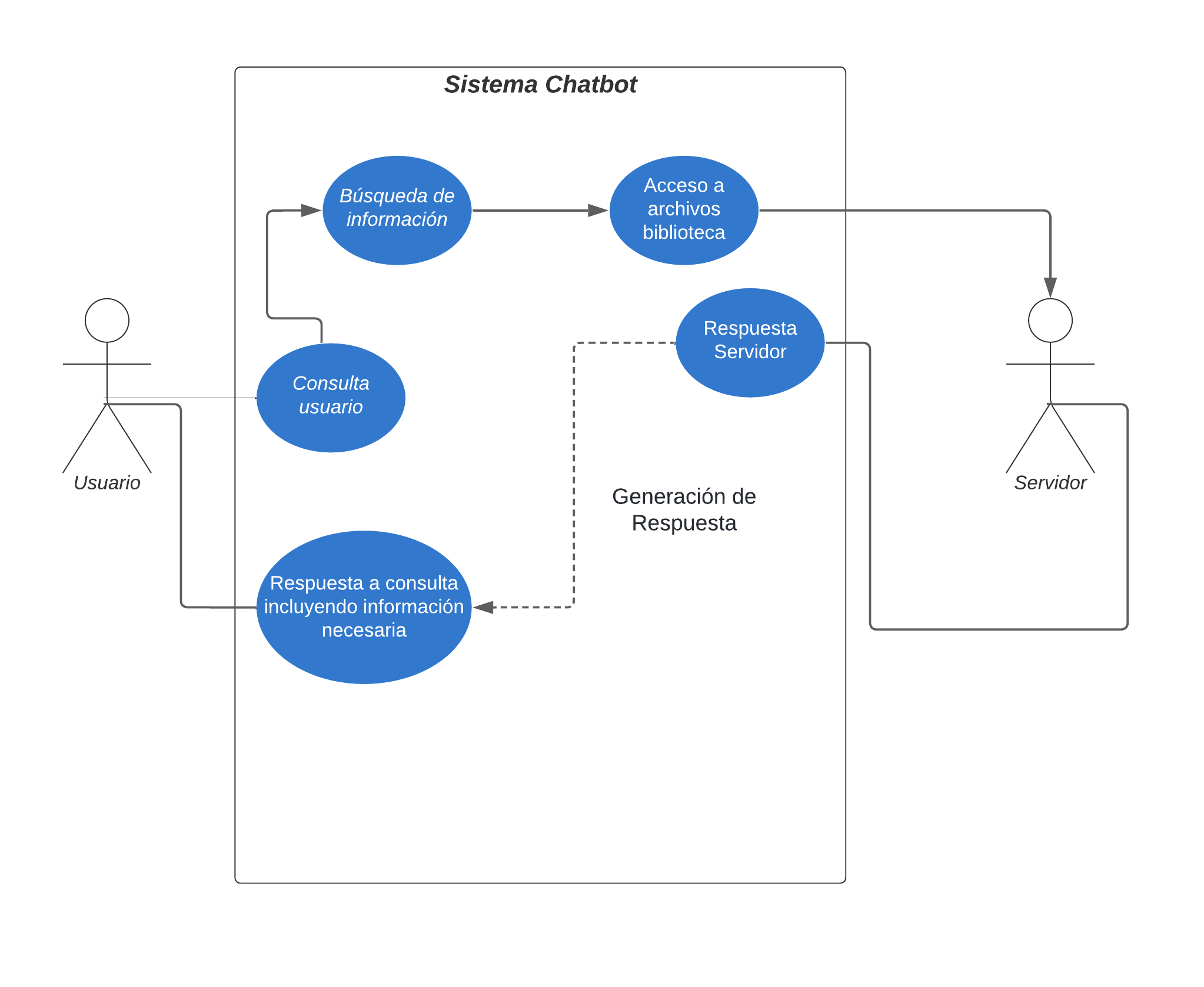
Llegar al resultado final contempla conectar los servicios de la Biblioteca Digital FIA al chatbot con la utilización de un Modelo de Lenguaje de Gran Tamaño (LLM) que permita agilizar las funciones de este chatbot, junto con la generación mejorada por recuperación (RAG) para optimizar la generación de las respuestas.

