OPEN VOICE NETWORK
Voice worthy of user trust

THE LINUX FOUNDATION

This is one of an introductory overview series of current research, hypotheses, and development methodologies leading to the proposal of global standards for conversational technologies by the Open Voice Network openvoicenetwork.org, an open-source community of The Linux Foundation.

PRIVACY PRINCIPLES AND CAPABILITIES UNIQUE TO VOICE

Executive Summary

This whitepaper by the Open Voice Network Privacy and Security Workgroup is a review of privacy risks and regulations unique to voice technologies. It is written for enterprise decision-makers, voice technologists, and for the awareness and assistance of all users of voice assistance whose privacy is being discarded. The document:

- Provides an overview of voice assistant architecture
- Reviews privacy regulations and identifies the primary privacy issues/risks, i.e., data acquisition and data use, that users face when they are using natural language processing (NLP) technologies
- Details four principles essential to the protection of privacy in a voice strategy — transparency, consent, limited collection and control
- Suggests specific actions for supporting and maintaining user privacy to effectively navigate the data privacy risks and potential harms specific to using voice technology.

For the latest, detailed information on the work of the Open Voice Network, please visit our website at openvoicenetwork.org and the Open Voice Network GitHub Repository at github.com/open-voice-network/docs .

Introduction

Voice Assistants Increase the Stakes for Personal Privacy

As their capabilities continue to improve, voice assistants' popularity is rising. According to a Vixen Labs and Open Voice Network 2021 survey (Open Voice Network, Vixen Labs, & Delineate, 2021), daily usage of voice assistants is now over 30% with weekly usage being almost 50% in the US, UK, and Germany. However, user trust about the collection and control of the personal data collected, used and perhaps shared during voice assistant transactions is a major issue for voice assistant users.

Since the introduction of Apple's Siri consumer voice assistant in 2011, the issue of voice privacy has been consistently cited as a user concern and an impediment to trial and adoption by researchers, consumer advocates, and enterprise decision-makers (Vixen Labs, 2021, p.12). However, the Opus Deepgram State of Voice Technology report for 2022, which surveyed 400 decision-makers, found that 75% expect to increase spending on speech technology solutions in the next 12 months, and 92% thought that widespread use of voice-enabled experiences is less than five years away.

A new dimension to existing privacy rights discussions

Consumers and innovators require protection not just for the information they provide during a voice transaction, but also for the unauthorized use of the rich personal information present in their voices in three main areas: *Biometric data* from voiceprint analysis can identify an individual; *biomarkers*, garnered from voiceprint analysis against a baseline, can be used to diagnose a growing list of physical and mental ailments, as well as general indicators of height/weight/body mass; and by conducting a comparative analysis of voices against a substantial database, a form of *cohort inference* can be used to derive information on gender, ethnicity, region, level of education, and household income.

Further compounding the risk of sharing highly personal information while using a voice application is that human voices can be overheard, misunderstood, or impersonated without any physical contact. As this paper makes clear, voice privacy is an issue that goes well beyond the "*it's always listening"* concerns of smart speaker users. The primary privacy issues explored in this paper are data acquisition and data use, and the elements of a privacy-preserving approach that is consistent with global regulation and legislation, and scalable to the inevitable advances in voice technology.

By addressing critical topics, such as Transparency, Consent, Limited Collection, and Control, we seek to minimize unintentional harms, such as systemic bias and privacy violation, and help to narrow, or close, the trust gap that exists between voice assistants and their users. The principles outlined in this paper are not exhaustive. However, the Open Voice Network believes that the recommended principles and plan of action contained within this paper are a solid first step toward raising awareness and addressing voice-specific data privacy concerns and questions with standards as the industry evolves.

Table of Contents

A Common Understanding of Voice Assistants: Architecture and Definitions

Voice assistants are part of human-to-machine interactions, as opposed to a human-to-human telephone exchange. Known by several names, such as "voice-enabled devices," "conversational agents," and the more common "voice assistant," their voice-activated software performs various tasks and can act both as a platform for voice-enabled applications and a user interface.

> Typically, voice assistants have several core components: One that transcribes the user speech (Automated Speech Recognition, or ASR), one that understands the transcribed utterances (Natural Language Understanding, or NLU), one that makes decisions (Decision-Making, or DM), and one that produces the output speech (Text to Speech, or TTS) (Attwater, 2022).

Currently, voice assistants are found in smartphones, wearables, smart home hardware and are embedded in vehicles, public infrastructure, and Internet of Things (IoT) products. Voice-enabled environments is a broader term for areas that contain hardware devices such as smart speakers or smartphones that are visible to the user. This also includes IoT smart devices that feature a speaker and microphone but may not be visible or obvious to a user. The hands-free option in a voice-enabled environment can accommodate accessibility issues and enhance use cases such as serving as a language translator or helping people who literally have their hands full — in a factory, office, vehicle, at home, or walking down a busy street.

