



Universidad Politécnica
de Madrid

**Escuela Técnica Superior de
Ingenieros Informáticos**



Grado en Ingeniería Informática

Trabajo Fin de Grado

Plan de Trabajo

Autor: Francisco de Asís Chicote Martín

Tutor: Daniel Garijo Verdejo y Pablo Calleja Ibañez

Madrid, Octubre - 2025

Este Trabajo Fin de Grado se ha depositado en la ETSI Informáticos de la Universidad Politécnica de Madrid para su defensa.

Trabajo Fin de Grado
Grado en Ingeniería Informática

Título: Plan de Trabajo

Octubre - 2025

Autor: Francisco de Asís Chicote Martín

Tutor: Daniel Garijo Verdejo y Pablo Calleja Ibañez
Departamento de Inteligencia Artificial
Escuela Técnica Superior de Ingenieros Informáticos
Universidad Politécnica de Madrid

Descripción y Objetivos

El objetivo principal del proyecto es el desarrollo de un sistema inteligente de búsqueda de correos electrónicos combinando las técnicas de recuperación RAG y los modelos de lenguaje(LLMs), todo ello ejecutado de forma local y con vistas a la futura creación de una extensión para el cliente de correo Thunderbird.

El sistema deberá permitir al usuario que mediante consultas escritas en un chat acceda de forma natural a la información contenida en su historial de correo, sin hacer distinciones en el idioma o el tiempo que haya pasado desde el mismo. Este sistema deberá responder a preguntas del estilo: "*¿Cuánto me costó el viaje a Marruecos de hace unos años?*", y dar una respuesta coherente a la pregunta, además de localizar y facilitar los correos que guardan relación con la consulta.

Los principales objetivos de este Trabajo de Fin de Grado son:

- Diseñar un pipeline para procesar los correos
- Desarrollar un módulo para indexar y dividir en chunks los correos en vectores semánticos
- Implementar un sistema RAG multilingüe
- Integrar un modelo de LLM ligero
- Crear un Benchmark de evaluación usando datasets reales
- Desarrollar una interfaz CLI inicialmente y posteriormente una extensión para Thunderbird
- Asegurar la indexación de nuevos correos

Tareas

Para poder desarrollar este Trabajo de Fin de Grado de forma estructurada se ha dividido el proyecto en diversas tareas que cubren desde el análisis inicial del problema hasta la elaboración de la documentación final. Cada tarea representa una fase del futuro desarrollo, buscando así un flujo coherente.

El objetivo es poder realizar un seguimiento del proceso y asegurar que todos los requisitos del proyecto se cumplen siguiendo los plazos establecidos.

A continuación podemos ver las distintas tareas que conforman el desarrollo del proyecto junto con su representación temporal en el diagrama de Gantt.

T1 - Análisis del problema

T2 - Preprocesamiento

T3 - Chunking e indexación

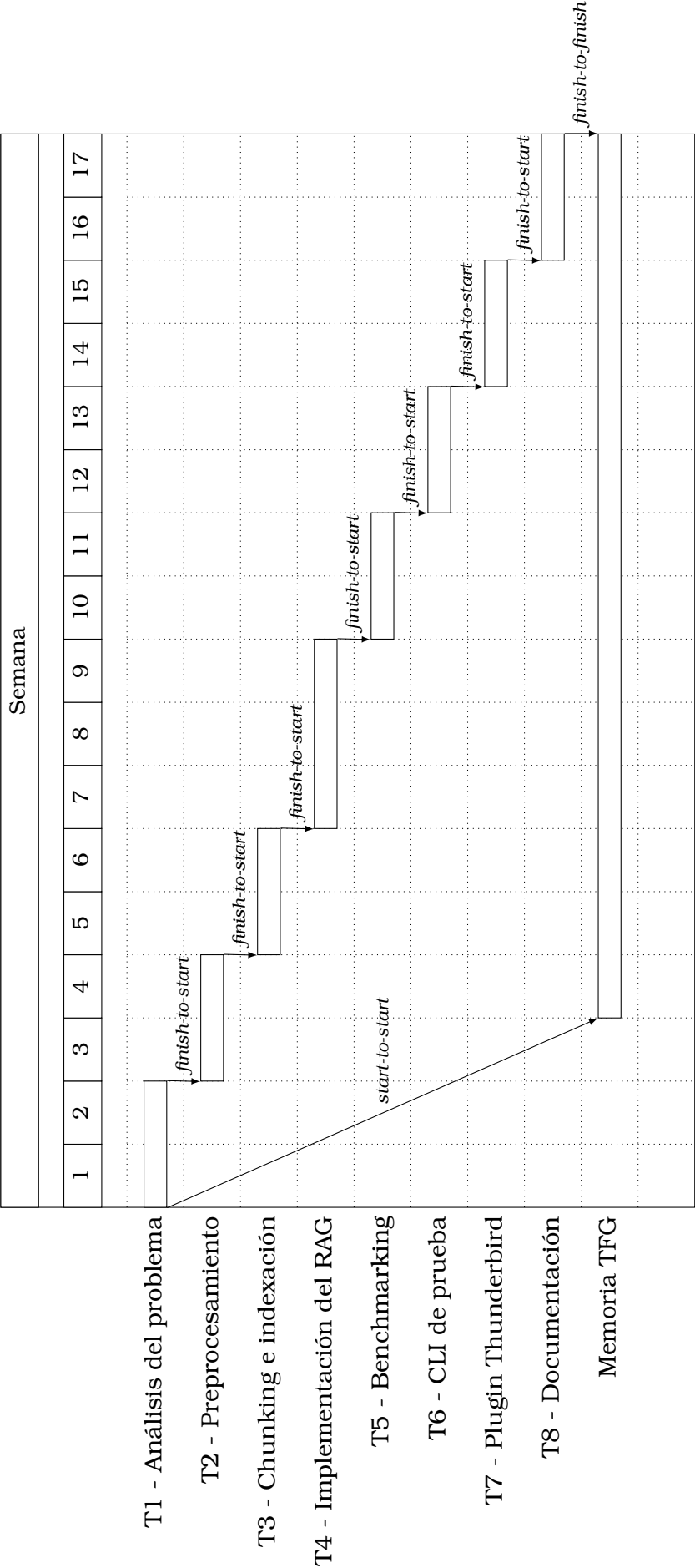
T4 - Implementación del RAG

T5 - Benchmarking

T6 - CLI de prueba

T7 - Plugin para Thunderbird

T8 - Documentación



Propuesta de Trabajo Oficial

Resumen general del trabajo:

La búsqueda sobre correos electrónicos se basa en filtros y palabras clave que dificultan encontrar el correo o la conversación deseada. En este trabajo el alumno explorará la utilización de Large Language Models (LLMs) de pequeño tamaño y tecnologías de Retrieval Augmented Generation (RAG) para facilitar la localización de correos electrónicos guardados en un equipo a nivel local. Además, se explorará si la creación de grafos de conocimiento mejora la respuesta del sistema.

Lista de objetivos concretos

- Diseño de la interacción con RAG y LLM por parte del usuario
- Diseño de una evaluación para probar el sistema
- Diseño de la aplicación y arquitectura del sistema
- Implementación y evaluación de la aplicación diseñada
- Encapsulación en contenedores del sistema