The NSA’s voice-recognition system raises hard questions for Echo and Google Home

Are Amazon and Google doing enough to keep spies out?

By Russell Brandom@russellbrandom

SHARE ORE

Suppose you're looking for a single person, somewhere in the world. (We'll call him Waldo.) You know who he is, nearly everything about him, but you don't know where he's hiding. How do you find him?

The scale is just too great for anything but a computerized scan. The first chance is facial recognition — scan his face against cameras at airports or photos on social media — although you'll be counting on Waldo walking past a friendly camera and giving it a good view. But his voice could be even better: How long could Waldo go without making a phone call on public lines? And even if he's careful about phone calls, the world is full of microphones — how long before he gets picked up in the background while his friend talks to her Echo?

As it turns out, the NSA had roughly the same idea. In an *Intercept* piece on Friday, reporter Ava Kofman detailed the secret history of the NSA's speaker recognition systems, dating back as far as 2004. One of the programs was a system known as Voice RT, which was able to match speakers to a given voiceprint (essentially solving the Waldo problem), along with generating basic transcriptions. According to classified documents, the system was deployed in 2009 to track the Pakistani army's chief of staff, although officials expressed concern that there were too few voice clips to build a viable model. The same systems scanned voice traffic to more than 100 Iranian delegates' phones when President Mahmoud Ahmadinejad visited New York City in 2007.

We've seen voice recognition systems like this before — most recently with the Coast Guard— but there's never been one as far-reaching as the Voice RT, and it raises difficult new questions about voice recordings. The NSA has always had broad access to US phone infrastructure, something driven home by the early Snowden documents, but the last few years have seen an explosion of voice assistants like the Amazon Echo and Google Home, each of which floods more voice audio into the cloud where it could be vulnerable to NSA interception. Is home assistant data a target for the NSA's voice scanning program? And if so, are Google and Amazon doing enough to protect users?

"ARE GOOGLE AND AMAZON DOING ENOUGH TO PROTECT USERS?"

In previous cases, law enforcement has chiefly been interested in obtaining specific incriminating data picked up by a home assistant. In the Bentonville murder case last year, police sought recordings or transcripts from a specific Echo, hoping the device might have triggered accidentally during a pivotal moment. If that tactic worked consistently, it might be a privacy concern for Echo and Google Home owners — but it almost never does. Devices like the Echo and Google Home only retain data after hearing their wake word ("Okay Google" or "Alexa"), which means all police would get is a list of intentional commands. Security researchers have been trying to break past that wake-word safeguard for years, but so far, they can’t do it without an in-person firmware hack, at which point you might as well just install your own microphone.

But the NSA’s tool would be after a person’s voice instead of any particular words, which would make the wake-word safeguard much less of an issue. If you can get all the voice commands sent back to Google or Amazon servers, you’re guaranteed a full profile of the device owner’s voice, and you might even get an errant houseguest in the background. And because speech-to-text algorithms are still relatively new, both Google and Amazon keep audio files in the cloud as a way to catalog transcription errors. It’s a lot of data, and *The Intercept* is right to think that it would make a tempting target for the NSA.

> ***"“TO THE EXTENT PLATFORMS STORE BIOMETRICS, THEY ARE VULNERABLE TO GOVERNMENT DEMANDS FOR ACCESS AND DISCLOSURE.”"***

When police try to collect recordings from a voice assistant, they have to play by roughly the same warrant rules as your email or Dropbox files — but the NSA might have a way to get around the warrant too. Collecting the data would still require a court order (in the NSA’s case, one approved by the FISA court), but the data wouldn’t necessarily need to be collected. In theory, the NSA could appeal to platforms to scan their own archives, arguing they would be helping to locate a dangerous terrorist. It would be similar to the scans companies already run for child abuse, terrorism or copyright-protected material on their networks, all of which are largely voluntary. If companies complied, the issue could be kept out of conventional courts entirely.

Albert Gidari, director of privacy at the Stanford Center for Internet and Society, says that kind of standoff is an inherent problem when platforms are storing biometric-friendly data. After years of sealed litigation, it’s still unclear how much help the government has a right to compel. “To the extent platforms store biometrics, they are vulnerable to government demands for access and disclosure,” says Gidari. “I think the government could obtain a technical assistance order to facilitate the scan, and under [the technical assistance provision in] FISA, perhaps to build the tool, too.”

> ***"AMAZON “WILL NOT RELEASE CUSTOMER INFORMATION WITHOUT A VALID AND BINDING LEGAL DEMAND PROPERLY SERVED ON US.”"***

We still don’t have any real evidence that those orders are being served. All *TheIntercept* article speaks to is how the program worked within the NSA, and no one at Google or Amazon has ever suggested something like this might be possible. But there’s still good reason to be suspicious: if such order were delivered to a tech company, it would probably come with a gag order preventing them from talking about what they’d done.

So far, there’s been little transparency about how much data agencies are getting from personal voice assistants, if any. Amazon has been noticeably shifty about listing requests for Echo data in its transparency report. Google treats the voice recordings as general user data, and doesn’t break out requests that are specific to Google Home. Reached for comment, an Amazon representative said the company “will not release customer information without a valid and binding legal demand properly served on us.”

The most ominous sign is how much data personal assistants are still retaining. There’s no technical reason to store audio of every request by default, particularly if it poses a privacy risk. If Google and Amazon wanted to decrease the threat, they could stop logging requests under specific users, tying them instead to an anonymous identifier as Siri does. Failing that, they could retain text instead of audio, or even process the speech-to-text conversion on the device itself.

But the Echo and the Home weren’t made with the NSA in mind. Google and Amazon were trying to build useful assistants, and they likely didn’t consider that it could also be a tool of surveillance. Even more, they didn’t consider that a person’s voice might be something they would have to protect. Like ad-targeting and cloud hosting itself, what started as information technology is turning into a system of surveillance and control. What happens next is up to Google, Amazon, and their customers.