For many voice assistants, especially those that aspire to conversational-level complexity, one or more of these components resides in a proprietary cloud. However, many non-platform providers of voice assistance are now developing localized, edge-based architectures for voice assistants for dialogues of limited complexity and greater data control and privacy.

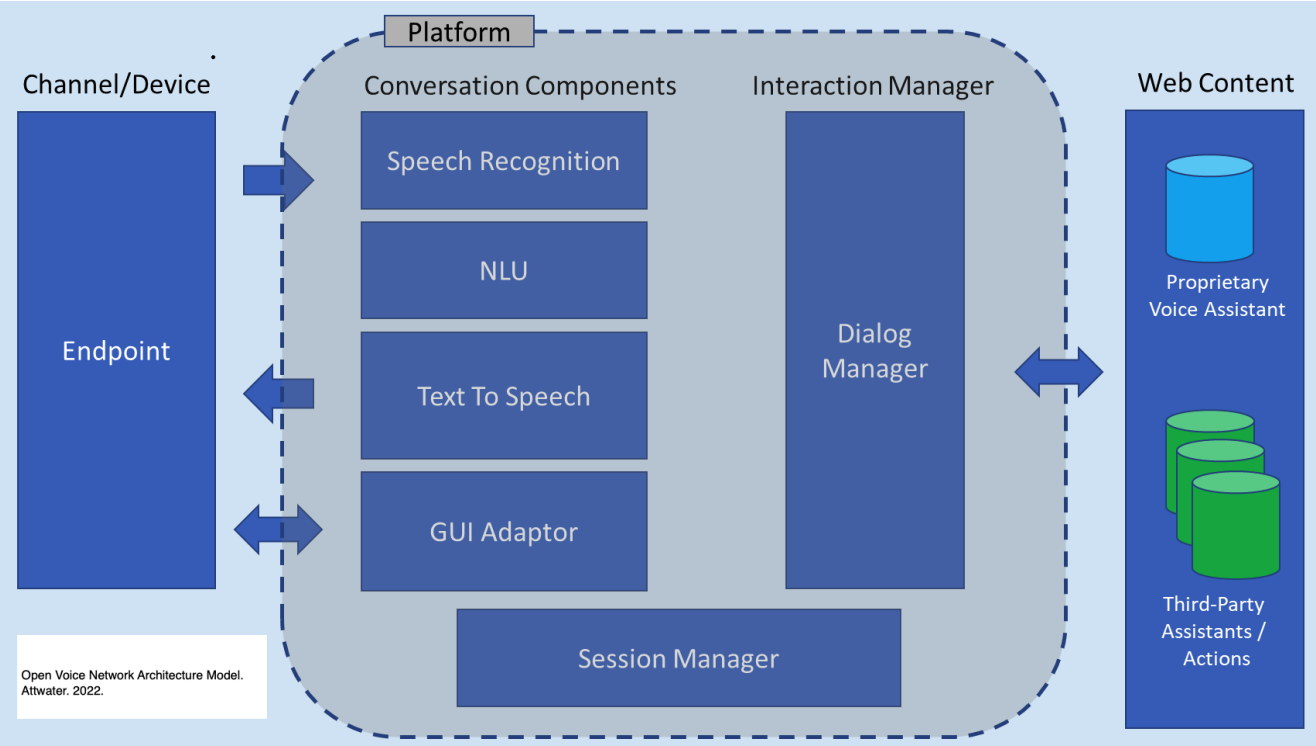

Figure 1: Voice Architecture Model Image: Attwater. Open Voice Network Architecture Committee. 2022.

The Open Voice Network Working Definition of Privacy

"Privacy" lacks a standard global definition — academics have compared it to "usability" and "security," holistic properties of interactive systems (Iachello & Hong, 2007).

User privacy can be viewed from two perspectives:

1. Privacy issues unique to voice-to-machine environments.
2. Privacy risks between users and third parties when human speech is the primary conduit in voice-to-machine environments:
    i. An individual owns their own voice data unless specifically stated otherwise.
    ii. Voice data is classified as personal information, subject to personal information laws and regulations.

If you'd like to learn more about the ethical implications of voice user privacy, see our paper "Ethical Guidelines for Voice Experiences: A Case for Inclusivity and Trustworthiness". Please visit our website at https://openvoicenetwork.org.

