

# **Análisis Automatizado de Documentos Legales y Financieros**

**Merger and Acquisition**

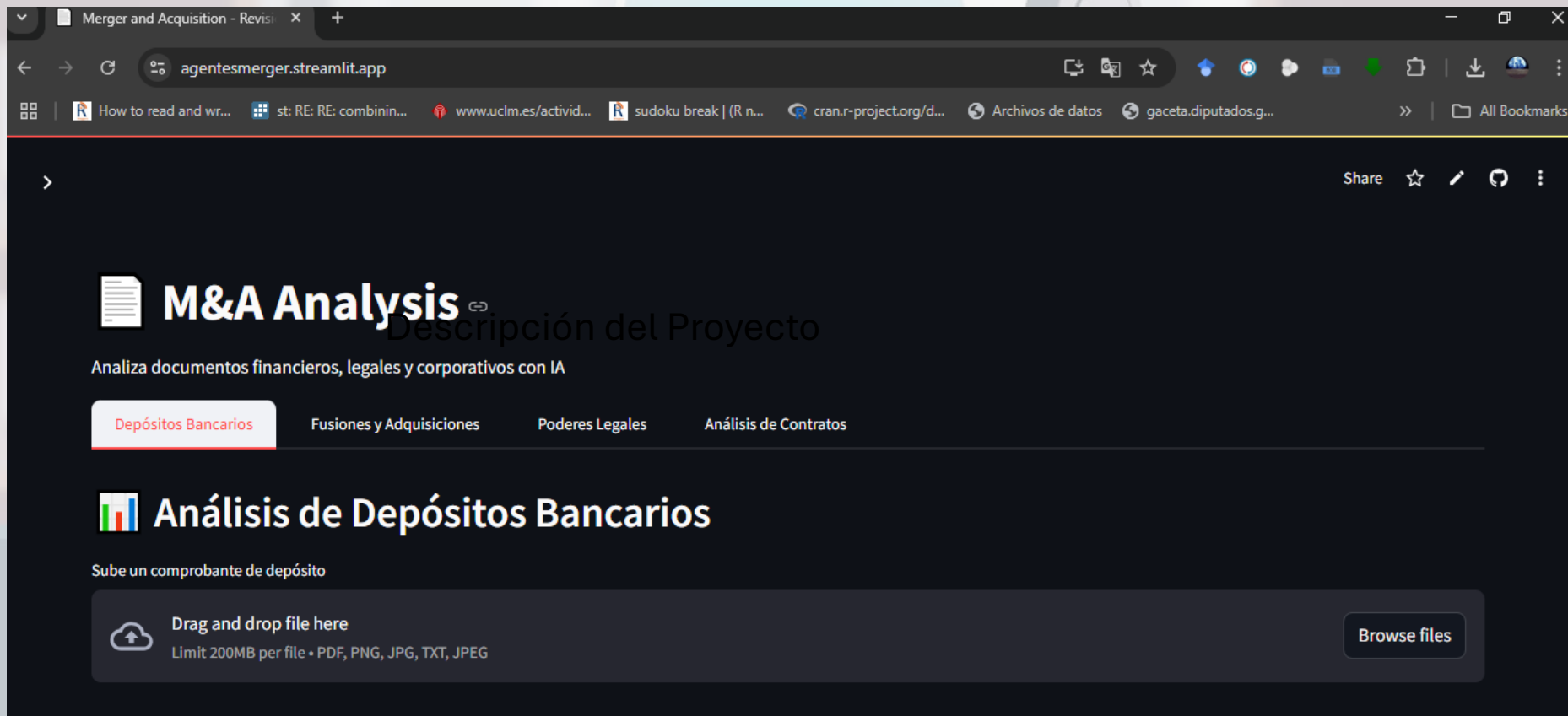
**ITam**

# Descripción del Proyecto

Este proyecto permite extraer, analizar y estructurar información clave de documentos PDF, imágenes o archivos de texto relacionados con:

- **Comprobantes de depósitos bancarios**
- **Operaciones de fusión y adquisición**
- **Poderees legales**
- **Contratos legales**

# Interfaz del Proyecto



A background image showing a white robot with a human-like face and large headphones sitting at a desk. The robot is using a laptop. On the desk, there are several books, a small potted plant, and a container of pens. The scene is brightly lit, suggesting a window in the background.