# Parámetros e información específica del Desafío

## Levantamiento de requerimientos de organización

### Requerimientos No funcionales

1. **Acceso Eficiente a la Información:**
   * Los usuarios necesitan acceder rápidamente a documentos digitales, imágenes y datos de investigación relacionados con el sector silvoagropecuario y la cadena agroalimentaria.



1. **Búsqueda avanzada:**
   * Los usuarios tanto internos como externos a FIA podrán realizar consultas de búsqueda bajos diferentes parámetros, obteniendo información precisa al respecto.

Diagrama

Descripción generada automáticamente

1. **Seguridad de la información de FIA:**
   * El desarrollo debe ser realizado bajo estándares de seguridad que garanticen a FIA la integridad de sus datos y sistemas.

Diagrama

Descripción generada automáticamente

1. **Optimización de Rendimiento**:
   * Asegurar un rendimiento óptimo del sistema, especialmente durante períodos de alta demanda.

Diagrama

Descripción generada automáticamente

1. **Escalabilidad y Mantenibilidad:**
   * Diagrama

     Descripción generada automáticamenteLa solución debe ser escalable para manejar un crecimiento futuro en la cantidad de contenido y usuarios, y debe ser fácilmente mantenible para realizar actualizaciones y correcciones.
2. **Respuesta y generación de texto:**
   * El chatbot debe responder con claridad para asegurar que sea comprensible y útil para el usuario (persona natural que interactúa con el chatbot).

Diagrama

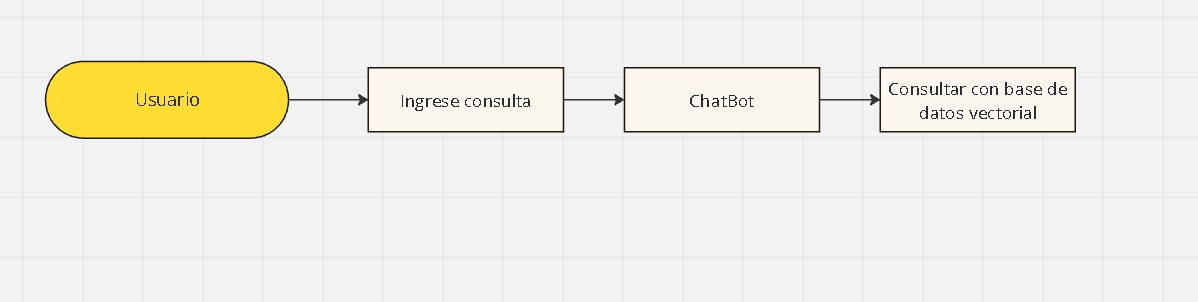
Descripción generada automáticamente

1. **Interfaz Intuitiva:**
   * Debe poseer una interfaz que sea amigable y entendible para el usuario, de manera que su experiencia sea agradable y fácil de utilizar.

Diagrama

Descripción generada automáticamente

1. **Conexión Base de Datos Vectorial:**
   * El chatbot debe estar conectado a una base de datos vectorial que posee el texto de los archivos de los proyectos vectorizados, para el entendimiento del chatbot.



