

---

Implementación de soluciones en la nube  
para el análisis de datos públicos a través  
de modelos de inteligencia artificial

-

Implementation of cloud solutions for  
public data analysis through artificial  
intelligence models

---



Trabajo de Fin de Master  
Curso 2024–2025

Autor

Cristian Molina Muñoz

Director

Jose Luis Vazquez-Poletti

Rubén Fuentes-Fernández

Máster en Ingeniería Informática

Facultad de Informática

Universidad Complutense de Madrid



Implementación de soluciones en la  
nube para el análisis de datos  
públicos a través de modelos de  
inteligencia artificial

-

Implementation of cloud solutions  
for public data analysis through  
artificial intelligence models

**Trabajo de Fin de máster en Ingeniería Informática**

**Autor**

**Cristian Molina Muñoz**

**Director**

**Jose Luis Vazquez-Poletti**

**Rubén Fuentes-Fernández**

**Convocatoria:** *Septiembre 2025*

**Calificación:**

**Máster en Ingeniería Informática**

**Facultad de Informática**

**Universidad Complutense de Madrid**

**21 de agosto de 2025**



# Autorización de difusión

los abajo firmantes, matriculados en el Master en Ingeniería en Informática de la Facultad de Informática, autorizan a la Universidad Complutense de Madrid (UCM) a difundir y utilizar con fines académicos, no comerciales y mencionando expresamente a sus autores el presente Trabajo Fin de Master: “Implementación de soluciones en la nube para el análisis de datos públicos a través de modelos de inteligencia artificial”, realizado durante el curso académico 2024-2025 bajo la dirección de Jose Luis Vazquez-Poletti y Rubén Fuentes-Fernández, y a la Biblioteca de la UCM a depositarlo en el Archivo Institucional E-Prints Complutense con el objeto de incrementar la difusión, uso e impacto del trabajo en Internet y garantizar su preservación y acceso a largo plazo.

Cristian Molina Muñoz

21 de agosto de 2025



# Dedicatoria

A mis padres, por hacer posible todo esto. Por su esfuerzo





# Agradecimientos

Muchas gracias a todos los profesores y compañeros que nos han acompañado en este viaje y de los que tanto hemos aprendido



# Resumen

[TODO 250 palabras sobre datos, cloud y IA] [ Se redacta en pasado y no debe incluir abreviaturas, referencias a figuras o tablas ni citas bibliográficas. Tampoco se debe incluir información que no aparezca en el proyecto.]

Los ficheros de GitHub se encuentran en el siguiente repositorio:

<https://github.com/crismo04/TFM-cloud-solutions-to-public-data/>

## Palabras clave

Tratamiento de datos, Cloud, Big data, inteligencia Artificial, [TODO mas sobre clouds, se mencionan en orden alfabético]



# Abstract

[ TODO three paragraphs on data, cloud and AI].

The GitHub files can be found in the following repository:

<https://github.com/crismo04/TFM-cloud-soliutions-to-public-data/>

## **Keywords**

Data processing, Cloud, Big data, Artificial intelligence, TODO something else about clouds.



# Índice

<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	1
1.2. Plan de trabajo . . . . .	1
<b>1. Introduction</b>	<b>3</b>
1.1. Motivation . . . . .	3
1.2. Work plan . . . . .	3
<b>2. Estado de la Cuestión</b>	<b>5</b>
2.1. Datos . . . . .	5
2.2. Nubes . . . . .	5
2.3. Inteligencia Artificial . . . . .	5
2.4. Trabajos anteriores . . . . .	6
2.4.1. Trabajos dirigidos a ... . . . .	6
<b>3. Materiales y métodos</b>	<b>7</b>
3.1. Materiales . . . . .	7
3.1.1. Lenguajes . . . . .	7
3.1.2. Herramientas . . . . .	8
3.1.3. Herramientas descartadas . . . . .	9
3.2. Metodos . . . . .	9
3.2.1. Utilización de la solución . . . . .	10
<b>4. Resultados</b>	<b>11</b>
<b>5. Manual de usuario y casos de uso</b>	<b>13</b>
<b>6. Conclusiones y Trabajo Futuro</b>	<b>15</b>
<b>6. Conclusions and Future Work</b>	<b>17</b>