Privacy Regulation — A Comparative Overview

Voice strategies are defined, regulated, and implemented differently in the United States than in the European Union (EU). Well-established doctrine addresses several aspects of personal privacy and digital technology.[1] However, the doctrine surrounding human voice data privacy in voice-to-machine environments is nascent and untested in court.

The EU GDPR is closely aligned with the Open Voice Network privacy principles and implementation with some gaps as noted below.

United States Personal Data Privacy Example

The United States does not have a comprehensive federal data privacy law. All US states have data breach notification rules, but only a few such as California (California Privacy Act, 2018) and CPRA (California Privacy Rights Act, 2020), Virginia CPRA (Virginia Consumer Data Privacy Act, 2021), and Illinois PIPA (Personal Information Protection Act, 2020) have passed data privacy laws. A handful of states (BCLP, 2021) have taken a different tact - passing biometrics statutes that govern the handling and storage (e.g., voice prints). In the US, data privacy is included in regulations, but the Open Voice Network privacy principles **Transparency**, **Consent**, **Limited Collection**, and **Control** are not uniformly addressed.

European Union Personal Data Privacy Example

In contrast, the EU emphasizes the primacy of individual rights (EU Convention on Human Rights [61, Article 8]). EU policy for online identity matters as much as offline identity. The EU European Data Protection Board (EDPB) ensures General Data Protection Regulation (GDPR) and the Data Protection Law Enforcement Directive are consistently applied across all 27 member countries. In 2021, the EDPB issued two guidelines that specifically addressed voice and privacy:

1. "Virtual Voice Assistants (VVA) v2 (European Data Protection Board, 2021)". The EDPB applied GDPR to virtual voice assistants – which was, from the Open Voice Network perspective, the first full review of voice and privacy. However, the paper failed to address a number of forward-looking issues.
2. "Sector Inquiry into Consumer Internet of Things (European Commision, 2021) ." The IoT paper noted the gatekeeper and data acquisition role played by proprietary Big Tech voice platforms, and the economic implications of that role.

[1] In the US, Federal Wiretap Act prohibits interception of or eavesdropping on electronic communications. The Stored Communications Act protects accessing stored electronic communications without authorization.

The Open Voice Network Voice Assistant Privacy Principles

The Open Voice Network has identified four key principles to govern human voice data privacy strategies:

Transparency, Consent, Limited Collection, and Control.

The Open Voice Network asserts:

- User privacy is a foundational element in the establishment and maintenance of trust.
- Trust is crucial for the continued adoption and maturation of voice-enabled environments.
- User privacy can be viewed from two perspectives:
    1. Privacy issues unique to voice-to-machine environments.
    2. Privacy risks between users and third parties when human speech is the primary conduit in voice-to-machine environments:
        i. An individual owns their own voice data unless specifically stated otherwise.
        ii. Voice data is classified as personal information, subject to personal information laws and regulations.

The Open Voice Network Privacy Problem Statement

What is the problem that needs to be solved?

The personal privacy of voice users is being discarded, often to enable what author, Shoshana Zuboff, the Charles Edward Wilson Professor Emerita at Harvard Business School, identified as Surveillance Capitalism, the commodification of personal data with a core purpose of making a profit. And yet, if the number accepting Apple's option to opt out of surveillance is any indication, a majority seem to agree that privacy is an issue and informed privacy is a goal. The growing development ecosystem for voice spotlights privacy concerns for enterprise and individual users. Currently, there are 1) no industrywide guidelines that inform consumers as to expectations, 2) no industrywide guidelines for developers, and 3) no industrywide protocols or processes that will allow users to choose the extent to which the a) personal data of their voice, b) the destination of their conversation, and c) the content of the conversations will be shared with providers of voice services. The process to change privacy settings is not standardized across voice assistants, is time consuming, and an administrative burden resulting in a diminished user experience. There is a trend for the implicit agreement of users to give personal data in return for perceived personalization services, information, and advantages.

Why is it a problem?

This problem violates human privacy and ethical use expectations, can potentially harm certain individuals through unauthorized use of highly personal information, and causes a generalized lack of trust in privacy protection efficacy for all voice assistants. Despite these potential harms and risks, the global voice interface market is predicted to grow at a 21.5% compound annual growth

rate (CAGR) between 2021 and 2030.[2] However, public opinion is shifting as voice users understand they are unable to easily audit or modify how their personal information is collected, protected, or used by third parties.

Where is the problem observed?

