

**Innovación en la Difusión de Información Agropecuaria con un Chatbot**

Integrantes: Matias Araneda

Ramses Calderón

Matias Villalobos

Ignacio Valdés

Asignatura:  
Taller de Empresa

Tabla de contenido

[1. Introducción 3](#_Toc170433479)

[2. Descripción del Proyecto 3](#_Toc170433480)

[Problemática 3](#_Toc170433481)

[Solución Propuesta 3](#_Toc170433482)

[Cómo Ataca la Problemática 4](#_Toc170433483)

[3. Enfoque Kanban 5](#_Toc170433484)

[Desglose de Tareas: 5](#_Toc170433485)

[4. Plan Detallado de Infraestructura 8](#_Toc170433486)

[Proceso de Embedding y Construcción del Índice Semántico 8](#_Toc170433487)

[Diagrama de Infraestructura General 9](#_Toc170433488)

[Diagrama de Secuencia 10](#_Toc170433489)

[5. Componentes y Analogías 11](#_Toc170433490)

[6. Requerimientos Funcionales y No Funcionales 14](#_Toc170433491)

[Requerimientos Funcionales (RF) 14](#_Toc170433492)

[Tabla detalle Requerimientos Funcionales (RF) 16](#_Toc170433493)

[Requerimientos No Funcionales (RNF) 19](#_Toc170433494)

[Tabla detalle Requerimientos No Funcionales (RFN) 20](#_Toc170433495)

[6. Conclusión 23](#_Toc170433496)

[7. Producto Mínimo Viable 24](#_Toc170433497)

# Introducción

**La Fundación para la Innovación Agraria (FIA)** es la agencia de innovación del Ministerio de Agricultura de Chile.

**Su misión es impulsar la innovación en el sector silvoagropecuario** y la cadena agroalimentaria nacional.

**FIA ofrece información a través de su portal** “Observatorio para la innovación Silvoagropecuaria y la cadena Agroalimentaria (OPIA) Fundación para la innovación Agropecuaria (FIA)”

**Los boletines generados contienen noticias, publicaciones científicas, patentes, proyectos, políticas públicas, información de mercado y eventos relevantes para cada tópico de estudio.**

# Descripción del Proyecto

## Problemática

La problemática que tiene FIA actualmente es el tiempo que demoran sus usuarios en buscar algún tipo de información específica en su portal de OPIA. Esto se debe a que tienen demasiada información sobre innovación silvoagropecuaria y cadenas agroalimentarias entonces se vuelve tedioso buscar algo en particular.

## Solución Propuesta

La solución que proponemos es la implementación de un sistema de chatbot integrado con una infraestructura robusta que permita a sus usuarios interactuar con este en español para así obtener información detallada y traducida de las publicaciones originales.

### Cómo Ataca la Problemática

1. **Traducción Automática**: Utilizando modelos de traducción avanzados como MarianMT, el sistema traducirá automáticamente los documentos y publicaciones al español, asegurando que la información completa esté disponible para los usuarios.
2. **Interacción Natural**: A través de Olama WebUI y APIs de embedding (SBERT), las consultas de los usuarios serán procesadas de manera eficiente, permitiendo una interacción natural y fluida con el chatbot.
3. **Búsqueda y Recuperación de Información**: Empleando tecnologías como Langchain y agentes especializados, el sistema buscará y recuperará información relevante de una base de datos (ChromaDB) actualizada trimestralmente. Esto asegurará que los usuarios reciban respuestas precisas y actualizadas.
4. **Escalabilidad y Seguridad**: Utilizando Kubernetes para la gestión de contenedores y NGINX para el balanceo de carga, la solución será escalable y capaz de manejar un alto volumen de consultas. Además, MinIO y scripts de backup garantizarán la seguridad y disponibilidad de los datos.

# Enfoque Kanban

**Visión General**: El proyecto se gestiona mediante la metodología Kanban, que proporciona una visualización clara del flujo de trabajo y la flexibilidad necesaria para adaptarse a cambios.Escala de tiempo