## Levantamiento de requerimientos del desafío

### Requerimiento Funcionales

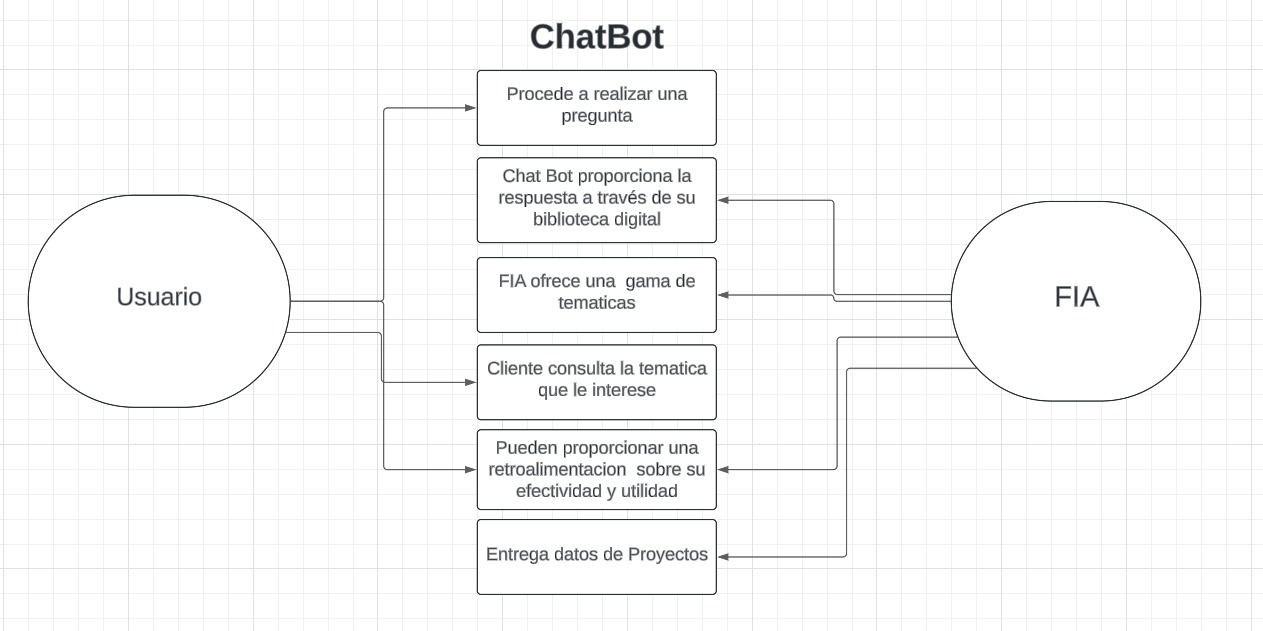
1. **Búsqueda de Información:**
   * El chatbot debe ser capaz de interpretar las consultas del usuario para buscar información precisa de los proyectos que contiene la biblioteca en la colección de iniciativas.

Diagrama

Descripción generada automáticamente

1. **Optimizar Información:**
   * Diagrama

     Descripción generada automáticamenteEl chatbot debe ser capaz de leer la información de los proyectos y desarrollar su respuesta en torno a la consulta, entendiendo las palabras que utiliza el usuario para dar una respuesta con respecto a esa misma información, mejorando la experiencia de búsqueda que existe actualmente.
2. **Calidad de respuestas:**
   * La IA debe ser capaz de generar respuestas y resúmenes en castellano considerando consultas o búsqueda de autores, búsquedas en base a temáticas específicas como rubros agrícolas, especies, región del proyecto, año, entre otros, y desplegar la respuesta de acuerdo con los parámetros definidos como fechas, buscar puntos claves, transcribir idioma, comparación con otros textos, ofrecer recomendaciones, citar la fuente de sus respuestas, etc.

****

# Definición y caracterización del habilitador tecnológico

## Inteligencia Artificial (IA)

La Inteligencia artificial es un conjunto de habilidades intelectuales expresadas por sistemas informáticos o combinaciones de algoritmos los cuales tienen el propósito de crear máquinas que imiten la inteligencia humana para realizar tareas, y que pueden mejorar conforme recopilen la información necesaria.

En su forma más simple, la inteligencia artificial es un campo que combina la informática y los conjuntos de datos para permitir la resolución de problemas. También los campos del aprendizaje automático y el aprendizaje profundo.

La IA se puede clasificar en función de la racionalidad y el pensamiento:

* Enfoque humano: Sistemas que piensan como humanos, sistemas que actúan como humanos.
* Enfoque ideal: Sistemas que piensan racionalmente, sistemas que actúan racionalmente.

En este proyecto será utilizado para manejar una biblioteca online (FIA) y servir como un asistente por lo que tendrá un enfoque más humano.

Por otra parte, lo que busca FIA en este proyecto, es la implementación de una inteligencia artificial dentro de su función de buscador, en el apartado de su biblioteca, donde se encuentran todos los proyectos que han finalizado su ejecución, para que, de esta manera, la búsqueda de información sea más focalizada a lo que quiere la persona.

Dentro de la biblioteca de FIA, existen aproximadamente más de 3600 documentos de texto, y alrededor de 700 documentos multimedia, es por esto, que es importante para esta organización, que su buscador funcione de manera eficiente, y que pueda realizar tareas automáticas, agilizando el proceso de búsqueda y mejorando el tiempo de obtención de respuesta para extraer información relevante.

