pdfplumber bibtexparser langchain sentence-transformers chromadb streamlit pyperclip openai