The problem is in every voice-enabled environment. Problem severity varies depending on the user's interest and knowledge of voice-enabled environment device limitations, their needs, circumstances, and other factors that may affect audio-signal-to-noise ratio for voice assistants. Generally, it is cumbersome for users to control, access, and manage their own personal privacy or their legal dependents when using voice assistants. The default privacy settings on voice assistants in voice-enabled environments are ambiguous and can often be confusing for the user to understand and navigate effectively (Fowler, 2018). Users are frustrated when they are unable to consent to partial terms and conditions for themselves or their legal dependents when read by a voice assistant.

Who is affected?

This problem affects all active voice assistant users and potentially people speaking nearby as it reduces control over personal information for the user, people speaking nearby, and third parties.

When was the problem first observed?

The lack of control over privacy for voice-to-machine interactions has existed since its invention. Voice-enabled assistants gained public attention in 2020 during the COVID-19 pandemic. They provided a touchless alternative means for COVID-19 screening (Chu, 2020). They were also used to detect physical and neurological issues and as a biometric means of authenticating people and identifying imposters. Their rise in popularity also opened them to press scrutiny which reported on voice privacy harms in consumer e-commerce and voice-enabled vehicles. In 2021, the GDPR published two guidelines that specifically called out voice-enabled environment privacy concerns.

How is the problem observed?

The problem is observed by user complaints, posted in social media channels, reported to the FTC, and vendors, about the difficulty in navigating system-setting menus, understanding terms, and adjusting voice assistant privacy policies.

How often is the problem observed?

The problem is encountered when initially setting up a new voice assistant or any time it is rebooted to factory settings. It also is observed when a user password or credit card account number associated with a voice assistant is updated. The problem may also surface belatedly after a case of bias or discrimination is identified due to a relaxed or insufficient voice assistant privacy policy.

[2] Allied Analytics LLP, November 2021.

Enterprise Privacy Policy Review

Our comparative study analyzed the privacy policies of major North American and European voice assistant platform providers and technology firms currently working with NLP-enabled technologies in some capacity. Major firms ("data processors") included in the study were Amazon, Google, Apple, Microsoft [an Open Voice Network sponsor], and Facebook. This study was completed to gather information and understand the terms and conditions of data collection and the policies pertaining to the sharing and use of personal information. Major conclusions drawn from the study include but are not limited to:

- Data processors collect vast amounts of personal data from consumers.
- Data processors collect vast amounts of personal data from consumers under the protection of "implied consent."
- Ambiguous language allows each data processor a high degree of limited liability when it comes to ethical data collection, data use, and data privacy.

Voice Privacy Is a Multiplayer Game

There are many stakeholders in privacy specific to voice assistance. A reason to establish privacy guidelines—and, in time, technology standards—for voice assistants is to privacy protection in voice assistants requires *the collaborative participation of multiple stakeholders:*

- The user, or data owner — the individual creating data through spoken utterances.
- The conversational agent — a presence, most often with a unique persona, with which the user converses.
- The provider of the conversational agent — the combination of components that enables the operation and management of one or more conversational agents, and which may reside (in whole, or in parts) on-premises, at the edge, or in the cloud.
- The provider of conversational sub-agents — agents accessible only through a specific platform (often known as "Skills" or "Actions").

The need to collaborate points to a need to establish privacy guidelines and, in time, technology standards for voice assistance.

Within a given dialogue — even a simple user request and yes-no response from the provider — raw and processed data will most likely flow from the user utterance through the conversational agent and subagent to and through the conversational platform. It is possible — and increasingly likely — that the provider/owner of the agent or subagent will be different from the provider/owner of the conversational platform.

Voice Assistant Privacy Risks and Related Interactions

The approach taken by the Open Voice Network to develop privacy principles and capabilities involved the steps described below:

1. Review current privacy principles, policies, and regulatory ecosystem.
2. Evaluation of user privacy risks and threats/harms relevant to voice assistants.
3. Exploration of voice-specific stages to determine voice privacy concerns and risks.
4. Definition, discussion, and prioritization of voice privacy risks.
5. Identification and discussion of voice privacy values, guidelines, and required technical capabilities.

The Open Voice Network identified two classes of privacy risks. The first class of risk relates to personal data collected during a voice assistant conversation without consent from the user. The second class of risk is personal data inadvertently revealed to third parties during confirmation and playback by a voice assistant which may be used/shared in ways that may cause harm to the consumer.

At a minimum, all voice assistants include microphone and speaker components to listen and speak with users. To be responsive to a user, voice assistants continuously monitor ambient acoustic patterns to detect when the wake word has been spoken.

Privacy Risk (1): Personal data collected during a voice assistant conversation without consent from the user that may cause harm to the user.

1. Voice assistant devices continuous listening and/or record raw data that is concurrently processed to provide user interaction, services, and conversation

2. Voice assistant devices process raw data through AI-powered models to infer user actions and intents

3. Voice assistant audio data collection usually occurs in the most sensitive of places such as the home where privacy expectations are highest.

