????????????? 2??????????

????????????? 2??????????

Centific ??



Tom Cruise and Val Kilmer are not the only stars of the hit movie *Top Gun: Maverick*. Artificial intelligence is right up alongside Maverick and Iceman to make the drama real and compelling for movie goers around the world.

Top Gun: Maverick is the sequel to the 1986 iconic movie *Top Gun*, in which Cruise and Kilmer played rival students at the U.S. Navy's Fighter Weapons School. Between 1986 and 2022, when *Top Gun: Maverick* was released both Cruise and Kilmer enjoyed high successful film careers. But unfortunately, Kilmer lost the use of his voice after a battle with throat cancer. Here is where AI played a role.

In one crucial scene in *Top Gun: Maverick*, Kilmer's character "Iceman" Kazansky and Cruise's character "Maverick" Mitchell have a brief but pivotal conversation. For that to happen, Kilmer's voice needed to sound authentic and believable to fans who are familiar with his past film roles. Trying to replicate Kilmer's voice as it sounds today (he can barely speak above a slight rasp) was

not an option. But the movie's producers had access to audio files from Kilmer's many movies. Could those be used to re-create his voice? A London-based startup, Sonantic -- which has been using AI to help Kilmer communicate since 2021 -- was given the job of creating Iceman's words onscreen.

Sonantic used a "voice engine" to teach an AI voice model how to speak like Kilmer based on old audio recordings of the actor. That model would literally speak his lines. But even with audio files to work from, the model had 10 times less data than it would have been given in a typical project – which demonstrates the enormous amount of data that AI models require. The team needed to create a new algorithm that could produce a higher-quality voice model using the available data. From there, Sonantic generated more than 40 different voice models and selected the one with the highest fidelity. After that, creative teams input text and fine-tune the performance. Iceman speaks!

Kilmer's artificially generated voice is known as synthetic voice, and the approach that Sonantic used to create Kilmer's synthetic voice is referred to as voice cloning. Over the past few years, businesses learned how to use AI to generate synthetic voices for applications such as corporate videos, digital assistants, and video-game characters. Oftentimes, businesses have relied on text-to-speech (TTS) services that can convey written text to spoken word with the right tone and voice to reflect a brand's personality. The goal is to create spoken content such as narration faster and less expensively by relying on machines instead of actors. One of the challenges of TTS has been creating synthetic speech that re-creates nuances in how human beings speak, such as tone. But advances such as voice cloning are improving synthetic voice vastly.

Improvements with deep learning have made it possible for synthetic voice to convey many of the subtleties of human speech. Voices pause and breathe when a listener would expect them to, and they change their style or emotion. Just as impressively, synthetic voices (unlike a recording of a human voice actor) can update their script in real time, which creates opportunities for personalizing the spoken word for different audiences and applications.

Synthetic voices matter to businesses for many reasons beyond making movies. One of them is the growing importance of sonic branding, or differentiating a brand through sound. One category of sonic branding involves businesses creating chimes or musical interstitials, such as the little ta-dum sound that any Netflix viewer recognizes instantly when they stream Netflix. In addition, KFC has used synthetic voice to bring to life the personality of corporate mascot Harlan Sanders for customer service applications.

For years, brands have famous actors to provide narration for ads, which creates a sense of familiarity and imparts a desired tone. But actors can be expensive, and their voices usually have a limited shelf life for commercial use. A synthetic voice offers an alternative.

But synthetic voice has a long way to go. As noted, the team that created a replica of Val Kilmer's voice was hampered by a lack of available data to work from. Fortunately, another type of approach for developing synthetic voice, synthetic data sets, has been gaining momentum. Synthetic data sets involve the use of synthetic data for audio, images, and text, which people use to help train voice recognition AI, optical character recognition, and natural language processing models. All this makes it possible for an AI application to learn faster and more accurately.

Synthetic data mimics what real data would look like, but smart engineering and/or AI with humans in the loop are used to make the data instead, either starting from "hints" of good data from AI, or already collected, and then engineering it to obtain the expected result. For example, Centific recently

helped a client rely on synthetic data sets for a voice application to learn about new and evolving concepts such as electric vehicles to provide more relevant results for a voice assistant that answers search queries.

Another challenge with synthetic voice is authentically reflecting the nuances of speech in local cultures and different languages. Even with language translation, synthetic voice is in the early stages of learning – an example being mastering different dialects in Spain, Italy, the United States, the United Kingdom, and so on.

In addition, synthetic voice needs to adapt to each language and culture simultaneously. Voice apps such as Alexa are still learning how to do that. Only by collecting data and how voices sound to local people can businesses make synthetic voice more inclusive globally, and you need someone with a knowledge of local nuances – not just the words you use but how you convey those words in the right tone.

How to Get Started

Many businesses are interested in synthetic voice but are not quite sure how to get started. They're hearing about other companies succeeding with synthetic voice, understand intuitively its value, but are not sure where synthetic voice can play a role. In branding? Employee training? And where else?

These are understandable questions. Tools such as design sprints can help. A design sprint consist of four-day test-and-learn process in which a team identifies a business problem with no clear-cut and easy solution and develops a prototype for a solution. For example, a business might ask, "How might we improve customer loyalty with AI technologies such as voice?" We use design sprints as part of our FUEL methodology for unlocking innovation.

Many other businesses have moved past the "how do we get started?" and are asking "How do we get better?" For those businesses, we also tap into our global AI expertise – including AI localization -- to help businesses improve. We combine both a diverse, global team of people, techniques such as training with synthetic data, and a platform (OneForma) to scale AI-based applications such as synthetic voice.

Contact us for more insight on how we can help you.

Image source: David Mark on Pixabay

-
-
-

-
-