Computer & Communications Industry Association
Tech Advocacy Since 1972

Draft Guidelines on Virtual Voice Assistants
CCIA Europe comments

Executive summary

The Computer & Communications Industry Association ('CCIA Europe') welcomes the opportunity to comment on the draft guidelines on Virtual Voice Assistants (02/2021) published by the European Data Protection Board ('EDPB').

CCIA Europe represents several developers and distributors of virtual assistant services that typically offer voice and touch user interfaces accessible through a wide range of devices (smartphones, speakers, car infotainment, TV, PCs, smartwatch, etc.).

Generally, CCIA Europe wonders why the draft guidelines focus exclusively on the voice User Interfaces ('UI') of virtual assistant services. In most cases, providers of so-called "Virtual Voice Assistants" ('VVA') provide at least both a voice and touch UI, and the same service can be activated and performed in different ways. The EDPB may wish to review the scope of its guidelines in this respect. At a minimum, we invite the EDPB to consider to what extent the recommendations and conclusions of the guidelines apply to touch-based UI, and consider providing justifications if the EDPB considers that these recommendations should not apply to touch-based UI.

More specifically, we are concerned that the draft guidelines reduce the range of legal bases to process personal data to two grounds: "contract" for processing that is necessary to execute user requests, and consent (including consent from multiple users) for everything else, including service improvement or the provision of new features which users would reasonably expect from VAAs. The draft guidelines appear to justify this restriction on the basis of a mischaracterisation, or perhaps a misunderstanding, of the processing stages involved to activate and perform a virtual assistant service, and an overly broad interpretation of Article 5(3) of the e-Privacy Directive (EPD).

You will find below our detailed comments on the draft guidelines. CCIA Europe also respectfully invites the EDPB to further consult with leading VVA service providers to ensure that the final guidelines reflect an accurate understanding of how VVA technologies work in practice and provide meaningful and workable recommendations to VVA providers.

1. EPD should not apply to VVAs which do not store or access data from end-user devices

Section 2.1 Basic characteristics of Virtual Voice Assistants; Section 3.1 Legal framework; Section 3.8.2 Specific considerations when processing biometric data

Paragraphs 16(2), (3), and (4) and paragraph 28 suggest that VVA services retrieve audio content from user devices and "locally" compare the recording with the wakeword. If they match, the VVA opens a listening channel and the audio content is immediately transmitted. The user's request is then transmitted to the VVA provider. Paragraph 6 also states that VVAs are "switched on by default".

However, most VVA services do not require storing or accessing information from the user device at any point in time, be it at the moment of activation or the processing and performance of the request. In addition, VVA services are only pre-installed on certain devices, but are never "switched on by default" contrary to what Paragraph 6 states.

Activation: it should be clear that VVA services can be activated in different ways e.g. pressing the microphone button on a display, squeezing a phone or using the wake word. Second, whatever the way users choose to activate their VAA, the user sends that information to the VVA cloud computing servers without the VVA service provider gaining access to the device. When using the wake word, the VVA software recognises on-device wakeword and initiates a connection to the VVA server, but this doesn't mean that the server gains access to the user's equipment. By way of comparison, the same process occurs for every user command expression, e.g. when email service users initiate the connection with their email service provider's servers when clicking the send button.

Request recognition and performance: the draft guidelines rightly recognise in paragraph 16 that "most voice related processing is currently performed in remote servers." Indeed, most of the processing operations that are necessary to effectively provide the VVA services, including the matching of voice templates, rely on cloud processing, rather than the limited device processing capabilities. CCIA Europe has therefore reservations on paragraph 133 of the draft guidelines which recommends that "[v]oice templates should be generated, stored and matched exclusively on the local device, not in remote servers." This is not how VVA technology works, nor is it always desirable from a security standpoint. We remind the EDPB that cloud storage of voice templates increases the integrity of users' data in certain circumstances, e.g. if a device is stolen or it changes ownership, there is no risk to expose biometric data stored on the device. Storage in the cloud prevents the unauthorised reading, copying, modification or removal of biometric data should one of those events occur. Furthermore, referring exclusively to device storage of voice template data appears inconsistent with the WP29 guidance on biometrics[1] which confirms that "identification can only be achieved by storing the reference data in a centralised database, because the system, in order to ascertain the identity of the data subject, must compare his/her templates or raw data (image) with the templates or raw data of all persons whose data are already centrally stored".

