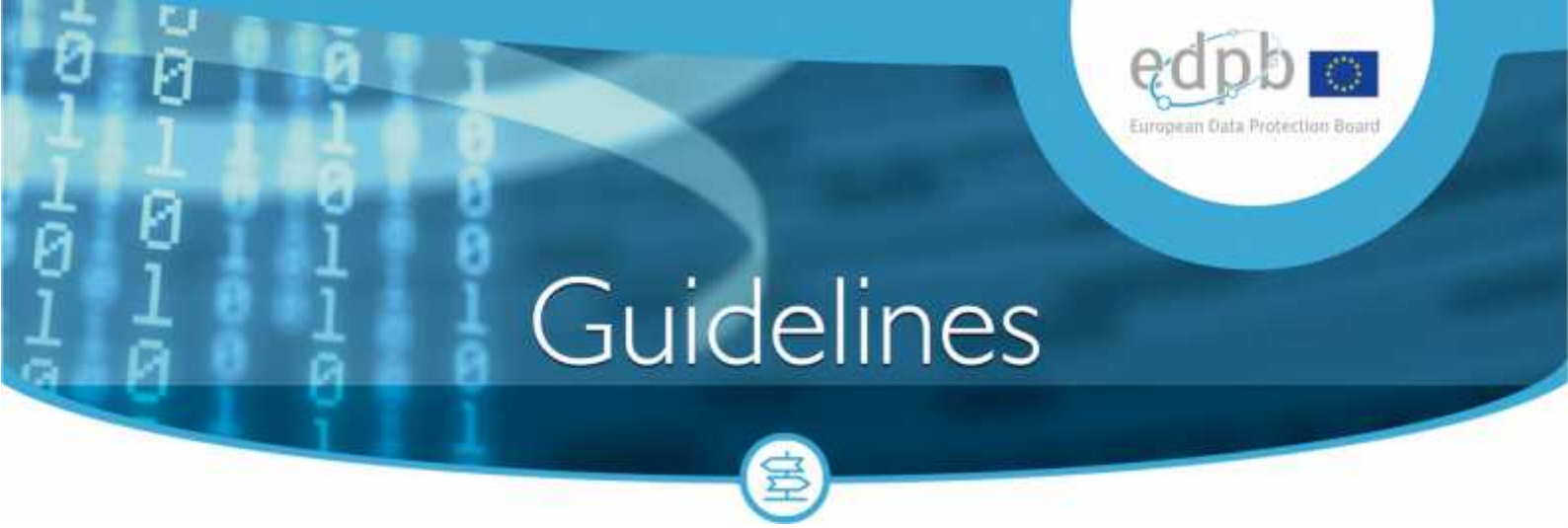

Guidelines 02/2021 on Virtual Voice Assistants

Version 1.0

Adopted on 9 March 2021

EXECUTIVE SUMMARY

A virtual voice assistant (VVA) is a service that understands voice commands and executes them or mediates with other IT systems if needed. VVAs are currently available on most smartphones and tablets, traditional computers, and, in the recent years, even standalone devices like smart speakers.

VVAs act as interface between users and their computing devices and online services such as search engines or online shops. Due to their role, VVAs have access to a huge amount of personal data including all users' commands (e.g. browsing or search history) and answers (e.g. appointments in the agenda).

The vast majority of VVA services have been designed by few VVA designers. However, VVAs can work jointly with applications programmed by third parties (VVA application developers) to provide more sophisticated commands.

To run properly, a VVA needs a terminal device provided with microphones and speakers. The device stores voice and other data that current VVAs transfer to remote VVA servers.

Data controllers providing VVA services and their processors have therefore to consider both the GDPR[1] and the e-Privacy Directive[2].

These guidelines identify some of the most relevant compliance challenges and provide recommendations to relevant stakeholders on how to address them.

Data controllers providing VVA services through screenless terminal devices must still inform users according to the GDPR when setting up the VVA or installing, or using a VVA app for the first time. Consequently, we recommend to VVA providers/designers and developers to develop voice-based interfaces to facilitate the mandatory information.

Currently, all VVAs require at least one user to register in the service. Following the obligation of data protection by design and by default, VVA providers/designers and developers should consider the necessity of having a registered user for each of their functionalities.

The user account employed by many VVA designers bundle the VVA service with other services such as email or video streaming. The EDPB considers that data controllers should refrain from such practices as they involve the use of lengthy and complex privacy policies that would not comply with the GDPR's transparency principle.

The guidelines consider four of the most common purposes for which VVAs process personal data: executing requests, improving the VVA machine learning model, biometric identification and profiling for personalized content or advertising.

Insofar the VVA data is processed in order to execute the user's requests, i.e. as strictly necessary in order to provide a service requested by the user, data controllers are exempted from the requirement of prior consent under Article 5(3) e-Privacy Directive. Conversely, such consent as required by Article

[1] Regulation 2016/679/EU of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC], (hereinafter "GDPR")

[2] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) as amended by Directive 2006/24/EC and Directive 2009/136/EC (hereinafter "e-Privacy Directive").

5(3) e-Privacy Directive would be necessary for the storing or gaining of access to information for any purpose other than executing users' request.

Some VVA services retain personal data until their users require their deletion. This is not in line with the storage limitation principle. VVAs should store data for no longer than is necessary for the purposes for which the personal data are processed.

If a data controller becomes aware (e.g. due to quality review processes) of the accidental collection of personal data, they should verify that there is a valid legal basis for each purpose of processing of such data. Otherwise, the accidentally collected data should be deleted.

VVAs may process data of multiple data subjects. VVA providers/designers should therefore implement access control mechanisms to ensure personal data confidentiality, integrity and availability. However, some traditional access control mechanisms such as passwords are not fit for the VVA context since they would have to by spoken aloud. The guidelines provide some considerations in this regard, including a section specific to the processing special categories of data for biometric identification.

VVA providers/designers should consider that when collecting user's voice, the recording might contain other individuals' voice or data such as background noise that is not necessary for the service. VVA designers should therefore consider technologies filtering the unnecessary data and ensuring that only the user's voice is recorded.

When evaluating the need for a Data Protection Impact Assessment (DPIA), the EDPB considers that it is very likely that VVA services fall into the categories and conditions identified as requiring a DPIA.

Data controllers providing VVA services should ensure users can exercise their data subject rights using easy-to-follow voice commands. VVA providers/designers, as well as app developers should at the end of the process inform users that their rights have been duly factored, by voice or by providing a writing notification to the user's mobile, account or any other mean chosen by the user.

Table of contents

The European Data Protection Board

Having regard to [Article 70 (1j) and (1e) of the Regulation 2016/679/EU of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC], (hereinafter "GDPR"),

Having regard to the EEA Agreement and in particular to Annex XI and Protocol 37 thereof, as amended by the Decision of the EEA joint Committee No 154/2018 of 6 July 2018[3],

Having regard to Article 12 and Article 22 of its Rules of Procedure,

HAS ADOPTED THE FOLLOWING GUIDELINES

1 GENERAL

1. Recent technological advances have greatly increased the accuracy and popularity of virtual voice assistants (VVA). Among other devices, VVAs have been integrated in smartphones, connected vehicles, smart speakers and smart TVs. This integration has given the VVAs access to information of an intimate nature that could, if not properly managed, harm the individuals' rights to data protection and privacy. Consequently, VVAs and the devices integrating them have been under the scrutiny of different data protection authorities.

2. There are several advantages to using speech-based interactions such as: the naturalness of the interaction which does not involve specific learning from the users, the speed of execution of the command and the extension of the field of action which can allow faster access to information. However, relying on speech also brings in difficulties in interpreting the message correctly: variability of the audio signal between different speakers, acoustic environment, ambiguity of the language, etc.

3. In practice, the fluidity or simplification of tasks remains the primary motivation for equipping oneself with VVAs. This may involve, for example, placing/answering a call, setting a timer, etc., especially when the users have their hands unavailable. Home automation is the major application put forward by the designers of VVAs. By proposing to simplify the execution of tasks (turning on the light, adjusting the heating, lowering the shutters, etc.) and to centralize them through a single tool that can be easily activated remotely, they fit into the discourse as a domestic facilitator. In addition to personal or domestic use, voice commands can be of interest in professional environments where it is difficult to handle computer tools and use written commands (e.g. manufacturing work).

4. In theory, the main beneficiaries of the voice interface could be people with disabilities or dependency for whom the use of traditional interfaces is problematic. Virtual voice assistance can provide easier access to information and computer resources and thus promote inclusive logics as the use of the voice makes it possible to overcome the

[3] References to "Member States" made throughout this document should be understood as references to "EEA Member States".

difficulties associated with the written word, which can be found among certain classes of users.

5. Finally, health is also an area where there are many cases of use for conversational agents, vocal or not. For instance, during the Covid-19 pandemic, various callbots were deployed to offer a pre-diagnosis to calling users. In the long term some anticipate that the entire patient care process could be impacted by human/assistant interactions: not only for well-being and prevention, but also for treatment and support.

6. There are currently more than 3 billion smartphones and all of them have integrated VVAs, most of them switched on by default. Some of the most widespread operating systems in personal computers and laptops also integrate VVAs. The recent rise of smart speakers (147 million were sold in 2019[4]) is bringing VVAs to millions of homes and offices. However, current VVA designs do not offer by default authentication or access control mechanisms.

7. This document seeks to provide guidance as to the application of the GDPR in the context of the VVAs.

2 TECHNOLOGY BACKGROUND

2.1 Basic characteristics of Virtual Voice Assistants

8. A VVA can be defined as a software application that provides capabilities for oral dialogue with a user in natural language.

9. Natural language has a semantics specific to human language. Depending on the characteristics of the language and the diversity of the lexicon, the same instruction can be formulated in multiple ways, whereas some commands may seem similar but relate to two different objects. Inference mechanisms are then frequently used to resolve these ambiguities, for example, depending on what has been said previously, the time when the instruction was given, the place, the person's interests, etc.

10. A VVA can be broken down into modules allowing to perform different tasks: sound capture and restitution, automatic speech transcription (speech to text), automatic language processing, dialogue strategies, access to ontologies (data sets and structured concepts related to a given domain) and external knowledge sources, language generation, voice synthesis (text to speech), etc. Concretely, the assistant should allow interaction in order to perform actions (e.g. "turn on the radio", "turn off the light") or to access knowledge (e.g. "what will the weather be like tomorrow?", "is the 7:43 a.m. train running?"). It thus plays the role of intermediary and orchestrator who is supposed to facilitate the accomplishment of the user's tasks.

11. In practice, a VVA is not a smart speaker but a smart speaker can be equipped with a voice assistant. It is common to confuse both of them, however, the latter is only a material incarnation of the former. A VVA can be deployed in a smartphone, a smart speaker, a connected watch, a vehicle, a household appliance, etc.

