July 27, 2018

The Honorable Jeff Flake
Chairman, Subcommittee on Privacy, Technology and the Law
United States Senate Judiciary Committee
224 Dirksen Senate Office Building
Washington, D.C. 20510

The Honorable Chris Coons
Ranking Member, Subcommittee on Privacy, Technology and the Law
United States Senate Judiciary Committee
224 Dirksen Senate Office Building
Washington, D.C. 20510

Dear Chairman Flake and Ranking Member Coons,

Thank you for your June 11, 2018 letter regarding Amazon's Echo device and Alexa voice service.

Alexa is a cloud-based voice service that lets customers play music, ask questions, make calls, send and receive messages, get information, news, sports scores, weather, and more. Alexa is available through a wide range of products, including Amazon's Echo family of devices, other Amazon products such as our Fire TV and Fire tablet devices, and devices developed by third party manufacturers participating in our Alexa Voice Service program. Alexa operates in a similar manner across the range of products on which it is available, although customers access Alexa differently based on the type of Alexa-enabled product they use.

Getting privacy right takes careful attention. From early-stage development, we built privacy deeply into the hardware and service by design, and with Alexa and Amazon's Alexa-enabled products we strive to put the control with our customers.

On our Echo family of devices, customers speak to Alexa by saying the "wake word" (Alexa, Amazon, Echo, or Computer) or, on some Echo devices, by pressing the action button on the top of the device. Echo devices use "on-device keyword spotting" technology designed to detect the wake word. The technology analyzes acoustic patterns to detect when the wake word has been spoken using a short, on-device buffer that is continuously overwritten. This on-device buffer exists only in temporary memory (RAM); no audio is ever recorded to any on-device storage. The device does not stream audio to the cloud unless the wake word is detected or the action button is pressed. The user experience also provides customers with a clear indication of when audio is being streamed. When the wake word is detected or the action button is pressed, a visual indicator appears on the device to clearly indicate to the customer that it is streaming audio to the Amazon cloud (e.g., a blue light ring on the Echo device and a blue bar on the Echo Show's screen). We also offer a setting where customers can choose to hear an audible tone when their Echo device begins and ends streaming audio to the cloud.

Next, the audio is streamed to the Amazon cloud, where our systems for "automatic speech recognition" (converting audio to text) and "natural language understanding" (interpreting the meaning of text) then determine the meaning of the customer's request so that Alexa can respond appropriately. Amazon encrypts all communication between Echo devices and Amazon's servers, and stores all customer data securely on our servers.

We also give customers control of their voice recordings in the cloud. Not only are customers able to see and play back the voice recordings associated with their account, customers can also delete those voice recordings one-by-one or all at once.

Echo devices come with a microphone off button that enables customers to manually control when their device's microphone is on. When the button is pressed to turn the microphones off, the microphones are electrically disconnected and a dedicated red LED is illuminated to indicate the microphones are off. As an additional safeguard, we designed the circuitry of Echo devices so that power can only be provided either to this dedicated red LED *or* to the device microphones, not to both at the same time. As a result, if the dedicated red LED is illuminated, the microphones are off and cannot stream audio to the cloud.

"Far-field voice recognition," where the user may be far from the microphone, poses substantial technical challenges – the environment may be noisy, different homes and different locations within homes can have widely varying acoustic characteristics, and customers have a diverse range of speech patterns, vocabulary, accents, and personal preferences. We have a large team of scientists and engineers dedicated to these challenges. We have designed our wake word detection, automatic speech recognition, and natural language understanding technologies to continuously get better, and minimize false responses, and as customers use their devices we constantly train and optimize our technologies. An essential part of that optimization is minimizing the amount of background noise and other speech not intended for Alexa that may be streamed to our cloud. Processing such extraneous audio is costly, provides no value to Amazon, and is detrimental to our customers' Alexa experience.

Echo devices also use a technology called "beamforming" to ensure high-accuracy audio processing. Beamforming isolates the customer's voice from background noise by emphasizing audio from the direction of the customer and suppressing audio from other directions. Customers see beamforming take place on an Echo device with a visual cue—the lightest blue color on the light indicator points to the source of the audio.

Customers talk to Alexa billions of times per month. In rare cases, Echo will wake up due to a word in background conversation sounding like Alexa (or the alternative wake word selected by the customer). But the service is designed to stop processing the audio if it determines the speech is not intended for Alexa.