[1] Available on https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2003/wp80_en.pdf.

CCIA Europe invites the EDPB to clarify that Article 5(3) EPD does not apply to VAA providers which do not store or access data stored from since the directive only governs "the storing of information, or the gaining of access to information already stored, in the terminal equipment of a subscriber or user."

2. Scope of EPD and availability of grounds for processing

a) General comments

Section 3.1 - Legal framework

In paragraph 24, the draft guidelines state that the e-Privacy Directive applies to VVA services to the extent that it "sets a specific standard for all actors who wish to store or access information stored in a terminal equipment of a subscriber or user in the EEA", and that Article 5(3) of said Directive would apply because "current VVA services require access to the voice data stored by the VVA device".

As explained in the section above, most VVA services do not require access to the voice data stored in users' terminal equipment, and we invite the EDPB to clarify that the EPD should not apply to them.

Furthermore, CCIA Europe does not believe that the EPD was intended to apply to the limited scenarios where a VVA does require access to voice data stored on the user device. Indeed, according to Recital 24, Article 5(3) has been construed to protect users' private sphere from so-called "spyware, web bugs, hidden identifiers and other similar devices (...) [which] can enter the user's terminal without their knowledge in order to gain access to information, to store hidden information or to trace the activities of the user". VAAs do not involve the use of spyware, web bugs, hidden identifiers, or other similar devices.

By applying the EPD to processing performed by VVA services, the draft guidelines discount legitimate interest as a ground for processing, only allow the use of contract in very narrow circumstances ("out of the box" execution of voice commands) and require consent for processing which VVA users would normally expect. This appears to ignore user expectations of VVA services and negate the principle recognised in various existing guidelines that the appropriate lawful basis must be determined on a case-by-case basis by the data controller.[2]

Having consent as the only option cannot be the intended effect of applying Article 5(3) of the e-Privacy Directive to the provision of VVA services. Other lawful bases, such as "performance of a contract" or "legitimate interest", provide in most instances for a strong, protective and more appropriate lawful basis. "Performance of a contract" under Article 6(1)(b) GDPR constitutes the appropriate lawful basis, for instance, for service improvement purposes or for certain content personalization purposes, as the EDPB states (par. 85 of the draft Guidelines), because those purposes

[2] E.g. WP29 guidelines on DPIAs endorsed by the EDPB wp248rev.01, page 18; Opinion 06/2014 on legitimate interest also provide that, page 43; and 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

participate in the fundamental objective of a VVA service, namely ensuring that customers can effectively and securely interact with the VVA service.

The pursuance of a legitimate interest can also constitute a valid lawful basis for data processing that is done by VVA service providers. The service provider, for instance, has a legitimate interest in offering an attractive product that continues to improve and add new features over time, just as the customers have a legitimate interest in receiving new desirable features. With appropriate safeguards in place, such as an unconditional and easily accessible way to opt-out from the processing, Article 6(1)(f) GDPR can indeed be an appropriate lawful basis.

CCIA Europe further submits the following recommendations for each specific purpose identified in the draft guidelines.

b) Service improvement

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

According to paragraph 80 of the draft guidelines, processing necessary for service improvement would be subject to consent if it is not strictly necessary to perform the contract and requires consent because "VVAs are already functional when they come out of the box and can already perform as (strictly) necessary for the performance of the contract."

CCIA Europe respectfully disagrees with the premise of the argument put forward by the EDPB. Paragraph 80 ignores the fact that VVAs are dynamic, ever-evolving services that must be able to respond to different linguistic characteristics, accents, speech patterns and take into account terminological evolutions over time. From a user/customer point of view, a VVA service that does not take into account individual characteristics and linguistic evolutions that fail to efficiently respond to their queries is a defective product by all means. That is why we believe at least contract performance should be a valid legal ground to process data for service improvement purposes.