[4] For example, see a press release of 1 August 2019 by the Hamburg Data Protection and Information authority: https://datenschutz-hamburg.de/pressemitteilungen/2019/08/2019-08-01-google-assistant

12. The organization of the underlying data processing may involve multiple information flow patterns. It is possible to isolate three main entities:

    The physical instance: the hardware element in which the assistant is embodied (smartphone, speaker, smart TV, etc.) and which carries microphones, speakers and network and computing capacities (more or less developed depending on the case).

    The software instance: the part implementing the human-machine interaction strictly speaking and which integrates the modules for automatic speech recognition, natural language processing, dialogue and speech synthesis. This can be operated directly within the physical equipment, but in many cases is performed remotely.

    The resources: external data such as content databases, ontologies or business applications that provide knowledge (e.g. "tell the time on the West Coast of the United States", "read my emails") or enable the requested action to be carried out in a concrete way (e.g. "increase the temperature by 1.5°C").

13. VVAs allow the installation of third party components or apps that expand their core functionalities. Each VVA name the components differently but all involve the exchange of users' personal data between the VVA designer and the app developer.

14. Although most VVAs do not share the voice snippet with the app developers, these actors still process personal data. Moreover, depending on the nature of the functionality provided, the app developer receives intentions and slots which could include sensitive information like health data.

2.2 Actors in the VVA ecosystem

15. A VVA may involve a great number of actors and intermediaries throughout the execution chain. In practice, up to five different actors can be identified. Depending on business models and technological choices, some actors may however take on several combinations of roles, for example, designer and integrator or designer and application developer:

    a. **The VVA provider (or designer):** responsible for the development of the VVA, designs and defines its possibilities and default functionalities: activation modalities, choice of architecture, data access, record management, hardware specifications, etc.

    b. **The VVA application developer:** as for mobile applications, creates applications extending the VVA's default functionalities. To do this, it is necessary to respect the development constraints imposed by the designer.

    c. **The integrator:** manufacturer of connected objects, who wishes to equip them with a VVA. It should respect the requirements defined by the designer.

    d. **The owner:** in charge of physical spaces receiving people (accommodation places, professional environments, rental vehicles, etc.) he/she wishes to provide a VVA to his/her audience (possibly with dedicated applications).

    e. **The user:** final link in the VVA value chain, who can use it on various devices (speaker, TV, smartphone, watch, etc.) depending on how and where the VVA has been deployed and set up.

2.3 Step-by-step description

16. In order for a VVA to carry out an action or to access information, a succession of tasks is carried out:

1) Deployed within a piece of equipment (smartphone, loudspeaker, vehicle), the VVA is on standby. To be precise, it is constantly listening. However, until a specific wake-up expression has been detected no audio is transmitted out of the device receiving the voice and no other operation than wake-up expression detection is performed. For this purpose a buffer of a few seconds is used (see following section for more details).

2) The user says the wake-up expression and the VVA locally compares the audio with the wake-up expression. If they match, the VVA opens a listening channel and the audio content is immediately transmitted.

3) In many cases, if the processing of the command is done remotely, a second check of the keyword pronunciation is done on the server side to limit unwanted activations.

4) The user states his request that is transmitted on the fly to the VVA provider. The sequence of speech spoken is then automatically transcribed (speech to text).

5) Using natural language processing (NLP) technologies, the command is interpreted. The intentions of the message are extracted and information variables (slots) are identified. A dialogue manager is then used to specify the interaction scenario to be implemented with the user by providing the appropriate response scheme.

6) If the command involves a functionality provided by a third party app (skill, action, shortcut, etc.), the VVA provider sends to the app developer the intentions and information variables (slots) of the message.

7) A response adapted to the user's request is identified – at least supposedly, the answer "I don't have the answer to your question" being an adapted response in the case the VVA was not able to correctly interpret the request. If necessary, remote resources are used: publicly accessible knowledge databases (online encyclopaedia, etc.) or by authentication (bank account, music application, customer account for online purchase, etc.) and the information variables (slots) are filled with the recovered knowledge.

8) An answer phrase is created and/or an action is identified (lowering the blinds, raising the temperature, playing a piece of music, answering a question, etc.). The sentence is synthesized (text to speech) and/or the action to be performed is sent to the equipment executed.

9) The VVA returns to standby.

Please note that while currently most voice related processing is performed in remote servers, some VVA providers are developing systems that could perform part of this processing locally[5].

[5] This has been reported, for example, here: https://www.amazon.science/blog/alexas-new-speech-recognition-abilities-showcased-at-interspeech.

2.4 Wake-up expressions

17. In order to be used, a VVA should be "awake". This means that the assistant switches to an active listening mode in order to receive orders and commands from its user. While this wake-up can also sometimes be achieved by a physical action (e.g. by pressing a button, pressing the smart speaker, etc.), almost all VVAs on the market are based on the detection of a wake-up expression or word to switch to active listening mode (also known as activation word or wake-up word / hot word).

18. To do this, the assistant relies on the use of the microphone and slight computational capabilities to detect whether the keyword has been spoken. This analysis, which takes place continuously from the moment the VVA is on, is carried out exclusively locally. Only when the keyword has been recognised are the audio recordings processed for interpretation and execution of the command, which in many cases means sending them to remote servers via the Internet. Keyword detection is based on machine learning techniques. The major challenge in using such methods is that the detection is probabilistic. Thus, for each word or expression pronounced, the system provides a confidence score as to whether the keyword has actually been pronounced. If this score turns out to be higher than a predefined threshold value, this is considered to be the case. Such a system is therefore not free of errors: in some cases activation may not be detected even though the keyword has been said (false rejection) and in other cases activation may be detected even though the user has not said the keyword (false acceptance).

19. In practice, an acceptable compromise should be found between these two types of errors to define the threshold value. However, since the consequence of a false detection of the keyword might be the sending of audio recordings, unexpected and unwanted transmissions of data are likely to occur. Very often, VVA providers implementing remote processing use a two-pass mechanism for this detection: a first pass embedded locally at the equipment level and a second one performed on remote servers where the next data processing are taking place. In this case, developers tend to set up a relatively low threshold in order to enhance the user experience and ensure that when the user says the keyword, it is almost always recognized - even if this means "over-detecting" it - and then implement a second detection pass on the server side, which is more restrictive.

2.5 Voice snippets and machine learning

20. VVAs rely on machine learning methods to perform a wide range of tasks (keyword detection, automatic speech recognition, natural language processing, speech synthesis, etc.) and thus necessitate large datasets to be collected, selected, labelled, etc.

21. The over- or under-representations of certain statistical characteristics can influence the development of machine learning-based tasks and subsequently reflect it in its calculations, and thus in its way of functioning. Thus, just as much as its quantity, the quality of the data plays a major role in the finesse and accuracy of the learning process.

22. In order to increase the quality of the VVA and improve the machine learning methods deployed, VVA designers might wish to have access to data relating to the use of the device in real conditions – i.e. voice snippets – in order to work on its improvement.

23. Whether it is to qualify the learning database or to correct errors made when the algorithm is deployed, learning and training of artificial intelligence systems necessarily require human intervention. This part of the work, known as digital labor, raises questions about

both working conditions and safety. In this context, news media have also reported data transfers between VVA designers and subcontractors allegedly without the necessary privacy protection guarantees.

3 ELEMENTS OF DATA PROTECTION

3.1 Legal framework

24. The relevant EU legal framework for VVAs is in the first instance the GDPR, as processing of personal data belongs to the core function of the VVAs. In addition to the GDPR, the e-Privacy Directive[6] sets a specific standard for all actors who wish to store or access information stored in the terminal equipment of a subscriber or user in the EEA.

25. In accordance with the definition of "*terminal equipment*"[7], smartphones, smart TVs and similar IoT devices are examples for terminal equipment. **Consequently, VVAs should be considered as "terminal equipment" and the provisions of Article 5(3) e-Privacy Directive apply whenever information in the VVA is stored or accessed**.[8]

26. Any processing operations of personal data following the aforementioned processing operations, including processing personal data obtained by accessing information in the terminal equipment, must also have a legal basis under Article 6 GDPR in order to be lawful.[9]

27. Since the controller, when seeking consent for storing or gaining of access to information pursuant to Article 5(3) e-Privacy Directive, will have to inform the data subject about all the purposes of the processing (meaning the "subsequent processing") – including any processing following the aforementioned operations – consent under Article 6 GDPR will generally be the most adequate legal basis to cover the subsequent processing of the personal data. Hence, consent will likely constitute the legal basis both for the storing and gaining of access to information already stored and the processing of personal data following the aforementioned processing operations. Indeed, when assessing compliance with Article 6 GDPR, one should take into account that the processing as a whole involves specific activities for which the EU legislature has sought to provide additional protection.[10] Moreover, controllers must take into account the impact on data subjects' rights when identifying the appropriate lawful basis in order to respect the principle of

[6] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) as amended by Directive 2006/24/EC and Directive 2009/136/EC (hereinafter "e-Privacy Directive").

[7] Article 1 of Commission Directive 2008/63/EC of 20 June 2008 on competition in the markets in telecommunications terminal equipment, defines "*terminal equipment*" as (a) an "*equipment directly or indirectly connected to the interface of a public telecommunications network to send, process or receive information; in either case (direct or indirect), the connection may be made by wire, optical fibre or electromagnetically; a connection is indirect if equipment is placed between the terminal and the interface of the network; (b) satellite earth station equipment*";

[8] See EDPB Guidelines 1/2020 paragraph 12 for similar reasoning regarding connected vehicles (hereinafter "EDPB Guidelines 1/2020"). See also EDPB, Opinion 5/2019 on the interplay between the e-Privacy Directive and the GDPR, in particular regarding to competence, tasks and powers of data protection authorities.

[9] Opinion 5/2019, paragraph 41.

[10] Opinion 5/2019, paragraph 41.

fairness.[11] The bottom line is that Article 6 GDPR cannot be relied upon by controllers in order to lower the additional protection provided by Article 5(3) e-Privacy Directive.

28. As shown in section 2.3 (steps 2 and 3), current VVAs require access to the voice data stored by the VVA device.[12] Therefore, Article 5(3) e-Privacy Directive applies. The applicability of Article 5(3) e-Privacy Directive means that the storing of information as well as the accessing to information already stored in a VVA requires, as a rule, end-user's prior consent[13] but allows for two exceptions: first, carrying out or facilitating the transmission of a communication over an electronic communications network, or, second, as strictly necessary in order to provide an information society service explicitly requested by the subscriber or user.