4. Voice assistant technology involves a broad scope of data collection with raw data recordings that may contain all manner of information – from background noise and audio context, to biometrics, biomarkers, emotion, dialects, and sentiment – that may not be intended to be shared. Explicit information is also collected, such as PII, account and payment information, personal preferences

5. Voice assistant technology may inadvertently include non-consenting and vulnerable populations in data collection (e.g., children, disabled persons, elderly persons).

6. Voice assistant technology involves multiple parties collecting, creating, and sharing data — and the channel — at the same time.

7. Voice assistant technology involves multiple devices in the same environment; potentially listening to, interacting with, and/or controlling each other, and/or controlling all manner of sensitive devices (eg., a vehicle, medical device, environmental systems).

8. Voice assistant technology privacy policies are difficult to communicate and clarify effectively over TTS playback.

Privacy Risk (2): Personal and sensitive data is inadvertently revealed to others during confirmation and playback and shared and used in ways that may cause harm to the consumer.

9. Voice assistant technology involves voice device responses for confirmation and playback

Recommended Principles to Address Voice Assistant Privacy

Transparency, **Consent**, **Limited Collection**, and **Control** are the four key principles identified by the Open Voice Network to govern human voice data privacy strategies.

1. Transparency

1.1 A voice-initiated user interface (VUI) must be made easily accessible and readily available to the consumer for the purpose of providing user notification of data collection practices, general data usage, and data-sharing policies.

Example: A standard invocation phrase and application service is available on a voice assistant, and when invoked, the privacy policy notification is communicated to the user, and user consent is audibly obtained. (e.g., User: "OK Service, what is your privacy policy?" Device: "Privacy policy is . . . Do you consent to this policy? User: "Yes").

1.2 A VUI must be made easily accessible and readily available to the consumer for the purpose of providing user notification of general data processing and AI inference routines e.g., if there is an AI inference routine for emotion, this should be disclosed in the notification.

Example: As a subroutine of the privacy policy notification, a voice assistant asks if the user would like to know what inference routines are utilized. The User responds, "yes," and the voice assistant notifies the user of the voice inference routines being used (e.g.,"voice speech-to-text only." But if the voice assistant is also doing sentiment analysis, the response would be "voice speech-to-text and voice sentiment analysis").

1.3 Privacy policy notifications should be adapted for text-to-speech playback, as well as made available in other modes that are accessible and adapted to the users and devices involved.

Example: When a voice assistant is asked "What is your privacy policy?" the response should be tailored to text-to-speech playback, including menu navigation and smaller segments delivered interactively; Device: "Our privacy policy has multiple sections; 1) Information collected; 2) Information usage; 3) Information sharing; 4) How to review and request changes to your information; … Please choose which option you would like to hear the details about."

2. Consent

Consent is defined as a "voluntary agreement to or in acquiescence in what another proposes or desires" (Oxford English Dictionary, 1989).

2.1 Data subjects ("users") must be allowed to give explicit, unambiguous consent before the collection and processing of personal data.

Example: With the initial use of any voice assistant, the voice assistant proactively communicates the privacy policy notification, addressing the collection and processing of personal data, and then audibly captures user consent prior to any collection and processing of personal data.

2.2 A VUI must be made easily accessible and readily available to the consumer for the purpose of accepting explicit consent for data collection, usage, and sharing policies.

Example: A standard invocation phrase and application service is available on a voice assistant, and when invoked, the privacy policy notification is communicated to the user, and user consent is audibly obtained. (e.g., User: "OK Service, what is your privacy policy?" Voice Assistant: "Privacy policy is . . . Do you consent to this policy? User: "Yes").

2.3 When there is an explicit third-party provider request (i.e., a request naming a specific party other than the platform provider) data subjects must be allowed to give explicit, unambiguous consent before the collection and processing of personal data by the third party.

Example: User: "OK Service, please connect me to XYZ company service provider to service a request." Voice Assistant: "Privacy policy of XYZ company is . . . Do you consent to this policy? User: "Yes"

2.4 Data must not be used for voice inference training without consent from the user.

Example: Upon initial use the voice assistant asks for voice training for invocation words and commands, and explicitly notifies the user it is in training mode and asks for permission for using the data captures for training. All other interactions that may be used for training are treated similarly with explicit notification and ask for permission before data is retained and/or used for training.

2.5 Increased protections should be afforded to children and vulnerable groups not capable of making their own privacy decisions. Reasonable steps should be taken to verify the age and cognitive abilities of the user. Parental (or guardian) consent is required before collecting, using, or sharing personal information of a child under 16 (United Nations, 1990).