Furthermore, because service improvement is essential for VVAs to perform and truly act as an assistant for each individual user, CCIA Europe wonders how VVA providers can permit users to refuse to consent to processing for such essential features and all the while comply with the conformity requirements laid down in the Digital Content Directive.[3] As a reminder, under the said directive, the service must (i) be supplied as described in the contract,[4] "be fit for the purposes for which digital service of the same type would normally be used,"[5] "possess the qualities and performance features, including in relation to functionality, compatibility, accessibility, continuity and security, normal for digital content or digital services of the same type and which the consumer may reasonably expect, given the nature of the digital content or digital service,"[6] and "is in conformity throughout the duration of a continuous supply."[7]

[3] Directive (EU) 2019/770 on certain aspects concerning contracts for the supply of digital content and digital services
[4] Article 7(a) Ibid. 2
[5] Article 8(1)(a) Ibid. 2
[6] Article 8(1)(b) Ibid. 2
[7] Article 8(1)(d) Ibid. 2

Service improvements are also crucial in ensuring the security of the service. For instance, accuracy improvements help reduce so-called false wakes, where the VVA service starts listening upon erroneously detecting a certain word as a wake word selected by the customer. Providing a secure product is clearly part of the fundamental objective of the contract between the customer and the VVA service provider. Service providers must therefore be able to rely on the performance of the contract under Article 6(1)(b) of the GDPR to learn from so-called false wakes to improve activation accuracy of the VVA service and reduce their recurrence and to rely on legitimate interests where it fits.

c) Profiling

Section 3.4.4 - User profiling for personalized content or advertising

The draft guidelines explicitly state in paragraph 88 that processing for advertising purposes "is never considered as a service explicitly requested by the end-user", and therefore "users' consent should be systematically collected".

As discussed in section 2(a) above, we respectfully dispute the applicability of the EPD to processing performed by VVA providers because most VVAs do not require access to the voice data stored in users' terminal equipment and the processing fall outside the purview of Article 5(3) EPD. For those VVAs that do require access to voice data stored in users' terminal equipment, we remind the EDPB that VAAs do not involve any form of intrusion described in Recital 24 of the e-Privacy Directive e.g. the use of spyware, web bugs, hidden identifiers, or other similar devices.

Instead, the GDPR should be the only applicable framework for VVAs. For this reason, we encourage the EDPB to clarify that processing for profiling purposes is only subject to user consent when it involves automated individual decision-making which produces legal effects concerning the user or similarly significantly affects the user, or where processing special category data for advertising purposes (Articles 9(2) and 22(4) GDPR). Profiling can be performed based on legitimate interests providing that data subjects can exercise their right to object to the processing in line with Article 21(1).

3. Qualifications: "terminal equipment", biometric data, controller-processor

a) VVAs are software services, not "terminal equipment"

Section 3.1 - Legal framework

Paragraph 25 states that "VVAs should be considered as "terminal equipment" and the provisions of Article 5(3) EPD apply whenever information in the VVA is stored or accessed," while paragraph 28 refer to "VVA device".

Equating VVAs to "terminal equipment" is factually and terminologically incorrect, and appears inconsistent with other sections of the draft guidelines.[8]

A Virtual Voice Assistant is a service providing an audio (and touch) User Interface which performs only by way of software commands, regardless of the device used. It is clear from Article 1(1) of Directive 2008/63/EC and English dictionaries[9] that the term "terminal equipment" only comprises physical objects. Like browsers, Operating Systems, or any software application pre-installed on a device, the mere fact that some VVA services may be pre-installed on user devices does not suddenly qualify VVA services as a physical object, or " terminal equipment."

