

Pretty sure Google's new talking AI just beat the Turing test

During the on-stage demonstration, Google played calls to a number of businesses, including a hair salon and a Chinese restaurant. At no point did either of the people on the other end of the line appear to suspect that the entity they were interacting with was a bot. And how could they when the Assistant would even throw in random "ums," "ahhs" and other verbal fillers people use when they're in the middle of a thought? According to the company, it's already generated hundreds of similar interactions over the course of the technology's development.

This robo-vocalization breakthrough comes as the result of Google's Duplex AI system, which itself grew out of earlier Deep Learning projects such as [WaveNet](#). As with Google's other AI programs, like [AlphaGo](#), Duplex is designed to perform a narrowly defined task but do it better than any human. In this case, that task is talking to people over the phone.



Duplex's initial application, according to the company, will be in automated customer-service centers. So rather than repeatedly shouting "operator" into your handset to get competent help the next time you call your bank, cable provider or power company, the digital agent will be able to assist you directly. "For such tasks, the system makes the conversational experience as natural as possible, allowing people to speak normally, like they would to another person, without having to adapt to a machine," the release read.

But don't expect to be able to ask the agent any random question that pops into your head. "Duplex can only carry out natural conversations after being deeply trained in such domains," a Google release points out. "It cannot carry out general conversations." Fingers crossed that the company trained Duplex to at least say "does not compute" in such situations.

This news comes as the customer-service industry continues to lean on automation and robotics to reduce operating costs and streamline wait times. Everyone from [DoNotPay](#) and [Kodak](#) to [Microsoft](#) and [Facebook](#) are looking to leverage chatbots and other automated respondents to their advantage. It's not even that hard to simply [code your own](#). However, like every technology before it, Duplex can only be as useful as we make it. And unfortunately, we've not had a particularly sterling record when it comes to training AIs to be upstanding members of online society. Yes, [Tay, I'm looking right at you](#). And [your racist chatbot buddy, Zo](#), too.

Of course, casual racism is only one of a number of ethical and societal challenges that such emerging technologies face. As always, there's the issue of privacy and data collection. "To obtain its high precision, we trained Duplex's RNN on a corpus of anonymized phone-conversation data," Google's release explained. This AI wasn't taught how to talk like people just by conversing with its developers (we're still years away from that level of machine-learning, unfortunately). No, it was trained on hundreds, if not thousands, of hours of recorded phone conversations skimmed from opted-in customer calls. Which brings us, yet again, back to the same debate we've been having for the better part of a decade between [maintaining personal data privacy](#) and advancing the boundaries of technological convenience.

Existential solutions to the [tragedy of the public commons](#) aside, the emergence of Duplex AI exposes a number of tangible issues that should probably be addressed before we start talking to machines like we do other people. For one thing, WaveNet is really, *really* good. WaveNet is Google's voice synthesizing program, and unlike conventional text-to-speech engines -- which have a voice actor basically read through the dictionary in order to populate a database of words and speech fragments -- it relies on a smaller grouping of raw audio waveforms on which the computer's words and responses are built. This results in a system that is faster to train and boasts a broader, more natural-sounding response range.

To illustrate, remember when TomToms were still a thing and had installable voices like James Earl Jones or Morgan Freeman to give you driving directions? Those were generated using the conventional "read a dictionary" method and were unable to pass for the real person given their stiff intonations. The new [John Legend voice available for Google Assistant](#) sounds far more like the man himself thanks to WaveNet's capabilities.

But what's to prevent someone from illicitly leveraging recordings of a public figure to train an AI and generate falsified audio? I'm not saying we're facing a "[my name is my password](#)" situation just yet, but given that people are already able to fake images and video ([even porn](#)), being able to incorporate a layer of false audio will only make [Google](#) and [Facebook's](#) content-moderation jobs even harder. Especially since their omega-level objectionable-content censors are human -- exactly whom the technology is designed to fool.

That said, the mere possibility of misuse should not be a deal-breaker for this technology (or any other). Duplex AI has the potential to revolutionize how we interact with computer systems, effectively making tactile inputs like keyboards, mice and touchpads obsolete. Just, next time a "family member" calls out of the blue asking for your Social Security number, maybe try to independently confirm their identities first by asking them a series of unrelated questions.

Click [here](#) to catch up on the latest news from Google I/O 2018!

Images: Google (diagram)

In this article: [AI](#), [AlphaGo](#), [breaking news](#), [DuplexAI](#), [google](#), [GoogleIO2018](#), [internet](#), [MachineLearning](#), [opinion](#), [tomorrow](#), [WaveNet](#)

All products recommended by Engadget are selected by our editorial team, independent of our parent company. Some of our stories include affiliate links. If you buy something through one of these links, we may earn an affiliate commission.



Comments



8254 Shares