Uncovered: 1,000 phrases that incorrectly trigger Alexa, Siri, and Google Assistant

reader comments

120 with 106 posters participating

Share this story

As Alexa, Google Home, Siri, and other voice assistants have become fixtures in millions of homes, privacy advocates have grown concerned that their near-constant listening to nearby conversations could pose more risk than benefit to users. New research suggests the privacy threat may be greater than previously thought.

The findings demonstrate how common it is for dialog in TV shows and other sources to produce false triggers that cause the devices to turn on, sometimes sending nearby sounds to Amazon, Apple, Google, or other manufacturers. In all, researchers uncovered more than 1,000 word sequences—including those from *Game of Thrones*, *Modern Family*, *House of Cards*, and news broadcasts—that incorrectly trigger the devices.

“The devices are intentionally programmed in a somewhat forgiving manner, because they are supposed to be able to understand their humans,” one of the researchers, Dorothea Kolossa, said. “Therefore, they are more likely to start up once too often rather than not at all.”

That which must not be said

Examples of words or word sequences that provide false triggers include

- Alexa: “unacceptable,” “election,” and “a letter”
- Google Home: “OK, cool,” and “Okay, who is reading”
- Siri: “a city” and “hey jerry”
- Microsoft Cortana: “Montana”

The two videos below show a *GoT* character saying “a letter” and *Modern Family* character uttering “hey Jerry” and activating Alexa and Siri, respectively.

[embedded content]
Accidental Trigger #1 – Alexa – Cloud
[embedded content]
Accidental Trigger #3 – Hey Siri – Cloud
In both cases, the phrases activate the device locally, where algorithms analyze the phrases; after mistakenly concluding that these are likely a wake word, the devices then send the audio to remote servers where more robust checking mechanisms also mistake the words for wake terms. In other cases, the words or phrases trick only the local wake word detection but not algorithms in the cloud.

Unacceptable privacy intrusion

When devices wake, the researchers said, they record a portion of what’s said and transmit it to the manufacturer. The audio may then be transcribed and checked by employees in an attempt to improve word recognition. The result: fragments of potentially private conversations can end up in the company logs.

The risk to privacy isn’t solely theoretical. In 2016, law enforcement authorities investigating a murder subpoenaed Amazon for Alexa data transmitted in the moments leading up to

the crime. Last year, The Guardian reported that Apple employees sometimes transcribe sensitive conversations overheard by Siri. They include private discussions between doctors and patients, business deals, seemingly criminal dealings, and sexual encounters.

The research paper, titled "Unacceptable, where is my privacy?," is the product of Lea Schönherr, Maximilian Golla, Jan Wiele, Thorsten Eisenhofer, Dorothea Kolossa, and Thorsten Holz of Ruhr University Bochum and Max Planck Institute for Security and Privacy. In a brief write-up of the findings, they wrote:

> *Our setup was able to identify more than 1,000 sequences that incorrectly trigger smart speakers. For example, we found that depending on the pronunciation, «Alexa» reacts to the words "unacceptable" and "election," while «Google» often triggers to "OK, cool." «Siri» can be fooled by "a city," «Cortana» by "Montana," «Computer» by "Peter," «Amazon» by "and the zone," and «Echo» by "tobacco." See videos with examples of such accidental triggers here.*
>
> *In our paper, we analyze a diverse set of audio sources, explore gender and language biases, and measure the reproducibility of the identified triggers. To better understand accidental triggers, we describe a method to craft them artificially. By reverse-engineering the communication channel of an Amazon Echo, we are able to provide novel insights on how commercial companies deal with such problematic triggers in practice. Finally, we analyze the privacy implications of accidental triggers and discuss potential mechanisms to improve the privacy of smart speakers.*

The researchers analyzed voice assistants from Amazon, Apple, Google, Microsoft, and Deutsche Telekom, as well as three Chinese models by Xiaomi, Baidu, and Tencent. Results

published on Tuesday focused on the first four. Representatives from Amazon, Apple, Google, and Microsoft didn’t immediately respond to a request for comment.

The full paper hasn’t yet been published, and the researchers declined to provide a copy ahead of schedule. The general findings, however, already provide further evidence that voice assistants can intrude on users’ privacy even when people don’t think their devices are listening. For those concerned about the issue, it may make sense to keep voice assistants unplugged, turned off, or blocked from listening except when needed—or to forgo using them at all.

https://arstechnica.com/?p=1688628