

Whisper Spracheingabe für Android von Lichtbote Martin

Nachdem ich mein erstes Programm für Windows 11 fertig geschrieben habe (Whisper Spracheingabe von Lichtbote Martin), wollte ich das gleiche auch für Android haben. Auch hier hatte ich eine Offline-Version geplant. Ich hatte dabei zuerst wieder das gleiche Modell verwendet, welches ich bereits zuvor für Windows 11 genutzt hatte. Allerdings zeigte sich schnell, dass das Handy dieses Modell technisch nicht richtig verarbeiten konnte. Der Arbeitsspeicher und die Prozessorleistung vieler Smartphones reichten dafür nicht aus, wodurch es zu langen Ladezeiten, Abstürzen und Leistungsproblemen kam.

Daraufhin hatte ich anschließend das kleinste Whisper-Modell verwendet, welches deutlich weniger Speicher und Rechenleistung benötigt. Dieses lief zwar wesentlich besser auf dem Handy, allerdings hatten die kleinen Versionen zu wenig Leistung bezüglich der Genauigkeit, was die Erkennung anging.

Die Google-KI schlug mir dann vor, das Ganze über eine API zu lösen. Außerdem wurden mir verschiedene API-Anbieter genannt. Ich habe mich letztlich für die zwei bekanntesten Anbieter entschieden. Das Geniale daran ist, dass einer dieser Anbieter bis zu 2000 Anfragen pro Tag kostenlos zur Verfügung stellt.

Was bedeuten 2.000 Anfragen pro Tag?

Da die Spracheingabe in einzelne Abschnitte von jeweils 28 Sekunden aufgeteilt wird, entspricht jede Anfrage einem solchen Sprachblock. Bei insgesamt 2.000 möglichen Anfragen pro Tag ergibt sich dadurch eine gesamte Aufnahmezeit von rund 56.000 Sekunden, genauer gesagt über 15 Stunden reinem Diktat pro Tag. Selbst bei intensiver Nutzung ist dieses Limit für normale Anwender daher kaum vollständig erreichbar und reicht problemlos für lange Texte, Notizen und den alltäglichen Gebrauch aus.

API Einrichtung

Der Link für die Einrichtung der API befindet sich innerhalb der App-Einrichtung. Sollte es später Probleme mit dem kostenlosen Anbieter geben oder der kostenlose Zugang eingeschränkt oder eingestellt werden, kann problemlos auf die API von OpenAI gewechselt werden. Die Whisper-API von OpenAI arbeitet nach einem nutzungsabhängigen Modell und kostet aktuell ungefähr 0,006 US-Dollar pro Minute Spracheingabe. Dadurch bleiben die Kosten selbst bei längerer Nutzung vergleichsweise gering.

Die App

Ausgehend von meinen ersten Erfahrungen würde ich sagen, dass die App genauer arbeitet als meine erste Windows-11-Version. Meine App befindet sich derzeit noch in der Erprobung und kann gegen eine faire Gegenleistung zur Verfügung gestellt werden.

<https://herunterladen.lichtbote-martin.de/>