Customer trust is of the utmost importance to our continued success, and we take that responsibility most seriously. The answers to your specific questions are as follows:

1. **Please indicate the number of complaints you have received from consumers reporting that an Amazon Echo device has improperly interpreted a command.**

Customers talk to Alexa billions of times per month. At times, factors like background noise, the location of the Echo device in the room (such as in a corner or next to a window), or the customer's speech patterns, vocabulary, or accent may cause Alexa to interpret a command improperly (e.g., Alexa mistakenly setting a timer in a busy kitchen for 13 minutes instead of 30 minutes). The customer normally hears immediately if Alexa has misunderstood a command based on Alexa's response and can choose to repeat or restate it. In particular, before Alexa takes an important action, like sending a message, Alexa confirms the action with the customer. As a result, events like the one described in your letter are exceedingly rare. In that case, the couple's Echo device woke up due to a word in background conversation sounding like the wake word, Alexa heard the subsequent conversation as a "send message" request and asked aloud "To whom?." The background conversation was interpreted as a name in the customers contact list and Alexa again asked aloud, "[contact name], right?" and then interpreted background conversation as "right." We are aware of fewer than ten customer complaints involving Alexa misinterpreting a conversation as a request to send a message.

2. **Regarding the technical design elements of Echo devices and Alexa software:**
    a. **Is an Echo device designed to send a user's voice data from the device to an Amazon-controlled server?**
        i. **When and how frequently does the device send voice data to Amazon?**
        ii. **How long does Amazon store and retain voice data?**
        iii. **Please explain any and all technical design elements that Amazon has created to anonymize user data collected and transmitted to Amazon-controlled servers by Echo devices, including how each of those elements is designed to protect consumers' privacy.**

Echo devices use on-device keyword spotting to detect when a customer says their selected wake word (Alexa, Amazon, Echo, or Computer) – the device does not stream audio to the cloud unless the wake word is detected or the action button is pressed. Once the customer activates Alexa with the wake word or presses the action button, a stream to the cloud is opened to begin processing the request. If Alexa is activated using the wake word, the audio streamed to the cloud includes a fraction of a second of audio before the wake word. The first step that occurs when the stream reaches the cloud is that the audio is reanalyzed using the more powerful processing capabilities of the cloud to verify the wake word was spoken. If this cloud software verification is unable to confirm the wake word was spoken, the service stops processing the audio and the audio stream to the cloud is shut off. If the wake word is verified (or if Alexa was activated using the action button), our systems for speech recognition and natural language understanding then determine the meaning of the customer's request so that Alexa can respond appropriately. As our speech recognition system analyzes the audio stream, the system continually attempts to determine when the customer's request to Alexa has ended and then immediately ends the audio stream. Because Alexa runs in the cloud, the service is always getting smarter—adding new features and improving accuracy. The more a customer uses Alexa, the more the service adapts to their speech patterns, vocabulary, and personal preferences. All audio streamed to the Amazon cloud is encrypted in transit and securely stored in the Amazon cloud.

We associate a customer's requests with their Amazon account to allow them to access other Amazon services (e.g., so they can ask Alexa to read their Kindle books and play audiobooks from Audible), to make the recordings available to the customer to view, and to provide the customer with a more personalized experience. For example, keeping track of the songs a customer has listened to helps Alexa choose what songs to play when the customer says, "Alexa, play music."

To allow customers to review their voice recordings, we retain the recordings until the customer chooses to delete them. Customers can review and listen to the voice recordings associated with their account in the Alexa app, and delete them individually or all at once, which also deletes them from our server.

b. Before being activated by a wake word, is an Echo device designed to be always listening for a wake word?

Echo devices use on-device keyword spotting technology designed to detect the wake word and only the wake word. Thus the device is not always listening, but instead is always available to be engaged via the wake word, which performs as the audio equivalent of a physical "on" switch. The technology analyzes acoustic patterns to detect when the wake word has been spoken using a short, on-device buffer that is continuously overwritten. This on-device buffer exists only in temporary memory (RAM); no audio is ever recorded to any on-device storage. Customers can turn the microphones off by pushing the microphone button on the top of the device, which will turn on a red LED to indicate that the mics are off (as described above, when this LED is on, the circuity of Echo devices is designed to disconnect power to the microphones so the microphones cannot record and stream audio to the cloud).

c. After being activated by a wake word, how long is the Echo device designed to listen for a command to perform?