<b>A. Manual de usuario y casos de uso</b>	<b>19</b>
<b>Bibliografía</b>	<b>21</b>



# Índice de figuras



# Capítulo 1

## Introducción

*“We can only see a short distance ahead, but we can see plenty there that needs to be done.”*  
— Alan Turing

### 1.1. Motivación

Empezaremos por el principio, definiendo que son los tres principales elementos de este proyecto [TODO]

El alcance de este proyecto es, por un lado [TODO]

### 1.2. Plan de trabajo

Una vez definido el alcance, me gustaría destacar las cinco fases en las que se ha dividido el proyecto, que se han ido iterando para la creación de varios prototipos funcionales:

1. **Fase de investigación:** Búsqueda de información a cerca de diferentes fuentes publicas de datos, tecnologías en la nube y modelos o herramientas de IA que nos ayuden a tratar, filtrar y entender todos los datos públicos recopilados.
2. **Fase de análisis de requisitos:** [TODO]
3. **Fase de implementación:** [TODO]
4. **Fase de pruebas:** [TODO]

5. **Memoria:** Elaboración de este documento, plasmando las fases anteriores en texto y especificando el desarrollo del proyecto y los resultados del mismo.

# Chapter 1

## Introduction

*“We can only see a short distance ahead, but we can see plenty there that needs to be done.”*  
— Alan Turing

### 1.1. Motivation

We will start at the beginning by defining what the three main elements of this project are, [TODO]

The scope of this project is, on the one hand [TODO]

### 1.2. Work plan

Having defined the scope, I would like to highlight the five phases in which the project has been partitioned, which have been iterated for the creation of several functional prototypes:

1. **Research phase.** Search for information about different public data sources, cloud technologies and AI models or tools that help us to treat, filter and understand all the public data collected.
2. **Requirements analysis phase:** [TODO].
3. **Implementation phase:** [TODO]
4. **Testing phase:** [TODO]

5. **Memory phase:** [TODO] The elaboration of this document, translating the previous phases into text and specifying the development of the project and its results.

# Capítulo 2

## Estado de la Cuestión

En este apartado expondremos el estado actual de los puntos principales de nuestro proyecto según las investigaciones realizadas, así como los trabajos o artículos relacionados con los temas a tratar. Estos son, entre otros, trabajos relacionados con los principales proveedores Cloud y su comparación, trabajos que traten con grandes volúmenes de datos públicos, o trabajos que utilicen diferentes IAs para el tratamiento de datos y la obtención de conclusiones a partir de estos

### 2.1. Datos

[TODO]

### 2.2. Nubes

[TODO]

### 2.3. Inteligencia Artificial

[TODO]

## **2.4. Trabajos anteriores**

[TODO]

### **2.4.1. Trabajos dirigidos a ...**

[TODO]



# Capítulo 3

## Materiales y métodos

[TODO, importante a tener en cuenta: - detallarse cada paso que se ha dado para llegar a los resultados describiendo, en orden lógico y expresado con claridad, los materiales y recursos empleados. - No avanzar resultados y redactarse en pasado

|

En este capítulo vamos a describir el proceso que se ha seguido en la realización del trabajo, las distintas tecnologías, lenguajes de programación y herramientas, así como las que se ha valorado pero descartado. También se definirán los métodos de desarrollo, aplicaciones e incluso modelo de trabajo.

### 3.1. Materiales

[TODO, herramientas, programas y material utilizado, incluyendo por ejemplo los tipos de IA]

#### 3.1.1. Lenguajes

##### **PYTHON**

Python es un lenguaje de programación interpretado y centrado en la legibilidad de su código. Se trata de un lenguaje de programación multiparadigma, ya que soporta parcialmente la orientación a objetos, programación imperativa y, en menor medida, programación funcional.

## SQL

SQL es un lenguaje de dominio específico utilizado en programación, diseñado para administrar, y recuperar información de sistemas de gestión de bases de datos relacionales. Es un sistema que facilita el tratamiento de datos, así como la separación de estos datos del programa principal, permitiendo tener más modularidad. Utilizamos SQL para almacenar información, así como para extraer esta misma información, tratarla y almacenarla ya tratada en la base de datos.