29. The second exception ("strictly necessary in order to provide an information society service explicitly requested by the subscriber or user") would allow a VVA service provider to process users' data to execute users' requests (see par. 72 in section 3.4.1) without the consent foreseen in Article 5(3) e-Privacy Directive. Conversely, such **consent as required by Article 5(3) e-Privacy Directive would be necessary for** the storing or gaining of access to information for **any purpose other than executing users' request** (e.g. user profiling). Data controllers would need to attribute consent to specific users. Consequently, data controllers should only process non-registered users data to execute their requests.

30. VVAs can accidentally capture audio of individuals who did not intend to use a VVA service. First, to a certain extent and depending on the VVAs, the wake-up expression can be changed. Individuals who are not aware of this change could accidentally use the updated wake-up expression. Second, VVAs can detect the wake-up expression by mistake or by error. It is highly unlikely that either of the exceptions foreseen in Article 5(3) e-Privacy Directive are applicable in the event of an accidental activation. Furthermore, consent as defined in the GDPR must be the "unambiguous indication of the data subject's wishes". Thus, it is highly unlikely that an accidental activation could be interpreted as a valid consent. If data controllers become aware (e.g. through automated or human review) that the VVA service has accidentally processed personal data, they should verify that there is a valid legal basis for each purpose of processing of such data. Otherwise, the accidentally collected data should be deleted.

31. Moreover, it should be noted that the personal data processed by VVAs may be highly sensitive in nature. It may carry personal data both in its content (meaning of the spoken text) and its meta-information (sex or age of the speaker etc.). The EDPB recalls that voice data is inherently biometric personal data.[14] As a result, when such data is processed for the purpose of uniquely identifying a natural person or is inherently or determined to be special category personal data, the processing must have a valid legal basis in Article 6 and be accompanied by a derogation from Article 9 GDPR (see section 3.8 below).

[11] EDPB Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects, Version 2.0, 8 October 2019, paragraph 1.

[12] It is possible that future VVA devices will adopt the edge computing paradigm and be capable of providing some service locally. In such event, it will be necessary to reassess the applicability of the e-Privacy Directive.

[13] See also EDPB Guidelines 1/2020, paragraph 14.

[14] Article 4(14) GDPR defines biometric data as 'personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data'

3.2 Identification of data processing and stakeholders

32. Considering the multiple possibilities of assistance that a VVA can provide in so many different environments of a data subject's daily life,[15] it is worth noting that careful consideration should be taken with the processing of personal data, which can also be impacted by different stakeholders.

3.2.1 Processing of personal data

33. From a personal data protection perspective, several constants can be observed irrespective of the VVA type (i.e. type of device, functionalities, services or combination of them) that can be used by a data subject. Such constants relate to the plurality of personal data, data subjects and data processing at stake.

Plurality of personal data types

34. The definition of personal data under Article 4(1) GDPR includes a wide variety of different data and applies in a technologically neutral context to any information that relates "to an identified or identifiable natural person".[16] Any interaction of a data subject with a VVA can fall under the scope of this definition. Once the interaction takes place, diverse range of personal data may be processed throughout the operation of the VVA as described in section 2.4.

35. From the initial request to the related answer, action or follow-up (e.g. setting up a weekly alert), the first personal data input will therefore generate subsequent personal data. This includes primary data (e.g. account data, voice recordings, requests history), observed data (e.g. device data that relates to a data subject, activity logs, online activities), as well as inferred or derived data (e.g. user profiling). VVAs use speech to mediate between users and all the connected services (e.g. a search engine, an online shop or a music streaming service) but unlike other intermediaries, VVAs may have full access to the requests' content and consequently provide the VVA designer with a wide variety of personal data depending on the purposes of the processing.

36. The plurality of personal data processed when using a VVA, also refers to a plurality of personal data categories for which attention should be paid (see below section 3.8). The EDPB recalls that when special categories of data[17] are processed, Article 9 GDPR requires the controller to identify a valid exemption from the prohibition to processing in Article 9(1) and a valid legal basis under Article 6(1), using an appropriate means identified under Article 9(2). Explicit consent may be one of the appropriate derogations where consent is the legal basis relied on under Article 6(1). Article 9 also notes (in detail) that the Member

[15] For example: at home, in a vehicle, in the street, at work or in any other private, public or professional spaces or a combination of these spaces.

[16] Article 4(1) GDPR also specifies that "*an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person*".

[17] Article 9(1) GDPR defines special categories of personal as "*personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited.*"

States may introduce further conditions to processing of biometric or other special categories data.

Plurality of data subjects

37. When using a VVA, personal data are processed from the first interaction with the VVA. For some data subjects this refers to the purchase of a VVA and/or the configuration of a user account (i.e. registered users). For other data subjects it refers to the first time they knowingly interact with the VVA of another data subject who purchased and/or configured this VVA (i.e. non-registered users). Besides these two categories of data subjects, there is a third one: accidental users who, registered or not, unknowingly make requests to the VVA (e.g. saying the correct wake-up expression without knowing the VVA is active, or saying other words that are mistakenly identified by the VVA as the wake-up expression).

38. The term plurality of data subjects also refers to multiple users for one VVA (e.g. device shared between registered and non-registered users, between colleagues, in a family, at school) and different types of users based on their condition (e.g. an adult, a child, an elder or a disable person). While a VVA can offer easier interaction with a digital tool and many benefits for some categories of data subjects, it is important to take into consideration the specificities of each category of data subjects and the context of use of the VVA.

Plurality of data processing

39. The technologies used to provide a VVA also have an impact on the amount of the processed data and the types of processing. The more a VVA provides services or features and is connected to other devices or services managed by other parties, the more the amount of personal data being processed and repurposing processing increases. This results in a plurality of processing carried out by automated means as described in section 2. Besides automated means, some processing may also involve human means. This is the case for example, when the implemented technology involves human intervention, such as the review of transcription of voices into texts, or the provision of annotations on personal data that can be used to insert new models in a machine-learning technology. This is also the case when humans analyse personal data (e.g. metadata) in order to improve the service provided by a VVA.

3.2.2 Processing by data controllers and processors

40. Data subjects should be in a position to understand and identify the roles at stake and should be able to contact or act with each stakeholder as required under the GDPR. The distribution of roles should not be to the detriment of the data subjects, even though scenarios can be complicated or evolving. In order to assess their roles, stakeholders are referred to the EDPB Guidelines 7/2020 on the concepts of controller and processor in the GDPR.[18]

41. As indicated in paragraph 16, the main stakeholders can be identified under the role of a provider or designer, an application developer, an integrator, an owner, or a combination of them. Different scenarios are possible, depending on who is doing what in the stakeholders' business relationship, on the user's request, the personal data, the data processing activities and their purposes. They should clearly decide and inform data

[18] EDPB Guidelines 07/2020 on the concepts of controller and processor in the GDPR, V1.0, adopted on 2 September 2020.

subjects on the conditions under which each of them will act and comply with the resulting roles of controllers, joint-controllers or processors as provided for by the GDPR.[19] Each of them may take on one or several roles, as they may be a unique data controller, a joint-controller, or a data processor for one data processing whereas carrying out another role for another data processing.

42. From a high-level perspective, the designer may act as a data controller when determining the purposes and means of a processing, but may intervene as a data processor when processing personal data on behalf of other parties, such as an application developer. The VVA user would therefore be subject to several data controllers: the application developer and the designer. It is also possible that the designer, the integrator and the developer are grouped into a single body acting as a unique data controller. In any case, the applicable qualifications have to be established on a case-by-case analysis.

Example 1:

The designer of the VVA processes user data for many purposes, including improving the VVA voice comprehension skills and respond accurately to requests. Therefore, and although this purpose may lead to the processing of data resulting from the use of applications provided by third parties, there is only one data controller: the designer of the VVA, on whose behalf and for whose purposes the processing is performed.

Example 2:

A bank offers to its customers an application that can be directly queried via the VVA in order to manage their accounts.

Two actors are involved in the processing of personal data: the designer of the VVA and the developer of the banking application.

In the scenario presented, the bank is the data controller for the provision of the service since it determines the purposes and essential means of processing related to the application allowing interaction with the assistant. Indeed, it offers a dedicated application that allows the user, a customer of the bank, to manage his/her accounts remotely. In addition, it decides on the means of processing by choosing appropriate processor, which is the designer of the VVA and can play an important role in assisting with its expertise to determining these means (for example, it can operate the development platform that allows third-party applications to be integrated into the VVA and, therefore, sets the framework and conditions to be respected by application developers).

43. On the data subject side, it is worth noting that several stakeholders may process the same personal data, even if the data subject does not really expect other parties than the VVA provider to be involved in the processing chain. So when a data subject acts with the VVA provider in relation to his/her personal data (e.g. exercise of data subject's rights), this does not automatically mean that this action will apply to the same personal data that is processed by another stakeholder. When these stakeholders are independent controllers, it is important that a clear information notice is given to the data subjects, explaining the various stages and actors of the processing. Moreover, in cases of joint controllership, it

[19] GDPR, Articles 12-14, Article 26.

should be made clear if every controller is competent to comply with all data subject's rights or which controller is competent for which right.[20]

> Example 3:
>
> In this scenario, the designer of the VVA wishes to use the data collected and processed for the service provided by the bank in order to improve its voice recognition system. The designer of the VVA, who processes the data for its own purposes, will then have the status of controller for this specific processing.

44. As many stakeholders may be involved in the processing chain, and respectively many staff, risky situations may occur if no appropriate measures and safeguards are in place. Controllers are accountable for them and therefore should focus on protecting personal data, notably by choosing appropriate business partners and data processors, applying privacy by default and by design principles,[21] implementing adequate security and other GDPR tools such as audits and legal agreements (e.g. Article 26 for joint controllers or Article 28 GDPR for processors).

45. VVA ecosystem is a complex one, where potentially many actors could exchange and process personal data as data controllers or processors. It is of utmost importance to clarify the role of each actor in respect of each processing and to follow the data minimisation principle also in respect of the data exchange.

46. In addition, controllers should be vigilant on personal data transfers and guarantee the required level of protection throughout the processing chain, in particular when they use services located outside of the EEA.

3.3 Transparency

47. Since VVAs process personal data (e.g. users' voice, location or the content of the communication), they must comply with the transparency requirements of the GDPR as regulated in Article 5(1)(a) as well as Article 12 and Article 13 (enlightened by Recital 58). Data controllers are obliged to inform users of the processing of their personal data in a concise, transparent, intelligible form, and in an easily accessible way.

48. Failure to provide necessary information is a breach of obligations that may affect the legitimacy of the data processing. Complying with the transparency requirement is an imperative, since it serves as a control mechanism over the data processing and allows users to exercise their rights. Informing users properly on how their personal data is being used makes more difficult for data controllers to misuse the VVA for purposes that go far beyond user expectations. For example, patented technologies aim to infer health status and emotional states from a user's voice and adapt the services provided accordingly.

