Google is reducing how much audio it saves for human review

Google is making some changes to its audio data retention policies in the coming months. Most importantly for those concerned about humans reviewers listening to you, it plans on asking every user to re-affirm their choice to opt-in to that program — which is "paused globally" pending an EU investigation.

The company is also making other changes, including a new sensitivity option for "Hey Google" hotword detection, so that users who want to can make it less likely that their smart speakers will pick up unintended audio.

The changes come in the wake of a summer where every major smart assistant is under renewed scrutiny over how data is stored, for how long, and who gets to listen to it. Alexa, Siri, and Google Assistant were all storing user utterances and having human reviewers listen to that audio to rate how accurately the assistants transcribed them. That data often included audio that wasn't intended for the smart speakers at all, but was nevertheless accidentally recorded.

Amazon, Google, and Apple have all had to make data privacy changes

Google, Amazon, and Apple all reacted differently to their individual scandals. Amazon, for example, added clearer privacy controls inside the Alexa app and made it possible to ask Alexa to delete your data. Apple issued a rare privacy-related apology and updated its policy to strengthen controls and prevent third-party contractors from hearing user audio.

Google says that having your voice recordings stored and reviewed was always opt-in and that voice recordings that humans reviewed were also always disassociated with user accounts. However, the so-called “Voice & Audio Activity (VAA)” setting wasn’t very clear about what was happening when you agreed. Going forward, Google will explicitly mention human review for the VAA setting and, just as importantly, present that new, clearer screen to *all* Google Assistant users so they can choose if they want to opt-in or not.

The company also says that it’s going to “vastly reduce the amount of audio data we store,” promising to “delete the vast majority of audio data associated with your account that’s older than a few months” for people who have opted in to VAA.

Google also made a vague promise to add “an extra layer of privacy filters” to the audio transcription process, which we are told involves filtering out certain classes of audio data. It’s not entirely clear what that means, however Google does say it intends on being more aggressive at automatically deleting accidental recordings.

The Google Assistant has never met a problem it hasn’t tried to fix with a new setting

Google has a penchant for solving all problems with more settings — especially with the Google Assistant — and it’s doing it yet again now. Soon it will add a hotword sensitivity option, which means you’ll be able to choose how clearly you have to enunciate “Hey Google” in order to turn on the smart speaker. If you are worried about accidental recordings, you’ll be able to turn it up, if you’re not, you can set it to be a little more forgiving. (Personally, I’d vote for a hotword that’s easier to say, but that’s another discussion.)

Before now, the Google Assistant seems to do the best job of giving you full use of everything the assistant can do without requiring you to opt-in to voice recordings. Specifically, the only functionality that Google says will be degraded if you

opt out is hotword detection. After Apple changed its policy, it no longer collects audio data by default.

Google and Amazon also do a good job of giving you web and app portals with controls that show you what data is being collected and options for deleting it. Apple has no such portal, though it collects vastly less data in the first place.

Google hasn’t provided clear dates for when these new policies will go through, though it did mention some of it will go into effect “later this year.”

https://www.theverge.com/2019/9/23/20878710/google-assistant-audio-recording-policy-hotword-human-review