CCIA Europe calls on the EDPB to clarify its position and ensure consistency throughout the guidelines to avoid confusion and prevent unintended and unforeseen regulatory consequences elsewhere.

b) Voice data does not necessarily constitute biometric data

Section 3.1 - Legal framework; Section 3.4.3 - User identification

Paragraph 31 states that "voice data is inherently biometric personal data". Paragraph 81 also appears to infer the same conclusion.

While the data that results after some type of specific technical processing to extract unique characteristics of a natural person that allow or confirm identification of that person might be considered biometric data, only recordings of voices are not biometric data, e.g., a voicemail, is not biometric data in the way that a picture is not biometric data. The guidelines should clarify that the mere processing of voice data does not constitute processing of biometric data.

This is for instance confirmed in the opinion released last year by the EDPB on video devices, in which not all recorded data related to voice or image qualifies as sensitive data as per EDPB's own words. In EDPB's guidance released on 29 January 2020 on video devices (§79) of its 3/ 2019 opinion, the EDPB outlines: "To qualify as biometric data as defined in the GDPR, processing of raw data, such as the physical, physiological or behavioural characteristics of a natural person, must imply a measurement of this characteristics. Since biometric data is the result of such measurements, the GDPR states in its Article 4.14 that it is "[…] 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 […]". The video footage of an individual cannot however in itself be considered as biometric data under Article 9, if it has not been specifically technically processed in order to contribute to the identification of an individual". We would respectfully suggest that, for the sake of consistency and legal certainty, the EDPB applies the same reasoning to voice footage.

---

8 For instance, paragraph 8 states that "*a VVA can be defined as a software application that provides capabilities for oral dialogue with a user in natural language*" and "*A virtual voice assistant (VVA) is a service that understands voice commands and executes them or mediates with other IT systems if needed.*" Paragraph 16(1) also implies a clear and welcome distinction between a VVA and a *piece of equipment* such as a smartphone, loudspeaker, or vehicle.

9 For instance, Collins Dictionary defines the term *equipment* as "*things* which are used for a particular purpose".

c) Controller-processor

The draft Guidelines provide that "[f]rom 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 processor when processing personal data on behalf of other parties, such as an application developer" (par. 42 of the draft Guidelines). We welcome that the EDPB does not adopt a rigid position on the qualification of the different stakeholders involved in the provision of a VVA service. As the EDPB rightfully states, "the applicable qualifications have to be established on a case-by-case analysis" (id.).

However, the draft guidelines appear to rescind its case-by-case approach in the illustrative example of the relationship between a VVA service provider and a bank offering an application. In this example, the draft guidelines determine that the VVA service provider qualifies as the bank's processor without consideration of other potential qualifications and the fact that the VVA service provider may independently determine both the purpose and the means of the processing. For instance, the purpose of processing may consist in interpreting and responding to the customer's spoken requests, and performing VVA contracts with the customer. The VVA provider may also likely determine which data are collected (i.e., the customer's voice input and other relevant personal data) and transferred to the bank. Finally, the service provider may also independently be responsible for complying with its obligations to customers in respect of these processing activities (e.g., informing customers of the processing activities or responding to data subject requests).

By way of analogy, the relationship between a VVA provider and the bank would be more akin to the controller-controller relationship described in the *ExploreMore* example of the guidelines 07/2020 whereby a travel agency (in this case, the VVA provider) sends personal data to the airline (in this case, the bank) in order for them to carry out its service.[10]

We invite the EDPB to reaffirm a case-by-case approach when determining when a VVA provider may act as a controller and when it may act as a processor.

4. Device-level vs user-level consent

Section 3.1 Legal framework

The draft guidelines state in paragraph 29 that consent under Article 5(3) EPD needs to be connected to an individual user (as opposed to relating to a specific terminal equipment).

Where consent should be deemed is the appropriate legal basis, consent "per user" raises severe practical challenges, would most likely require more data than what the service actually needs to perform the service, and appears inconsistent with the wording of Article 5(3) EPD. Instead, CCIA Europe recommends that the guidelines recognise that consent relating to a specific terminal equipment is a more appropriate form of consent.