49. Complying with the transparency requirements can be particularly difficult for the VVA service provider or any other entity acting as data controller. Given the specific nature of VVAs, data controllers face several obstacles to comply with the GDPR's transparency requirements:

[20] EDPB Guidelines 07/2020 on the concepts of controller and processor in the GDPR, V1.0, adopted on 2 September 2020, par. 162 (hereinafter "Guidelines 7/2020").
[21] See EDPB 4/2019 Guidelines on Article 25 Data Protection by Design and by Default, Version 2.0, adopted on 20 October 2020.

- **Multiple users**: data controllers should inform all users (registered, non-registered and accidental users), not only the user setting up the VVA.
- **Ecosystem complexity**: as explained in the technology background section, the identities and roles of those processing personal data when using a VVA is far from evident for the users.
- **Specificities of the vocal interface**: digital systems are not yet fit for voice-only interactions as the almost systemic use of a companion screen proves. However, adapting to the vocal interface and being able to inform the user clearly and correctly through this means is a necessity.

50. VVAs can be regarded as finite states machines going through a number of states during their ordinary functioning. They can be listening locally for the detection of wake-up expressions, or interacting with a remote server to resolve a command, but they can assume many other states depending on the context (e.g. if there is background environmental sound) or the user talking to them (e.g. they may talk to an identified or unknown user). Unfortunately, these situations take place in a substantial asymmetry of information with the user, who is hardly aware if the device is listening, and even less on the status in which it lies.

51. It is highly recommended that VVA designers and developers take adequate steps to fill those asymmetries, making the functioning of VVAs more interactive. Users should be informed of the status in which the device currently lies. This enhancement in transparency can be achieved both making the dialogue man-machine more interactive (e.g. the device might acknowledge in some way the reception of a vocal command), or broadcasting the status of the machine with specific signals. There are many options that can be explored in this regard, ranging from the use of specific vocal acknowledgements and visible icons or lights, or the use of displays on the device.

52. These issues are especially relevant considering the plurality of users and the presence among them of vulnerable categories of individuals, such as children, elderly people or users with audio-visual disabilities.

53. Two important questions become evident from the issues above: what is it the most feasible way to inform users and when is the appropriate time to inform them? These issues should be further examined in two different situations, depending on whether the VVA has only one user (such as a personal smart phone) or potentially multiple users (e.g. smart home device). Using the VVA technology, a subversion of these two basic settings could also occur, e.g. when a user has a personal smart phone and connects this to a car. The VVA of the smart phone, which could reasonably be expected to be used by that user only, is now "extended" to the others in the car.

54. Currently, all VVAs are connected to a user account and/or are set up by an application that requires one. The question of how data controllers could consider informing these users about the privacy policy while setting up the VVA should be addressed as described in the Article 29 Working Party Guidelines on transparency. Apps should make the necessary information available in an online store prior to download[22]. This way the information is given the earliest time possible and at the latest, at the time the personal

[22] Guidelines on transparency under Regulation 2016/679, WP260 rev. 01, endorsed by EDPB (hereinafter "WP29 Guidelines WP260"), paragraph 11.

data is obtained. Some VVA providers include third-party apps in the VVA default setup so these apps can execute those apps by using specific wake up expressions. VVAs using this third-party app deployment strategy should ensure that users get the necessary information also on the third-party processing.

55. However, many VVA designers require VVA user accounts that bundle the VVA service with multiple other services like email, video streaming or purchases to name a few. The decision by the VVA designer of linking the account to many different services has the effect of requiring very lengthy and complex privacy policies. The length and complexity of such privacy policies greatly hinder fulfilling the transparency principle.

> Example 4:
>
> A VVA designer requires its users to have an account to access the VVA service. This user account is not specific to the VVA service and can be used for other services offered by the VVA designer such as email, cloud storage and social media. To create the account, users have to read and accept a 30 pages long privacy policy. The policy includes information on the processing of personal data by all the services that could be linked with the account.
>
> The information provided by the VVA designer in this case should not be deemed as concise and its complexity reduces the required transparency. Therefore, the VVA designer would not be complying with the transparency requirements as set out in Articles 12 and 13 GDPR.

56. Although the most common way to provide necessary information is in writing, the GDPR allows "other means" as well. Recital 58 explicitly states that the information can be given in electronic form, e.g. through a website. In addition, when choosing the appropriate method to inform the data subjects, account should be taken to the specific circumstances, such as the manner the data controller and the data subject otherwise interact with each other.[23] An option for devices without a screen could be to provide a link which is easy to understand, either directly or in an e-mail. Already existing solutions could serve as example for the information, e.g. call centres' practices of notifying the caller about a phone call being recorded and directing them to their privacy policies. The constraints of screen less VVA does not exempt the data controller from providing the necessary information according to the GDPR when setting up the VVA or installing or using a VVA app. VVA providers and developers should develop voice-based interfaces to facilitate the mandatory information.

57. VVAs could be of great interest for users with impaired vision since they provide an alternative interaction means with the IT services that traditionally rely on visual information. According to Article 12 (1) GDPR providing the necessary information orally is possible exclusively if requested so by the data subject, but not as the default method. However, the constraints of screen less VVAs would require automated oral information means that could be augmented by written means. When using audio to inform data subjects, data controllers should provide the necessary information in a way that is concise and clear. Furthermore, data subjects should be able to re-listen[24].

58. Taking the appropriate measures to comply with GDPR transparency requirements is more complex when there are multiple users of the VVA other than the owner of the device.

[23] WP29 Guidelines WP260, paragraph 19.
[24] WP29 Guidelines WP260, paragraph 21.

VVA designers must consider how to properly inform non-registered and accidental users when their personal data is processed. When consent is the legal basis for processing users' data, users must be properly informed for the consent to be valid [25].

59. In order to comply with the GDPR, data controllers should find a way to inform not only registered users, but also non-registered users and accidental VVA users. These users should be informed at the earliest time possible **and at the latest, at the time of** the processing. This condition could be especially difficult to fulfil in practice.

60. Certain corporate specificities should also not be detrimental to the data subjects. As many stakeholders are global companies or are well known for a specific business activity (e.g. telecommunication, e-commerce, information technologies, web activities), the way they provide a VVA service should be clear. Adequate information should make the data subjects understand whether or not their use of the VVA will be linked to other processing activities managed by the VVA service provider (e.g. telecommunication, e-commerce, information technologies or web activities) apart from the strict use of the VVA.

> Example 5:
>
> To use its assistant, a VVA designer which also provides a social media platform and a search engine, requires the user to link his/her account to the assistant. By linking his/her account to the use of the VVA, the designer can thus enhance the profile of its users through the use of the assistant, the applications (or skills) that are installed, the orders placed, etc. Hence assistant interactions are a new source of information attached to a user. VVA designer should provide users' with clear information as to how their data will be processed for each service and with controls allowing the user to choose if the data will be used or not for profiling.

Recommendations

61. When users are informed about the VVA processing of personal data using a user account's privacy policy and the account is linked to other independent services (e.g. email or online purchases), the EDPB recommends the privacy policy to have a clearly separated section regarding the VVA processing of personal data.

62. The information provided to the user should match the exact collection and processing that is carried out. While some meta-information is contained in a voice sample (e.g. stress level of the speaker), it is not automatically clear, whether such analysis is performed. It is crucial that controllers are transparent on what specific aspects of the raw data they process.

63. Furthermore it should at all times be apparent which state the VVA is in. Users should be able to determine whether a VVA is currently listening on its closed-loop circuit and especially whether it is streaming information to its back-end. This information should also be accessible for people with disabilities such as colour blindness (daltonism), deafness (anaccousia). Specific care should be given to the fact that VVAs suggest a usage scenario where eye-contact to the device is not necessary. So, all user feedback, including state changes should be available in visual and acoustic form at least.

[25] GDPR, Article 4 (11).

64. Particular consideration should be applied if devices allow adding third party functionality ("apps" for VVAs). While some general information can be given to the user when they are the ones adding such functionality (given that it is the user's choice), during normal use of the device, the boundaries between the various controllers involved can be much less clear, i.e. the user might be not sufficiently informed how and by whom their data is processed (and to which extent) in a specific query.

65. All information about processing based on data collected and derived from the processing of recorded voice should also be available to users according to Article 12 GDPR.

66. VVA controllers should make transparent what kind of information a VVA can derive about its surroundings, such as but not limited to other people in the room, music running in the background, any processing of the voice for medical or marketing other reasons, pets, etc.

3.4 Purpose limitation and legal basis

67. The processing of voice requests by VVAs has an evident purpose, the execution of the request. However, there often are additional purposes which are not so evident, like the improvement of the VVA natural language understanding capacities by training VVA model with machine learning techniques. Among the most common purposes for processing personal data by VVAs we find:

- Execute users' requests
- VVA improvement by training of the machine learning model and human review and labelling of voice transcriptions
- User identification (using voice data)
- User profiling for personalized content or advertising

68. Due to their role as intermediaries and the way they are designed, VVAs process a wide variety of personal and non-personal data. This allows processing personal data for many purposes that go beyond answering the users' requests and that could go totally unnoticed. By analysing data collected via VVAs, it is possible to know or infer user interests, schedules, driving routes or habits. This could enable personal data processing for unforeseen purposes (e.g. sentiment analysis or health condition assessment[26]), which would be far beyond the reasonable user expectations.

69. Data controllers should clearly specify their purpose(s) in relation to the context in which the VVA is used, so that they are clearly understood by the data subjects (e.g. presenting purposes in categories). In line with Article 5(1) GDPR, the personal data should be collected for specific, explicit and legitimate purposes and not further processed in a manner that is incompatible with these purposes.

3.4.1 Execute users' requests

70. The main use for a VVA is to issue voice commands that need to be executed by the VVA or an associated app or service (e.g. a music streaming service, a mapping service or an electronic lock). The user's voice and potentially other data (e.g. the user's position when requesting a route for a certain destination) might therefore be processed.

[26] Eoghan Furey, Juanita Blue, "Alexa, Emotion, Privacy and GDPR", Conference paper, Human Computer Interaction Conference, July [2018].

Example 6:

The passenger of a smart car including a VVA requests a route to the closest gas station. The VVA processes the user voice to understand the command and the car's position to find the route and sends it to the smart component to show it in the car's screen.

71. Insofar as the processing of voice commands involves the storage or access to information stored in the terminal devices of the end-user, Article 5(3) of the e-Privacy Directive must be complied with. While Article 5(3) includes the general principle that such storage or access requires the prior consent of the end-user, it also provides for an exemption to the consent requirement where it is "strictly necessary in order for the provider of an information society service explicitly requested by the subscriber or user to provide the service". Insofar as the voice data is processed in order to execute the user's requests, it is exempted from the requirement of prior consent.