La solución a esta problemática, viene de la mano del desarrollo de una inteligencia artificial, la cual será una especie de chatbot, con funcionalidades como las de ChatGPT, Gemini o Copilot, ésta será implementada en el buscador, y podrá rescatar información importante de textos, como fechas, autores, localización del proyecto, entre otros, además de poder contar con otras capacidades más complejas de este tipo de inteligencia, como hacer resúmenes de los textos, extraer puntos claves dentro de los textos, explicar diferentes tipos de iniciativas en determinado tiempo o localización, u otras funciones que el usuario necesite o requiera.

El funcionamiento de esta inteligencia artificial será desarrollada y entrenada, específicamente con la base de datos que tiene FIA, sobre todos los documentos de proyectos que tiene su biblioteca, por lo que su funcionalidad será limitada simplemente a los datos que tiene esta organización en su base de datos.

# Propuesta de solución

## Idea principal

La implementación de un chatbot será una gran innovación dentro de la plataforma de FIA, esto permitirá a los usuarios tener una mejor experiencia a la hora de interactuar con la página y documentos almacenados en su biblioteca virtual, reduciendo el tiempo que demoren los usuarios a la hora de encontrar lo que necesitan, y haciendo la interacción más personal y agradable.

Una de las mayores preocupaciones a la hora de la seguridad de las respuestas del chatbot, es que no de información confidencial, o que no esté relacionada con la FIA, por lo que, en este caso, el chatbot estará completamente relacionado e informado con la documentación que cuenta de los documentos de FIA. Por lo que, solo responderá en base a lo que conoce, que son los documentos dentro de la biblioteca virtual de FIA, es decir, no está capacitado para responder otra pregunta que no sea la documentación, o algo que este programado directamente dentro de su modelo.

Una vez aclarado esto, podemos seguir hablando sobre las innovaciones con las que contará este modelo:

* Multilenguaje: Este modelo soporta la transcripción de idiomas, por lo que el lenguaje en el que se le hagan consultas no será una barrera.
* Consultas de usuarios: El chatbot, está capacitado para hacer búsquedas generales, por lo que, si se le pregunta sobre algo en general o algún tópico en específico, podrá buscar a través de su base de datos, y dar un resultado acorde a lo que se pregunta, dando respuestas certeras, relacionadas con la pregunta.
* Resúmenes: El modelo está capacitado para analizar un documento, y realizar un resumen de este, para que así, el usuario pueda tener una idea sobre el documento que está buscando.
* Extracción de información específica: El chatbot tiene la capacidad de poder extraer información en específico, por lo que, si se le pregunta sobre un autor, un proyecto, nombre, ubicación, rubro o tópico en específico, el modelo será capaz de responder en base a lo preguntado.

Además de estas funcionalidades, la idea de este chatbot es que sea sencillo de usar, permitiendo que cualquier persona pueda utilizarla sin problemas. La interfaz de usuario será algo clave, que sea intuitiva, y que permita a los usuarios interactuar fácilmente con ella, permitiendo obtener la información que busca.

## Recursos

* Acceso a información de la FIA a través de una base de datos vectorial Chroma.
* Documentación Ollama.
* Documentación ofrecida por Hugging Face (<https://huggingface.co/>) para el desarrollo del chatbot.
* Documentación Langchain.

## Datos

* El acceso a los datos será a través de la carga de documentos que proporciona FIA a una base de datos vectorial.

