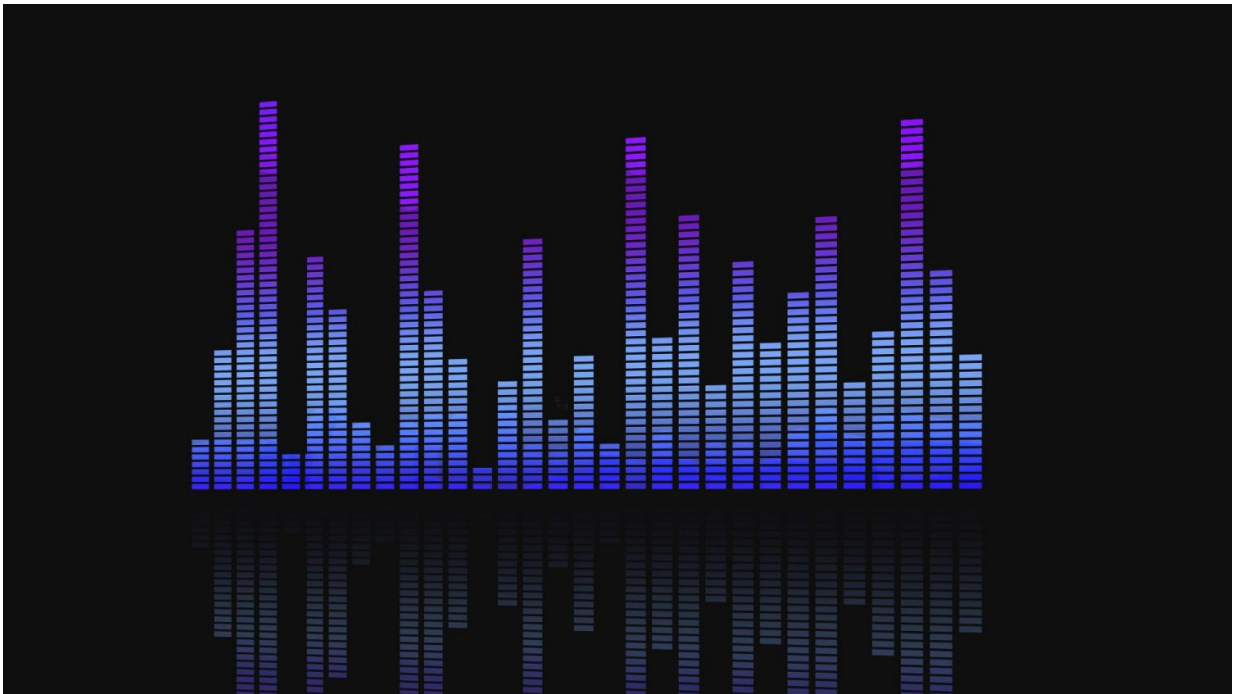# Lyrebird to offer tech to copy the voice of anyone

April 26 2017, by Nancy Owano

Credit: CC0 Public Domain

(Tech Xplore)—Technology via an application programming interface can copy any person's voice just from a one-minute audio recording.

The team behind it issued a press announcement on Monday. They are Montreal-based startup Lyrebird. The announcement said, "Lyrebird is going a step further in the development of AI applications by offering to

companies and developers new speech synthesis solutions."

What they want to talk about is a voice-imitation algorithm that can mimic a person's voice and have it read any text with a given emotion, based on the analysis of just "a few dozen seconds" of audio recording.

Lyrebird is a group focused on developing speech synthesis technologies. Apparently their choice of the name Lyrebird is not casual. Nature's fascinating lyrebird, which can mimic sounds it hears, including car alarms.

As for their speech technology, it is the result of their research conducted at the Montreal Institute for Learning Algorithms (MILA) lab at University of Montréal. They said that their GPU clusters generate 1,000 sentences in less than half a second.

The target users would be developers interested in recreating a person's voice using an audio recording. Lyrebird posted some interesting (and entertaining) audio examples using well known voices of famous people including Obama and Trump to show the capabilities.

"On the website lyrebird.ai, samples using the voices of Donald Trump, Barack Obama and Hillary Clinton illustrate the accuracy and effectiveness of the technology."

Samples indicate that the team found success in voice mimicry and also in offering imitations of different emotions. You want a voice showing signs of stress, or anger? You've got it.

"Control the emotion. Anger, Sympathy, Stress. Lyrebird allows to control the emotion of the generated voice."

This would not be the first time people would be made more aware of

how [voice technology](#) has advanced. There was lots of interest in last year's Adobe debut of Project VoCo technology. Think photo editing but this time with audio.

The software showed one could take an audio recording and change it to include words and phrases the original speaker never said but in what sounded like their voice.

Contrasting the two this week, however, reports said that VoCo needed to 'hear' at least 20 minutes of original audio to do the job in speech synthesis while the Lyrebird tech just needs about a minute.

So what's next for this team? Their API is still under development. "We believe that vocal human-computer interfaces will become more and more widespread in the future and we want to lead the race."

Their site is inviting anyone interested to subscribe to become a beta tester or get informed of the [launch](#).

Reactions? Andy Weir on Monday in *Neowin*: "it's easy to envisage how the technology will be refined to eventually enable the creation of digital voices that sound realistic enough to fool the listener into believing that they're hearing a real [person](#)."

Meanwhile, that ability to fool the listener could possibly raise concerns that mischief makers with bad motives could tamper with audio to mislead people.

Nonetheless, the team had in mind beneficial applications, such as the technology being used for personal assistants, and for people with disabilities.

Well aware of concerns about how this technology may be used, the

team stated that they hope to accomplish something positive from their work. "By releasing our technology publicly and making it available to anyone, we want to ensure that there will be no such risks. We hope that everyone will soon be aware that such technology exists and that copying the voice of someone else is possible. More generally, we want to raise attention about the lack of evidence that audio recordings may represent in the near future."

More information: lyrebird.ai/
lyrebird.ai/demo

© 2017 Tech Xplore