Descripción generada automáticamente

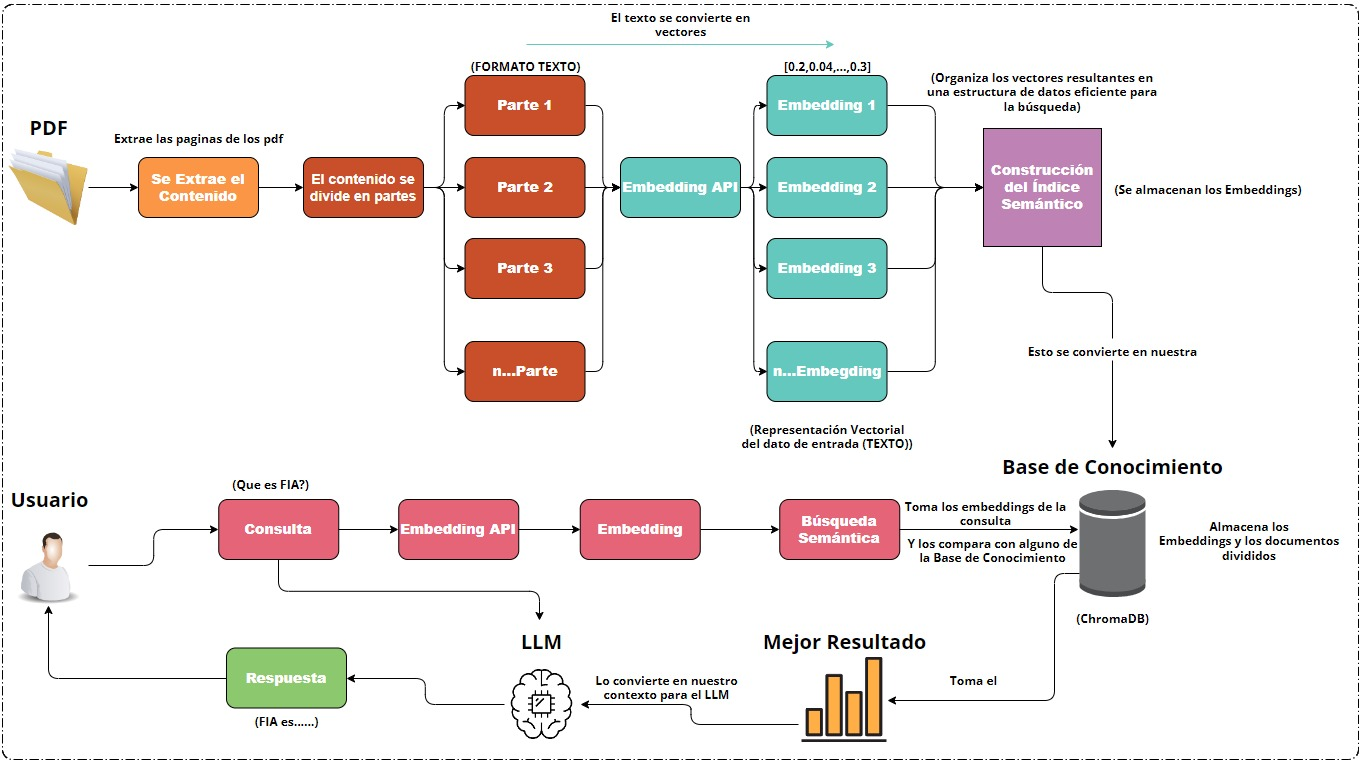
## Desglose de Tareas:

|  |
| --- |
| * Identificar y documentar requerimientos |
| * Recopilar requisitos funcionales y no funcionales del sistema. |
| * Realizar reuniones con FIA para validar requisitos. |
| * Documentar los requisitos en un documento formal. |
| * Diseñar y documentar la solución propuesta |
| * Esquema de solución propuesta |
| * Desarrollar un diagrama de infraestructura general. |
| * Documentar la arquitectura del sistema. |
| * Elaborar un plan de infraestructura detallado. |
| * Crear Diagramas UML relevantes |
| * Diagrama de casos de uso. |
| * Diagrama de clases. |
| * Diagrama de secuencia. |
| * Diagrama de componentes. |
| * Diagrama de despliegue. |
| * Desarrollar y configurar la IA |
| * Configurar el entorno de desarrollo (Docker, Kubernetes). |
| * Desarrollar el chatbot utilizando Olama WebUI y SBERT Embedding API. |
| * Implementar el modelo de traducción automática (MarianMT). |
| * Integrar Langchain y agentes especializados. |
| * Configurar y conectar la base de datos ChromaDB. |
| * Establecer los scripts de backup y MinIO para almacenamiento seguro. |
| * Ejecutar pruebas y corregir errores |
| * Desarrollar casos de prueba para cada componente del sistema. |
| * Realizar pruebas unitarias y de integración. |
| * Ejecutar pruebas de rendimiento y carga. |
| * Documentar y corregir errores encontrados durante las pruebas. |
| * Validar la solución con usuarios finales y ajustar según feedback. |