## Herramientas y conceptos

* Flask: Framework web escrito en Python proporciona las herramientas y librerías necesarias para crear endpoints (rutas) que pueden manejar solicitudes HTTP, como GET y POST, y devolver respuestas.
* Langchain: Es una biblioteca diseñada para trabajar con cadenas de procesamiento de lenguaje natural (NLP). Proporciona herramientas para dividir, procesar, y combinar documentos, así como para interactuar con modelos de lenguaje. Se utiliza para crear cadenas que manejan el flujo de datos desde la entrada del usuario hasta la respuesta del modelo de lenguaje.
* Ollama: Es una implementación específica de un modelo de lenguaje dentro de Langchain. Es un modelo preentrenado que puede procesar y generar texto en base a una consulta dada. En el contexto de este proyecto, Ollama se utiliza para manejar las consultas del usuario y generar respuestas.
* Llama 3: Es un modelo de lenguaje avanzado. En este proyecto, "llama3" se refiere al modelo de lenguaje específico que se utiliza dentro de Ollama para procesar las consultas.
* JSON (JavaScript Object Notation): Es un formato de texto para el intercambio de datos. En este proyecto, JSON se utiliza para enviar y recibir datos en las solicitudes y respuestas entre el cliente y el servidor.
* POST y GET: Son métodos HTTP utilizados para enviar y recibir datos desde un servidor web.
  + GET: Se utiliza para solicitar datos de un servidor.
  + POST: Se utiliza para enviar datos al servidor para crear o actualizar un recurso.
* Embeddings: En el contexto de NLP, los embeddings son representaciones vectoriales de palabras o documentos. Transforman texto en números para que los modelos de aprendizaje automático puedan procesarlos. En este proyecto, los embeddings se utilizan para convertir fragmentos de documentos en vectores que pueden ser comparados y almacenados en una base de datos vectorial.
* ChromaDB: Es una base de datos vectorial que almacena embeddings. Permite realizar búsquedas eficientes y recuperar documentos relevantes basándose en la similitud de los embeddings. En este proyecto, ChromaDB se utiliza para almacenar y buscar fragmentos de documentos PDF.
* Chunk: Un chunk (fragmento) es una porción de un documento más grande. Dividir documentos en chunks más pequeños facilita el procesamiento y la búsqueda. En este proyecto, los documentos PDF se dividen en chunks para ser indexados y almacenados en la base de datos vectorial.
* Chain: Una chain (cadena) en Langchain se refiere a una secuencia de pasos o procesos que transforman y manejan datos. Puede incluir la división de documentos, la búsqueda de información, y la generación de respuestas por parte del modelo de lenguaje. En este proyecto, las cadenas se utilizan para conectar la búsqueda de documentos relevantes con la generación de respuestas a las consultas de los usuarios.

## 5.Metodología

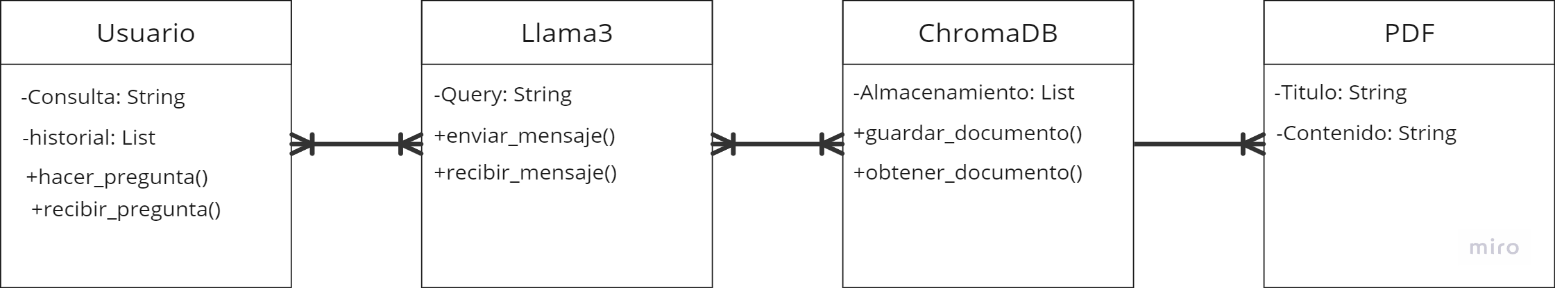
* Se utiliza la metodología Scrum como guía para ver el rendimiento del chatbot.

