Detailed Model Research for "Beyond QWERTY" Project.

Below is a detailed **model research breakdown**.

These models use cloud servers to process speech, ensuring **high accuracy** and **real-time capabilities** but requiring an **internet connection**.

❖ **Azure OpenAI Whisper (Azure Speech-to-Text)**

- **Accuracy:** High (5star)
- **Pros:**
    - Native integration with **Azure ecosystem**.
    - Supports over **100 languages** for transcription and translation.
    - **Real-time** and **batch processing** support.
    - **Custom speech models** for domain-specific terms.
- **Cons:**
    - Requires **internet**.
    - **Paid service** (cost depends on usage).
- **Use Case: Best for online deployments** with Azure-based infrastructure.

❖ **Google Cloud Speech-to-Text**

- **Accuracy:** High (4 star)
- **Pros:**
    - Supports **125+ languages** with auto-detection.
    - Can adapt models for **domain-specific** terminology.
    - Works well for **noisy environments**.
- **Cons:**
    - **Requires internet**.
    - **Pricing** can be high for large-scale use.
- **Use Case:** Plan to integrate with **Google Cloud**, this is a great option.

These models can process **speech-to-text without an internet connection**.

❖ **OpenAI Whisper (Open-Source)**

- **Accuracy:** High (5 star)
- **Pros:**
    - Best accuracy for **multilingual speech**.
    - Works **fully offline** when deployed locally.
    - Can be **fine-tuned** for specific accents and domains.
- **Cons:**
    - Requires **high computational power** (GPU recommended).
    - **Large model size** (not ideal for mobile devices).
- **Use Case: Best for high-end devices** running offline applications.

❖ **Vosk (Kaldi-Based)**

- **Accuracy:** Low
- **Pros:**
    - **Very lightweight** (works on mobile, Raspberry Pi).
    - **Works fully offline**.
    - Supports **20+ languages**.
- **Cons:**
    - **Lower accuracy** than Whisper.
    - Requires additional **post-processing** for better results.
- **Use Case: Best for low-resource environments** where Whisper is too heavy.
