

August 26th 2019

Automatic speech assistant systems - Implementation of data protection measures by Google and other providers required

Following media reports on the practice of transcribing audio recordings by the Google Assistant Google announced that it has stopped this practice for at least three months. Last week the Hamburg Commissioner for Data Protection and Freedom of Information (HmbBfDI) met with Google representatives to clarify the legal requirements to be implemented before the practice under scrutiny is resumed.

Johannes Caspar, the Hamburg Commissioner for Data Protection and Freedom of Information: "We have made clear to Google representatives that essential data protection requirements for the operation of the Google Assistant are currently not met. This applies not only to the practice of transcription, but also to the processing of audio data generated by the operation of the speech assistance system in general."

In particular, the following requirements shall be addressed:

- As long as the transcription and evaluation of audio recordings of speech assistant systems by humans do not comply with the GDPR standards, this practice won't be continued.
- The legal basis for the storing of audio recordings is to obtain an informed consent of the user (opt-in). This applies to regular operation, even without transcription and evaluation of malfunctions by humans.
- Speech assistance systems are incorrectly activated in an undisclosed percentage of cases. Audio recordings made without the knowledge or intent of users pose a high risk to the privacy of users and other individuals such as visitors or children. Transparent information about the risk of false activation is therefore a central prerequisite for the processing of audio data.
- The transcription of audio recordings by humans reinforces the impact on the personal rights of users. The evaluation of audio snippets by contractors or employees to improve speech assistance systems without additional informed consent to this practice violates the privacy rights and freedoms of users.
- Users must be informed that the data protection rights and freedoms of other persons may be adversely affected by the use of speech assistant systems. This is particularly important when considering the possibility that audio recordings of non-users may be incorrectly processed. The use of technologies such as voice recognition can protect the rights of non-users, in particular by preventing the storage of audio recordings of their voices.

Google agreed to keep the HmbBfDI informed of future measures and changes - as prerequisites for lawful, fair and transparent data processing of audio recordings.

Johannes Caspar: "Our discussions have shown that Google is willing to implement changes before resuming the practice of transcription in order to improve machine learning-based speech assistant systems. This must then be checked. For the time being, no supervisory measures will be taken under the urgency procedure. If it turns out that transcription is resumed contrary to the legal requirements of the GDPR, urgent measures can be taken at any time to protect the privacy rights of the users."

It should also be noted that the discussion on the protection of rights and freedoms of data subjects concerns not only Google but all other providers of speech assistant services. Globally operating companies such as Apple, Amazon and Microsoft - for which the HmbBfDI has no competence to issue urgent regulatory measures - are also prompted to meet the legal requirements by GDPR. In particular, this also applies to Facebook Inc. Manual evaluation was used in Facebook Messenger to optimize the transcription function. In so far, not only human-to-machine communication, but also human-to-human communication were analysed. This is currently the subject of a separate investigation."

Press contact:

Martin Schemm

Phone: +49 40 428 54-4044

E-mail: presse@datenschutz.hamburg.de