Once activated by the wake word or the action button, the Echo device opens an audio stream to the cloud and sends the request to Alexa to respond accordingly. The audio stream closes immediately once our automatic speech recognition system determines the customer has stopped speaking the request. A blue light illuminates on the Echo device to indicate when audio is being streamed to the cloud, and customers can also enable an audible tone that plays when their Echo device begins and ends streaming audio to the cloud.

d. Is the Echo designed to record background conversations while it listens for a command?

i. If so, for how long?

The Echo device is designed not to record background conversations. As we have described, Echo devices use on-device keyword spotting technology designed to detect the wake word and only the wake word; the on-device buffer merely serves as a technological filter for the wake word to turn on Alexa functionality, and that buffer is continuously overwritten.

e. Is the Echo designed to record speech after it has identified a command?

i. If so, for how long?

No. Once activated by the wake word, the Echo device opens an audio stream to the cloud and sends the request to Alexa to respond accordingly. As Alexa analyzes the audio stream, the service continually attempts to determine when the customer's request to Alexa has ended and then immediately ends the audio stream. In some circumstances, in response to customer commands, the stream will open again for a customer to follow up, including if the customer's request involves multiple interactions (e.g., if the customer asks Alexa to set a timer without indicating the duration, Alexa will ask for the length of the timer) or the customer has elected to enable a setting that allows Alexa to respond to a series of requests in rapid succession without the customer needing to repeat the wake word for each request.

We have also implemented technical measures to minimize the background noise Echo devices stream to the cloud. Processing background noise provides no benefit to Amazon, is costly, and is detrimental to our customers' Alexa experience, so we have designed our Echo devices with a goal of avoiding it. Echo devices use a technology called beamforming to ensure high-accuracy audio processing. Beamforming isolates the customer's voice from background noise by emphasizing audio from the direction of the customer and suppressing audio from other directions. Customers see beamforming take place on an Echo device with a visual cue—the lightest blue color on the light indicator points to the source of the audio. As discussed above, Alexa is designed to continually attempt to determine when the customer's request to Alexa has ended and then immediately end the audio stream.

f. Is the Echo designed to store pre-command or command recordings, either on the device or in a remote location like the cloud?

i. If so, for how long?

No. As described above, the Echo device is not designed to store audio on the device.

Once the Echo device detects the wake word, it streams the audio to the cloud, including a fraction of a second of audio before the wake word. As discussed above, customers' audio recordings are then stored securely in the cloud until a customer chooses to delete an individual audio recording, or all of the audio recordings associated with their account.

g. Is the Echo designed to allow the consumer to delete stored recordings?

i. If so, please describe what steps the consumer must take to do so.

Yes. Customers can view, play back, and delete specific voice recordings associated with their account by going to History in Settings in the Alexa App and drilling down for a specific entry. Or, the customer can delete all voice recordings associated with their account by selecting their Echo device at the Manage Your Content and Devices page at www.amazon.com/mycd. Customers can also contact customer service to request deletion of their voice recordings.

h. Is the Alexa software designed to use artificial intelligence or machine learning to better identify and understand a consumer's commands?

i. If so, does Alexa rely on a comprehensive collection of stored recordings associated with that device to do so?

Alexa is designed to get smarter every day—this is accomplished through Amazon's cloud services, including machine learning software. The more a customer uses Alexa, the more the service adapts to the customer's speech patterns, vocabulary, and personal preferences. When the customer says the wake word, their subsequent request to Alexa is processed and stored in the cloud to respond to the customer's request and to improve the customer's experience and our services, including training our speech recognition and natural language understanding systems so Alexa can better understand customers' requests. The more data we use to train these systems, the better Alexa works, and training Alexa with voice recordings from a diverse range of customers helps ensure Alexa works well for everyone.

i. After Alexa has performed a command, is the device designed to power off or deactivate its listening feature?

i. If so, how long after the command has been performed does the device remain powered on and activated?