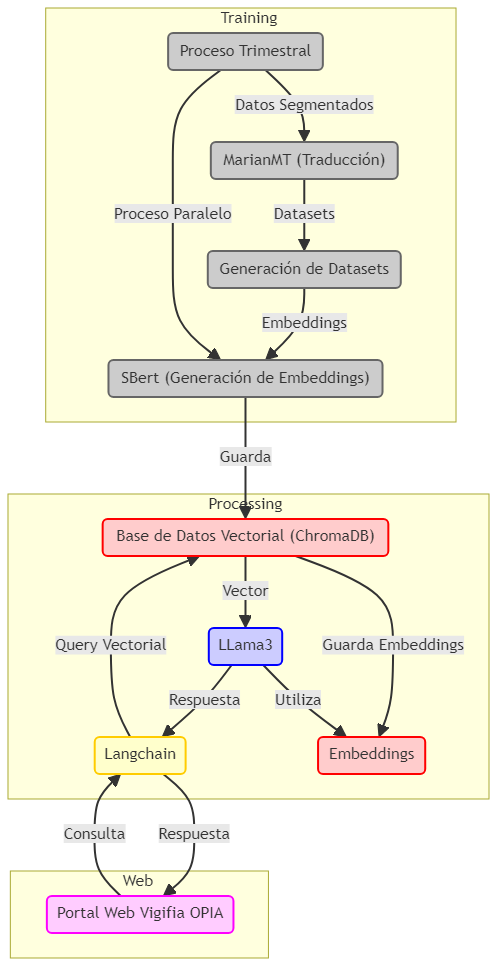


# Plan Detallado de Infraestructura

## Proceso de Embedding y Construcción del Índice Semántico



## Diagrama de Infraestructura General



## Diagrama de SecuenciaDiagrama Descripción generada automáticamente

# Componentes y Analogías

1. **Olama WebUI**
   * **Qué es**: Una interfaz dentro del sistema que recibe las consultas desde el portal web de FIA y las envía para su procesamiento.
   * **Analogía**: Similar a un recepcionista en una biblioteca que toma las preguntas de los visitantes y las prepara para los expertos.
2. **Embedding API (SBERT)**
   * **Qué es**: Una herramienta que convierte las consultas de los usuarios en un formato que el sistema pueda entender y procesar.
   * **Analogía**: Un asistente que traduce las preguntas de los visitantes a un lenguaje especial comprensible para los expertos.
3. **Contenedores Docker**
   * **Qué son**: Tecnología que permite empaquetar aplicaciones y sus dependencias en unidades llamadas contenedores. Estos contenedores aseguran que las aplicaciones funcionen de manera consistente en cualquier entorno.
   * **Analogía**: Cajas mágicas que contienen a los asistentes y expertos, asegurando que siempre funcionen de la misma manera, sin importar dónde estén ubicadas.
4. **Langchain y Agentes Especializados**
   * **Langchain**:
     + **Qué es**: Una herramienta que ayuda a buscar información relevante en la base de datos.
     + **Analogía**: Una lupa mágica utilizada por los bibliotecarios para encontrar rápidamente los documentos más relevantes.
   * **Agentes Especializados**:
     + **Qué son**: Modelos entrenados en temas específicos como "Adaptación y Mitigación al Cambio Climático", "Gestión Sostenible de Recursos Hídricos" y "Sistemas Alimentarios Sostenibles".
     + **Analogía**: Bibliotecarios especializados que saben todo sobre sus temas y pueden encontrar la información más precisa y relevante.
5. **ChromaDB y Proceso de Entrenamiento**
   * **ChromaDB (Base de Conocimiento)**:
     + **Qué es**: Una base de datos que almacena todos los documentos y datos procesados.
     + **Analogía**: El archivo central de la biblioteca donde se guarda toda la información importante.
   * **Proceso de Entrenamiento**:
     + **Qué es**: Un proceso que actualiza la base de datos con nueva información cada tres meses.
     + **Analogía**: Cursos de actualización periódicos para los bibliotecarios para que siempre tengan la información más reciente.
6. **Modelo de Traducción (MarianMT)**
   * **Qué es**: Un modelo que traduce documentos a español antes de ser almacenados en la base de datos.
   * **Analogía**: Un traductor que asegura que todos los documentos en la biblioteca estén en español, facilitando su consulta y procesamiento.
7. **Kubernetes y NGINX**
   * **Kubernetes**:
     + **Qué es**: Una herramienta que organiza y gestiona los contenedores Docker.
     + **Analogía**: El gerente de la biblioteca que asegura el funcionamiento eficiente y la utilización adecuada de los recursos.
   * **NGINX**:
     + **Qué es**: Un servidor que distribuye el tráfico entre los diferentes componentes del sistema.
     + **Analogía**: El portero que dirige a los visitantes (consultas) al bibliotecario adecuado.
8. **MinIO y Backup Scripts**
   * **MinIO**:
     + **Qué es**: Una solución de almacenamiento que gestiona y guarda documentos y datos.
     + **Analogía**: Un archivista que asegura el almacenamiento seguro y accesible de todos los documentos.
   * **Backup Scripts**:
     + **Qué son**: Programas que realizan copias de seguridad de los datos periódicamente.
     + **Analogía**: Archivistas que hacen copias de seguridad de todos los documentos importantes para prevenir la pérdida de información.