# Extracción de Texto

- Soporte para PDF, imágenes (OCR con Tesseract) y archivos .txt
- OCR automático cuando el PDF no contiene texto legible



# Análisis de Depósitos

- **Extrae JSON** con los campos: beneficiario, ordenante, monto, fecha\_hora
- Utiliza el modelo deepseek-chat para interpretar comprobantes
- Arroja la hora de pago, el que deposita y el que recibe el dinero



# Análisis de Fusiones

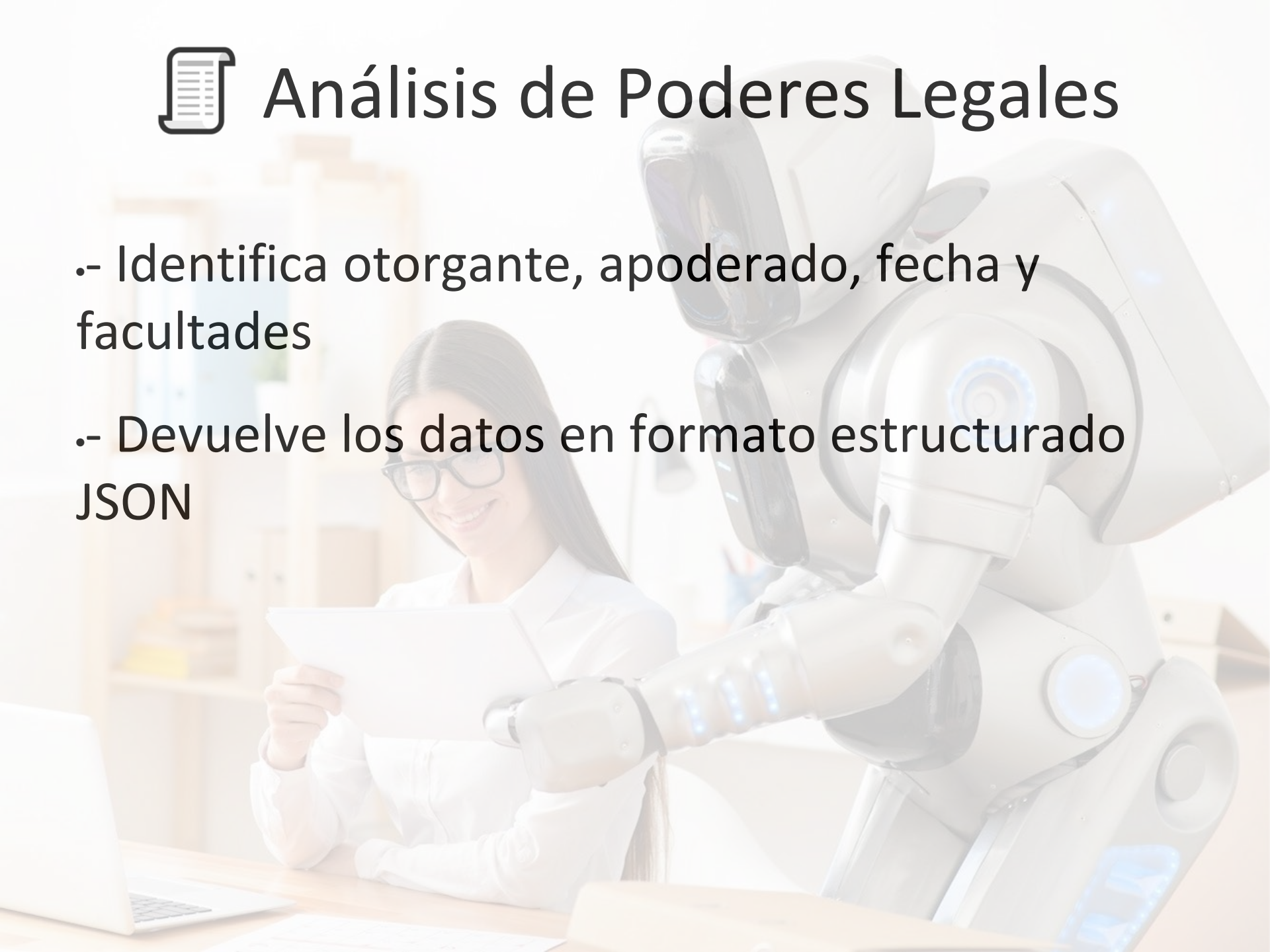
- Extrae partes relevantes de una operación: compradores, vendedores, monto, traslapes
- **Genera narrativa ejecutiva con GPT-4 Turbo**
- Permite generar un dictamen formal en formato .docx