72. As indicated earlier, any processing operations of personal data subsequent to the storage or access to information in the terminal device of end-users must have a legal basis under Article 6 GDPR to be lawful.

73. There are two consecutive processing operations taking place on the VVA. As mentioned above the first one requires access to the VVA (and therefore the conditions of Article 5(3) e-Privacy Directive must be met). In addition to the conditions of Article 5(3) e-Privacy Directive, this second step also requires a legal basis under Article 6 GDPR.

74. When an individual takes the decision to use a VVA, this action generally implies that the initial user first needs to register an account to activate the VVA. In other words, this situation refers to a contractual relationship[27] between the registered user and the VVA controller. In view of its substance and fundamental objective, the core purpose of this contract is to use the VVA in order to execute the user's request of assistance.

75. Any personal data processing that is necessary to execute the user's request can therefore rely on the legal basis of the performance of the contract[28]. Such processing notably includes the capture of the user's voice request, its transcription to text, its interpretation, the information exchanged with knowledge sources to prepare the reply and then, the transcription to a vocal final reply that ends the user's request.

76. Performance of a contract can be a legal basis for processing personal data using machine learning (ML) when it is necessary for the provision of the service. Processing personal data using ML for other purposes which are not necessary such as service improvement should not rely on that legal basis.

77. Last but not least, the legal bases of the performance of the contract and consent under the GDPR should not be confused. The consent provided for entering into, i. e. agreeing to the contract is part of the validity of this contract and does not refer to the specific meaning of the consent under the GDPR[29].

[27] Provided that "*the contract is valid pursuant to applicable national contract laws*", extract from Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects ("Guidelines 2/2019"), §26.

[28] In accordance with the Guidelines 2/2019, which moreover states that the Opinion 06/2014 remains relevant to Article 6(1)(b) and the GDPR (see in particular pages 11, 16, 17, 18 and 55 in this Opinion 06/2014).

[29] See Guidelines 2/2019, respectively §18, §19, §20, §21 and §27.

78. When using a VVA does not require to previously configure a user account to the VVA, consent could be a possible legal basis.

3.4.2 Improve the VVA by training the ML systems and manually reviewing of the voice and transcripts

79. The accents and variations of human speech are vast. While all VVAs are functional once out of the box, their performance can improve by adjusting them to the specific characteristics of users' speech. As mentioned in section 2.6, this adjustment process relies on machine learning methods and consists of two processes: adding to the VVA training dataset new data collected from its users and the human review of the data processed for the execution of a fraction of the requests.

> Example 7:
>
> A VVA user has to issue three times the same voice command due to the VVA not understanding it. The three voice commands and the associated transcriptions are passed to human reviewers to review and correct the transcriptions. The voice commands and reviewed transcriptions are added to the VVA training dataset to improve its performance.

80. The processing activities described in the example should not be considered as (strictly) "necessary for the performance of a contract" within the meaning of Article 6(1)(b) GDPR, and therefore require another legal basis from Article 6 GDPR. The main reason being that VVAs are already functional when they come out of the box and can already perform as (strictly) necessary for the performance of the contract. The EDPB does not consider that Article 6(1)(b) would generally be an appropriate lawful basis for processing for the purposes of improving a service or developing new functions within an existing service. In most cases, a user enters into a contract to avail of an existing service. While the possibility of improvements and modifications to a service may routinely be included in contractual terms, such processing usually cannot be regarded as being objectively necessary for the performance of the contract with the user.

3.4.3 User identification[30] (using voice data)

81. The use of voice data for user identification implies the processing of biometric data as defined in Article 4.14 of the GDPR. Consequently, the data controller will need to identify an exemption under Article 9 of the GDPR in addition to the identification of a legal basis under Article 6 of the GDPR[31].

82. Of the exemptions listed in Article 9 of the GDPR, only data subjects' explicit consent seems applicable for this specific purpose.

---

[30] Technically, the notion of identification has to be distinguished from verification (authentication). Identification is a one-to-many (1: N) search and comparison and requires in principle a database in which several individuals are listed. Differently, the processing for verification purposes is a one-to-one (1:1) comparison and is used to verify and to confirm by a biometric comparison whether an individual is the same person as the one from whom the biometric data originates. To the knowledge of the EDPB, VVA on the market rely on the sole use of speaker identification technologies.

[31] GDPR considers that the mere nature of data is not always sufficient to determine if it qualifies as special categories of data since "the processing of photographs […] are covered by the definition of biometric data only when processed through a specific technical means allowing the unique identification or authentication of a natural person." (recital 51) The same reasoning applies to voice.

83. However, since this purpose requires to apply the specific legal regime of Article 9 of the GDPR further details follow in the section 3.8, related to the processing of special categories of data.

3.4.4 User profiling for personalized content or advertising

84. As mentioned above, VVAs have access to the content of all voice commands even when they are aimed at services provided by third parties. This access would enable the VVA designer to construct very accurate user profiles that could be used to offer personalized services or advertisements.

> Example 8:
>
> Each time a VVA user makes an Internet search, the VVA adds labels signalling topics of interest to the user profile. The results for each new search are presented to the user ordered taking into account those labels.
>
> Example 9:
>
> Each time a VVA user makes a purchase from an e-commerce service, the VVA stores a record of the purchase order. The VVA provider enables third parties to target the VVA user with targeted advertisements on the basis of past purchases.

85. Personalisation of content may (but does not always) constitute an intrinsic and expected element of a VVA. Whether such processing can be regarded as an intrinsic aspect of the VVA service will depend on the precise nature of the service provided, the expectations of the average data subject in light not only of the terms of service but also the way the service is promoted to users, and whether the service can be provided without personalisation.[32]

86. Where personalisation takes place in the context of a contractual relationship and as part of a service explicitly request by the end-user (and the processing is limited to what is strictly necessary to provide this service), such processing may be based on article 6(1)(b) of the GDPR.

87. If processing is not strictly "necessary for the performance of a contract" within the meaning of Article 6(1)(b) GDPR, the VVA provider must, in principle, seek the consent of the data subject. Indeed, because consent will be required under Article 5(3) of the e-Privacy directive for the storing or gaining of access to information (see paragraphs 29-30 above), consent under Article 6(1)(a) GDPR will also, in principle, be the appropriate legal basis for the processing of personal data following those operations as reliance on legitimate interest could, in certain cases, risk undermining the additional level of protection provided by Article 5(3) of the e-Privacy directive.

88. Regarding user profiling for advertisement, it should be noted that this purpose is never considered as a service explicitly requested by the end-user. Thus, in case of processing for this purpose users' consent should be systematically collected.

[32] See also Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects Version 2.0 8 October 2019, paragraph 57.

Recommendations

89. Users should be informed of the purpose of processing personal data and that purpose should accord with their expectations of the device they purchase. In case of a VVA, that purpose – from a user's point of view – clearly is the processing of their voice for the sole purpose of interpreting their query and provide meaningful responses (be that answers to a query or other reactions like remote-controlling a light switch).

90. When the processing of personal data is based on consent, such consent "should be given in relation to *one or more specific* purposes and that a data subject has a choice in relation to each of them". Moreover, "a controller that seeks consent for various different purposes should provide a separate opt-in for each purpose, to allow users to give specific consent for specific purposes"[33]. For example, users should be able to separately consent or not for the manual review and labelling of voice transcriptions or the use of their voice data for user identification/authentication (see section 3.8).

3.5 Processing of children's data

91. Children can also interact with the VVAs or can create their own profiles connected to the ones of the adults. Some VVAs are embedded in devices which are specifically aimed at children.

92. When the legal basis for the processing is the performance of a contract, the conditions for processing children data will depend on national contract laws.

93. When the legal basis for the processing is consent and according to Article 8(1) GDPR, processing of children's data is only lawful "*where the child is at least 16 years old. Where the child is below the age of 16 years, such processing should be lawful only if and to the extent that consent is given or authorized by the holder of parental responsibility over the child*". Consequently, to comply with the GDPR, when consent is the legal basis, explicit permission should be sought from parents or guardians to collect, process, and store children's data (voice, transcripts, etc.).

94. Parental controls are available to a certain degree but in their current form they are not user friendly (e.g. it is necessary to sign in a new service) or have limited capacities. Data controllers should invest in developing means for parents or guardians to control children use of VVAs.

3.6 Data retention

95. VVAs process and generate a wide variety of personal data like voice, transcriptions of voice, metadata or system logs. These types of data could be processed for a wide range of purposes like provision of a service, NLP improvement, personalization or scientific research. Following the GDPR data storage limitation principle, VVAs should store data for no longer than is necessary for the purposes for which the personal data are processed. Therefore, the data retention periods should be tied to different processing purposes. VVA service providers or third-parties providing services through VVAs should assess the maximum retention period for each data set and purpose.

[33] See Guidelines on consent under Regulation 2016/679, section 3.2.

96. The data minimization principle is closely related to the data storage limitation principle. Not only do data controllers need to limit the data storage period, but also the type and quantity of data.

97. Data controllers should ask themselves, among others the following questions: Is it necessary to store all voice recordings or all transcriptions to achieve the purpose X? Is it necessary to store the voice data once the transcription has been stored? In that case, for what purpose? How long is voice or transcription data necessary for each purpose? The answer to these and other similar questions will define the retention periods that should be part of the information available to the data subjects.

98. Some VVA store by default personal data like voice snippets or transcriptions for an undefined period while providing users means to delete such data. Retaining personal data indefinitely goes against the storage limitation principle. Providing data subjects with means to delete their personal data does not remove the data controller's responsibility of defining and enforcing a data retention policy.

99. VVA design needs to take into account users' controls to delete their personal data in their devices and in all remote storage systems. These controls may be required to resolve different kind of users' requests, for example, a request of erasure or the withdrawal of previously given consent. The design of some VVAs did not take into account this requirement.[34]

100. As in other contexts, data controllers may need to retain personal data as evidence of a service provided to a user to comply with a legal obligation. The data controller may retain personal data based on that basis. However, the data retained should remain the minimum necessary to comply with such a legal obligation and for the minimum amount of time. Of course, the data retained for the purpose of complying with a legal obligation should not be used for any other purposes without a legal basis under Article 6 GDPR.

> Example 10:
>
> A user purchases a TV in an e-commerce service using a voice command issued to a VVA. Even if the user requests afterwards the deletion of their data, the VVA provider or developer could still retain some data on the grounds of their legal obligation set by tax regulation to keep purchase evidence. However, the data stored for this purpose should not exceed the minimum necessary to comply with the legal obligation and could not be processed for any other purposes without a legal basis under Article 6 GDPR.