## L<sup>A</sup>T<sub>E</sub>X

L<sup>A</sup>T<sub>E</sub>X es un sistema de composición tipográfica de alta calidad que incluye funcionalidades diseñadas para la producción de documentación técnica y científica. Es el estándar de facto para la comunicación y publicación de documentos científicos, el cual nos ha permitido desarrollar una memoria profesional y facilitar el diseño sin tener que preocuparnos por la forma cada vez que añadíamos cambios. Hemos usado L<sup>A</sup>T<sub>E</sub>X para desarrollar este documento en la aplicación de TeXstudio y el compilador MikTeX.

### 3.1.1.1. Lenguajes descartados

[TODO]

### 3.1.2. Herramientas

#### Visual Studio Code

Visual Studio Code es un editor de código fuente desarrollado por Microsoft para Windows, Linux y MacOS. Incluye soporte para la depuración, control integrado de Git, resaltado de sintaxis, finalización inteligente de código, fragmentos y refactorización de código.

Utilizamos Visual Studio Code como entorno de desarrollo software, ya que comparándolo con otras alternativas que nos brindan en la carrera, lo creemos bastante más útil, sobre todo para la programación en Python o React.

Lo que nos ha hecho decantarnos por él por encima del resto, es la gran comunidad que tiene detrás, la cual cuenta con un gran número de tutoriales y extensiones que nos facilitan mucho la programación y la integración con otras aplicaciones como Github. También destacar su intérprete, para probar pequeños fragmentos de código, lo cual nos ha ahorrado tiempo en depuración de errores.

## **Github**

GitHub es una forja para alojar proyectos utilizando el sistema de control de versiones Git. Se utiliza principalmente para la creación de código fuente de programas de ordenador.

Utilizamos GitHub como sistema de control de versiones y repositorio de código por su tremenda utilidad para comunicarnos y trabajar en paralelo, lo cual ha sido una necesidad en los tiempos de pandemia en los que este proyecto se ha realizado. Esto nos ha asegurado no perder nada de progreso y llevar un control del avance del proyecto en todo momento. Además nuestros tutores nos facilitaron un repositorio privado en el que nos asegurábamos la seguridad del código, por lo que su uso era casi una obligación frente a otras alternativas.

## **TeXstudio y MiKTeX**

TeXstudio es un editor de  $\text{\LaTeX}$  de código abierto y Multiplataforma con una interfaz similar a Texmaker. TeXstudio es un IDE de  $\text{\LaTeX}$  que proporciona un soporte moderno de escritura, como la corrección ortográfica interactiva, plegado de código y resaltado de sintaxis, por lo que lo hemos considerado ideal para la elaboración de este documento. Mientras que MiKTeX es el gestor de paquetes integrado, que instala los paquetes que hacen falta para el correcto funcionamiento de TeXstudio y para la creación de este documento.

### **3.1.3. Herramientas descartadas**

[TODO]

## **3.2. Metodos**

Para llegar a nuestro objetivo de diseño, hemos dividido la implementación en diferentes módulos:

- **búsqueda y almacenamiento de datos.**
- **Tratamiento básico de los datos.**
- **Estudio con modelos de IA en diferentes nubes**
- **Comparación y estudio de resultados.**

### 3.2.1. Utilización de la solución

# Capítulo 4

## Resultados

[TODO, importante a tener en cuenta: Aquí se recogen los nuevos conocimientos que el proyecto aporta al conocimiento científico, redactarse en pasado. utilizando recursos gráficos. ]



# Capítulo 5

## Manual de usuario y casos de uso





# Capítulo 6

## Conclusiones y Trabajo Futuro

[TODO, importante a tener en cuenta: Señalar los principios y relaciones que indican los resultados (qué es lo que se ha sacado en claro con la investigación, futuras implicaciones que se pueden extraer, etc.). · Relacionar los resultados con otros trabajos publicados. · Hay que mencionar también las excepciones, faltas de correlación o aspectos no resueltos. · Indicar futuras líneas de trabajo ]

### Trabajo futuro



# Chapter 6

## Conclusions and Future Work

[TODO]



# Apéndice **A**

## Manual de usuario y casos de uso



## Bibliografía

