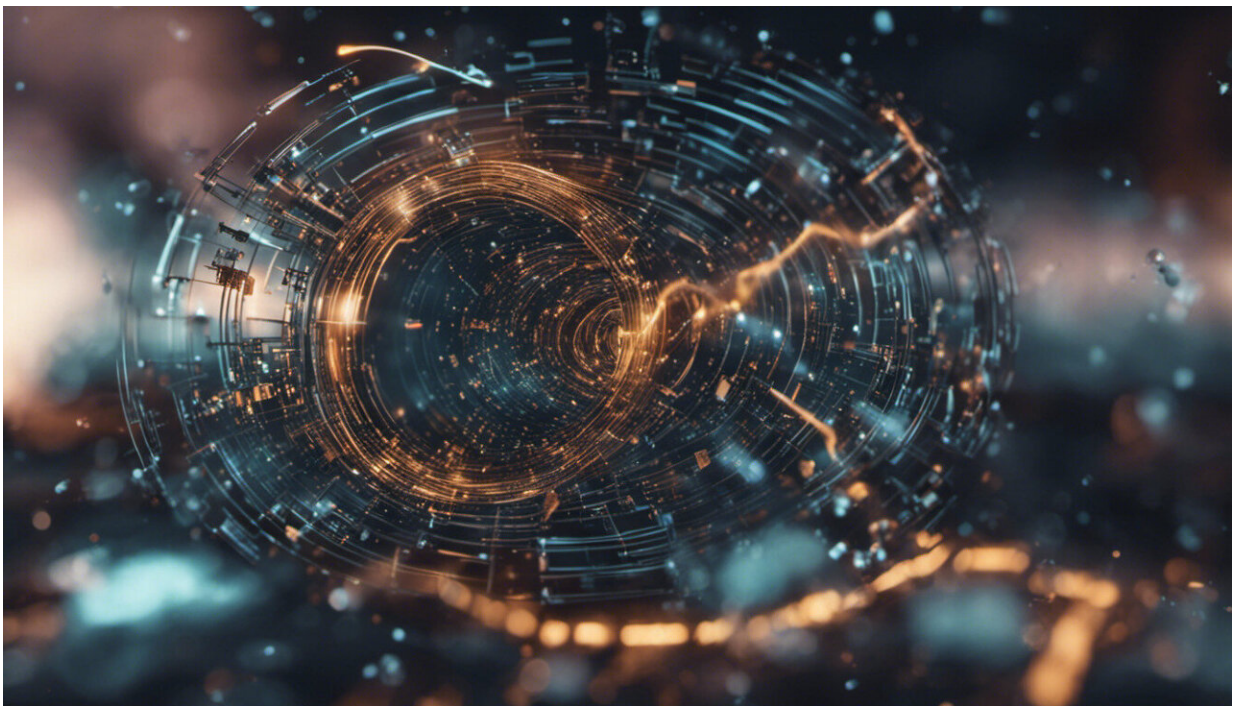# Text-to-audio generation is here: One of the next big AI disruptions could be in the music industry

February 3 2023, by Oliver Bown

Credit: AI-generated image (disclaimer)

The past few years have seen an explosion in applications of artificial intelligence to creative fields. A new generation of image and text generators is delivering impressive results. Now AI has also found applications in music, too.

Last week, a group of researchers at Google released MusicLM—an AI-based music generator that can convert text prompts into audio segments. It's another example of the rapid pace of innovation in an incredible few years for creative AI.

With the music industry still adjusting to disruptions caused by the internet and streaming services, there's a lot of interest in how AI might change the way we create and experience music.

Automating music creation

A number of AI tools now allow users to automatically generate musical sequences or audio segments. Many are free and open source, such as Google's Magenta toolkit.

Two of the most familiar approaches in AI music generation are:

1. **continuation**, where the AI continues a sequence of notes or waveform data, and

2. **harmonization or accompaniment**, where the AI generates something to complement the input, such as chords to go with a melody.

Similar to text- and image-generating AI, music AI systems can be trained on a number of different data sets. You could, for example, extend a melody by Chopin using a system trained in the style of Bon Jovi—as beautifully demonstrated in OpenAI's MuseNet.

