



The Role of Ethics in Voice Development

February 26, 2020

When was the last time a computer spoke to you? If you're like many consumers, it was probably this morning at home while checking the weather via a smart speaker, or on your way to work getting navigation tips from your in-car assistant.

As voice assistants become more prevalent and text-to-speech (TTS) technologies become more advanced – think [voice clones](#), [genderless voices](#), and more – we must continue to consider ethical standards when we develop innovative products. Here, I answer a few of the most frequently asked questions I receive on this topic:

Can people be fooled or misled by TTS?

Users can be misled if an application using computer-generated speech is poorly designed. The quality of computer speech is becoming hard to distinguish from human speech; therefore, applications must clearly identify themselves as a computer system to avoid confusion. Cerence User Interface (UI) experts conduct a variety of user studies and support Cerence customers with best practices in UI design to meet this requirement.

Can TTS be used for inappropriate purposes?

As the quality of computer speech improves, it's true that bad actors can more easily create an application to mislead or harm people. For example, a perpetrator may obtain recordings of a victim and impersonate them using TTS. This will create changes in society where you need additional caution before trusting a recording or a person you interact with over a loudspeaker. It is similar to the impact of digital image editing on the perceived authenticity of photos and videos. Voice impersonations are not new, but they will become more accessible to people with different intentions. Voice technology can also help against inappropriate use, which is the topic of the next question.

What does Cerence do to prevent unethical use of its TTS?

We are closely aligned with our customers, all the world's leading automakers, and support them to make their use of TTS technology beneficial to users in full consideration of ethical standards. We typically do not make our technology available to individual developers, where there is a higher chance that it might be used unethically.

We also offer voice biometric solutions that can not only identify a person through their unique voice, but also detect the use of synthetic speech. Further, we're developing audio watermarking, inaudible changes embedded into our TTS to make sure our biometric solutions can always distinguish it. In the future, governments may require TTS systems to always embed an audio watermark.

Opportunities in the world of text-to-speech and computer-generated voices are abundant, but as with any technological innovation, we must keep ethical considerations for the end user at the forefront. We at Cerence look forward to staying on the leading edge.