## Lépések:

Létrehoztam a github repot, majd a virtuális környezetet a projektnek venv-el, majd a poetry segítségével fogom kezelni a függőségeket, hogy meglegyenek egy helyen.

A megfelelő függőségeket leszedtem, illetve az ASR rendszernek a az openAI Whisperét választottem, amelyből van több is, de én a tinyt használom, a gyorsaság miatt, nem lesz probléma ha apróbb hibák lesznek a leiratozásban. Építettem egy gradio web appot, amelybe képesek vagyunk feltölteni hanganyagot, vagy a mikrofonon keresztül beszélni, és ezeket az előbb említett whisper modell leiratozza és elmentjük az eredményt a további lépésekhez.

Fontos lépés a megkapott leiratozás eredményének a vektorizációja, amelyet egy embedding modell végez el. Ezt először az Ollamán elérhező nomic embeddert használtam, ami azonban minden embedeléshez egy HTTP kérést küldene és emiatt nagyon lelassulna a teljes folyamat. Ezek miatt átálltam a Huggingfacen elérhető offline beágyazó modellt a sentence transformersen belül a "all-MiniLM-L6-v2"-t, amelynél használtam a batchelésnek a lehetőségét, hogy egyszerre 32 chunkot képes embeddelni így ha nagy dokumentumokat kapunk akkor sem lassul le nagyon a rendszer.