Informe Proyecto APT – AI Email Parser

Contenido

[**Proyecto APT** 3](#_Toc207571134)

[Nombre del proyecto 3](#_Toc207571135)

[Abstract (English) 3](#_Toc207571136)

[Abstract (Spanish) 3](#_Toc207571137)

[Contexto del proyecto 3](#_Toc207571138)

[Descripción del proyecto 4](#_Toc207571139)

[Perfil de egreso 4](#_Toc207571140)

[Competencias 4](#_Toc207571141)

[Relación con los intereses profesionales 4](#_Toc207571142)

[Objetivos del proyecto 5](#_Toc207571143)

[Factibilidad del proyecto 5](#_Toc207571144)

[Objetivo General 5](#_Toc207571145)

[Objetivos específicos 5](#_Toc207571146)

[Metodología de trabajo 6](#_Toc207571147)

[Plan de trabajo 6](#_Toc207571148)

[Evidencias de logro 8](#_Toc207571149)

[Conclusiones 8](#_Toc207571150)

[Individual conclusion (Sebastián Lara) 8](#_Toc207571151)

[Reflection 8](#_Toc207571152)

[Individual conclusión () 9](#_Toc207571153)

[Reflection 9](#_Toc207571154)

# **Proyecto APT**

## Nombre del proyecto

**AI Email Parser.**

## Abstract (English)

This project aims to automate the reading, análisis and organization of institutional emails using artificial intelligence. Integrated with Gmail API, Google Cloud Platform Pub/Sub, Ollama3 AI, PostreSQL and a React Frontend. The system optimizes email management, ensuring relevant information is not overlooked and improving administrative efficiency.

## Abstract (Spanish)

El proyecto AI Email Parser busca automatizar la lectura, análisis y organización de correos electrónicos institucionales mediante inteligencia artificial. Integrado con Gmail API, Pub/Sub de Google Cloud Platfrom, IA Ollama3, PostgreSSQL y un frontend en React. El sistema optimiza la gestión de correos, asegurando que la información relevante no se pierda y mejorando la eficiencia administrativa.

## Contexto del proyecto

En el contexto académico, los docentes que imparten múltiples secciones de uno o más ramos, suelen recibir gran cantidad de correos por parte de estudiantes. El tema de los correos puede contener consultas sobre evaluaciones, solicitud de revisión de notas, ayuda para feedback, posible cambio de grupos, entre muchos otros. La alta carga de correos entrantes dificulta una gestión eficiente de tiempo generando estrés administrativo y la omisión de solicitudes importantes. Este problema crece en temporada de evaluaciones o cierres de semestre. El proyecto busca abordar esta problemática a través del análisis de correos utilizando inteligencia artificial. El sistema extrae información clave de los correos y muestra dicha información en un panel dinámico, accesible y ordenado.

## Descripción del proyecto

El proyecto busca automatizar la lectura, análisis y organización de correos electrónicos institucionales mediante un sistema basado en inteligencia artificial.

El sistema funcionará de la siguiente manera:

* Integración con **Gmail API** y **OAuth 2.0** para la lectura segura de correos.
* Procesamiento en **FastAPI** con notificaciones en tiempo real a través de Pub/Sub, herramienta nativa de **Google Cloud Platform**.
* Uso de IA **Ollama3** para extraer datos relevantes como fechas, recordatorios y eventos.
* Almacenamiento estructurado en **PostgreSQL**.
* Interfaz en **React** para la visualización de correos procesados.

**Funcionalidad principal:** Recibir un correo, extraer remitente, asunto y cuerpo, enviar la información al modelo IA Ollama para extraer lo mas relevante, almacenar los datos y mostrar los resultados en un panel visual.

# Perfil de egreso

## Competencias

Este proyecto aplica varias competencias del perfil de egreso de un Ingeniero en Informática:

* Integración de sistemas tecnológicos en soluciones informáticas.
* Diseño y desarrollo de soluciones informáticas.
* Aplicación de metodologías.
* Uso de herramientas de inteligencia artificial.
* Gestión de bases de datos.
* Desarrollo backend y frontend

Estas competencias de egreso permiten abordar problemas reales, como la sobrecarga de información y la automatización de tareas, aplicando tecnologías de vanguardia.

## Relación con los intereses profesionales

Mis intereses profesionales principalmente se centran en:

* Automatización de procesos.
* Desarrollo de Software.
* Análisis de datos.

Este proyecto sea alinea con estos intereses al permitir la creación de un sistema para tareas repetitivas facilitando la toma de decisiones mediante un análisis automatizado de correos electrónicos.

# Objetivos del proyecto

## Factibilidad del proyecto

El proyecto es factible de desarrollar dentro de la duración del semestre ya que las horas asignadas a la asignatura permiten distribuir el trabajo de manera equilibrada entre la investigación, el desarrollo y las pruebas, complementándose con horas de trabajo autónomo.

En cuanto a materiales requeridos, el proyecto utilizara software libre y recursos de la nube, lo que reduce los costos y facilita el acceso. Además, se cuenta con computadores personales y conexión a internet, que son suficientes para el desarrollo e implementación del sistema.