*Example: Upon creation of a user account by a minor, the voice assistant obtains verifiable parental consent. Confirmation of parental consent is obtained by providing knowledge-based challenge questions that would be difficult for someone other than the parent to answer. (*Privo, 2021.)

2.6 Users need a voice "privacy policy navigator" for (1) listening to fragments of the privacy policy, and (2) locating answers to specific questions about the privacy policy. In addition to the start/stop/fast forward/rewind commands for listening to audio files, a privacy policy navigator should present the relevant sections of the privacy policy in response to verbal requests.

Example: "What is the policy on sharing my information with others?" or, "How may I correct an error in the information you have about me?"

3. Limited Collection and Use

3.1 The collection and analysis of raw and processed voice data — beyond that necessary for immediate dialog functionality — should be used only for its stated and creator consent-given purpose.

Example: Voice providers that want further analysis of raw and processed voice data must provide "simple" transparency (as defined above) as to the intent of the analysis, and request simple, transparent consent to that use.

If no additional data analysis is desired or required by the provider, the voice assistant, once raw audio capture and Automated Speech Recognition module (ASR) is completed, must delete the raw audio capture data (once it has successfully been passed to the ASR module).

3.2 Voice data should be stored by providers no longer than is necessary for its stated and creator consent-given analytical use.

Example: An estimate of time requirement must be included in the "simple" transparency provided by providers to creators for the purpose of obtaining consent. Should voice providers desire more time, additional consent must be requested and obtained.

3.3 Providers should collect *the minimum amount* of raw and processed data necessary for the stated and consent-given purpose.

Example: Providers are expected to identify the intent of desired data analysis within the "simple" transparency and consent request. Should research goals change, additional consent must be requested and obtained.

3.4 Providers must *minimize the number of individuals and provider partners* granted access to raw and processed data.

Example: Raw and processed voice data must be maintained in a secure environment and transmitted through secure methods; only those with direct responsibilities for data analysis should be granted access to the secure environment. Provider partners must not be granted access to creator data without "simple" transparency and request for consent.

4. Control

Control, in the context of this paper, is defined as the ability of the user to easily access, rectify, suppress, limit, oppose, and transport their data.

4.1 Any data that is not deleted must be accessible by the user with the ability to review, correct, or securely delete their data.

Example*: The voice technology platform has an interface that allows the user to review all collected and stored data, providing capabilities for the user to review, correct, or securely delete their data.*

4.2 Voice confirmation and playback routines should only include information provided in a related request; if additional information is necessary, explicit prompting and explicit affirmation must occur before confirmation and playback of this information.

Example*: A user provides information to a voice assistant to refill a prescription with a reference to a prescription number. The voice assistant confirms the request but does not playback the name or description of the prescription medicine unless the user requests this explicitly.*

4.3 The consumer has the right to portability of voice data with the right to reuse it for their own purposes. The voice technology platform must provide the consumer at their request any data that is held by the platform, in a structured, commonly used, and machine-readable format. The consumer is free to either store the data for personal use or to transfer it to another entity (Irwin, 2020, June 9).

Example*: The consumer desires to transfer their prescription to a new pharmacy. To transfer the prescription for continued refills, rather than providing the same information to the new provider, data portability makes it easy to provide relevant information to the new provider's voice assistant.*

4.4 Data collection across multi-platform, multi-device requires the consumer to have control and to grant consent regarding the legitimate collection and processing of data on each platform.

Example*: A user provides information to a voice assistant to begin travel arrangements to an international conference. To complete booking of the trip requires a visa. Connection to the visa voice assistant to begin the visa application process requires the user's consent to share the necessary personal data between the travel and visa agents. As proof of negative COVID-19 test is required three days prior to the trip, connection to a COVID-19 test scheduler is made which will require collection of sensitive medical data during pre-screening.*

4.5 The provision of transparency and the need to obtain consent is an ongoing process. A blanket, one-time consent is generally unacceptable.

Example*: Though a fully enunciated legal consent agreement is always required, users must be presented with a consent request that is clearly and simply stated. Such a request must state what*

is being affected ("your voice data") for what purpose ("for building a better data model for speakers of your language") with access by whom ("our speech scientists") for what period ("an estimated 60 days") after which the data will be erased. The request will then guide the user to the full legal consent agreement

Recommended Plan of Action: Privacy Principles & Capabilities

Along with identifying and recommending the aforementioned privacy principles that address voice-specific privacy risks, this white paper also lays the groundwork for a more comprehensive plan of action. This plan of action outlines a set of resources and next-step action items to further support the Open Voice Network's privacy principles and provide enterprise decision makers with a set of resources and tools to help them effectively navigate the privacy risks facing voice assistance. The Open Voice Network recognizes that there is still a great deal of work ahead concerning voice-specific data privacy, but we believe the following plan of action provides a solid framework to begin this important work.