101. As mentioned in section 2, VVAs voice recognition capacity improves by training machine learning systems with users' data. If users do not consent or withdraw their consent to the use of their data for such purpose, their data could not be lawfully used to train any more model and should be deleted by the data controller, assuming that there is no other purpose justifying the continued retention. However, there is evidence that there may be risks of re-identification in some machine learning models.[35]

[34] See Amazon's letter of 28 June 2019 in response to US Senator Christopher Coons: https://www.coons.senate.gov/imo/media/doc/Amazon%20Senator%20Coons Response%20Letter 6.28.19 [3].pdf.

[35] Veale Michael, Binns Reuben and Edwards Lilian 2018 "Algorithms that remember: model inversion attacks and data protection law" Phil. Trans. R. Soc. A.37620180083, doi: 10.1098/rsta.2018.0083.

102. Data controllers and processors should use models which do not restrict their ability to stop processing if an individual revokes their consent, nor should they use models which restrict their ability to facilitate data subject rights. Controllers and processors should apply mitigation measures to reduce the re-identification risk to an acceptable threshold.

103. In the event that the user withdraws his or her consent, the data collected from the user can no longer be used for further training of the model. Nevertheless, the model previously trained using this data does not have to be deleted. The EDPB however highlights that there is evidence that there may be risks of personal data leaking in some machine learning models In particular, numerous studies showed that reconstruction as well as membership inference attacks can be performed, allowing attackers to retrieve information about individuals.[36] Data controllers and processors should therefore apply mitigation measures to reduce the re-identification risk to an acceptable threshold to make sure they use models which do not contain personal data.

104. Data subjects should not be nudged to keep their data indefinitely. While deleting stored voice data or transcriptions might have an impact on the service performance, such impact should be explained to users in a clear and measurable way. VVA service providers should avoid making general statements on the degradation of the service after personal data is deleted.

105. Anonymizing voice recordings is specially challenging, as it is possible to identify users through the content of the message itself and the characteristics of voice itself. Nevertheless, some research[37] is being conducted on techniques that could allow to remove situational information like background noises and anonymize the voice.

Recommendations

106. From a user's perspective, the main purpose of processing their data is querying and receiving responses and/or triggering actions like playing music or turning on or off lights. After a query has been answered or a command executed, the personal data should be deleted unless the VVA designer or developer has a valid legal basis to retain them for a specific purpose.

107. Before considering anonymization as means for fulfilling the data storage limitation principle, VVA providers and developers should check the anonymization process renders the voice unidentifiable.

108. Configuration defaults should reflect these requirements by defaulting to an absolute minimum of stored user information. If these options are presented as part of a setup wizard, the default setting should reflect this, and all options should be presented as equal possibilities without visual discrimination.

---

[36] N. Carlini et al, "Extracting Training Data from Large Language Models", December 2020.

[37] J. Qian, H. Du, J. Hou, L. Chen, T. Jung and X. Li, "Speech Sanitizer: Speech Content Desensitization and Voice Anonymization," in *IEEE Transactions on Dependable and Secure Computing*, doi: 10.1109/TDSC.2019.2960239.
A. Cohen-Hadria, M. Cartwright, B. McFee and J. P. Bello, "Voice Anonymization in Urban Sound Recordings," *2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)*, Pittsburgh, PA, USA, 2019, pp. 1-6, doi: 10.1109/MLSP.2019.8918913.
See, for example, VoicePrivacy (https://www.voiceprivacychallenge.org), an initiative to develop privacy preservation solutions for speech technology.

109. When during the review process the VVA provider or developer detects a recording originated on a mistaken activation, the recording and all the associated data should be immediately deleted and not used for any purpose.

3.7 Security

110. To securely process personal data, VVAs should protect their confidentiality, integrity and availability. Apart from risks stemming from elements on the VVA ecosystem, using voice as communication means creates a new set of security risks.

111. VVAs are multiuser. They may allow for more than one registered user and anyone on their surroundings can issue commands and use their services. Any VVA service requiring confidentiality will involve some access control mechanism and user authentication. Without access control, anyone able to issue voice commands to the VVA could access, modify or delete any users' personal data it (e.g. ask for received messages, user's address or calendar events). Issuing voice commands to the VVA does not require being physically close to it since they can be manipulated, for example, via signal broadcasting[38] (e.g. radio or TV). Some of the known methods to remotely issue commands to VVAs like laser[39] or ultrasonic (inaudible) waves[40] are not even detectable by human senses.

112. User authentication can rely on one or more of the following factors: something you know (such as a password), something you have (such as a smart card) or something you are (such as a voice fingerprint). A closer look at these authentication factors in the VVA context shows that:

- Authentication using something the user knows is problematic. The secret that would allow users to prove their identity should be spoken aloud, exposing it to anyone in the surroundings. VVAs communication channel is the surrounding air, a channel type that cannot be fortified in the way traditional channels are (e.g. by limiting the access to the channel or encrypting its content).
- Authentication using something the user has would require the VVA service providers to create, distribute and manage "tokens" that could be used as proof of identity.
- Authentication using something the user is implies the use of biometric data for the purpose of uniquely identifying a natural person (see section 3.8 below).

113. VVA user accounts are associated to the devices in which the service is provided. Often the same account used to manage the VVA is used to manage other services. For example, owners of an Android mobile phone and a Google Home speaker can and most likely associate their Google account to both devices. Most VVAs do not require or offer an identification or authentication mechanism when a device providing a VVA service has just one user account.

114. When there is more than one user account associated with the device, some VVAs offer an optional basic access control in the form of PIN number with no real user

[38] X. Yuan et al., "All Your Alexa Are Belong to Us: A Remote Voice Control Attack against Echo" 2018 IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 2018, pp. 1-6, doi: 10.1109/GLOCOM.2018.8647762.
[39] See, for example, https://lightcommands.com.
[40] See, for example, https://surfingattack.github.io.

authentication. Some other VVAs have the option to use voice fingerprint recognition as identification mechanism.

115. Although user identification or authentication might not be necessary to access all VVA services, it will definitely be for some. Without an identification or authentication mechanism, anyone could access other users' data and modify or erase them at will. For example, anyone close to a smart speaker could delete other users' playlists from the music streaming service, commands from the command history or contacts from the contact list.

116. Most VVAs blindly trust their local networks. Any compromised device in the same network could change the settings of the smart speaker or allow the installation of malware or to associate fake apps/skills to it without the user's knowledge or agreement.[41]

117. VVAs, like any other software, are subject to software vulnerabilities. However, due to the VVA market concentration[42] any vulnerability could affect millions of VVA users. If working as currently designed, VVAs do not send any information to the speech recognition cloud service until the wake-up expression is detected. However, software vulnerabilities could allow an attacker to bypass the VVA set-up and security measures. It could then be possible for example to get a copy of all data sent to the VVA cloud and forward it to a server controlled by the attacker.

118. Data lawfully processed or derived by VVAs allow building a fairly accurate profile of their users as VVA know or can infer the location, the relations and the interests of their users. VVAs are increasingly present in users' homes and smartphones. This circumstance increases the risk of mass surveillance and mass profiling. Consequently, the security measures to protect the data both in transit and at rest, in the devices and in the Cloud, should match those risks.

119. The increasing use of VVA in conjunction with not adequately balanced access rights by law enforcement authorities could induce a chilling effect that would undermine fundamental rights like freedom of speech.

120. Law enforcement authorities, both in[43] and out[44] of the EU, have already expressed their interest in accessing the voice snippets captured by VVAs. Access to data processed or derived by VVAs in the EU should comply with the existing EU data protection and privacy regulation framework. In case some Member States consider issuing specific legislation restricting the fundamental rights to privacy and data protection, such restrictions should always comply with the requirement set out in Article 23 of the GDPR[45].

121. The human review of voice recordings and associated data to improve VVA service quality is a common practice among VVA providers. Due to the sensitive nature of the data that is

[41] See, for example, Deepak Kumar et al., *Skill Squatting Attacks on Amazon Alexa*, USENIX Security Symposium, August 2018, https://www.usenix.org/conference/usenixsecurity18/presentation/kumar
Security Research Labs, *Smart Spies: Alexa and Google Home expose users to vishing and* eavesdropping, November 2019, https://srlabs.de/bites/smart-spies
[42] The VVA market is currently shared among less than a dozen service providers.
[43] See, for example, https://www.ft.com/content/ad765972-87a2-11e9-a028-86cea8523dc2
[44] See, for example, https://cdt.org/insights/alexa-is-law-enforcement-listening
[45] See also EDPB, Guidelines 10/2020 on restrictions under Article 23 GDPR.

processed by this human reviewers and the fact that this process is often subcontracted to a processors, it is of utmost relevance that adequate security measures are put in place.

Recommendations

122. VVA designers and application developers should provide secure state-of-the-art authentication procedures to users.

123. Human reviewers should always receive the strictly necessary pseudonymised data. The legal agreements governing the review should expressly forbid any processing that could lead to the identification of the data subject.

124. If emergency calling is provided as a service through the VVA, a stable uptime[46] should be guaranteed.

3.8 Processing of special categories of data

125. As previously mentioned, VVAs have access to information of an intimate nature which can be protected under Article 9 of the GDPR (see section 3.8.1), such as biometric data (see section 3.8.2). Therefore, VVA designers and developers must carefully identify in which cases the processing implies special categories of data.

3.8.1 General considerations when processing special categories of data

126. VVAs may process special categories of data in different circumstances:

- As part of their own services, for example when managing medical appointments in users' agendas.
- When acting as interface for third-party services, VVA providers process the content of the commands. Depending on the type of service requested by the user, VVA providers could process special categories of data. An example could be when a user issues commands to a VVA to use a third-party app used to keep track of her ovulation.[47]
- When voice data is used for the purpose of uniquely identify the user, as developed below.

3.8.2 Specific considerations when processing biometric data

127. Some VVA have the capability of uniquely identifying their users merely based on their voice. This process is known as voice template recognition. During the enrolment phase of voice recognition, the VVA processes a user's voice to create a voice template (or voiceprint). During its regular use, the VVA can calculate the voice template of any users and compare it to the enrolled templates to uniquely identify the user who executed a command.

> Example 11:
>
> A group of users set up a VVA to use voice template recognition. After doing so, each of them enrol their voice templates.

[46] The time a device or a service can be left unattended without crashing, or needing to be rebooted for administrative or maintenance purposes.

[47] See for example, a product available here: https://www.amazon.com/Ethan-Fan-Ovulation-Period-Tracker/dp/B07CRLSHKY

Later, a user requests the VVA access to the meetings in his or her agenda. Since access to the agenda requires user identification, the VVA extracts the template from the request's voice, calculates its voice template and checks if it matches an enrolled user and if that specific user has access to the agenda.

