NUANCE | WHAT'S NEXT BLOG

Customer engagement, Authentication & fraud prevention

Should you be worried about OpenAI's new Voice Engine?

Brett Beranek | Vice President & General Manager, Security & Biometrics

April 2, 2024

Although a voice biometric engine on its own may be vulnerable to synthetic voice attacks, Nuance Gatekeeper employs highly effective anti-spoofing technology to detect synthetic voices and playback of recordings—a capability that we have extensively tested and enhanced over the past year and are continuing to develop as new voice tools hit the market.

On March 29th 2024, OpenAI shared preliminary testing results from their new synthetic voice tool, Voice Engine, which is capable of generating highly realistic synthetic voice audio based on a small sample of just 15 seconds from the original speaker. While Voice Engine is still not publicly available, its capabilities represent an impressive new entry into the already crowded category of voice "cloning" software.

The dawning of generative AI has introduced a whole host of new applications for synthetic voice content, from audio/video production to human-sounding voicebots and communication assistance for the disabled. However, as with any new technology, it also introduces possibilities for abuse.

Does Voice Engine present a threat to voice biometric

security?

Because of their ability to replicate human voices in a manner convincing to the human ear, tools like Voice Engine have raised concerns about the reliability of voice biometrics as a secure authentication method for high-risk scenarios like accessing bank accounts. Thankfully, Nuance has remained two steps ahead on this issue.

Although a voice biometric engine on its own may be vulnerable to synthetic voice attacks, Nuance Gatekeeper employs highly effective anti-spoofing technology to detect synthetic voices and playback of recordings—a capability that we have extensively tested and enhanced over the past year and are continuing to develop as new voice tools hit the market. Considering this, and considering the myriad threats of the fraud landscape as a whole, our recommendation is that organizations **continue to invest in modern biometric security** while avoiding any voice authentication solutions that do not include strong, proven synthetic speech detection.

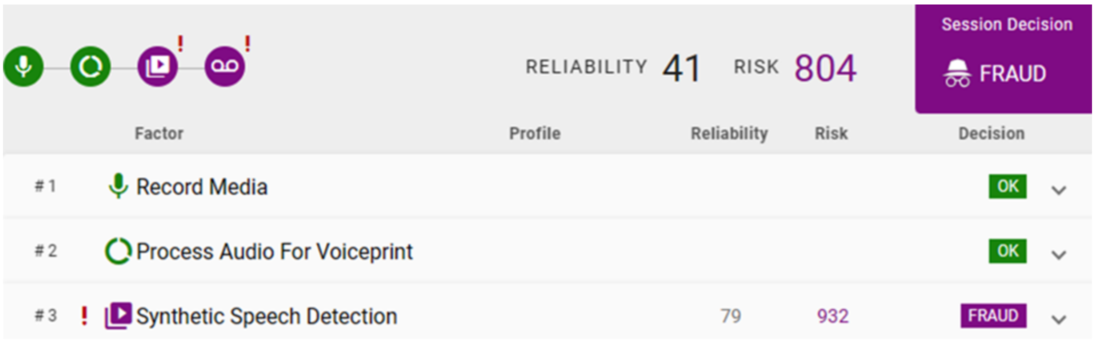
There are two more important details to keep in mind as you evaluate your organization’s defenses against emerging fraud vectors:

1. Voice Engine is not a publicly available tool for consumers (yet).
2. Voice biometrics are unique in their ability to identify the actual person behind the interaction based on intrinsic characteristics. The alternative is reverting back to knowledge-based and possession-based factors like security questions and two-factor PINs, which have long ago proven a favorite playground of professional fraudsters.

Can Gatekeeper detect Voice Engine audio as synthetic?

The audio samples shared by OpenAI this weekend are no doubt impressive to the human ear, arguably even more so than those produced by other voice AI tools on the market. But can they fool Gatekeeper’s multiple layers of biometric matching, anti-spoofing, and AI risk scoring?

In our initial testing, Gatekeeper was easily able to detect Voice Engine audio as synthetic, showcased in the screenshot below. The synthetic speech score in this case measured 932, which is very high on a scale from 0 to 1,000. This indicates that the OpenAI voice is without a doubt synthetic and not a real person. The overall risk score measured 804 out of 1,000, which would trigger a “suspected fraud” alert in real time during any authentication attempt. Subsequent tests produced similar results.



RELIABILITY 41 RISK 804

Session Decision

FRAUD

	Factor	Profile	Reliability	Risk	Decision
# 1	Record Media				OK
# 2	Process Audio For Voiceprint				OK
# 3	! Synthetic Speech Detection		79	932	FRAUD

Now imagine what would happen without Gatekeeper, if a fraudster used a convincing deepfake to conceal their identity while socially engineering one of your contact center agents. Would they recognize the threat?

The long-term importance of comprehensive voice authentication

This news is yet another reminder for the need for a comprehensive voice authentication and fraud prevention solution that provides multiple layers of security to mitigate attack vectors such as brute force attacks and presentation attacks, be they recordings or synthetic voices. The Gatekeeper team continues to invest in enhancing our AI algorithms to ensure even the most convincing attacks to the human ear can be detected in real time.

As the security community witnessed with the Hong Kong deepfake attack that led to the loss of $26 million earlier this year, the need to protect all voice-based interactions is crucial given the sophistication of modern AI tools in creating synthetic voices and video.

For more information on how Gatekeeper detects synthetic voices, please read my blog post on this topic from last year.

Tags: Voice biometrics, Fraud prevention, Nuance Gatekeeper, Synthetic voice detection, Deepfakes

About Brett Beranek

Brett Beranek is responsible for overseeing the security and biometric line of business at Nuance, a Microsoft company. In this role for the past 12 years, Beranek has brought Nuance to a leadership position in the biometric authentication and biometric fraud prevention space. A thought leader in the field of biometrics, Beranek is a frequent contributor in industry events and the media on the topic of AI technology and it’s use by the fraud community, and how society can mitigate against these evolving threats. Prior to Nuance, he held various leadership positions in the biometrics and security industry. He has earned a Bachelor of Commerce, Information Systems Major, from McGill University as well as an Executive Marketing certificate from Massachusetts Institute of Technology’s Sloan School of Management. Beranek is also a certified Master Fraud Prevention Black Belt professional.

View all posts by Brett Beranek