Yes. Please see answers to questions 2(d) and 2(e).

j. Is the Alexa software able to be modified by a third-party software developer?

i. If so, in what ways can a third-party developer currently modify the Alexa software?

ii. Does Amazon plan to limit a third-party developer's ability to modify the software?

iii. Does Amazon plan to employ a third-party developer to modify or improve certain software design choices?

iv. What are some of the software design choices that Amazon does not intend to modify?

No. We do not make Alexa, our cloud-based voice service, or our Echo device software available to third-party companies to modify. We work with third-party software developers in other capacities, but those third parties do not and cannot modify our Alexa or Echo software. For example, our Alexa Voice Service program provides third-party companies with software, hardware kits, and other tools to allow those third parties to enable access to Alexa in products they develop. Our Alexa "skills" program allows third-party developers (from brands like Jeopardy! and OpenTable) to offer the equivalent of apps that add to the features and capabilities we offer through Alexa. These skills may be made available through a combination of online services and device software provided by third-party skill developers that supplement, but do not modify, our software.

3. Please describe any and all purposes for which Amazon uses, stores, and retains consumer information, including voice data, collected and transmitted by an Echo device.

We use, store, and retain data collected and transmitted by an Echo device to deliver and improve our services, and only as described in our privacy notice. This includes responding to customers' requests, allowing customers to review and listen to their voice recordings, and training Alexa to better understand and interpret customers' requests.

4. Please describe any and all policies Amazon follows to ensure that customer information, including voice data, collected and transmitted by an Echo device is protected from misuse or abuse.

Amazon takes customer security seriously. We have developed technical measures, and we have full teams, dedicated to ensuring the safety and security of our products. These measures include rigorous security reviews, secure software development requirements, and encryption of communication between the Echo device and Amazon servers. Furthermore, our information security policies are based on industry standards and include policies on data classification and handling. Alexa voice recordings are classified as highly sensitive data, are subject to specialized handling requirements, and access to them is restricted only to select authorized systems and individuals. We require that individuals with access to our network complete annual security training, and we regularly conduct awareness campaigns throughout the year. Also, as discussed in response to question 6 below, we limit the amount of customers' personal information shared with skill developers (e.g., we do not share customers' voice recordings with skill developers) and we share it in a way that is transparent to customers.

5. What steps is Amazon taking to minimize the risk that consumer information, including voice data, is not misinterpreted or misused by Echo devices?

Customers talk to Alexa billions of times per month and, as discussed in response to question 1, events like the one you described in your letter are exceedingly rare. However, we take privacy very seriously at Amazon, and we continue to improve wake word accuracy and minimize false responses.

Far-field voice recognition raises a number of technical challenges – the environment may be noisy, different homes and different locations within homes can have widely varying acoustic characteristics, and customers have a diverse range of speech patterns, vocabulary, accents, and personal preferences. We have a large team of scientists and engineers dedicated to these challenges. We have designed our wake word detection, automatic speech recognition, and natural language understanding technologies to continuously get better, and minimize false responses– as customers use their devices, we use that data to further train and optimize those technologies. Our goal is to minimize the amount of background noise and other speech not intended for Alexa streamed to our cloud. Processing that audio is costly, provides no value to Amazon, and is detrimental to our customers' Alexa experience.

6. What steps is Amazon taking to ensure consumer information, including voice data, is not shared without consumers' consent?

Getting privacy right takes careful attention. We handle any personal information we receive in accordance with Amazon's Privacy Notice. From early-stage development, we built privacy deeply into the hardware and service by design, and with Alexa and Amazon's Alexa-enabled products we strive to put the control with our customers. For example, we've designed our Alexa skill program to share only a limited amount of customers' personal information with skill developers and to share that information in a way that is transparent to customers. For example, when a customer makes a request to a skill, the skill developer receives the content of the customer's request in text form so the developer can respond appropriately, but they do not receive a copy of the voice recording.

Some skills may request access to personal information the customer has provided Amazon to provide a service to the customer (e.g., an on-demand car service like Lyft may need access to the address the customer has set for their Echo device so they can send a car to that location). In those cases, we use a permission framework similar to the one used by mobile devices, which requires customers to grant permission to share that personal information with the skill developer, and we would only grant access to that information after the customer grants permission. As a result, customers are involved anytime we share their personal information with a skill developer.

Thank you again for your interest in Amazon's Echo device and Alexa voice service.

Sincerely,

Brian Huseman
Vice President, Public Policy