128. In the example above, the recognition of a the user's voice on the basis of a voice template amounts to the processing of special categories of personal data within the meaning of Article 9 GDPR (processing of biometric data for the purpose of uniquely identifying a natural person).[48] The processing of biometric data for the purpose of user identification as required in the example will require the explicit consent of the data subject(s) concerned (Article 9(2)(a) GDPR). Therefore, when obtaining users' consent, data controllers must comply with the conditions of Article 7 and recital 32 of the GDPR and should offer an alternative identification method to biometrics, with regard to the free nature of consent.

129. When using voice data for biometric identification or authentication, data controllers are required to make transparent where biometric identification is used and how voiceprints (biometric templates) are stored and propagated across devices. To fulfil this transparency requirement, the EDPB recommends to provide the answers to the following questions:

- Does the activation of voice identification on one device automatically activate this feature on all other devices running with the same account?
- Does the activation of voice identification propagate through the VVA controller's infrastructure to devices owned by other users?
- Where are biometric templates generated, stored and matched?
- Are biometric templates accessible to VVA providers, developers or others?

130. When the registered user configures the VVAs to identify the voice of its users, the voice of non-registered and accidental users will also be processed for the purpose of uniquely identifying them.

131. Indeed, detecting the voice of the right speaker also involves comparing it with that of other people in the assistant's vicinity. In other words, the speaker recognition functionality implemented in voice assistants may require the voice biometrics of people speaking in the household to be recorded, to allow the user's voice characteristics to be distinguished from those of the person who wishes to be recognised. Biometric identification may therefore have the consequence of subjecting uninformed persons to biometric processing, by registering their template and comparing it with that of the user wishing to be recognised.

132. In order to avoid such collection of biometric data without the knowledge of the data subjects while allowing a user to be recognized by the assistant, solutions based on the user's data alone should be given priority. In concrete terms, this means that biometric recognition is only activated at each use at the user's initiative, and not by a permanent analysis of the voices heard by the assistant. For instance, a specific keyword or question to the persons present could be provided in order to obtain their consent to trigger biometric processing. For example, the user can say "identification" or the assistant can

ask "do you wish to be identified" and wait for a positive response to activate biometric processing.

Example 12:
If the user wishes to set up biometric authentication for access to certain protected data such as his/her bank account, the voice assistant could activate speaker verification, when he/she launches the application only, and verify his/her identity in this way.

Recommendations

133. Voice templates should be generated, stored and matched exclusively on the local device, not in remote servers.

134. Due to the sensitiveness of the voiceprints, standards such as ISO/IEC 24745 and techniques of biometric template protection[49] should be thoroughly applied.

135. If a VVA uses voice based biometric identification VVA providers should:

- Ensure that the identification is accurate enough to reliably associate personal data to the right data subjects.
- Ensure that the accuracy is similar for all user groups by checking that there is no substantial bias towards different demographic groups.

3.9 Data minimization

136. Controllers should minimize the amount of data that is collected directly or indirectly and obtained by processing and analysis, e.g. not perform any analysis on the user's voice or other audible information to derive information about their mental state, possible disease or circumstances of their life.

137. Roll out settings by default that limit any data collecting and/or processing to a minimum required amount needed to provide the service.

138. Depending on the location, use context and microphone sensitivity, VVA could collect third parties' voice data as part of the background noise when collecting the users' voice. Even if background noise does not include voice data, it can still include situational data that could be processed to derive information about the subject (e.g. location).

Recommendations

139. In order to avoid recording background voices and situational information, VVA service providers should apply automated background-noise filtering.

[49] See, for example: Jain, Anil & Nandakumar, Karthik & Nagar, Abhishek. (2008). "*Biometric Template Security*". EURASIP Journal on Advances in Signal Processing. 2008. 10.1155/2008/579416.
S. K. Jami, S. R. Chalamala and A. K. Jindal, "*Biometric Template Protection Through Adversarial Learning*" *2019 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, USA, 2019, pp. 1-6, doi: 10.1109/ICCE.2019.8661905.

140. VVA designers should consider technologies deleting the background noise and conversations ensuring that only the user voice is recorded.[50]

3.10 Accountability

141. For any processing that is based on consent, controllers are obliged to be able to prove the consent of data subjects according to Article 7 (1) GDPR. Voice data can be used for accountability (e.g. to prove consent). The retention obligation for such voice data would then be dictated by the accountability requirements of the relevant specific legislation.

142. When evaluating the need for a Data Protection Impact Assessment (DPIA), the EDPB set out criteria[51] to be used by data protection authorities in creating lists of processing operations that require a mandatory DPIA and provide examples of processing that are likely to require a DPIA. It is very likely that VVA services fall into the categories and conditions identified as needing a DPIA. This includes considering if the device may be observing monitoring or controlling data subjects or systematically monitoring at large scale as per Article 35(3)(c), use of "new technology", or the processing of sensitive data and data concerning vulnerable data subjects.

143. All data collection and processing activities must be documented in accordance with Article 30 GDPR. That includes all processing involving voice data.

Recommendations

144. If voice messages are to be used to inform users according to Article 13, the data controllers should publish such messages on their website so they are accessible to the users and the data protection authorities.

3.11 Data protection by design and by default

145. VVA providers and developers should consider the necessity of having a registered user for each of their functionalities. While it is clear that it is necessary to have a registered user to manage an agenda or an address book, it is not so clear that making a phone call or an Internet search requires the VVA to have a registered user.

146. By default, services which do not require an identified user should not associate any of the VVA identified users to the commands. A privacy and data protection friendly default VVA would only process users' data for executing users' requests and would store neither voice data nor a register of executed commands.

147. While some devices can only run one VVA, others can choose among different VVAs. VVA providers should develop industry standards enabling data portability in accordance with Article 20 of the GDPR.

148. Some VVA providers alleged their VVAs could not delete all users' data even when requested by the data subject. VVA providers should ensure that all users' data can be erased at the user's request in accordance with Article 17 of the GDPR.

[50] J. Qian, H. Du, J. Hou, L. Chen, T. Jung and X. Li, "Speech Sanitizer: Speech Content Desensitization and Voice Anonymization," in *IEEE Transactions on Dependable and Secure Computing*, doi: 10.1109/TDSC.2019.2960239.
[51] Article 29 Working Party, Guidelines on Data Protection Impact Assessment (DPIA), wp248, rev.01, endorsed by EDPB.

4 MECHANISMS TO EXERCISE DATA SUBJECT RIGHTS

149. In compliance with the GDPR, data controllers providing VVA services must allow all users, registered and non-registered, to exercise their data subject rights.

150. VVA providers and developers should facilitate data subjects' control over their data during the entire processing period, in particular ease their right of access, rectification, erasure, their right to restrict the processing and, depending on the legal basis of the processing, their right to data portability and their right to object.

151. The data controller should provide information on the data subject's rights at the time when data subjects switch on a VVA and at the latest, at the time when the first user's voice request is processed.

152. Given that the main interaction means for VVAs is voice, VVA designers should ensure that users, registered or not, can exercise any data subject rights, using easy-to-follow voice commands. VVA designers, as well as app developers in case they are part of the solution, should at the end of the exercise process inform the user that his/her rights have been duly factored, by voice or by providing a writing notification to the user's mobile, account or any other mean chosen by the user.

153. At least, VVA designers and app developers, notably, should implement specific tools providing an effective and efficient way to exercise such rights. They should therefore propose for their devices, a way to exercise data subjects' rights as by providing the data subject with self-service tools, as a profile management system[52.] This could facilitate an efficient and timely handling of data subject's rights and will enable the data controller to include the identification mechanism in the self-service tool.

154. In regards of the exercise of data subjects rights in case of multiple users, when a user, registered or not, exercises one of his or her rights, he or she should do so without prejudice to any other users' rights. All the users, registered and non-registered can exercise their rights as long as the data controller is still processing the data. Data controller should set up a process ensuring that data subject rights are exercised.

4.1 Right to access

155. According to Article 12(1) GDPR, communication under Article 15 should be provided in writing, or by other means, including, where appropriate, by electronic means. As regards access to the personal data undergoing processing, Article 15(3) states that where the data subject makes the request by electronic means, and unless otherwise requested by the data subjects, the information should be provided in a commonly used electronic form. What could be considered as a commonly used electronic form should be based upon the reasonable expectations of the data subjects and not upon what format the data controller uses in its daily operations. The data subject should not be obliged to buy specific software or hardware in order to get access to the information.

[52] Profile management system is understood as a place within the VVA system, where users may, anytime, store its preferences, set modification and change easily his/her privacy settings.

156. On demand, data controllers should therefore send a copy of personal data, and audio data (including voice recordings and transcriptions) in particular, in a common format readable by the data subject.

157. When deciding about the type of format the information under Article 15 GDPR should be provided in, the data controller needs to keep in mind that the format should enable the information to be presented in a way that is both intelligible and easy accessible. Data controllers should also tailor the information to the specific situation of the data subject making the request.

> Example 13:
>
> A data controller providing a VVA service receives, from a user, both a request of access and a request for data portability. The data controller decides to provide the information under both Article 15 and Article 20 in a PDF file. In such a case, the data controller should not be considered to handle both requests in a correct manner. A PDF file technically fulfils the obligations on the data controller under Article 15, but does not fulfil the obligations on the data controller under Article 20.[53]
>
> It should be noted that simply referring users to a history of their interactions with the voice assistant does not appear to enable the data controller to meet all its obligations under the right of access, as the accessible data generally represents only part of the information processed in the context of providing the service.

158. The right of access should not be used to counter / to get around the principles of minimisation and data retention.

4.2 Right to rectification

159. To facilitate data rectification, users, registered or not, should be able to manage and update, at any time, their data by voice directly from the VVA device, as described above. Furthermore self-service tool should be implemented inside the device or an application in order to help them to rectify easily their personal data. Users should be notified by voice, or by writing of the update.

160. More generally, the right to rectification applies to any opinions and inferences[54] of the data controller, including profiling, and should consider the vast majority of data is highly subjective.[55]

4.3 Right to erasure

161. Users, registered or not, should be able, at any time, by voice from the VVA device, or from a self-service tool integrated into any device associated to the VVA, to delete data concerning them. In this respect, the personal data can be deleted by a data subject as easily as it is submitted. Due to the inherent difficulties of anonymising voice data and the

[53] WP29 Guidelines on the right to data portability - endorsed by the EDPB, p. 18.
[54] The fact that opinions and inferences can qualify as personal data has been confirmed by the CJEU, which noted that the term 'any information' in the definition of personal data includes information that is 'not only objective but also subjective, in the form of opinions and assessments, provided that it 'relates' to the data subject' - Case C-434/16 *Peter Nowak v Data Protection Commissioner* ECLI:EU:C:2017:994 [34].
[55] Getting Data Subject Rights Right, A submission to the EDPB from data protection academics, November 2019.

wide variety of personal data collected from, observed and inferred about the data subject,[56] in this context the right to erasure could be hardly accommodated by anonymising personal datasets. As the GDPR is technology neutral and technology evolves rapidly, it will nevertheless not be excluded that right to erasure may be made effective through anonymization.