Such tools can be great inspiration for artists with "blank page syndrome," even if the artist themselves provide the final push. Creative stimulation is one of the immediate applications of creative AI tools

today.

But where these tools may one day be even more useful is in extending musical expertise. Many people can write a tune, but fewer know how to adeptly manipulate chords to evoke emotions, or how to write music in a range of styles.

Although music AI tools have a way to go to reliably do the work of talented musicians, a handful of companies are developing AI platforms for music generation.

Boomy takes the minimalist path: users with no musical experience can create a song with a few clicks and then rearrange it. Aiva has a similar approach, but allows finer control; artists can edit the generated music note-by-note in a custom editor.

There is a catch, however. Machine learning techniques are famously hard to control, and generating music using AI is a bit of a lucky dip for now; you might occasionally strike gold while using these tools, but you may not know why.

An ongoing challenge for people creating these AI tools is to allow more precise and deliberate control over what the generative algorithms produce.

New ways to manipulate style and sound

Music AI tools also allow users to transform a musical sequence or audio segment. Google Magenta's Differentiable Digital Signal Processing library technology, for example, performs timbre transfer.

Timbre is the technical term for the texture of the sound—the difference between a car engine and a whistle. Using timbre transfer, the timbre of

a segment of audio can be changed.

Such tools are a great example of how AI can help musicians compose rich orchestrations and achieve completely new sounds. In the first [AI Song Contest](), held in 2020, Sydney-based music studio [Uncanny Valley]() (with whom I collaborate), used timbre transfer to bring singing koalas into the mix.

Timbre transfer has joined a long history of synthesis techniques that have become instruments in themselves.

Taking music apart

Music generation and transformation are just one part of the equation. A longstanding problem in audio work is that of "source separation." This means being able to break an audio recording of a track into its separate instruments.

Although it's not perfect, AI-powered source separation has come a long way. Its use is likely to be a big deal for artists; some of whom won't like that others can "pick the lock" on their compositions.

Meanwhile, DJs and mashup artists will gain unprecedented control over how they mix and remix tracks. Source separation start-up [Audioshake]() claims this will provide new revenue streams for artists who allow their music to be adapted more easily, such as for TV and film.

Artists may have to accept this Pandora's box has been opened, as was the case when synthesizers and drum machines first arrived and, in some circumstances, replaced the need for musicians in certain contexts.

But watch this space, because copyright laws do offer artists protection from the unauthorized manipulation of their work. This is likely to

become another gray area in the music industry, and regulation may struggle to keep up.

New musical experiences

Playlist popularity has revealed how much we like to listen to music that has some "functional" utility, such as to focus, relax, fall asleep, or work out to.

The start-up Endel has made AI-powered functional music its business model, creating infinite streams to help maximize certain cognitive states.

Endel's music can be hooked up to physiological data such as a listener's heart rate. Its manifesto draws heavily on practices of mindfulness and makes the bold proposal we can use "new technology to help our bodies and brains adapt to the new world," with its hectic and anxiety-inducing pace.

Other start-ups are also exploring functional music. Aimi is examining how individual electronic music producers can turn their music into infinite and interactive streams.

Aimi's listener app invites fans to manipulate the system's generative parameters such as "intensity" or "texture," or deciding when a drop happens. The listener engages with the music rather than listening passively.

It's hard to say how much heavy lifting AI is doing in these applications—potentially little. Even so, such advances are guiding companies' visions of how musical experience might evolve in the future.

The future of music

The initiatives mentioned above are in conflict with several long-established conventions, laws and cultural values regarding how we create and share music.

Will copyright laws be tightened to ensure companies training AI systems on artists' works compensate those artists? And what would that compensation be for? Will new rules apply to source separation? Will musicians using AI spend less time making music, or make more music than ever before?

If there's one thing that's certain, it's change. As a new generation of musicians grows up immersed in AI's creative possibilities, they'll find new ways of working with these tools.

Such turbulence is nothing new in the history of music technology, and neither powerful technologies nor standing conventions should dictate our creative future.

This article is republished from The Conversation under a Creative Commons license. Read the original article.

Provided by The Conversation