Voice-to-Text script:
- Automates transcription of trending podcast episodes using OpenAI's Whisper ASR.
- Fetches trending episodes from Podcast Index (based on recent engagement, not just listens).
- Downloads audio and transcribes to text, saving results as JSON files.

Whisper ASR
- Converts audio waveforms to text transcripts.
- Model sizes: tiny (fast, less accurate), base (balanced), up to large (slowest, most accurate).
- Runs locally, no API needed
- Online APIs (e.g., AssemblyAI) are faster and more accurate but have free usage limits.

Inputs & Outputs
- Inputs:
  - top_podcasts.json — podcast metadata (title, description, id, author)
  - top_episodes.json — episodes with audio_url, podcast, episode_title
- Outputs:
  - Audio files saved in audio/
  - Transcripts saved in transcripts/ as JSON

Process flow
1. Fetch trending podcast episodes list (engagement-based).
2. Download episode audio if not already saved (using yt-dlp).
3. Convert audio and transcribe with Whisper.
4. Save transcript as JSON.
5. Skip episodes if the transcript already exists.