# Análisis de Poderes Legales

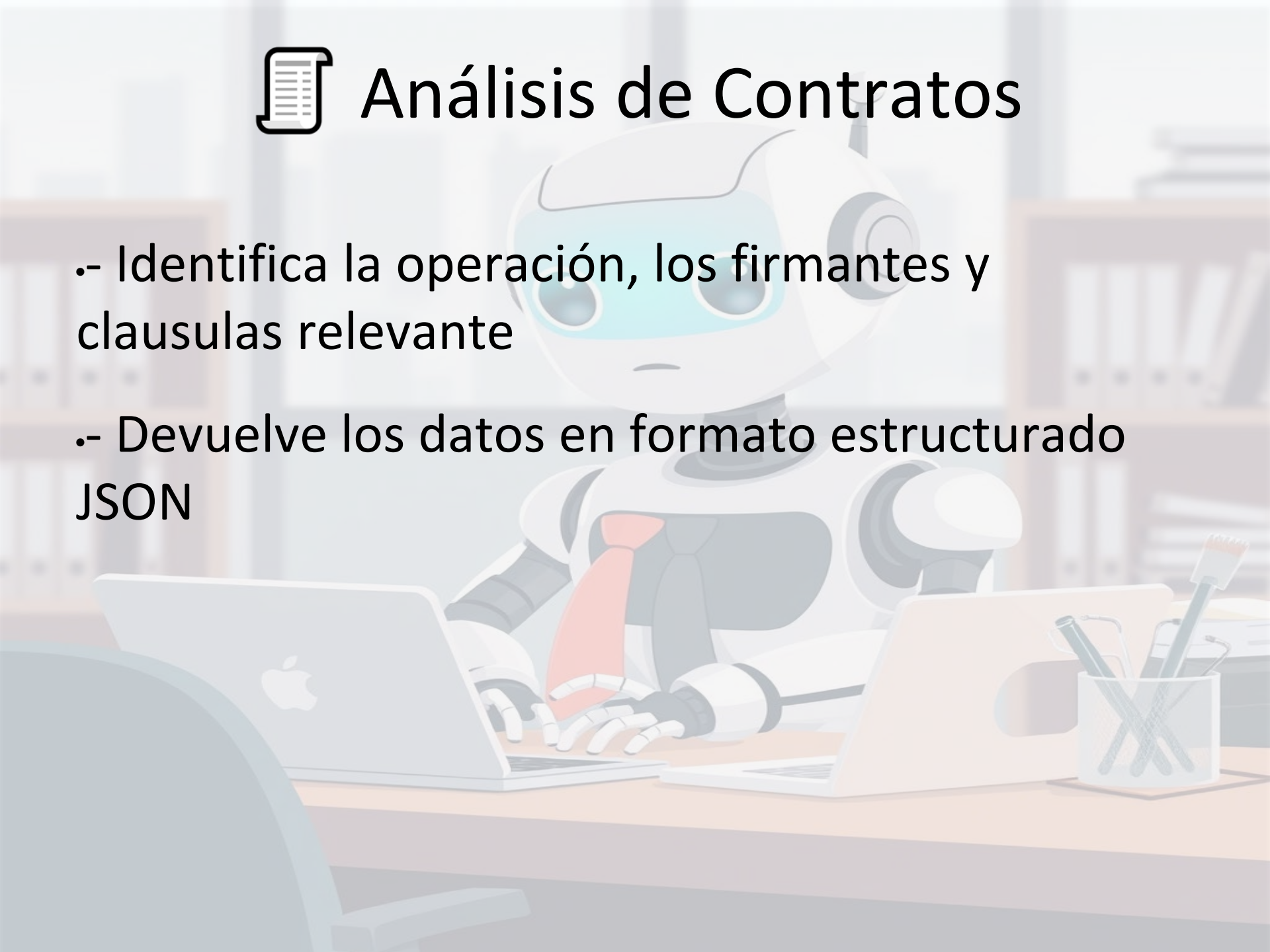
- Identifica otorgante, apoderado, fecha y facultades
- Devuelve los datos en formato estructurado JSON