## Diagrama de componentes

Imagen que contiene Diagrama

Descripción generada automáticamente

## Diagrama de Clases



## Arquitectura solución-diagrama de contexto

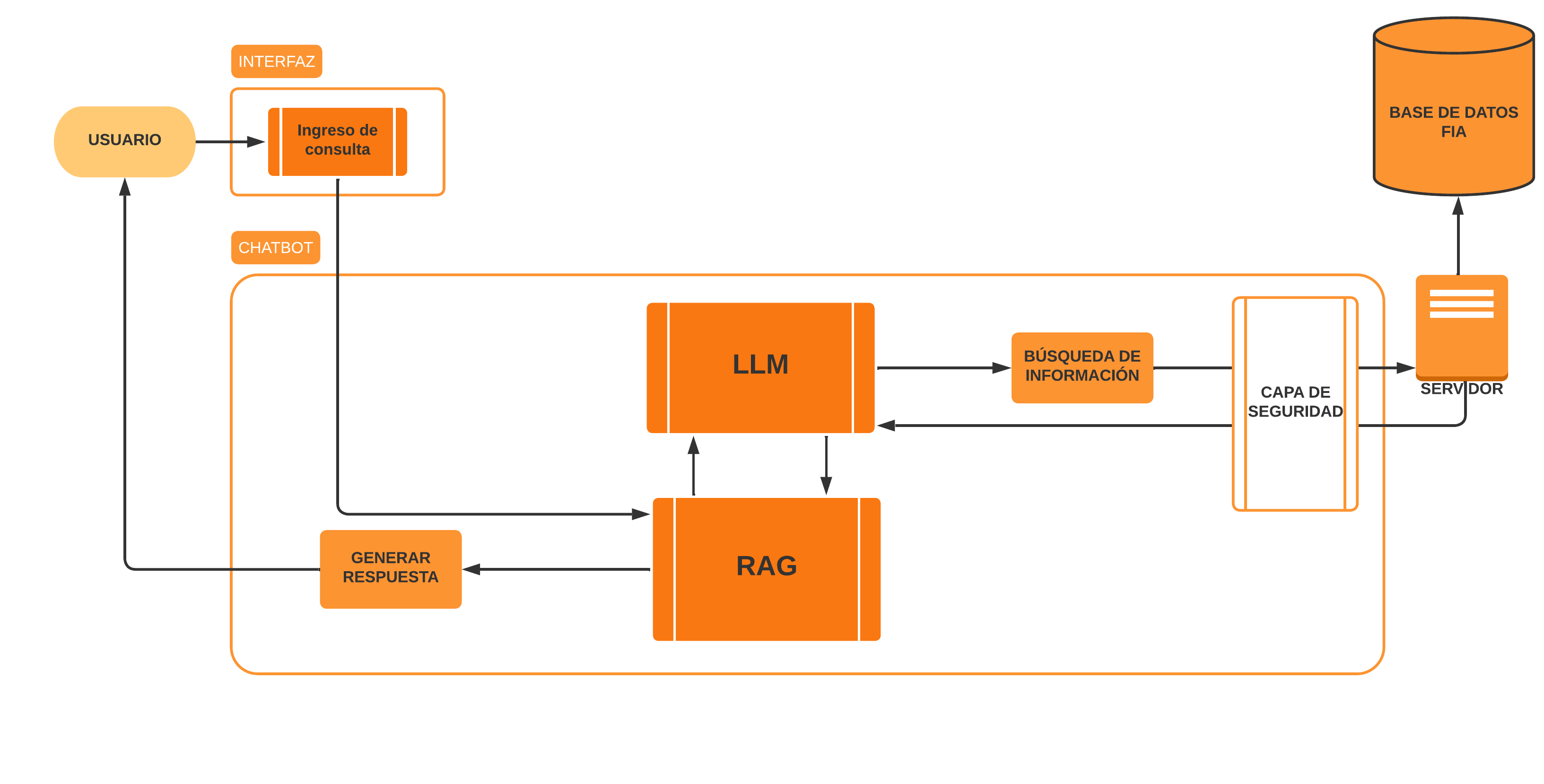
Diagrama

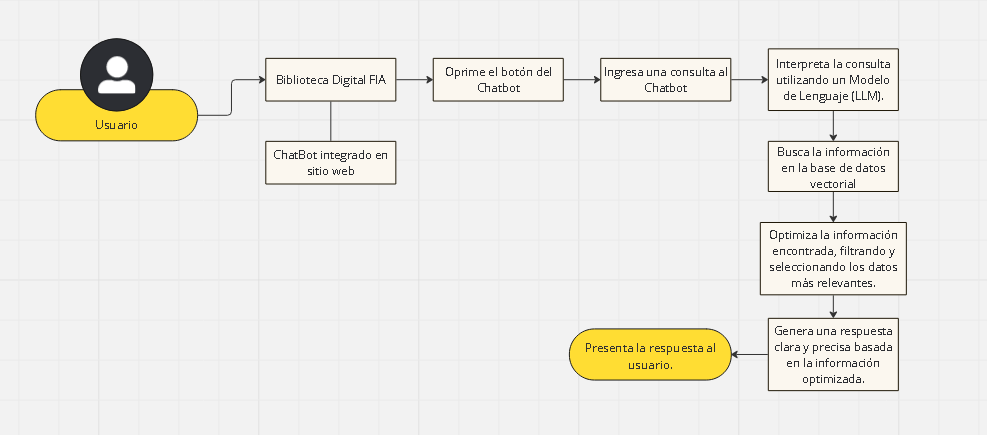
Descripción generada automáticamente

# Diseño de Solución

1. **Funcionalidades del chatbot.**
   1. Búsqueda de Información.
      * El chatbot busca información relacionada a la consulta que se le entrega para generar su respuesta.
   2. Optimización de Información.
      * El chatbot accede a los distintos tipos de archivos que posee la Biblioteca Digital de la FIA, diferenciando su formato y seleccionando aquellos que cumplen con la información para la generación de su respuesta.
   3. Generación de respuestas.
      * El chatbot genera la respuesta de manera clara y precisa, entregando la información requerida por el usuario.
   4. Autoaprendizaje.
      * El chatbot genera un aprendizaje automático luego de cada consulta que se le realiza, permitiendo mejorar sus búsquedas de información, tiempo y respuestas.
   5. Interpretar consultas.
      * El chatbot interpreta las consultas ingresadas en lenguaje natural a través de la identificación de palabras claves que le ayudan a realizar la búsqueda de información de manera específica.
2. **Diseño de Arquitectura Prototipo.**
   1. Interfaz.
      * El chatbot poseerá una interfaz en la cual el usuario interactúa para realizar las consultas, la cual se despliega como un chat común.
   2. LLM.
      * La capa de LLM (Modelo de Lenguaje de gran tamaño) es la encargada de procesar y comprender las consultas de los usuarios y convertirlas en comandos comprensibles para el sistema. Así mismo, genera la respuesta en lenguaje natural para el usuario.
   3. RAG.
      * Esta capa es la encargada de optimizar la salida del LLM, de manera que genere una respuesta basada en datos específicos de la FIA para un resultado preciso y útil.
