Bloomberg

Amazon's Alexa Team Can Access Users' Home Addresses

By Matt Day, Giles Turner, and Natalia Drozdiak
24 avril 2019 à 18:21 UTC+2

- Some members of Alexa Data Services see latitude and longitude
- The team is charged with helping Alexa improve its performance

An Amazon.com Inc. team auditing Alexa users' commands has access to location data and can, in some cases, easily find a customer's home address, according to five employees familiar with the program.

The team, spread across three continents, transcribes, annotates and analyzes a portion of the voice recordings picked up by Alexa. The program, whose existence Bloomberg revealed earlier this month, was set up to help Amazon's digital voice assistant get better at understanding and responding to commands.

Team members with access to Alexa users' geographic coordinates can easily type them into third-party mapping software and find home residences, according to the employees, who signed nondisclosure agreements barring them from speaking publicly about the program.

While there's no indication Amazon employees with access to the data have attempted to track down individual users, two members of the Alexa team expressed concern to Bloomberg that Amazon was granting unnecessarily broad access to customer data that would make it easy to identify a device's owner.

Location data is more sensitive than many other categories of user information, said Lindsey Barrett, a staff attorney and teaching fellow at Georgetown Law's Communications and Technology Clinic.

"Anytime someone is collecting where you are, that means it could go to someone else who could find you when you don't want to be found," she said. Widespread access to location data associated with Alexa user recordings "would set up a big red flag for me."

In an April 10 statement acknowledging the Alexa auditing program, Amazon said "employees do not have direct access to information that can identify the person or account as part of this workflow."

In a new statement responding to this story, Amazon said "access to internal tools is highly controlled, and is only granted to a limited number of employees who require these tools to train and improve the service by processing an extremely small sample of interactions. Our policies strictly prohibit employee access to or use of customer data for any other reason, and we have a zero tolerance policy for abuse of our systems. We regularly audit employee access to internal tools and limit access whenever and wherever possible."

Amazon's Alexa Data Services team, which manages the scads of recordings of human speech and other data that helps train the voice software, numbers in the thousands of employees and contractors, spread across work sites from Boston to Romania and India.

Some of the workers charged with analyzing recordings of Alexa customers use an Amazon tool that displays audio clips alongside data about the device that captured the recording. Much of the information stored by the software, including a device ID and customer identification number, can't be easily linked back to a user.

However, Amazon also collects location data so Alexa can more accurately answer requests, for example suggesting a local restaurant or giving the weather in nearby Ashland, Oregon, instead of distant Ashland, Michigan.

In a demonstration seen by Bloomberg, an Amazon team member pasted a user's coordinates, stored in the system as latitude and longitude, into Google Maps. In less than a minute, the employee had jumped from a recording of a person's Alexa command to what appeared to be an image of their house and corresponding address.

It's unclear how many people have access to that system. Two Amazon employees said they believed the vast majority of workers in the Alexa Data Services group were, until recently, able to use the software.

Sometimes Amazon scoops up data by default. As recently as last year, the first time a customer asked an Echo smart speaker a question related to location, the company often used the device's internet connection to get its approximate location. More recently, the company has started using the shipping address associated with a customer's account as the Echo's default location.

Amazon's location data is not always precise, and it doesn't always refer to the location of an Echo. The Alexa smartphone app prompts users to enter a home address when they set up a smart speaker and also asks for permission to use smartphone location data.

In a list of frequently asked questions about Alexa, Amazon says it uses mobile device location to provide more relevant answers and recommendations, and to enable features like reminders designed to trigger when a user reaches a certain place.

A second internal Amazon software tool, available to a smaller pool of workers who tag transcripts of voice recordings to help Alexa categorize requests, stores more personal data, according to one of the employees.

After punching in a customer ID number, those workers, called annotators and verifiers, can see the home and work addresses and phone numbers customers entered into the Alexa app when they set up the device, the employee said. If a user has chosen to share their contacts with Alexa, their names, numbers and email addresses also appear in the dashboard. That data is in the system so that if a customer says "Send a message to Laura," human reviewers can make sure transcribers wrote the name correctly so that the software learns to pair that request with the Laura in the contact list.

Amazon appears to have been restricting the level of access employees have to the system.

One employee said that, as recently as a year ago, an Amazon dashboard detailing a user's contacts displayed full phone numbers. Now, in that same panel, some digits are obscured.

Amazon further limited access to data after Bloomberg's April 10 report, two of the employees said. Some data associates, who transcribe, annotate and verify audio recordings, arrived for work to find that they no longer had access to software tools they had previously used in their jobs, these people said. As of press time, their access had not been restored.