# Análisis de Contratos

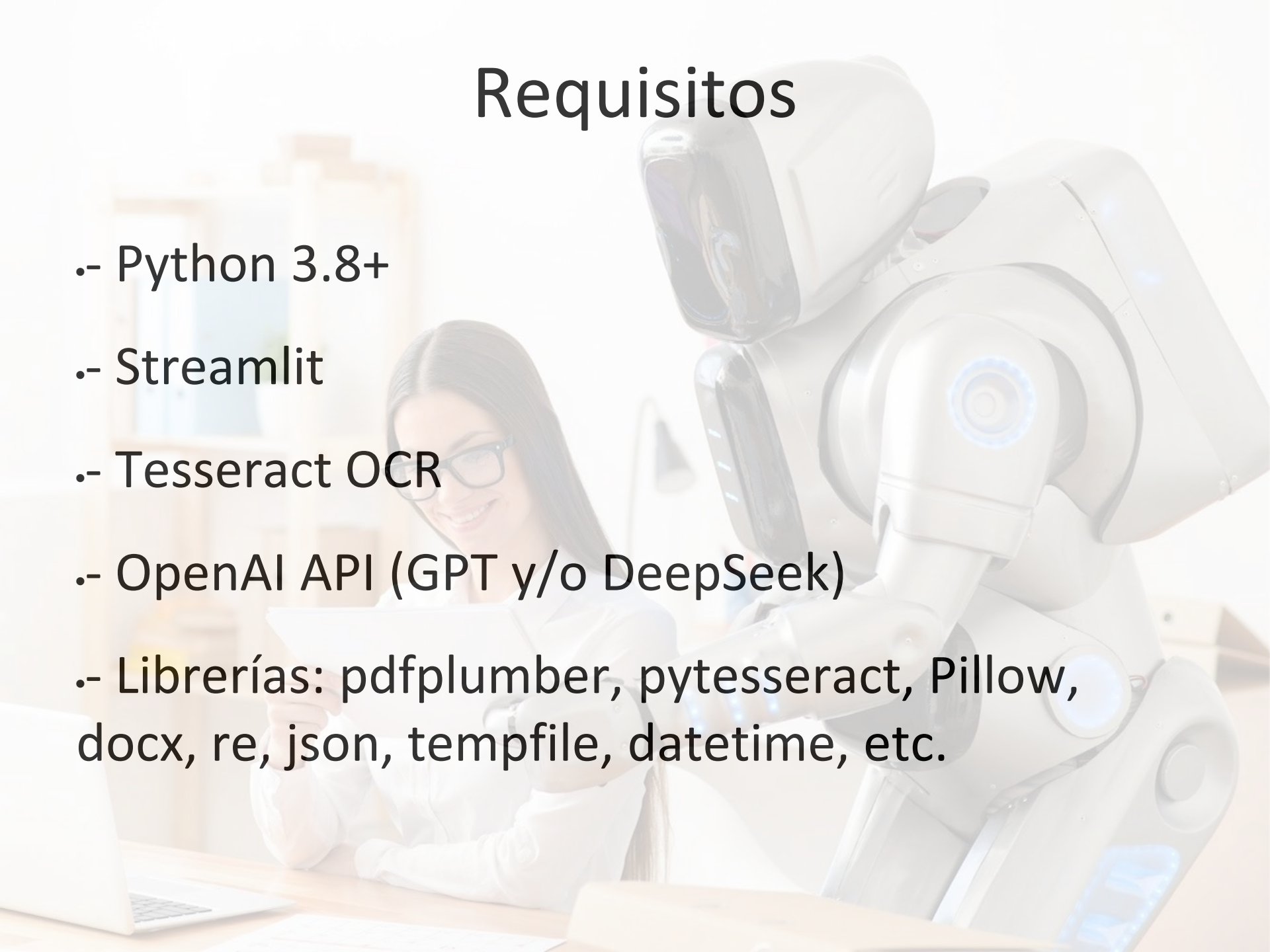
- Identifica la operación, los firmantes y cláusulas relevante
- Devuelve los datos en formato estructurado JSON





# Requisitos

- Python 3.8+
- Streamlit
- Tesseract OCR
- OpenAI API (GPT y/o DeepSeek)
- Librerías: pdfplumber, pytesseract, Pillow, docx, re, json, tempfile, datetime, etc.



# Cómo Usar



1. **Ejecuta la aplicación con:** `streamlit run app.py`
2. **Sube un documento PDF, imagen o texto**
3. **Selecciona el análisis deseado** (depósito, fusión o poder)
4. **Visualiza los resultados y descarga el JSON o dictamen generado**

# Créditos

- Desarrollado para facilitar el análisis documental en contextos legales, financieros y regulatorios por mí.
- .- Parte del código generado con ChatGPT (RAG), Deepseek (Streamlit), Mistral (Función poderes) y Anthropic (correcciones)