[10] Page 29 of the EDPB guidelines on the concepts of controller and processor in the GDPR Version 1.0

While Article 5(3) EPD refers to the " terminal equipment of a subscriber or user", the requirement has been interpreted to mean a consent provided by a user of a device is sufficient for a certain period, without the need to obtain a new consent or re-authenticate each time the website is visited from the same device . This means that different members of the same household may use a device and access the same website, without the website operators collecting consent or attempting to identify the relevant user upon each visit. The ePD ought to be applied consistently in order to not create a specific regime per device or service.

The logical consequence of the EDPB's draft Guidelines is that every time a user opens an application or visits a website, they should be re-authenticated to make absolutely certain it is the same user that previously consented. Such a requirement would materially impair the experience of using a VVA. By contrast, and provided that consent is actually required, a device-level consent on a VVA would enable the owner of the device to grant or decline consent on unboxing and setting up the device.

The EDPB recognised in paragraph 111 that these VVA services are unique in that they are multi-user. In addition the EDPB recognises in paragraphs 145 and 146 that log-in mechanisms may not be required for all VVA functionality. Therefore, the EDPB should recognise the value and preference for device level identity and device level consent.

Finally, should consent be the appropriate legal basis, the requirement to obtain consent for each user would require an authentication every time a user uses their VVA. This would hamper user experience and goes against the principles of privacy by design by, in effect, requiring biometric identification each time a voice command is issued. Users issue voice commands to VVA services as a convenience to avoid needing to use a keyboard or touchscreen. Methods of authentication involving passwords or PIN codes would be wholly inconsistent with how users interact with VVA services and, further, many VVA services do not have keyboards or screens. This would essentially mean processing special category data for each and every instruction issued by a user.

5. Processing of non-user voice data

Section 3.1 - Legal framework; Section 3.8.2 - Specific considerations when processing biometric data

According to paragraph 30 of the draft Guidelines, voice recognition may only be activated “at each use at the user’s initiative, and not by a permanent analysis of the voices heard by the assistant.” Paragraph 130 also explains that, when the VVA service is set to permanently listen for the voice of a registered user, “the voice of non-registered and accidental users will also be processed for the purpose of uniquely identifying them”.

This assumption is incorrect: only voices registered to the specific device can be identified. Where companies apply privacy-preserving measures for on-device processing, the voice-prints to identify registered voices are usually saved locally on device and the number of these voice-prints is limited (usually less than 10 different user profiles for multi-user devices like smart speakers). However, VVA providers are not capable of uniquely identifying or singling out non-registered or accidental users, and any voice that is not recognised is simply considered to be a non-registered user/ guest.

CCIA Europe therefore encourages the EDPB to amend its draft guidelines and remove the requirement for VVA service providers to be only activated at each use at the customer's initiative.

6. Noise filtering

Section 3.1 - Legal framework; Section 3.9 - Data minimization

Paragraphs 139 and 140 recommend VVA providers to apply automated background noise filtering to avoid recording background voices and situational information.

We invite the EDPB to remove this recommendation as there is currently no known reliable mechanism to do "background noise filtering" in the context of VVAs.

First, it is technically challenging to isolate the sound of human voices because other noises also happen at the same frequencies. On a spectrogram of speech signal, unwanted noise appears in the gaps between speech and overlapping with the speech. If a person's speech and noise overlap, it is very difficult to distinguish between the two. Instead, it would be necessary to train a neural network beforehand on what noise looks like and speech looks like. However, since Paragraph 80 mentioned above provides that contractual necessity is an inappropriate legal basis for product improvement, it would be difficult to justify reliance on machine learning technology for the training of real-time background noise technology.

Second, the paper cited in footnote 50 is about distorting voice to prevent identification, not about filtering background noise. In addition, the paper's summary cited a 19% drop in ASR accuracy, which is too significant.

Finally, we remind the EDPB that users can already choose to apply noise filtering technologies through their hardware devices, where available.

***

For further information, please contact Alexandre Roure, Senior Manager, Public Policy, CCIA Europe: aroure@ccianet.org