# Requerimientos Funcionales y No Funcionales

## Requerimientos Funcionales (RF)

1. **RF001 - Subida de documentos**
   * **Descripción**: Permitir la subida de documentos cada tres meses.
   * **Componentes**:
     + Portal Web FIA: Interfaz de usuario para subir documentos.
     + Docker Container: Asegura que el proceso de subida sea consistente.
     + ChromaDB: Almacena los documentos subidos.
     + MinIO: Solución de almacenamiento para gestionar y guardar documentos.
   * **Analogía**: Los usuarios suben documentos como visitantes entregan libros a la biblioteca, que son guardados en cajas mágicas (contenedores Docker) y archivados en el archivo central (ChromaDB).
2. **RF002 - Almacenamiento de documentos**
   * **Descripción**: Almacenar los documentos subidos en una base de datos segura.
   * **Componentes**:
     + ChromaDB: Base de datos central para el almacenamiento de documentos.
     + MinIO: Solución de almacenamiento que asegura la persistencia y accesibilidad de los documentos.
     + Backup Scripts: Aseguran la realización de copias de seguridad periódicas.
   * **Analogía**: Los documentos subidos son archivados de manera segura y copias de seguridad son realizadas periódicamente para prevenir la pérdida de información.
3. **RF003 - Traducción de documentos**
   * **Descripción**: Permitir la traducción de documentos entre inglés y español.
   * **Componentes**:
     + MarianMT: Modelo de traducción para convertir documentos.
   * **Analogía**: Un traductor asegura que todos los documentos estén en español, facilitando su consulta y procesamiento.
4. **RF004 - Interacción tipo ChatGPT**
   * **Descripción**: Permitir la interacción con el sistema a través de un chatbot similar a ChatGPT.
   * **Componentes**:
     + Olama WebUI: Recibe las consultas y las envía para su procesamiento.
     + SBERT Embedding API: Convierte las consultas a un formato comprensible para el sistema.
     + Langchain y Agentes Especializados: Buscan información relevante y generan respuestas.
   * **Analogía**: Los usuarios interactúan con un recepcionista en la biblioteca (Olama WebUI), que traduce sus preguntas (SBERT) y las envía a los bibliotecarios especializados (agentes entrenados) para obtener respuestas precisas.

### Tabla detalle Requerimientos Funcionales (RF)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Código** | **Requerimiento** | **Descripción** | **Entrada** | **Salida** | **Prioridad** | **Actor** |
| RF001 | Subida de documentos | Permitir la subida de documentos cada tres meses. | Documento (archivo PDF, link, etc.) | Confirmación de subida exitosa | Alta | Usuario |
| RF002 | Almacenamiento de documentos | Almacenar los documentos subidos en una base de datos segura. | Documento (archivo PDF, link, etc.) | Documento almacenado | Alta | Sistema |
| RF003 | Traducción de documentos | Permitir la traducción de documentos entre inglés y español. | Documento (texto en inglés o español) | Documento traducido | Alta | Usuario |
| RF004 | Interacción tipo ChatGPT | Permitir la interacción con el sistema a través de un chatbot tipo ChatGPT. | Pregunta del usuario | Respuesta generada | Alta | Usuario |
| RF005 | Respuesta automática a consultas | Generar respuestas automáticas a consultas sobre los documentos. | Pregunta del usuario | Respuesta generada | Alta | Usuario |
| RF006 | Integración con Ollama | Integrar el sistema con Ollama para usar Llama3. | - | - | Alta | Sistema |
| RF007 | Uso de RAG LangChain | Implementar RAG LangChain para mejorar la búsqueda de información. | - | - | Alta | Sistema |
| RF008 | Eliminación de documentos | Permitir la eliminación de documentos del sistema. | - | - | Media | Usuario |
| RF009 | Configuración de preferencias de idioma | Permitir la configuración de preferencias de idioma (inglés/español). | - | - | Media | Usuario |
| RF010 | Seguridad de documentos | Asegurar que los documentos estén protegidos y solo accesibles a usuarios autorizados. | - | - | Alta | Sistema |
| RF011 | Análisis de consultas | Realizar análisis de consultas para mejorar las respuestas del chat. | - | - | Media | Sistema |
| RF012 | Optimización del rendimiento | Optimizar el rendimiento del sistema para manejar grandes volúmenes de documentos. | - | - | Alta | Sistema |