162. In some cases, without a third party screen or the possibility of displaying data stored (e.g. a mobile application or a tabular device), it is difficult to have a preview of the recorded tracks, to judge the relevance of the suggestions. A dashboard (or an application) widely accessible to users in order to ease its use should be supplied with the voice assistant to delete the history of the requests asked and customize the tool according to user's needs.[57]

163. For any data processing and, in particular, when registered data subjects consent to the voice recordings to be transcribed and used by the provider for the improvement of its services, VVA providers should, on demand of the user, be able to delete the initial voice recording as well as any related transcription of the personal data.

164. The data controller should ensure that no more processing may occur, after the exercise of the right of erasure. In regards to previous actions, the right to erasure may meet some legal and technical limits, notably.

> Example 14:
>
> If prior to the deletion request, a user made an online purchase by means of his/her VVA, the VVA provider may delete the voice recording relating to the online purchase and ensure no more future further use. However, the purchase will still be effective as well as the vocal order or the written transcription processed by the e-commerce website (here the exemption is based of legal obligation of the e-commerce website).
>
> In the same vein, if prior to the deletion request, the user added a specific song to his/her playlist, by means of his/her VVA, the VVA providers will be able to delete the oral request, but not the past consequences of such request (the erasure will not impact the user's playlist).

165. Based on the above, in case the same personal data is processed for different processing purposes, data controllers should interpret erasure requests as a clear signal to stop processing of the data for all purposes that are not legally exempted.

166. In accordance with the conditions set out in Article 21(1) GDPR, data processed on the basis of legitimate interests of the VVA providers should not be an exemption to the right of erasure, in particular, because data subjects do not reasonably expect further processing of their personal data.

4.4 Right to data portability

167. The data processing made by the VVA providers falls under the scope of data portability, as processing operations are mainly based, on the data subject's consent (under Article 6(1)(a), or under Article 9(2)(a) when it comes to special categories of personal data) or, on a contract to which the data subject is a party under Article 6(1)(b).

[56] Article 29 Working Party Opinion 05/2014 on Anonymisation Techniques, adopted on 10 April 2014.
[57] "Assistants vocaux et enceintes connectées, l'impact de la voix sur l'offre et les usages culturels et médias", the French "Conseil Supérieur de l'Audiovisuel", May 2019.

168. In practice, the right to data portability should facilitate switching between different VVA providers. VVAs operating in a digital environment in particular and data subject's voice being recorded in an application or a platform, the right to data portability should be granted for all personal data provided by the data subject. Furthermore, the data controller should offer users the possibility of directly retrieving their personal data from their user area, as a self-service tool. The users should also be able to exercise this right through voice command.

169. VVA providers and developers should give to the data subjects an extensive control over the personal data concerning him or her, in order to allow them to transfer personal data from a VVA provider to another. Data subjects should, therefore, receive his/her personal data provided to the data controller, in a structured, commonly used and machine-readable format as well as from means[58] that contribute to answer data portability requests (such as download tools and Application Programming Interfaces)[59]. As stated in the Guidelines on the right to data portability, in case of large or complex personal data collection, that could be the case here, the data controller should provide an overview "in a concise, transparent, intelligible and easily accessible form, using clear and plain language" (see Article 12(1) GDPR) in such a way that data subjects should always have clear information of what data to download or transmit to another data controller in relation to a given purpose. For example, data subjects should be in a position to use software applications to easily identify, recognize and process specific data from it.

170. This right should allow the user to retrieve for his/her personal use, the data that he/she has communicated by means of his/her voice (e.g. the history of voice interactions) and within the framework of the creation of his/her user account (e.g.: name and first name), notably.

171. For the full application of this data subjects' right in a context of one digital single market, VVA designers and app developers, notably, should develop common machine-readable formats that ease interoperability of the data format between VVA systems[60], including the standard formats for voice data. Technologies should be structured in order to ensure that personal data, including voice data, processed are easily and fully reusable by the new controller[61].

---

[58] See as an illustration, the reasoning of the Article 29 Working Party in the Guidelines on the right to data portability - endorsed by the EDPB, p. 16:
"*On a technical level, data controllers should explore and assess two different and complimentary paths for making portable data available to the data subjects or to other data controllers:*
- a direct transmission of the overall dataset of portable data (or several extracts of parts of the global dataset);
- an automated tool that allows extraction of relevant data.
The second way may be preferred by data controllers in cases involving of complex and large data sets, as it allows for the extraction of any part of the data-set that is relevant for the data subject in the context of his or her request, may help minimizing risk, and possibly allows for use of data synchronization mechanisms (e.g. in the context of a regular communication between data controllers). It may be a better way to ensure compliance for the "new" data controller, and would constitute good practice in the reduction of privacy risks on the part of the initial data controller".

[59] In this respect: Article 29 Working Party Guidelines on the right to data portability - endorsed by the EDPB, p. 1.

[60] In this respect: recital 68 of the GDPR; WP29 Guidelines on the right to data portability - endorsed by the EDPB, p. 17.

[61] "In this respect, recital 68 encourages data controllers to develop interoperable formats that enable data portability but without creating an obligation for controllers to adopt or maintain processing systems which are

172. In regards to the format, VVA providers should provide personal data using commonly used open formats (e.g. mp3, wav, csv, gsm, etc.) along with suitable metadata used in order to accurately describe the meaning of exchanged information.[62]

5 ANNEX: AUTOMATIC SPEECH RECOGNITION, SPEECH SYNTHESIS AND NATURAL LANGUAGE PROCESSING

173. Following the theoretical foundations of signal processing, notably Claude Shannon's information and sampling theories, automatic speech processing has become a fundamental component of engineering sciences. At the crossroads of physics (acoustics, wave propagation), applied mathematics (modelling, statistics), computer science (algorithms, learning techniques) and human sciences (perception, reasoning), speech processing has rapidly been broken down into numerous subjects of study: speaker identification and verification, automatic speech recognition, voice synthesis, emotion detection, etc. Over the last fifteen years or so, the discipline as a whole has made very significant progress, with various factors contributing to this: improved methods, a significant increase in computing capacities and greater volumes of data available.

5.1 Automatic Speech Recognition (ASR)

174. The automatic speech recognition (also known as speech to text) used to involve three distinct stages aimed at: 1) determine which phonemes were said using an acoustic model; 2) determine which words were said using a phonetic dictionary; 3) transcribe the sequence of words (sentence) most likely to have been said using a language model. Today, with the progress made possible by deep learning (a machine learning technique), many systems offer "end to end" automatic speech recognition. This avoids the need to go through the complex training of three different models while offering better performance in terms of results and processing time. Almost all major digital players now offer their own ASR implementations that can be easily used by API systems, but open-source systems also exist (for example, DeepSpeech[63] or Kaldi[64]).

5.2 Natural Language Processing (NLP)

175. Natural Language Processing is a multidisciplinary field involving linguistics, computer science and artificial intelligence, which aims to create natural language processing tools for a variety of applications. The fields of research and applications are numerous: syntactic analysis, machine translation, automatic text generation and summarization, spell checking, question answering systems, text mining, named entity recognition, sentiment analysis, etc. Concretely, NLP's goal is to give computers the ability to read, understand and derive meaning from human languages. The development of NLP applications is challenging because computer tools traditionally require humans to interact with them in a programming language that is formal, meaning precise, unambiguous and highly structured. Human speech, however, is not always precise. It is often ambiguous

technically compatible. The GDPR does, however, prohibit controllers from establishing barriers to the transmission" - WP29 Guidelines on the right to data portability - endorsed by the EDPB, p. 5.

[62] EDPB strongly encourages cooperation between industry stakeholders and trade associations to work together on a common set of interoperable standards and formats to deliver the requirements of the right to data portability.

[63] https://github.com/mozilla/DeepSpeech

[64] https://github.com/kaldi-asr/kaldi

and the linguistic structure can depend on many complex variables, including slang, regional dialects and social context.

176. Syntax and semantic analysis are two main techniques used with NLP. Syntax is the arrangement of words in a sentence to make grammatical sense. NLP uses syntax to assess meaning from a language based on grammatical rules. Syntax techniques used include parsing (grammatical analysis for a sentence), word segmentation (which divides a large piece of text to units), sentence breaking (which places sentence boundaries in large texts), morphological segmentation (which divides words into groups) and stemming (which divides words with inflection in them to root forms). Semantics involves the use and meaning behind words. NLP applies algorithms to understand the meaning and structure of sentences. Techniques that NLP uses with semantics include word sense disambiguation (which derives meaning of a word based on context), named entity recognition (which determines words that can be categorized into groups), and natural language generation (which will use a database to determine semantics behind words). While earlier approaches to NLP involved rules-based approaches, where simple machine learning algorithms were told what words and phrases to look for in text and given specific responses when those phrases appeared, current approaches to NLP are based on deep learning, a type of AI that examines and uses patterns in data to improve a program's understanding.

5.3 Speech Synthesis

177. Speech synthesis is the artificial production of human speech. Speech synthesis has mainly been implemented by concatenation of vocal units that are stored in a database. This technique consists in selecting, from all the recordings of an actor previously transcribed into phonemes, syllables and words, the bricks of sound that correspond to the words that one wishes to have pronounced by the VVA and to assemble them one after the other to form an intelligible sentence with natural diction. Alternatively, a speech synthesizer can incorporate a model of the vocal tract and other human voice characteristics in order to model the parameters of a voice such as intonation, rhythm, and timbre, by generative statistical models (such as WaveNet[65], Tacotron[66] or DeepVoice[67]) and to create a completely synthetic voice output.

[65] Aäron van den Oord et Sander Dieleman, *WaveNet: A generative model for raw audio*, Deepmind blog, september 2016, https://deepmind.com/blog/article/wavenet-generative-model-raw-audio.

[66] Yuxuan Wang, *Expressive Speech Synthesis with Tacotron*, Google AI blog, March 2018, https://ai.googleblog.com/2018/03/expressive-speech-synthesis-with.html.

[67] *Deep Voice 3: 2000-Speaker Neural Text-to-Speech*, Baidu Research blog, October 2017 http://research.baidu.com/Blog/index-view?id=91.