

Implementación 1 - GPT4All

- Instalar GPT4All: <https://www.nomic.ai/gpt4all>
- Descargar modelo local (Llama 3.2 3B)
- Probar modelo local (Nuevo chat)
- Crear carpeta Localdocs
- Subir docs y hacer embeddings (mirar extensiones de archivos)
- Probar modelo con documentos locales
- Habilitar server y probar endpoint
 - <http://localhost:4891/v1/models>