The Open Voice Network recommends the following:

Develop a Privacy Policy Template

An effective voice assistance-specific privacy policy template can be used by any enterprise or independent voice assistant designer or developer that addresses voice-specific privacy risks. The template will be used as a framework for how voice assistant designers and developers inform users of their data collection policies, data use policies, and informed consent practices that adhere to the Open Voice Network privacy principles.

Privacy Legislation Review

This review shall include an overview of consumer-focused privacy regulations in North America, Europe, and Asia. The review will also include guidance for regulatory officials on how to incorporate voice assistance in the broader scope of consumer data privacy. The Open Voice Network has had discussions with data privacy-focused organizations such as the World Wide Web Consortium (W3C), Future of Privacy Forum, Ecommerce Europe, and the Stanford University Open Virtual Assistance Lab (OVAL).

Establish Open Voice Network Privacy Certification

The Open Voice Network Privacy Certification will allow enterprise voice assistant platform providers the ability to self-certify that their voice assistant platform meets the Open Voice Network privacy standards for voice assistance. Future deliverables will include defining an official certification body to administer and govern the privacy certification process.

Propose Open Voice Network Privacy Standards

This resource will summarize the future state of the Open Voice Network's voice-specific data privacy work. This summary will include a listing of proposed standards, guidelines, and action items that enterprise and independent voice assistant platform providers should make note of as they develop future voice assistant technologies and platforms.

Call to Action

Raise awareness about voice-specific data privacy and continuously educate yourself and others. Global consumer-focused privacy regulations in North America, Europe, and Asia are ever evolving. Help advocate for changes to bridge the gap of adoption of voice technologies due to lack of user trust.

Continuously monitor and assess the state of voice-specific privacy policies and guidelines within your organization. Review practices and policies about data acquisition and data use to measure adherence to our principles of transparency, consent, limited collection, and control.

Get involved with the Open Voice Network. The Open Voice Network will continue to monitor and update our guidance, and we hope you continue to monitor our guidelines and provide feedback as well. Provide feedback on the work that we are doing, share our work with others to amplify our message, attend our events, join our committees and workgroups, and reach out to collaborate with us at openvoicenetwork.org or connect with our ambassadors listed there.

About the Open Voice Network

The Open Voice Network (OVON) is a non-profit industry association dedicated to the development of standards for voice assistance transparency, consent, limited collection, and control of voice data that will make using voice technology worthy of user trust. In any reality, virtual or otherwise, we believe personal privacy should be respected as the default. The Open Voice Network operates as an open-source community within The Linux Foundation. It is independently funded and governed with participation from more than 120 voice practitioners and enterprise leaders from 12 countries.

The Open Voice Network community's work is open source. We seek inclusive input and like to share our insights. At present, our work is focused in four areas:

- **Interoperability**, defined as the ability for conversational agents to share dialogs (and accompanying context, control, and privacy),

- **Destination registration and management**, the ability of users to confidently find a destination of choice through specific requests, and for the providers of goods and services to register a verbal "brand" — similar to the Domain Name System (DNS) of the internet;
- **Privacy**, with voice-specific guidance for both the protection of individual user data and that of commercial users; and
- **Security**, with a focus on voice-specific threats and harms.

Please see our papers in 2022 and support the Open Voice Network by visiting openvoicenetwork.org.

About The Linux Foundation

Founded in 2000, The Linux Foundation is supported by more than 1,000 members and is the world's leading home for collaboration on open-source software, open standards, open data, and open hardware. Linux Foundation's projects are critical to the world's infrastructure including Linux, Kubernetes, Node.js, and more. The Linux Foundation's methodology focuses on leveraging best practices and addressing the needs of contributors, users, and solution providers to create sustainable models for open collaboration. For more information, please visit us at linuxfoundation.org.

Acknowledgements

Authored by the Open Voice Network with special thanks to the Privacy and Security Work Group of the Technical Committee contributors Peter Bentsen, Maria Brinas-Dobrowski, Oita Coleman, Ali Dalloul, Jonathan Eisenzopf, Mike Frazzini, John Iwasz, Roger Kibbe, Jim Larson, Maarten Lens-Fitzgerald, Brenda Leong, Lawrence Lin, Nick Myers, Michael Novak, Sara Oliver, Brian Owen, Shyamala Prayaga, Doug Rogers, Jon Stine, and John Trammell; edited by Janice Mandel.

