

Aplicaciones LLM

LangChain: Document Loaders

Versatilidad para la carga de datos

- Variedad de fuentes:
 - Páginas web.
 - Bases de datos.
 - YouTube, Twitter.
 - Excel, Pandas, Notion, Figma, HuggingFace, Github, etc.
- Variedad de formatos:
 - PDF.
 - HTML.
 - JSON.
 - Word, Powerpoint, etc.

Ejemplos

- Carga de documento PDF.
- Carga del audio de un video de YouTube.
 - La extracción de los datos es lenta.
 - Necesario instalar ffmpeg (en Mac, con Brew).
- Carga de los contenidos de una página web.
 - Tres opciones.
 - En este caso vemos la necesidad de limpiar y preparar los datos cargados antes de utilizarlos.

Otros

- Hay otros document loaders que puedes ver en la documentación de LangChain.