   4. Base de Datos
      * Almacenamiento de la información de la Biblioteca FIA.
   5. Servidor.
      * Encargado de intermediar en todos los procesos.
   6. Capa de Seguridad.
      * Garantiza la seguridad de la información y la integridad de los sistemas de la FIA.
3. **Interfaz de usuario (UI)** 
   1. Interfaz Intuitiva.
      * Interfaz entendible y fácil de usar para el usuario.
   2. Acceso al Chatbot.
      * La interfaz corresponde a toda el área visual del chatbot en la cual el usuario interactúa con él.
   3. Compatibilidad.
      * La interfaz es responsiva y compatible con cualquier tipo de dispositivo y navegador web.
4. **Seguridad de la Información.**
   1. Estándares.
      * Utilización de guía de Estándares internacionales ISO/IEC 42001 que especifica como establecer, implementar y mantener un sistema de IA en organizaciones, y también ISO/IEC 27001 que establece prácticas para la gestión de seguridad de información.
   2. Implementación.
      * Implementación de los estándares en el desarrollo de la solución.
5. **Fase de Pruebas.**
   1. Detección de problemas y/o Vulnerabilidades.
      * Realización de pruebas al chatbot para detectar posibles fallas.
   2. Corrección de problemas y/o Vulnerabilidades.
      * Corrección de las fallas.
   3. Experiencia de usuario (UX).
      * Prueba en la que usuarios seleccionados hacen uso del chatbot y entregan un feedback con su experiencia.
   4. Aplicación de mejoras según UX.
      * Implementar las mejoras correspondientes al feedback recibido por los usuarios.
6. **Implementación de la Solución.**
   1. Añadir Chatbot a sitio Biblioteca Digital FIA.
      * Integrar de manera oficial el chatbot al sitio de la Biblioteca Digital FIA.

# Diagramas UML Solución

* ****Diagrama de Flujo.
* Caso de Uso General.