Appendix A. Vocabulary Terms

- **Conversational Access Point:** A physical or virtual means of carrying a signal to a conversational platform.
- **Conversational Platform:** The combination of components that enables the operation and management of one or more conversational agents. The components may include AI-powered NLU, NLG, and dialog management.
- **Conversational Agent:** Is perceived by users to be a single conversation actor. It uses the infrastructure of the conversational platform, and one or more conversational capabilities to hold user conversations. It has continuity of knowledge and bounded identity, and a name by which it is addressed.
- **Conversational Sub-Agent:** A conversational agent that can only be invoked through another agent. It performs a delegated task on behalf of another conversational agent. Sub-agents have a name and their own discourse context. Sub-agents are often referred to as "skills" or "actions."
- **Conversational Capability:** Provides dialog functions that inform an agent. Capabilities do not have a defined name or voice, nor the continuity of discourse context.

Reference List

Attwater, "Open Voice Network Architecture Model" as adopted by the Open Voice Network Architecture Committee, 2022.

Bryan, Cave, Leighton, and Paisner. (2021). U.S. Biometric Laws & Pending Legislation Tracker. https://www.bclplaw.com/en-US/insights/us-biometric-laws-and-pending-legislation-tracker.html?utm_source=Mondaq&utm_medium=syndication&utm_campaign=LinkedIn-integration

California Consumer Privacy Act. (2018). *TITLE 1.81.5. California Consumer Privacy Act of 2018 [1798.100 - 1798.199.100].* https://www.oag.ca.gov/privacy/ccpa

California Privacy Rights Act. (2020). *Section 3, Title 1.81. 5 of the CCPA, added to Part 4 of Division 3 of the California Civil Code. [3] § 1798.185(a)(1)-(2), (4), (7) . [4] § 1798.140(c).* https://thecpra.org/

Chu, J. 2020, October 29. *Artificial Intelligence Model Detects Asymptomatic Covid-19 Infections through Cellphone-recorded Coughs.* MIT News. https://news.mit.edu/2020/covid-19-cough-cellphone-detection-1029

European Commission. 2021. Sector Inquiry Into Consumer Internet of Things. https://ec.europa.eu/competition-policy/system/files/2021-06/internet_of_things_preliminary_report.pdf

European Data Protection Board. 2021. Guidelines 02/2021 on Virtual Voice Assistants Version 2.0.

https://edpb.europa.eu/system/files/2021-07/edpb_guidelines_202102_on_vva_v2.0_adopted_en.pdf

Fowler, J. (2018, June 1). *Hands off my data! 15 default privacy settings you should change right now.* The Washington Post. https://www.washingtonpost.com/news/the-switch/wp/2018/06/01/hands-off-my-data-15-default-privacy-settings-you-should-change-right-now/

Iachello, Giovanni & Hong, Jason (2007). Foundations and Trends in Human-Computer Interaction Vol. 1, No. 1, pages 1-137, *End-User Privacy in Human-Computer Interaction*. https://www.cs.cmu.edu/~jasonh/publications/fnt-end-user-privacy-in-human-computer-interaction-final.pdf

Illinois Personal Information Protection Act. 2020. *P.A. 815 ILCS 530/ 94-36, eff. 1-1-06*. https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=2702&ChapterID=67

Irwin, L. 2020. *The GDPR: Understanding the right to data portability*. IT Governance. https://www.itgovernance.eu/blog/en/the-gdpr-understanding-the-right-to-data-portability

Markets and Markets (2020). Voice Assistant Application Market with COVID-19 Impact by Component, Deployment Mode, Organization Size, Channel Integration (Websites, Mobile Applications), Application Area (Smart Banking, Connected Healthcare), and Region - Global Forecast to 2026. https://www.marketsandmarkets.com/Market-Reports/voice-assistant-application-market-141810993.html

Open Voice Network, Vixen Labs, & Delineate (2021). *Voice Consumer Index 2021.* https://vixenlabs.co/voice-consumer-index

Oxford English Dictionary, Second Edition. 1989. https://www.oed.com/oed2/00047775#:~:text=1

Privo. 2021. *What is Verifiable Parental Consent*. https://www.privo.com/blog/what-is-verifiable-parental-consent

United Nations. (1990). *Convention on the Rights of the Child.* https://www.ohchr.org/EN/ProfessionalInterest/Pages/CRC.aspx

Virginia Consumer Data Protection Act. (2021). *Title 59.1 a chapter numbered 52, consisting of sections numbered 59.1-571 through 59.1-581, relating to Consumer Data Protection Act.* https://lis.virginia.gov/cgi-bin/legp604.exe?212+ful+HB2307ER

Licensing and Attribution