From Prompt to Production

Phlo Young

Introduction: Who am I?

- **Phlo:** Lifelong musician, artist, songwriter.
- **Background:** Experimentation, automation, self-taught engineer.
- **Goal:** Interactive, fun, tangible takeaways.

Who I am not.

- **Not** associated with Suno, Udio or any other music gen company.
- **Not** a credentialed professional.
- **Not** a paid shill.

What is AI Music?

Initial research showed blurred lines and varying definitions.

Let's categorize:

1. **Text-to-Music:** Generating audio (samples or full songs) from text prompts.
2. **Audio-to-Music:** Transforming existing sounds into music.
3. **Voice Conversion:** Applying one voice's characteristics to another voice. (Often called "AI Music" by the general public initially).

1. Voice Conversion

- **Concept:** Take existing vocals -> Change them to sound like someone else.
- **Process:** Often involves traditional songwriting (production, lyrics, recording original vocals) + AI voice model.

- **Early Example** - Roberto Nickson: AI Kanye voice (first demo of its kind).
- **Viral Example** - Ghostwriter: Drake/The Weeknd AI song ("Heart on My Sleeve") - Human-written/performed, AI voices applied. Went viral, DMCA'd.
- **State-of-the-Art Example** - Randy Travis: Artist-sanctioned use after voice loss. High quality, overwhelmingly positive public reception. Shows ethical potential & improved tech.

2. Text-to-Music

- **Concept:** Pure generation from text prompts/lyrics. No prior recording needed.
- **Examples**
  - **King Willonius:** "BBL Drizzy" meme song created in Udio from prompt/lyrics.
  - **Eleven Labs:** "Started to Sing" song about GPUs sung by an AI model.
- This is closer to true **"AI Button"** music creation.

3. Audio-to-Music

- **Concept:** Input an audio sample -> Get music generated based on it.
- **Examples**
  - **Stable Audio:** Transforming a basic sound into musical elements.
  - **Suno:** Finger drumming on a desk -> Full song with guitar, etc.
- **Potential:** Unlocks creative workflows for musicians using existing sounds/ideas.

Use Cases

- Background music for content
- Video game music
- Learning tool
  - Create songs to help my daughter memorize facts for school
- Custom songs that otherwise wouldn't get created
  - Theme songs for events, companies
    - Family friendly music at events
    - Custom birthday songs
- Combining unpopular genres to create hybrids
- Loops for music production

Music Generation Services

Udio

- "Songwriting Partner"
- Generates 30s chunks (extend, remix, scrap).
- In-painting (modify sections).
- Audio-to-Audio feature.
- Experimental longer models available.

Suno

- "In-house Music Producer"
- Outputs often sound polished, Top 40-ish.
- Improved song structure & timing recently.
- Custom mode for lyrics/structure control.

Riffusion

- "AI Sound Designer"
- Unique approach: Generates music by creating spectrogram images.
- Visualizes music generation process.
- Can modify existing sounds or generate from text.

Open Models and Tools

Stable Audio Open

- Specializes in short audio samples, sound effects, drum beats, and ambient sounds.
- Generates content up to 47 seconds in high-quality 44.1kHz stereo.
- Respects creator rights with careful licensing of training data.
- Ideal for creating sound effects, ambient sounds, and production elements.

YuE

- Open foundation model for full-song music generation with lyrics.
- Generates songs up to 5 minutes with both vocals and instrumentals.
- Supports multiple languages including English, Chinese, Japanese, and Korean.
- Built on LLaMA2 architecture with dual-track system for synchronized vocals/instrumentals.
- Features in-context learning for style transfer and music continuation.

DiffRhythm

- First open-source diffusion-based music generation model for complete songs.
- Generates full-length songs (up to 4m45s) in only 10 seconds.
- Creates both vocal and instrumental tracks simultaneously.
- Simple design with no complex cascading pipelines.
- Supports text-based style prompts and instrumental-only generation.

Other Notable Tools & Tech

- **Boomy:** Early text-to-music platform.
- **Eleven Labs:** Text-to-speech company, now also in text-to-music.
- **RVC (Retrieval-based Voice Conversion):** Considered OSS state-of-the-art for voice cloning. Can be run locally or in the cloud. Can blend voices.
- **UVR5:** Open-source local software to extract stems from *any* audio file.

How Do They Work?

(The Short Version)

- **Training Data:** Models are trained on massive datasets of music.
- **The Controversy:** Allegations (and lawsuits) that training data includes copyrighted material without permission.
- **Architectures:** Techniques like Transformers and Diffusion models are used. (technical papers linked in the notes at `aietalk.com/music`)
- **Content Filters:** Platforms often block generation based on specific artist names due to legal risks.

Prompting: The Art of AI Music Gen

- **Basics:** Describe genre, mood, instruments, tempo.
- **Getting Specific:** The more detail, the better (sometimes).
- **Artist Names:** Avoid direct names. Try describing the *style* ("in the style of...", "sounds like..."). Sometimes variations work ("Tandy Ravis").
- **Custom Lyrics:** Using Suno/Udio's custom modes with your own lyrics gives more control.
- **Rate Your Music hack:** Search RYM for sounds/artists -> Use the descriptive tags/genres RYM uses as prompts in Suno/Udio.

Beyond Generation: Editing & Control

- AI Music often needs tweaking.
- **Stems:** Separated tracks (vocals, bass, drums, etc.).
- **Tools for Stems:**
  - **Suno Tool:** Suno now has their own built-in tool for stems.
  - **UVR5:** More powerful, works on any audio, requires local setup/hardware.
- Allows for mixing, mastering, replacing parts, integrating with traditional DAWs.

The Future & Ethics

- **Public Perception:** Shifting? Still taboo for many, but positive examples (Randy Travis) exist.
- **Legal Landscape:** Highly uncertain. Ongoing lawsuits (RIAA) are critical. Copyright status of AI outputs is debated.
- **Opportunities:** Lowering the barrier to creation, tools for artists, new income streams.
- **Responsible Use:** A major ongoing discussion. Artist consent, attribution, compensation are key issues.

Resources & Final Thoughts

- Workshop notes page (`aietalk.com/music`) for:
  - Links to tools (Suno, Udio, etc.)
  - GPT Prompt Templates
  - List of Music Model Research Papers
  - Full Transcript & Recording
- **Experiment!** The best way to learn is to try generating music yourself.

Thank You!