* Historias de usuario.

|  |  |
| --- | --- |
| Historia de Usuario 1: Búsqueda de Información Específica | Criterios de Aceptación |
| Como usuario de la Biblioteca Digital FIA, quiero buscar información específica sobre un tema particular, para obtener datos precisos y actualizados para mi investigación. | El chatbot debe interpretar correctamente la consulta en lenguaje natural. El chatbot debe proporcionar una respuesta clara y precisa basada en los datos almacenados en la biblioteca. La respuesta debe incluir citas y referencias a las fuentes de información utilizadas. |

|  |  |
| --- | --- |
| Historia de Usuario 2: Optimización de Resultados de Búsqueda | Criterios de Aceptación |
| Como usuario de la Biblioteca Digital FIA, quiero que el chatbot optimice los resultados de mi búsqueda, para que la información presentada sea la más relevante y útil para mis necesidades. | El chatbot debe filtrar y seleccionar los datos más relevantes de la biblioteca. La respuesta optimizada debe ser clara y fácil de entender. El chatbot debe ser capaz de aprender de las interacciones para mejorar futuras búsquedas y respuestas. |

* Mockup Interfaz.

Interfaz de usuario gráfica, Sitio web

Descripción generada automáticamente



Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

# Prototipo

## Alcance del Prototipo

### Principales Funcionalidades

* Procesamiento y Respuesta de Consultas:
  + Descripción: Utiliza el modelo de lenguaje Llama 3 a través de Ollama para procesar consultas y generar respuestas.
  + Funcionalidad: El modelo responde a las consultas teniendo en cuenta las especificaciones y contextos definidos en un PromptTemplate.
* Carga y Procesamiento de Documentos PDF:
  + Descripción: Se cargan archivos PDF a través de un endpoint, y estos documentos se procesan y almacenan para su posterior recuperación.
  + Funcionalidad: Los documentos PDF se dividen en fragmentos (chunks) y se almacenan en una base de datos vectorial (ChromaDB) con embeddings generados para facilitar la búsqueda y recuperación basada en similitud.
* Búsqueda y Recuperación de Información:
  + Descripción: Permite la búsqueda de fragmentos de documentos relevantes utilizando una base de datos vectorial y la recuperación de información mediante un pipeline de Langchain.
  + Funcionalidad: Cuando un usuario envía una consulta, el sistema busca en los documentos almacenados los fragmentos más relevantes y utiliza una cadena de procesamiento para generar una respuesta basada en los datos encontrados.

# Aprendizaje

Durante el desarrollo de este proyecto profundizamos en el funcionamiento de los chatbots y asistentes virtuales, comprendiendo cómo la IA llega a facilitar la interacción con los usuarios y mejorar el acceso a la información, interiorizándonos en todo el funcionamiento y arquitectura que esto involucra en tiempos que este tipo de herramientas toman cada vez más fuerzas.

La aplicación de conocimientos que hemos adquirido durante nuestra carrera nos ayuda a pulir ciertos contenidos importantes en el desarrollo de un proyecto como es el uso de UML y la planificación en equipo para ejecutar tareas.

Es importante recalcar la importancia de las habilidades blandas a través de la interacción con el cliente, FIA en este caso, lo cual es una experiencia que nos brinda mejorar el entendimiento frente a una necesidad, y llevarlo al plano profesional como es el desarrollo de un proyecto, considerando el levantamiento de requerimientos y la comunicación para llegar a un acuerdo en el planteamiento de una solución.

# Conclusión

Durante esta fase del proyecto nos enfrentamos a desafíos técnicos y organizativos que nos pusieron a prueba. El proceso de diseño de la solución fue una tarea compleja, sin embargo, estos desafíos se convirtieron en oportunidades de aprendizaje, permitiéndonos mejorar nuestros conocimientos.

La inclusión del sistema de chatbot IA en la Biblioteca Digital de FIA tendrá un impacto positivo, debido a la mejora en la búsqueda de información por parte del usuario en la biblioteca y la eficiencia de ésta. Se espera que sea significativo en la innovación y el avance tecnológico dentro de la fundación. La implementación de esta herramienta de inteligencia artificial no sólo optimizará la búsqueda y el acceso a la información, sino que también marcará una mejora en la interacción usuario-sistema. Este proyecto nos demuestra cómo la tecnología puede transformar los servicios de información, haciendo que sean más accesibles y eficientes para todos los usuarios.