Existen diversos factores externos que facilitan el desarrollo de este proyecto, como herramientas de integración gratuita, amplia disponibilidad de documentación en línea y foros de apoyo.

No obstante, los factores externos que podrían dificultar el avance del proyecto se encuentra la curva de asociada a nuevas tecnologías, la disponibilidad de tiempo debido a otras asignaturas y la práctica profesional. Para mitigar estos riesgos, se propone un plan iterativo con entregas parciales, el uso de entornos virtualizados y herramientas de gestión de tareas como Jira o Trello.

## Objetivo General

El objetivo general es desarrollar un sistema automatizado que procese correos institucionales con inteligencia artificial, organizando la información clave de forma estructurada y accesible.

## Objetivos específicos

Los objetivos específicos del proyecto son los siguientes:

* Implementar integración con Gmail API usando el protocolo estándar de autorización llamado OAuth 2.0 para acceder a correos institucionales.
* Configurar un servicio de notificaciones en tiempo real con Pub/Sub de GCP para el manejo de nuevos correos.
* Aplicar IA en Ollama para identificar y extraer información relevante de los correos.
* Diseñar la base de datos en PostgreSQL para almacenar la información procesada.
* Implementar un frontend con React para la visualización de correos procesados.

Desarrollar el backend en FastAPI para la extracción y procesamiento de datos.

# Metodología de trabajo

La metodología será iterativa, combinando desarrollo incremental y pruebas continuas. Las etapas del proyecto serán:

1. Análisis y diseño: Definición de requerimientos, arquitectura del sistema y modelo de datos.
2. Desarrollo backend: Creación del servicio FastAPI, integración con Gmail API y Pub/Sub
3. Procesamiento con IA: Implementación de IA Ollama para la extracción de datos.
4. Almacenamiento: Configuración de PostgreSQL y migraciones con alembic.
5. Desarrollo Frontend: Creación de una interfaz para visualizar los correos y las respuestas generadas por la IA en un listado o tabla básica.
6. Pruebas e integración: Validación del sistema completo, ajustes y corrección de errores.

El enfoque iterativo permite entregar avances parciales y realizar ajustes.

# Plan de trabajo

El siguiente plan de trabajo organiza las actividades del proyecto de manera secuencial y ordenada, permitiendo que cada etapa se ejecute de forma efectiva y se pueda realizar un seguimiento de los avances. Se incluyen las tareas principales y los resultados esperados.

|  |  |  |  |
| --- | --- | --- | --- |
| Semana | Actividad | Descripción | Resultado esperado |
| 1 | Conexión con Gmail API y Pub/Sub | Configuración de la autenticación con OAuth 2.0 para acceder a correos institucionales y activación de notificaciones en tiempo real mediante Pub/Sub en GCP | Backend conectado correctamente con Gmail API, correos recibidos y notificaciones funcionando. |
| 2 a 3 | Integración con IA Ollama | Implementación de IA Ollama para procesar los correos y extraer información clave como remitente, asunto, fecha y mensaje del correo. Creación de prompt predefinido. | IA extrayendo datos de prueba correctamente y enviando resultados al backend. |
| 4 a 5 | Crear e implementar modelo de datos | Diseño y creación de la base de datos en PostgreSQL para almacenar los correos y la información extraída. Implementación de migraciones. | Base de datos estructurada y funcional. |
| 6 a 7 | Desarrollo del frontend con React | Creación de la interfaz para visualizar los correos procesados. | Interfaz funcional y navegable, mostrando datos procesados correctamente. |
| 8 a 10 | Integración y pruebas | Integración de backend, IA y frontend. Pruebas de flujo completo, ajustes de errores y validación del funcionamiento general del sistema. | Sistema integrado funcionando de forma estable. |
| 11 | Documentación y entrega | Elaboración de la documentación técnica. Preparación del informe final del proyecto. | Documentación completa con resultados obtenidos, instrucciones de uso y arquitectura final implementada. |

# Evidencias de logro

Las evidencias de logro corresponden a los resultados concretos que permitirán demostrar el cumplimiento de cada etapa del proyecto APT. Estas evidencias no solo validan el avance técnico, sino que también aseguran que el desarrollo del sistema se alinea con los objetivos planteados:

* Sistema funcional completo: Demostración del flujo desde la recepción del correo, hasta la visualización en frontend.
* Logs de backend y Pub/Sub: Evidencia que los correos se reciben y procesan correctamente.
* Resultados de IA Ollama: Registros de la información extraída por la IA.
* Base de datos operativa: Esquema, tablas y registros insertados en BD.
* Frontend funcional: Interfaz interactiva con visualización de correos y resultados.
* Documentación técnica completa: Manual de usuario, descripción de arquitectura, modelo de datos e informe final de proyecto.

# Conclusiones

## Individual conclusion (Sebastián Lara)

Through the definition and planning of the AI Email Parser Project, i have gained a clearer understanding of the technologies and methodologies required for its development, including backend and frontend design, database management, and the use of artificial intelligent. This planning has allowed me to has a better organize the steps needed to achieve the projects objectives.

## Reflection

This projects aligns with my profesional interests in data análisis and process automation. During this pase, i have identified some strengths as well as weaknesses that require further development and study.

## Individual conclusión ()

## Reflection