## Requerimientos No Funcionales (RNF)

1. **Escalabilidad**
   * **Descripción**: Capacidad del sistema para manejar un aumento en la carga de trabajo.
   * **Componentes**:
     + Kubernetes: Organiza y gestiona los contenedores Docker para asegurar una utilización adecuada de los recursos.
     + NGINX: Distribuye el tráfico entre los componentes del sistema.
   * **Analogía**: El gerente de la biblioteca (Kubernetes) asegura que los recursos sean utilizados eficientemente, y el portero (NGINX) dirige a los visitantes al bibliotecario adecuado.
2. **Seguridad**
   * **Descripción**: Protección de los datos almacenados y las interacciones del usuario.
   * **Componentes**:
     + MinIO: Proporciona una solución de almacenamiento segura.
     + Backup Scripts: Realizan copias de seguridad para prevenir la pérdida de datos.
   * **Analogía**: Un archivista asegura el almacenamiento seguro de los documentos y realiza copias de seguridad periódicas.
3. **Rendimiento**
   * **Descripción**: Velocidad y eficiencia del sistema en procesar consultas y manejar datos.
   * **Componentes**:
     + SBERT Embedding API: Optimiza la comprensión de las consultas.
     + Langchain: Facilita la búsqueda rápida de información relevante.
   * **Analogía**: El asistente que traduce preguntas y la lupa mágica que encuentra rápidamente los documentos más relevantes aseguran un rendimiento óptimo.

### Tabla detalle Requerimientos No Funcionales (RFN)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Código** | **Requerimiento** | **Descripción** | **Prioridad** | **Actor** |
| RNF001 | Tiempo de respuesta | El sistema debe responder a las consultas en menos de 2 segundos. | Alta | Sistema |
| RNF002 | Disponibilidad | El sistema debe estar disponible el 99.9% del tiempo. | Alta | Sistema |
| RNF003 | Escalabilidad | El sistema debe ser escalable para manejar hasta 1000 documentos. | Alta | Sistema |
| RNF004 | Seguridad de datos | El sistema debe asegurar que los datos están cifrados. | Alta | Sistema |
| RNF005 | Usabilidad | El sistema debe ser fácil de usar para usuarios no técnicos. | Alta | Usuario |
| RNF006 | Compatibilidad | El sistema debe ser compatible con los navegadores modernos. | Media | Sistema |
| RNF007 | Eficiencia | El sistema debe optimizar el uso de recursos para maximizar el rendimiento. | Alta | Sistema |
| RNF008 | Interfaz intuitiva | La interfaz de usuario debe ser intuitiva y fácil de navegar. | Alta | Usuario |
| RNF009 | Documentación | El sistema debe contar con documentación completa y actualizada. | Media | Usuario |
| RNF010 | Integridad de datos | El sistema debe garantizar la integridad de los datos durante las transacciones. | Alta | Sistema |
| RNF011 | Latencia mínima | El sistema debe tener una latencia mínima en la recuperación de documentos. | Alta | Sistema |
| RNF012 | Respaldo de datos | El sistema debe contar con un mecanismo de respaldo de datos diario. | Alta | Sistema |
| RNF013 | Flexibilidad | El sistema debe ser flexible para adaptarse a futuros requerimientos y tecnologías. | Media | Sistema |
| RNF014 | Rendimiento óptimo | El sistema debe mantener un rendimiento óptimo bajo cargas máximas. | Alta | Sistema |

# 6. Conclusión

En conclusión, este proyecto permitirá a FIA ofrecer un servicio robusto y escalable para responder preguntas sobre "Adaptación y Mitigación al Cambio Climático", "Gestión Sostenible de Recursos Hídricos" y "Sistemas Alimentarios Sostenibles", asegurando que la información proporcionada esté siempre actualizada y sea de alta calidad.

# 7. Producto Mínimo Viable

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente